TAYLOR & FRANCIS
Founded 1798

Pharmaceutical Formulation Development of Peptides and Proteins

edited by

Sven Frokjaer and Lars Hovgaard



Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 1
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

Pharmaceutical Formulation Development of Peptides and Proteins

Pharmaceutical Formulation Development of Peptides and Proteins

Edited by
SVEN FROKJAER AND LARS HOVGAARD



Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 3
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

First published 2000 by Taylor & Francis
11 New Fetter Lane, London EC4P 4EE

Simultaneously published in the USA and Canada
by Taylor & Francis Inc
325 Chestnut Street, 8th Floor, Philadelphia PA 19106

Taylor & Francis is an imprint of the Taylor & Francis Group

© 2000 Taylor & Francis Limited

Typeset in Times by Graphicraft Limited, Hong Kong
Printed and bound in Great Britain by TJ International Ltd, Padstow, Cornwall

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

Every effort has been made to ensure that the advice and information in this book is true and accurate at the time of going to press. However, neither the publisher nor the authors can accept any legal responsibility or liability for any errors or omissions that may be made. In the case of drug administration, any medical procedure or the use of technical equipment mentioned within this book, you are strongly advised to consult the manufacturers' guidelines.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
A catalog record for this book has been requested

ISBN 0-748-40745-6

Contents

List of figures *page* xi
List of tables xiii
Contributors xv
Preface xvii

1 Peptide Synthesis 1

Bernard A. Moss

1.1 Introduction 1
1.2 Chemical synthesis of peptides 2
1.2.1 Solution and solid phase peptide synthesis 4
1.2.2 Large-scale peptide synthesis 6
1.3 Concluding remarks 10
References and additional sources 10

2 Basics in Recombinant DNA Technology 12

Nanni Din and Jan Engberg

2.1 Introduction 12
2.2 General methods in gene technology 13
2.2.1 DNA cloning tools 13
2.2.2 Cloning of cDNA 15
2.2.3 PCR cloning and DNA database mining 16
2.3 Expression of recombinant proteins 18
2.3.1 Transcription, translation and protein modifications 18
2.3.2 Choice of expression system 19
2.4 Protein design 22
2.4.1 Protein variants 22
2.4.2 Protein chimeras 24
2.4.3 Epitope display libraries 24

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 5
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

2.5 Recombinant protein therapeutics – status and future trends 25
References 27

3 Protein Purification 29

Lars Hovgaard, Lars Skriver and Sven Frokjaer

3.1 Introduction 29
3.2 Fractionation strategies 30
3.2.1 Initial fractionation step 30
3.2.2 Intermediate purification step 31
3.2.3 Final polishing step 32
3.2.4 The finished product 32
3.3 Protein stability in downstream processing 33
3.3.1 Protein conformation stability 33
3.3.2 Protein instability 34
3.3.3 Essential process-related parameters 37
References 38

4 Peptide and Protein Characterization 41

Miroslav Baudyš and Sung Wan Kim

4.1 Introduction 41
4.2 Chromatography 43
4.2.1 Reversed phase chromatography 43
4.2.2 Hydrophobic interaction chromatography 46
4.2.3 Ion-exchange chromatography 47
4.2.4 Size-exclusion chromatography 47
4.3 Electrophoresis 48
4.3.1 Gel electrophoresis 48
4.3.2 Two-dimensional gel electrophoresis 50
4.3.3 Capillary electrophoresis 50
4.4 Structural characterization 51
4.4.1 Primary structure 52
4.4.2 Mass spectrometry 53
4.5 Secondary and tertiary structure 55
4.5.1 Absorption and fluorescence spectroscopy 56
4.5.2 Circular dichroism spectroscopy 57
4.5.3 Infrared spectroscopy 58
4.5.4 Other methods 59
4.6 Conclusion 60
References 60

5 Chemical Pathways of Peptide and Protein Degradation 70

Chimanlall Goolcharran, Mehrnaz Khossravi and Ronald T. Borchardt

5.1 Introduction 70
5.2 Hydrolytic pathways 71
5.2.1 Deamidation of Asn and Gln residues 71
5.2.2 Degradation of Asp residues 75

5.2.3 Degradation of N-terminal sequences containing penultimate Pro residues via diketopiperazine formation 78
5.3 Oxidation pathways 79
5.3.1 Autooxidation 80
5.3.2 Metal-catalysed oxidation 80
5.3.3 Photooxidation 82
5.3.4 Strategies to prevent oxidation 83
5.4 Other chemical pathways 84
5.4.1 β-Elimination reactions 84
5.4.2 Disulphide exchange reactions 84
5.5 Conclusion 85
References 86

6 Physical Stability of Proteins 89

Jens Brange

6.1 Introduction 89
6.2 Protein structure 90
6.2.1 Stabilizing interactions 92
6.2.2 Role of water in structure and stability 93
6.3 Protein destabilization (denaturation) 94
6.3.1 Unfolding intermediates (molten globule) 95
6.3.2 Temperature-induced changes 98
6.3.3 Influence of pH 98
6.3.4 Influence of pressure 99
6.4 Aggregation and precipitation 99
6.4.1 Mechanisms of aggregation 100
6.4.2 Precipitation and fibrillation phenomena 102
6.4.3 Factors influencing aggregation and precipitation 105
6.5 Surface adsorption 106
6.6 Solid phase stability 107
6.6.1 Lyophilization-induced aggregation 107
6.7 Stabilization of protein drugs 107
6.7.1 Stabilization strategies 108
References 109

7 Peptides and Proteins as Parenteral Suspensions: an Overview of Design, Development, and Manufacturing Considerations 113

Michael R. DeFelippis and Michael J. Akers

7.1 Introduction and scope 113
7.2 Rationale for suspension development 114
7.3 Types of suspensions and particle formation 116
7.3.1 *In situ* particle formation 116
7.3.2 Combination of particles and vehicle 122
7.4 Excipient selection 124
7.5 General requirements for suspension products 126
7.6 Testing and optimization of chemical, physical, and microbiological properties 127

7.6.1 Chemical properties 127
7.6.2 Physical properties 129
7.6.3 Microbiological properties 133
7.7 Techniques for characterizing and optimizing suspensions 133
7.8 Suspension manufacture 136
7.8.1 Scale-up 136
7.8.2 Manufacturing controls: special considerations for peptide and protein suspensions 136
7.9 Other related systems 138
7.10 Conclusions 139
Acknowledgements 139
References 139

8 Peptides and Proteins as Parenteral Solutions 145
Michael J. Akers and Michael R. DeFelippis
8.1 Overview and introduction 145
8.2 Optimizing hydrolytic stability 147
8.2.1 Buffers 150
8.2.2 Ionic strength 151
8.3 Optimizing oxidative stability 153
8.3.1 Antioxidants 154
8.3.2 Chelating agents 157
8.3.3 Inert gases 157
8.3.4 Packaging and oxidation 157
8.3.5 Other chemical stabilizers 158
8.4 Optimizing physical stability 158
8.4.1 Denaturation 159
8.4.2 Protein aggregation 160
8.4.3 Adsorption 162
8.4.4 Precipitation 163
8.4.5 Surfactants 163
8.4.6 Cyclodextrins 165
8.4.7 Albumin 166
8.4.8 Other physical complexing/stabilizing agents 167
8.5 Optimizing microbiological activity: antimicrobial preservatives (APs) 167
8.6 Osmolality (tonicity) agents 170
8.7 Packaging 170
8.8 Processing 171
8.9 Conclusion 171
References 172

9 Roles of Protein Conformation and Glassy State in the Storage Stability of Dried Protein Formulations 178
John F. Carpenter, Lotte Kreilgaard, S. Dean Allison and Theodore W. Randolph
9.1 Introduction 178
9.2 Infrared spectroscopy to study protein secondary structure 180

9.3 Physical factors affecting storage stability of dried protein formulations 181
9.4 Summary and conclusions 186
Acknowledgements 186
References 186

10 Peptide and Protein Drug Delivery Systems for Non-parenteral Routes of Administration 189

Mette Ingemann, Sven Frokjaer, Lars Hovgaard and Helle Brøndsted

10.1 Introduction 189
10.2 Non-parenteral routes of delivery for peptides and proteins 189
10.2.1 Barriers to non-parenteral administration of peptides and proteins 191
10.2.2 General approaches to bypass enzymatic and absorption barriers 192
10.3 Formulation principles for peptides and proteins 194
10.3.1 Entrapment and encapsulation 194
10.3.2 Covalent binding 198
10.4 Immobilized proteins intended for local effect in the GI tract – a case study 200
10.4.1 Oral enzyme supplementation therapy – phenylalanine ammonia-lyase 200
10.5 Future perspectives 202
10.6 Summary 203
References 203

11 Peptide and Protein Derivatives 206

Gitte Juel Friis

11.1 Introduction 206
11.2 4-Imidazolidinone prodrugs 207
11.3 Prodrugs of TRH 209
11.4 Derivatives of desmopressin 210
11.5 Derivatives of insulin 212
11.6 Cyclic prodrugs 213
11.7 Conclusions 214
References 215

12 Chemical and Pharmaceutical Documentation 220

Karen Fich and Deirdre Mannion

12.1 Introduction 220
12.2 Composition 221
12.3 Method of manufacture 222
12.4 Control of starting materials 222
12.4.1 Active substances 222
12.4.2 Excipients 226
12.4.3 Packaging materials 226

12.5 Intermediate products 227
12.6 Control tests on the finished product 227
12.7 Stability 228
12.7.1 Expert report 230
References 231

Index 232

Figures

1.1	Structural representation of an amino acid and a peptide chain (a tripeptide)	3
1.2	Structural representation of oxytocin	7
1.3	Outline of possible peptide by-products of oxytocin synthesis	9
2.1	Overview of the steps involved in the cloning of chromosomal DNA and cDNA fragments into a plasmid vector	14
2.2	Overview of the steps involved in the screening of a plasmid library with a radioactive DNA probe	16
2.3	The polymerase chain reaction using single-stranded cDNA as template	17
2.4	Oligonucleotide-directed mutagenesis by enzymatic primer extension	23
2.5	Phage-display of peptide epitopes	25
3.1	Outline of protein production from microbial and mammalian sources	31
3.2	Unfolding and aggregation of proteins	35
5.1	Major pathways of peptide and protein degradation	71
5.2	Pathways for the deamidation of Asn residues	72
5.3	Formation of pyroglutamate from N-terminal Gln residue	72
5.4	Pathways for the spontaneous fragmentation of Asn polypeptides	74
5.5	Pathways for the degradation of Asp-polypeptides	76
5.6	Degradation of Asp-peptides in acidic media	77
5.7	Degradation of N-terminal sequences containing penultimate Pro residues via diketopiperazine formation	79
5.8	Mechanism of β-elimination reactions of amino acid residues at alkali pH	85
6.1	Self-association of insulin from the monomer to the dimer, and of three dimers into a hexamer in the presence of zinc ions	91
6.2	Unfolding, aggregation and precipitation	95
6.3	Free-energy profiles for protein unfolding as a result of increasing temperature or increasing concentration of denaturant	96
6.4	Native and molten globule states of a hypothetical protein molecule	97
6.5	Relative population of native, molten globule and unfolded states during unfolding of a hypothetical protein via an intermediate state	97

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 11
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

6.6 Changes of entropy and Gibbs' free energy of protein unfolding as a result of change in temperature 99
6.7 Dissociation of the insulin hexamer as caused by dilution or addition of organic medium 101
6.8 Formation of protein aggregates (fibrils) as a function of time 103
6.9 Transmission electron micrographs of bovine insulin aggregates (fibrils) formed under various conditions 104
7.1 Solubility diagram for a hypothetical protein 117
7.2 Schematic representation of the NPH insulin crystallization process 119
7.3 Schematic representation of preparation of Ultralente suspension 120
7.4 Potential energy curves for particle interactions in coarse suspensions 129
7.5 Caking diagram 131
8.1 Protein reactions as a function of pH 148
8.2 Effect of pH on deamidation and polymerization of insulin 149
8.3 Inactivation of β-galactosidase 152
8.4 Effect of salt concentration on rAAT solution stability 152
8.5 Effect of ionic strength on rbSt solubility–pH profiles 153
8.6 Response surface for the oxidation of methionine 59 in hIGF-I in aqueous solution at 25°C 155
8.7 Methods to prevent temperature-induced oxidation of rhuMAb HER2 156
8.8 Aggregation scheme of potential folding pathways for an unfolded protein 159
8.9 Effect of Tween 80 concentration on precipation of rbSt as a function of thermal stress at 54°C 165
8.10 Time course of aggregate formation of 1.0 mg/ml rhIFN-γ 169
10.1 Routes for the transport of drugs across a mucous cell barrier 192
10.2 Introduction of drug: before crosslinking, and after hydrogel synthesis 196
10.3 PEG-activation of PEG-coupling reactions to proteins for cyanuric chloride, carbonylimidazole, succinimidyl active ester and tresyl chloride 199
10.4 PAL degradation in solutions of trypsin and chymotrypsin for free or SPA-PEG-conjugated PAL 202
11.1 Illustration of the analogue and prodrug approaches 207
11.2 (A) Enzymatic degradation of Leu-enkephalin; (B) example of the structure of a 4-imidazolidinone prodrug of Leu-enkephalin 208
11.3 (A) Enzymatic degradation of TRH; (B) structure of the N-octyloxycarbonyl TRH prodrug 209
11.4 Structure of desmopressin and the O-pivaloyl ester prodrug of desmopressin 211
11.5 Part of the structure of human insulin acylated with fatty acids at the ε-amino group of Lys^{B29} 212
11.6 Proposed pathway for the conversion of the acyloxyalkoxy-based cyclic prodrug to the parent peptide in esterase media 213

Tables

2.1	Examples of post-translational protein modifications	19
2.2	Recombinant protein drugs on the market	26
3.1	Examples of factors that may adversely affect the stability of proteins	32
3.2	The Hoffmeister series	36
3.3	Parameters in the downstream processing of proteins	38
4.1	Common methods used for chemical and physical characterization of pharmaceutical proteins and monitoring of structural alterations that occur during purification, formulation and storage	44–45
5.1	Common degradation products formed from metal-catalysed oxidation of amino acid residues	82
5.2	Common degradation products formed from reaction of singlet oxygen with amino acid residues	83
7.1	Examples of peptide or protein suspensions	115
7.2	Relative properties of deflocculated and flocculated particles in suspension	130
8.1	Development strategy for protein and peptide parenteral solution dosage forms	148
8.2	Buffers used in protein formulations	151
8.3	Antioxidants and chelating agents for protein formulations	155
8.4	Examples of commercial protein solution formulations containing surface active agents	164
8.5	Examples of commercial protein solution formulations containing albumin	166
8.6	Antimicrobial preservative agents for protein products	167
8.7	Comparison of USP 23 and EP 2 requirements for preservative efficacy testing	168
8.8	*D* values against *Staphylococcus aureus* for different antimicrobial preservative systems in insulin solutions	170
10.1	Non-parenteral routes of administration: absorption area and proteolytic barrier	190
10.2	Classes of absorption enhancers	193

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 13
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

10.3 Expected effect on the enzymatic and absorption barrier of the presented drug delivery systems for peptide and protein drug delivery 194
10.4 Remaining enzymatic activity after hydrogel entrapment and after acrylic monomer contact for PAL 201

Contributors

Michael J. Akers is with Akers Consulting and Training Services, Indianapolis.
S. Dean Allison is at the University of Colorado Health Science Centre, Denver.
Miroslav Baudyš is at the Biomedical Polymers Research Building, Salt Lake City.
Ronald T. Borchardt is at the University of Kansas, Lawrence.
Jens Brange is with Brange Consult, Klampenborg, Denmark.
Helle Brøndsted is at the Royal Danish School of Pharmacy, Copenhagen, Denmark.
John F. Carpenter is at the University of Colorado Health Science Centre, Denver.
Michael R. DeFelippis is with Eli Lilly & Co., Indianapolis.
Nanni Din is with Novo Nordisk A/S, Bagsvaerd, Denmark.
Jan Engberg is at the Royal Danish School of Pharmacy, Copenhagen, Denmark.
Karen Fich is with H. Lundbeck A/S, Valby, Denmark.
Gitte Juel Friis is at the Royal Danish School of Pharmacy, Copenhagen, Denmark.
Sven Frokjaer is at the Royal Danish School of Pharmacy, Copenhagen, Denmark.
Chimianlall Goolcharran is at the University of Kansas, Lawrence.
Lars Hovgaard is at the Royal Danish School of Pharmacy, Copenhagen, Denmark.
Mette Ingemann is with LCO Pharmaceutical Products A/S, Ballerun, Denmark.
Mehrnaz Khossravi is at the University of Kansas, Lawrence.
Sung Wan Kim is at the Biomedical Polymers Research Building, Salt Lake City.
Lotte Kreilgaard is with Immunex Corporation, Seattle.
Deirdre Mannion is with the Danish Medicines Agency, Brønshøj, Denmark.
Bernard A. Moss is with PolyPeptide Laboratories, Hillerød, Denmark.
Theodore W. Randolph is at the University of Colorado, Boulder.
Lars Skriver is with L & K Biosciences, Vedbaek, Denmark.

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 15
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

2

Basics in Recombinant DNA Technology

NANNI DIN[1] AND JAN ENGBERG[2]

[1]*Department of Molecular Genetics, Novo Nordisk A/S, Bagsvaerd, Denmark*
[2]*Department of Pharmacology, The Royal Danish School of Pharmacy, Copenhagen, Denmark*

2.1 Introduction
2.2 General methods in gene technology
  2.2.1 DNA cloning tools
  2.2.2 Cloning of cDNA
  2.2.3 PCR cloning and database mining
2.3 Expression of recombinant proteins
  2.3.1 Transcription, translation and protein modifications
  2.3.2 Choice of expression system
2.4 Protein design
  2.4.1 Protein variants
  2.4.2 Protein chimeras
  2.4.3 Epitope display libraries
2.5 Recombinant protein therapeutics – status and future trends
  References

2.1 Introduction

Gene technology is one of several terms used to describe the methods which have evolved as a result of research on the structure and function of genes. Other terms are 'recombinant DNA (rDNA) technology', 'genetic engineering' and 'gene splicing'. Gene technology involves taking genetic material from one source and recombining it *in vitro* with genetic material from another source followed by introduction of the recombined material into a host cell.

By the aid of genetic engineering it is possible to introduce a gene coding for a desired protein into a biological environment where this protein can be produced in large quantities. This possibility has been one of the main reasons for the interest of the pharmaceutical industry in applying genetic engineering methods. Using these methods, manufacturers can in principle produce unlimited quantities of biologically active proteins and peptides, as has already been convincingly demonstrated in many cases. In 1982, the first approved human therapeutics produced by the aid of gene technology appeared. This landmark drug was insulin, a peptide hormone used to treat diabetes. At the present time (1999), more than 20 different peptides or proteins produced by gene technology methods have

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 16
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

been approved as human therapeutics, and more are certain to follow. In contrast to insulin, for which a production method based on extraction from the pancreas of pig or ox had existed for many years, some of the other approved gene technology products have never been in regular use as drugs before, because they were impossible to produce in sufficient quantities. Thus, one of the advantages of gene technology is the possibility of producing substances that have hitherto been so elusive as to make their use as pharmaceuticals remote.

Another advantage of gene technology is that it allows introductions of modifications into proteins at desired positions. The modifications may be a few changes in the amino acid sequence, made by introducing specific mutations in the gene encoding the protein. Such changes may lead to desirable altered physicochemical characteristics, e.g. new solubility properties or a different physiological half-life of the active polypeptide. However, more profound changes giving rise to novel activities are also possible.

This chapter aims at describing the basic tools and concepts of gene technology as used in pharmaceutical research and production, and sums up the current status and future trends for recombinant protein therapeutics.

2.2 General methods in gene technology

2.2.1 DNA cloning tools

At the beginning of the 1970s, many new methods of significant importance for recombinant DNA technology were developed. Since then the methods have improved and new ones have emerged, but the principles are still based on the same tools: DNA modifying enzymes, DNA vectors and host cells (Sambrook *et al.*, 1989).

A key factor in the breakthrough of the 1970s was the discovery of two types of enzymes, the restriction enzymes and the DNA ligases. Restriction enzymes are endonucleases that recognize and cleave DNA at specific sequences, typically 4–8 base pairs (bp) long. As a result, DNA fragments of defined sizes are generated, some of which may contain a single complete gene. Restriction endonucleases can be isolated from a variety of microorganisms, and a large range of enzymes, each with a specific recognition sequence, is now commercially available. An example is the restriction enzyme *BamHI* (endonuclease I isolated from *Bacillus amyloliquefaciens H*) which cleaves DNA at the sequence GGATCC. In contrast to the restriction enzymes, DNA ligases have no specific sequence requirement but are capable of joining together compatible ends of DNA molecules. These features of the restriction enzymes and DNA ligases make it possible to perform *in vitro* cleavage and rejoining ('splicing') of DNA fragments as illustrated in Figure 2.1.

In order to clone a DNA molecule it must be introduced into a host cell in a form which allows it to be replicated. Replication requires the presence of a specific sequence (replication origin) which can be provided by circular 'minichromosomes' called plasmids. Therefore, to generate clonable DNA, a restriction fragment is joined to a plasmid DNA molecule which can replicate in a suitable host cell. A very important host–plasmid system is the bacterium *Escherichia coli* and its naturally occurring small (2–10 kbp) plasmids. Some of these plasmids have been modified so that they are particularly useful for cloning purposes, and they are hence called vectors. Apart from the replication origin, an important feature of the vector is the presence of a gene coding for a selectable marker, such as resistance to an antibiotic. When recombinant plasmids are introduced into *E. coli* (in a process called transformation) the resistance gene will allow the cells to

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 17
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

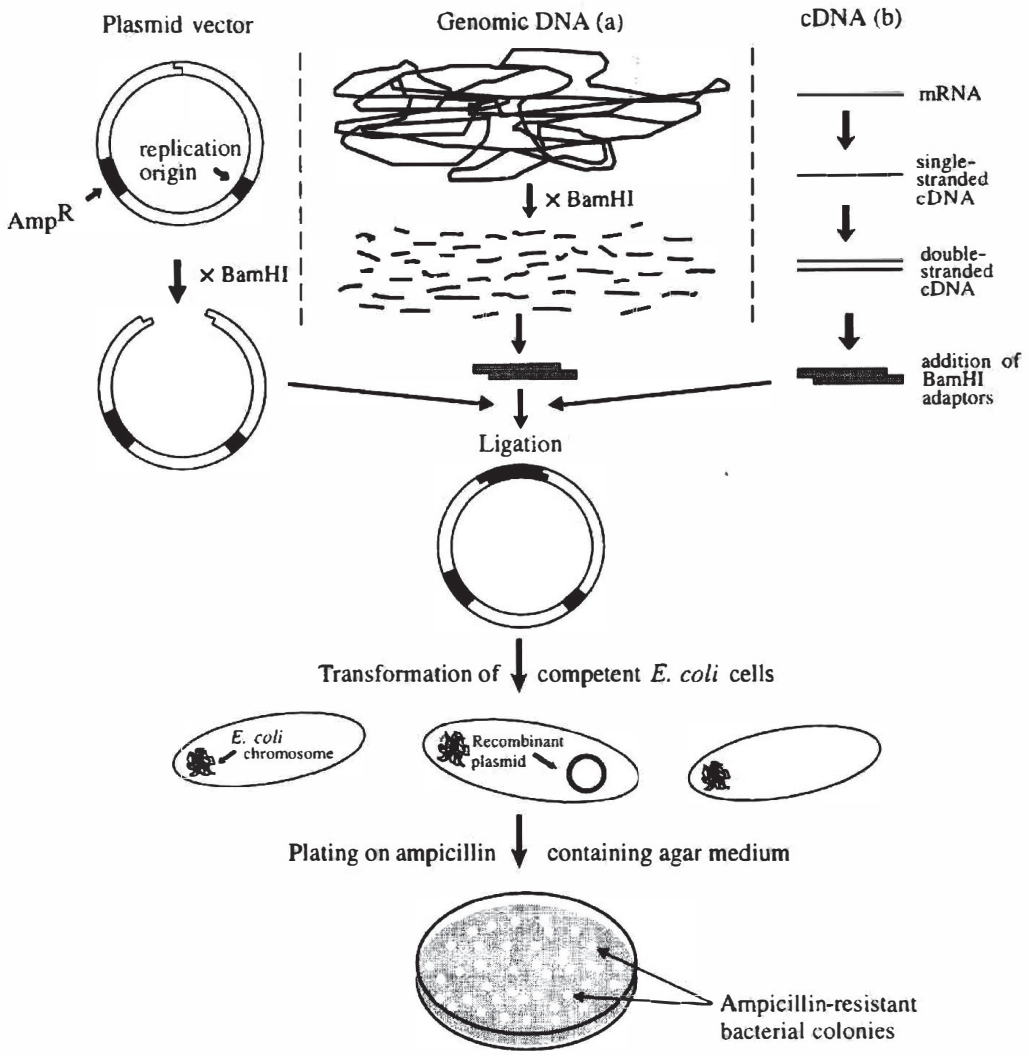


Figure 2.1 Overview of the steps involved in the cloning of (a) chromosomal DNA and (b) cDNA fragments into a plasmid vector. The plasmid vector carries a replication origin and a gene which confers resistance to the antibiotic ampicillin (*Amp*^{R}). Outside these regions, a unique site for the restriction endonuclease (*BamHI*) is present. In order to clone DNA fragments into this site, vector and foreign DNA are cleaved with *BamHI*, mixed and joined together with the aid of the enzyme DNA ligase. The mixture of ligated DNA molecules is transformed into *E. coli* cells. Only cells which have received a plasmid will grow in the presence of ampicillin. Generation of double-stranded cDNA from eukaryotic mRNA requires the following steps. A chemically synthesized oligo-dT primer is annealed to the 3′ polyadenylated tails of the mRNA molecules. Single-stranded cDNA is then synthesized by the aid of reverse transcriptase. Following removal of the mRNA molecules by hydrolysis with sodium hydroxide to give a population of single-stranded cDNA molecules, these are used as templates in a new round of DNA synthesis catalysed by the enzyme DNA polymerase. Synthetic DNA adaptors for *BamHI* are subsequently ligated to the double-stranded cDNA molecules and the resulting fragments can be inserted into the *BamHI* site of the vector.

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 18
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

grow in the presence of the antibiotic, while cells without the plasmid will be eliminated. A recombinant plasmid is able to multiply until each cell contains several copies (often in the range from 20 to 200), and as the doubling time for *E. coli* under optimal conditions is 20–30 minutes, it is possible to multiply a single copy of the recombinant plasmid to more than 10^{10} molecules within 12 hours, thereby making further manipulation of a particular gene possible.

2.2.2 Cloning of cDNA

Cloning and expression of human genes, which are of special interest within pharmaceutical contexts, present certain special challenges. This is because most human genes (in contrast to bacterial genes) contain internal regions, called introns, that do not code for any part of the corresponding protein. Introns can range in size from less than 100 bp to several kbp, and mammalian genes can be very large due to the presence of several introns. In human and other eukaryotic cells, intron-containing RNA are processed to generate mature mRNA without introns through processes that bacteria cannot perform. When human chromosomal DNA (also called genomic DNA) is cloned in *E. coli* (Figure 2.1a), each individually cloned restriction fragment rarely contains a complete gene and in most cases intron sequences will be present. Therefore, while cloning of chromosomal DNA in *E. coli* is very useful for characterization of human gene structure, it is less useful as a starting point for expression of a protein by a particular gene.

Fortunately, a method has been devised that allows cloning of a DNA fragment containing only the coding regions of a human gene: by an enzymatic *in vitro* reaction using an enzyme called reverse transcriptase, it is possible to make a DNA copy of the mature mRNA. This DNA copy (called cDNA for copy DNA or complementary DNA) contains an uninterrupted coding region which can be cloned and expressed in *E. coli* (Figure 2.1b).

To generate cDNA coding for a protein of interest, mRNA is isolated from specific tissues or organs known to produce this protein. The isolated mRNA consist of a population of molecules representing all the genes expressed in the particular tissue, and cloning of the complete cDNA population results in a so-called tissue-specific cDNA library. From this library, clones that contain the desired cDNA must be identified and isolated before further analysis. This can sometimes be a formidable task, since only a small proportion of the cDNA molecules generated by reverse transcription corresponds to the protein of interest. A standard identification method is the DNA hybridization technique illustrated in Figure 2.2. This method requires some knowledge about the sequence of the desired cDNA, e.g. a partial amino acid sequence of the encoded protein. Based on a known amino acid sequence, a short DNA fragment with the corresponding coding capacity can be designed and synthesized in a radioactive form. This radioactive DNA probe is able to bind (hybridize) to the desired cDNA through base pairing, and thereby give rise to a signal from the desired clone. If no sequence information is available, other methods must be used which usually require that the protein encoded by the cDNA becomes expressed by the host cell. If a specific antibody against the desired protein is available, a screening procedure similar to the hybridization technique outlined in Figure 2.2 can be used, except that the antibody is used instead of a DNA probe. In such cases, the detection signal (radioactivity or light emission) is normally produced by a secondary antibody that recognizes the specific antibody.

In some cases the biological activity of a protein may be the only means by which to identify the cDNA clone encoding it. Thus, it becomes necessary to use host cells which

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 19
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

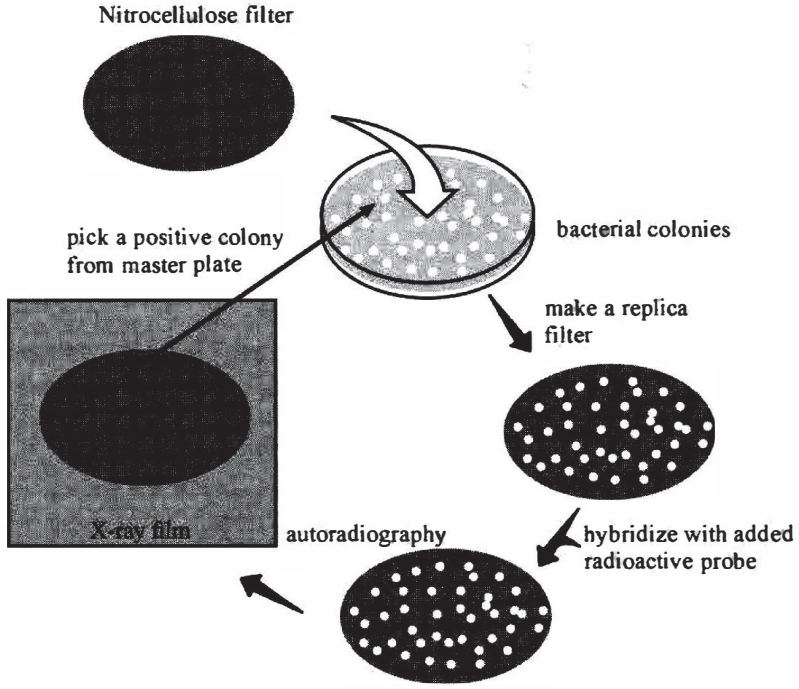


Figure 2.2 Overview of the steps involved in the screening of a plasmid library with a radioactive DNA probe. Plasmid libraries are typically screened by spreading several thousand plasmid-containing cells on each of a series of agar plates. After the bacterial colonies have grown to visible size, nitrocellulose filters are carefully laid onto the surface of the plates. Colonies from the plate adhere to the filter, creating a precise replica on the filter of the colonies on the plate. The filters are treated to solubilize the cells and immobilize their content of DNA to the filter surface. The filters are now incubated in a solution containing a radioactively labelled DNA probe complementary in sequence to a portion of the desired cloned gene. The filters are carefully washed to remove unbound probe, leaving behind only the probe molecules specifically bound to complementary sequences. The location of the bound probe is determined by exposing the filters to X-ray film (autoradiography). The position is represented by a spot of exposure on the film. By orienting the film with the original agar plate, the colony harbouring the desired clone can be identified and the gene isolated.

preserve this activity. An obvious choice is mammalian cells grown in culture, and many different mammalian host–vector systems now exist for such purposes. Other host cells, notably frog oocytes, have also proven very versatile in their ability to express complex functional proteins, such as mammalian membrane receptors or ion channels.

2.2.3 PCR cloning and DNA database mining

Since its invention in the early 1980s, the polymerase chain reaction (PCR) has become increasingly important as a cloning tool (Innis *et al.*, 1990). The PCR technique is based on enzymatic *in vitro* amplification of a specific DNA sequence. Figure 2.3 illustrates how a sequence, often called the target or template DNA, can be amplified from a population of single-stranded cDNA.

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 20
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

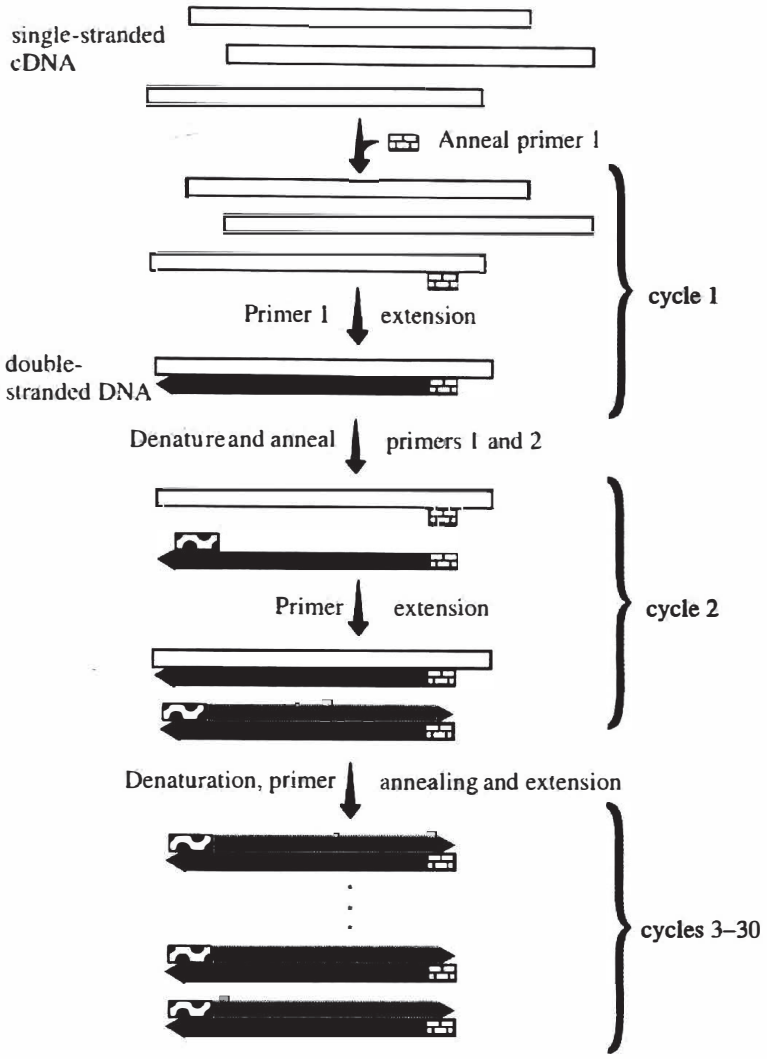


Figure 2.3 The polymerase chain reaction using single-stranded cDNA as template. The reaction mixture is heated and primer 1 is annealed to the mixture of single-stranded cDNA molecules. Added *Taq* polymerase synthesizes a new strand of DNA complementary to the template and extends to the end of the molecule. The reaction mixture is heated again and the original and newly synthesized DNA strands separate. Binding sites for primers 1 and 2 now become available and complementary strands are being synthesized using the original strand and its complementary strand as templates. Repeated rounds of denaturation, annealing and primer extension will produce double-stranded DNA molecules identical to the region defined by the two primers in the original target sequence.

Repeated rounds of DNA synthesis are catalysed by a thermostable DNA polymerase in a suitable buffer containing the cDNA template, a specific set of two short synthetic DNA primers and a mixture of the four deoxynucleotides. The primer set is designed so that one primer can bind near one end of the cDNA and the other can bind near the other end of the complementary strand. When the first primer anneals to its target it creates a starting point for the DNA polymerase and a complementary copy of the template is generated. By heating the reaction mixture, the two complementary strands are separated,

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 21
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

and targets for both primers are thereby exposed. When the temperature is lowered, both primers can anneal to their respective targets and both strands can now be copied by DNA polymerase. Each consecutive cycle consists of the same series of strand separation, primer annealing and DNA polymerization reactions, and after about 20 cycles the target sequence may become amplified more than a million-fold.

Some knowledge about the sequence of the template DNA is needed in order to design primers which are able to bind to the target. In addition to using protein sequence data as described previously, other types of data can be used for primer design. Thus, the sequences now available for a vast number of genes have made it clear that genes coding for particular functions have evolved as gene families which exhibit substantial sequence homology. Primers designed to recognize conserved regions from such a gene family may readily amplify cDNA for new members of the gene family. This procedure proved very successful and has expanded the knowledge of gene families

A dedicated effort to sequence the entire human genome of around 3×10^9 bp, including introns and other non-coding DNA stretches, was initiated around 1990, as part of the so-called Human Genome Project (HUGO). Initially, the sequence information accumulated very slowly, and a 15–20 year time frame was envisaged for its completion, but improved technology has enabled much faster data accumulation (Rowen *et al.*, 1997). Thus, it is now expected that 90% of the entire sequence will be known in the year 2000, and the remaining 10% will be completed a year or two after. The generated sequences have been made available to researchers in public databases. Partial sequences from a number of cDNA libraries – so-called expressed sequence tags (ESTs) – are also available in databases, and the EST sequences help in identifying coding from non-coding genome sequences. It is currently estimated that the human genome contains 50 000 to 100 000 genes, and all of these genes will soon be available in sequence databanks as well as in DNA clone banks. Thus, in the near future it will be possible to acquire the DNA encoding any human protein from such a bank. However, the selection of interesting genes will always require a vast background of biological expertise, and the experimental use of DNA will require knowledge of the basic recombinant DNA procedures.

2.3 Expression of recombinant proteins

Production of recombinant proteins and peptides to be used as pharmaceuticals requires efficient expression systems that generate proteins with the correct structure and function. In this section, some requirements for expression will be outlined together with a discussion of the relative merits and drawbacks of different host–vector expression systems (Hodgson, 1993).

2.3.1 Transcription, translation and protein modifications

The first step in gene expression is transcription into RNA. Transcription is initiated with the binding of the enzyme RNA polymerase to a specific DNA sequence at the beginning of the coding region, the so-called promoter. Many different promoters and other elements that influence transcription efficiency have been characterized in both prokaryotes and eukaryotes. Promoters are most often species-specific, and in higher organisms also tissue-specific. Transcription termination is also determined by specific DNA sequences, located downstream of the coding region. In prokaryotes, the RNA molecule produced by

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 22
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

Table 2.1 Examples of post-translational protein modifications

Modification	Occurrence
Cleavage of peptide bonds (maturation) Formation of disulphide bonds Phosphorylation of tyrosine, serine or threonine	Prokaryotes and eukaryotes
N-linked glycosylation of asparagine O-linked glycosylation of threonine or serine	Eukaryotes
Hydroxylation of proline or lysine Myristylation of N-terminal glycine Prenylation of cysteine	
γ-carboxylation of glutamic acid Amidation of C-terminal glycine Sulphation of tyrosine	Higher eukaryotes

transcription of protein-coding genes is used without any major modifications as mRNA to direct protein synthesis. However, in eukaryotes, complex enzymatic processing steps which remove intron regions from the primary transcription product are required for the generation of mature mRNA.

The next step in gene expression is translation of the mRNA into protein. The basic translation mechanisms are the same in prokaryotes and eukaryotes. Thus, the codon for translation initiation is with rare exceptions universal and leads to translation products containing methionine at their N-terminus. The N-terminal methionine is, however, often removed already during protein elongation by a ribosome-associated methionyl amino peptidase. Most proteins destined for secretion from cells are synthesized with a so-called signal peptide of 20–30 amino acids at their N-terminus. The signal sequence mediates contact between the membrane and the nascent protein, and it is usually cleaved off during the translocation over the membrane before the protein is fully synthesized. Compared to prokaryotes, eukaryotes have a very complex pathway for secretion of proteins, and many essential modifications of proteins occur in the secretory pathway. Therefore, maturation of the primary translation product to a protein with correct secondary and tertiary structure cannot always be obtained if *E. coli* is used for expression of a eukaryotic protein. To overcome this problem, several efficient eukaryotic expression systems have been developed in addition to the 'classical' *E. coli* systems. Examples of post-translational protein modifications in prokaryotes and eukaryotes are listed in Table 2.1 (Creighton, 1993).

2.3.2 *Choice of expression system*

E. coli

Since 1977, when a human protein (somatostatin) was expressed in *E. coli* in a functional form for the first time (Itakura *et al.*, 1977), many hundreds of eukaryotic genes have been expressed in this organism but very often not in a functional form. Because *E. coli* is incapable of performing many of the post-translational modifications that occur in mammalian cells, it is most suitable for production of proteins with no or limited modifications. However, its ease of handling and its safety as a production organism speak in

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 23
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

favour of maintaining *E. coli* as an host even for proteins which are not completely matured in this organism. In order to obtain functional proteins, various ingenious methods have been devised that allow the final maturation to occur *in vitro* and recent progress has also improved the *in vitro* processing capacity tremendously (Georgiou and Valax, 1996). A number of approved recombinant pharmaceuticals are produced in *E. coli*, e.g. growth hormone, interferon-α and interleukin 2.

Intensive work has been performed aimed at optimizing the signals involved in expression in *E. coli*. Promoters from highly transcribed bacterial genes, or synthetic promoters, are used for high production levels. Regulated promoters that can be turned on or off by altering the growth conditions (temperature, added chemicals, etc.) are often used for production of recombinant proteins that are harmful to the cell. Examples of such promoters are the *lac* promoter from the β-galactocidase operon, or the bacteriophage λ promoters P_L and P_R. The advantage of these promoters is that synthesis of recombinant proteins can be delayed until the cell density has reached a high level where a short protein production phase results in generation of large amounts of product.

Yeast

Saccharomyces cerevisiae (baker's yeast) is an eukaryotic microorganism that performs many post-translational modifications which also occur in humans. This makes it useful for production of peptides and proteins of medium complexity (Romanos *et al.*, 1992). The long tradition of using yeast for food production contributes significantly to its acceptability as a producer of pharmaceutical proteins.

S. cerevisiae contains a natural plasmid (2 μm DNA) which can be used as vector when equipped with a selectable marker. Instead of using a gene conferring resistance to an antibiotic, a yeast gene which complements a mutated gene in the host genome is used for plasmid selection. This principle (which can also be used in *E. coli* and other hosts) is preferable to antibiotic selection for large-scale work.

It is often an advantage to produce recombinant proteins from yeast in a secreted form, because secreted proteins may be subjected to desired post-translational modifications in the secretory pathway, and because they become easier to purify. However, undesirable changes may occur during passage through the secretory pathway, e.g. incorrect N- and O-linked glycosylation. Since the glycosyl side-chains made in yeast and humans are different, yeast-produced glycoproteins may be immunogenic in man. Thus, glycoproteins are generally not suitable for production in yeast. Insulin is an example of an approved recombinant protein produced in *S. cerevisiae*.

Mammalian cells

Mammalian systems are the obvious choice for production of glycoproteins or proteins with other complex post-translational modifications. Some of the modifications (e.g. sulphation, amidation and γ-carboxylation) are often essential for the biological activity of the proteins. For example, the activity of a number of blood coagulation factors (factors II, VII, IX and X) is completely dependent on γ-carboxylation of glutamic acid, a modification which occurs only in cells of higher eukaryotes.

Many different mammalian host–vector systems have become available for expression of recombinant proteins (Birch and Froud, 1994). Examples of immortalized mammalian cell lines, which can be grown in culture and therefore are suitable as host cells, are CHO (Chinese hamster ovary), BHK (baby hamster kidney), COS (African green monkey

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 24
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

kidney) and HeLa (human epitheloid carcinoma) cells. Mammalian viruses which replicate autonomously in cells can be used as vectors for recombinant DNA, but it is also possible to perpetuate introduced DNA in mammalian cells without an autonomously replicating vector, because transfected DNA molecules can integrate into mammalian chromosomes. Viruses which have been modified to be particularly useful as vectors include SV40 (simian virus), Epstein-Barr virus and retroviruses. The first two are self-replicating, while the third mediates cellular uptake and integration into the genome with significantly increased efficiency compared to DNA constructs without retroviral elements. To allow easy isolation of cells in which recombinant DNA has been taken up, a selectable marker gene such as resistance to the cytotoxic drug neomycin is introduced into the vector. The promoters used for gene expression in mammalian cells are often strong viral promoters such as the SV40 and CMV (cytomegalovirus) promoters.

It is important to realize that post-translational modifications may not be identical to those occurring in authentic protein, even when mammalian cells are used for production. In the body, proteins are produced in organs composed of specialized cells, and it has become increasingly clear that differences exist in the processing capabilities of different tissues. Also, immortalized cell lines may have diverged from their progenitors not only in their growth characteristics but also in other ways. Proteins with aberrant post-translational modification may be recognized as 'non-self' by the immune system and therefore evoke an immune response that will eliminate the protein. Therefore, careful characterization of recombinant proteins is equally important whether production occurs in mammalian cells or in microorganisms.

From an economic point of view, mammalian cell culture systems are less attractive than microbial systems. The low growth rate, moderate expression levels, complex medium requirements and the need for advanced cultivation equipment are the main reasons for the relatively high cost of producing biologicals from mammalian cell lines. However, production costs are usually an issue only if large amounts of the protein are necessary for therapeutic purposes. The coagulation factors FVIIa and FVIII are examples of marketed recombinant proteins produced in mammalian cell cultures.

An alternative to the expression of recombinant proteins in mammalian cell cultures is the use of so-called transgenic animals (Velander *et al.*, 1997). Transgenic animals can be generated by injection of a foreign gene into fertilized eggs by *in vitro* micromanipulation. The injected foreign gene integrates into the chromosome of the egg, normally at random locations. The eggs are then implanted into the oviduct of a foster mother which after a normal period of pregnancy will give birth to transgenic progeny with the foreign gene incorporated permanently into the genome of all cells. When using transgenic animals as producers of recombinant proteins, it is important that expression of the transgene is regulated, because the compound is usually biologically active and therefore should be produced and stored in organs or compartments where it does not affect the animal. The most attractive transgenic production systems employ mammary gland specific regulatory elements that target accumulation of the protein in the milk of the animal. Transgenic sheep have been made which express substantial levels of recombinant proteins in the milk, e.g. tissue plasminogen activator, α_1-antitrypsin and coagulation factor IX. None of the proteins produced by transgenic animals have yet been marketed as pharmaceuticals.

Insect cells

Insect cell expression systems have become very popular, especially for small-scale production of mammalian recombinant proteins. High expression levels of active proteins

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 25
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

can usually be obtained much faster than in mammalian cells, which is a key attraction for the systems.

Most vectors are based on a lytic insect virus belonging to the baculovirus family, and the foreign cDNA is inserted in the viral genome without interfering with the lytic life cycle of the virus. Protein production thus occurs during a lytic infection of insect cells with recombinant viruses, and usually less than a week is needed from infection to maximal product yield (Griffiths and Page, 1997). The fact that baculoviruses are non-infectious to vertebrates and their promoters inactive in mammalian cells gives insect systems a potential advantage over other systems when expressing oncogenes or other genes that are harmful to mammals. Thus, baculovirus has been used for expression of an HIV protein (gp160) which is at present in phase III clinical trial as a recombinant AIDS vaccine.

2.4 Protein design

Recombinant DNA technology makes it feasible not only to produce natural proteins, but also to design and produce new types of protein molecules. New proteins can be broadly classed in two categories: (1) natural protein variants harbouring replacements, insertions and deletions of small numbers of amino acids, and (2) chimeric proteins consisting of domains originating from different proteins. These approaches have been very useful in structure–function studies of many proteins, e.g. in the determination of which part of the protein is responsible for catalytic function, tissue targeting or stability. The techniques play an important role in the rational design of protein drugs with improved characteristics such as prolonged stability *in vivo* (Blundell, 1994). A third category comprising peptides and proteins designed from scratch has recently become a booming activity as the result of the introduction of the so-called epitope-display technology.

2.4.1 Protein variants

Protein variants can be generated by a genetic engineering method called site-directed mutagenesis. Using this method, it is possible to change, delete or insert one or a few nucleotides within a coding region (Figure 2.4). Expression of a mutated gene will result in a protein variant with a specific amino acid alteration, whose function is then tested in various ways. Engineering small changes in proteins can have a variety of objectives, such as changing the solubility or stability properties or the affinity for substrates or receptors.

For example, extensive work has been performed to design insulin molecules with improved characteristics for the replacement therapy necessary in diabetes treatment (Bristow, 1993). The insulin molecule contains four glutamic acid residues, and by changing one or more of these to the neutral amino acid glutamine the isoelectric point is shifted towards a higher pH, thus resulting in an insulin variant with a lower solubility at the physiological pH (around 7.3). Such insulin variants constitute improved alternatives to existing slow-release formulations made by complexing native insulin with protamine or zinc ions. Similarly, other amino acid substitutions can lead to new quick-acting insulin preparations. As anticipated, many variants have proved undesirable at various stages in the test programmes, e.g. because of altered affinity for the receptors for insulin and the related hormone insulin-like growth factor. Another major problem for the clinical application of the analogues is the risk of adverse reactions due to immunogenicity or altered physiological functions of degradation products. These risks are very serious, especially

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 26
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

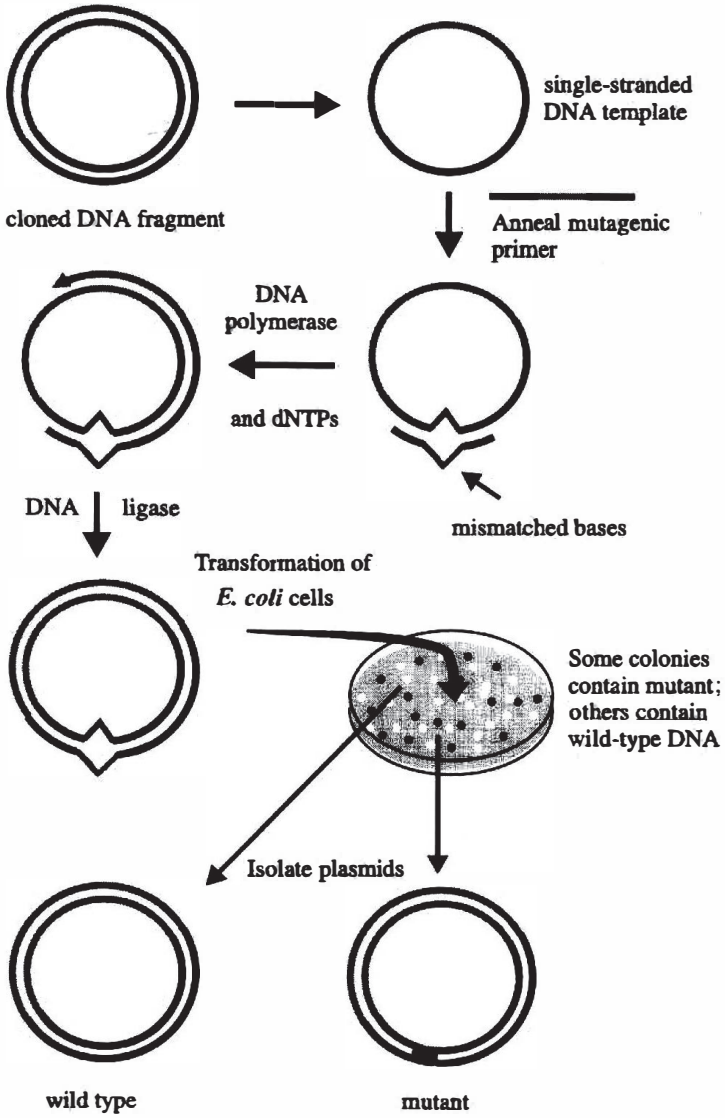


Figure 2.4 Oligonucleotide-directed mutagenesis by enzymatic primer extension. A 'mutagenic' oligonucleotide encoding the desired mutation embedded in wild-type flanking sequences is annealed to single-stranded DNA template. The sequence of the oligonucleotide is complementary to the template except for the nucleotides that define the mutation. The mutagenic oligonucleotide serves as a primer for DNA synthesis by DNA polymerase. Once the entire template has been copied, the ends of the newly synthesized strand are linked by DNA ligase. The heteroduplex DNA is transformed into *E. coli*. Replication results in segregation into separate wild-type and mutant plasmids and further repair mechanisms generate colonies containing either one or the other plasmid. Plasmid DNA is isolated from the resulting colonies and is screened to identify mutants.

for a pharmaceutical such as insulin which is administered daily for years, and extensive test programmes for adverse reactions are therefore necessary. However, several insulin analogues are now in phase III clinical trial, and one fast-acting analogue (Insulin Lispro, so named because the alteration involves the amino acids Lys and Pro) has already been marketed.

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 27
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

2.4.2 Protein chimeras

Chimeric proteins are made by fusion of the coding region for one protein (or protein domain) with that of another followed by expression of the recombined coding regions. The successful generation of a functional chimeric protein usually requires that structure–function relationships of the two starting proteins are well known, so that the desired active domains will be maintained within the new protein. However, the chimeric approach can also be used as a means to discover which regions of a protein are important for its activity.

The chimeric protein approach has a great potential for the production of so-called humanized monoclonal antibodies. Production of murine monoclonal antibodies based on the immortalization and proliferation of individual antibody-forming B-cells, the so-called hybridoma technique, has been very successful. However, in spite of major efforts, it has been difficult to establish similar production methods for human monoclonal antibodies. Instead, with access to cloned antibody genes of both human and murine origin, a production based on recombinant DNA expression is possible. In mice, expansion of antibody-producing cells with a required specificity can be obtained by immunization with any antigen, and the corresponding cDNA can be isolated. By combining the coding regions corresponding to the constant and the framework regions from human antibodies with the variable regions from a murine antibody, a chimeric antibody can be generated which is tolerated by the human immune system and has the desired antigenic specificity (Winter and Harris, 1993).

Such humanized antibodies may have numerous therapeutic applications, either alone or coupled to other biologically active substances. For example, they might be used in the elimination of toxic substances such as bacterially derived endotoxins, or as regulators of the humoral immune response by binding of excess levels of cytokines (interleukine 1, tumour necrosis factor, etc.). In cancer therapy, antibodies might be able to mediate selective destruction of tumour cells, not only via normal immunological mechanisms, but also via targeted toxin delivery. Thus, a tumour-cell specific antibody fused to parts of a cellular toxin, such as *Pseudomonas* exotoxin A, interacts selectively with and destroys tumour cells. This so-called 'magic bullet' approach is not limited to antibody–toxin fusions; chimeras between toxins and protein ligands for cell-type specific receptors are another possibility for targeted delivery (Vitetta *et al.*, 1993).

2.4.3 Epitope display libraries

Vast libraries of peptides without any predetermined homology to known proteins or peptides can be created through cloning complex mixtures of combinatorially synthesised oligonucleotides into specialized expression 'display' vectors. The filamentous phage display system, whereby the expressed peptides are displayed as fusion to phage coat proteins, has been effective in the discovery of many new ligands for various target proteins (Cortese, 1996). Affinity purification of the population of peptide phage-particles on the target protein is used to recover peptides with binding activity. Repeated rounds of binding selection and repropagation of bound eluted phage will enrich for tighter binders, and the identity of the selected peptide is simply revealed by sequencing of the appropriate segment of the DNA of each captured phage (Figure 2.5). An exceptionally elegant demonstration of the power of this technology was the identification of a small peptide that binds to and activates the receptor for erythropoietin (Wrighton *et al.*, 1996). Surprisingly,

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 28
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

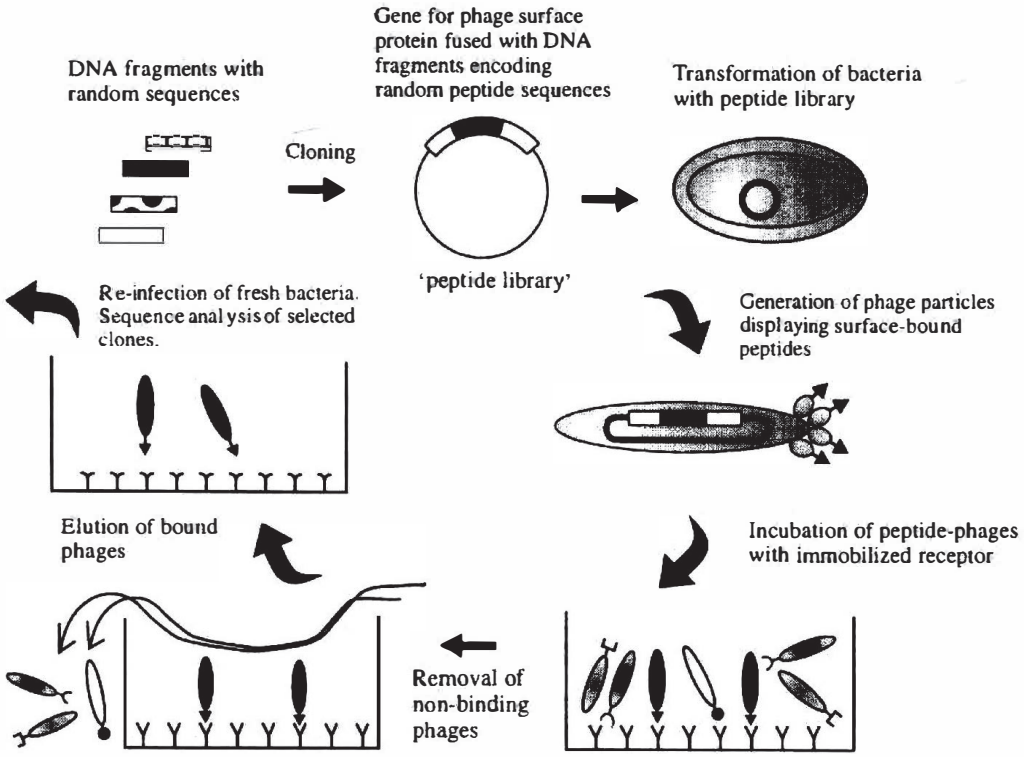


Figure 2.5 Phage-display of peptide epitopes. Cloning of random DNA sequences in frame with the coding region of a phage coat protein (gIIIp) results in the formation of phage particles each displaying one unique peptide epitope on the phage surface. Phage particles displaying an epitope recognizing a specific target (e.g. a receptor or an enzyme) can be isolated by repeated rounds of affinity purification and repropagation in *E. coli* of the eluted bound phages.

the peptide showed no sequence or structural homology to the cytokine it mimics. Peptides with improved binding characteristics can now be generated from the selected peptide phage-particle by random or directed mutagenesis of the different amino acid residues of the peptide by site-directed mutagenisis procedures. The size of the phage-displayed epitope is not limited to that of a small peptide. Thus, a flush of publications during the past few years have described the generation of phage libraries displaying antibody fragments (single chain Fv and Fab), cDNA derived proteins and proteins of random sequence fused to scaffold proteins designed to offer a endless diversity of molecular shapes (Dougall *et al.*, 1994).

2.5 Recombinant protein therapeutics – status and future trends

Table 2.2 lists the recombinant proteins which, to our knowledge, have been marketed as drugs at the time of writing. Only a few of these drugs were previously produced by non-recombinant means, and only one (insulin) was available in ample amounts. As can be seen, they represent a broad variety of biological substances from hormones and cytokines through enzymes and blood coagulation regulators to vaccines.

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 29
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

Table 2.2 Recombinant protein drugs on the market (PharmaProjects, 1997)

Product	Category	Indication	Year of introduction
Insulin	Hormone	Diabetes	1982
Growth hormone	Hormone	Growth defects	1985
Glucagon	Hormone	Insulin-induced hypoglycaemia	1993
Insulin-like growth factor I	Hormone	Diabetes and growth defects	1995
Atrial natriuretic peptide	Hormone	Acute heart failure	1995
Follicle stimulating hormone	Hormone	Infertility	1996
Interferon-α	Cytokine	Cancer therapy	1986
Interferon-β	Cytokine	Cancer therapy	1994
Interferon-γ	Cytokine	Cancer therapy	1990
Interleukin 2	Cytokine	Cancer therapy	1989
Granulocyte colony-stimulating factor (G-CSF)	Cytokine	Leucopenia	1991
Granulocyte–macrophage colony-stimulating factor (GM-CSF)	Cytokine	Leucopenia	1991
Erythropoetin	Cytokine	Anaemia	1989
Plasminogen activator	Antithrombotic	Thrombosis	1987
Factor VIII	Coagulation factor	Haemophilia A	1989
Factor VIIa	Coagulation factor	Haemophilia A and B	1996
Factor IX	Coagulation factor	Haemophilia B	1997
α_1-antitrypsin	Enzyme inhibitor	α_1-antitrypsin deficiency	1992
Hirudin	Enzyme inhibitor		
DNase I	Enzyme	Cystic fibrosis	1994
Glucocerebrosidase	Enzyme	Gaucher's disease	1994
Hepatitis B subunit vaccine	Vaccine	Hepatitis prevention	1988
Hepatitis A vaccine	Vaccine		
Pertussis vaccine	Vaccine	Pertussis prevention	
Cholera vaccine	Vaccine	Cholera prevention	

In addition to the compound classes listed in the Table 2.2, several recombinant antibodies have also appeared on the market. So far these have been used for diagnostic purposes rather than as therapeutics, but therapeutic antibodies are expected to appear soon, and will probably, together with recombinant vaccines, form the fastest growing compound class. According to various sources, there are at present around 250 new proteins in advanced clinical trials. A sizable proportion of these represent improved versions of already known proteins, where specific modifications have been introduced to improve, for example, their pharmacokinetic properties. However, around 100 represent truly novel pharmaceutical substances with no precedent in medical therapy. It is estimated that around 2000 gene technology-based drugs, comprising both proteins and non-proteins, are in early stages of development. The information coming out of the human genome and EST projects is a very important resource for the selection of new potential protein drugs, and so-called DNA database mining has become a standard procedure in drug discovery both in major pharmaceutical companies and in small biotech companies.

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 30
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

Protein therapeutics are traditionally viewed as less convenient drugs than low-molecular therapeutics, from both the production and the user points of view. For the producer, the complexity of proteins poses challenges for purification, stability and administration. For the user, administration of protein therapeutics is more cumbersome because, unlike low-molecular weight drugs, proteins cannot yet be administered orally and usually have to be injected. Thus, at the same time as many resources are spent on discovering new therapeutic proteins, pharmaceutical companies are also trying to find low-molecular weight analogues that can supplant the known protein therapeutics. At a more basic research level, gene therapy is being explored as another substitute for protein therapeutics. In the long run, gene therapy may be able to replace protein therapeutics, but currently there are considerable technical obstacles (Sandhu *et al.*, 1997). Thus, in the near future protein therapeutics are expected to retain and increase their importance. With improvements in formulation and administration routes they may never become obsolete.

References

BIRCH, J.R. and FROUD, S.J., 1994, Mammalian cell culture systems for recombinant protein production, *Biologicals*, **22**(2), 127–133.

BLUNDELL, T.L., 1994, Problems and solutions in protein engineering – towards rational design, *Tibtech*, **12**, 145–148.

BRISTOW, A.F., 1993, Recombinant-DNA-derived insulin analogues as potentially useful therapeutic agents, *Tibtech*, **11**, 301–305.

CORTESE, R. (ed.), 1996, Combinatorial Libraries: Synthesis, Screening and Application Potential. Berlin, New York: Walter de Gruyter.

CREIGHTON, T.E., 1993, *Proteins: Structures and Molecular Properties*, 2nd edn. New York: W.H. Freeman & Co.

DOUGALL, W.C., PETERSON, N.C. and GREENE, M.I., 1994, Antibody-structure-based design of pharmacological agents, *Tibtech*, **12**, 372–379.

GEORGIOU, G. and VALAX, P., 1996, Expression of correctly folded proteins in *Escherichia coli*, *Current Opinion in Biotechnology*, **7**(2), 190–197.

GRIFFITHS, C.M. and PAGE, M.J., 1997, Production of heterologous proteins using the baculovirus/insect expression system, *Methods in Molecular Biology*, **75**, 427–440.

HODGSON, J. 1993, Expression systems: a user's guide, *Biotechnology*, **11**, 887–893.

INNIS, M.A., GELFAND, D.H., SNINSKY, J.J. and WHITE, T.J., 1990, *PCR Protocols. A Guide to Methods and Applications*. London: Academic Press.

ITAKURA, K., HIROSE, T., CREA, R., RIGGS, A.D., HEYNEKER, H.L., BOLIVER, F. and BOYER, H.-W., 1977, Expression in *E. coli* of a chemically synthesized gene for the hormone somatostatin, *Science*, **198**, 1056.

PharmaProjects, 1997, Richmond: PJB Publications.

ROMANOS, M.A., SCORER, C.A. and CLARE, J.J., 1992, Foreigh gene expression in yeast: a re-view, *Yeast*, **8**, 423–488.

ROWEN, L., MAHAIRAS, G. and HOOD, L., 1997, Sequencing the human genome, *Science*, **278**, 605–607.

SAMBROOK, J., FRITSCH, E.F. and MANIATIS, T., 1989, *Molecular Cloning. A Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

SANDHU, J.S., KEATING, A. and HOZUMI, N., 1997, Human gene-therapy, *Critical Reviews in Biotechnology*, **17**(4), 307–326.

VELANDER, W.H., LUBON, H. and DROHAN, W.N., 1997, Transgenic livestock as drug factories, *Scientific American*, **276**, 70–74.

VITETTA, E.S., THORPE, P.E. and UHR, J.W., 1993, Immunotoxins: magic bullets or misguided missiles? *Immunology Today*, **14**(6), 252–259.

WINTER, G. and HARRIS, W.J., 1993, Humanized antibodies, *Immunology Today*, **14**(6), 243–246.

WRIGHTON, N.C., FARRELL, F.X., CHANG, R., KASHYAP, A.K., BARBONE, F.P., MULCAHY, L.S., JOHNSON, D.L., BARRETT, R.W., JOLIFFE, L.K. and DOWER, W.J., 1996, Small peptides as potent mimetics of the protein hormone erythropoietin, *Science*, **273**, 458–463.

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 32
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

5

Chemical Pathways of Peptide and Protein Degradation

CHIMANLALL GOOLCHARRAN, MEHRNAZ KHOSSRAVI AND RONALD T. BORCHARDT

Department of Pharmaceutical Chemistry, The University of Kansas, Lawrence, Kansas, USA

5.1 Introduction
5.2 Hydrolytic pathways
5.2.1 Deamidation of Asn and Gln residues
5.2.2 Degradation of Asp residues
5.2.3 Degradation of N-terminal sequences containing penultimate Pro residues via diketopiperazine formation
5.3 Oxidation pathways
5.3.1 Autooxidation
5.3.2 Metal-catalysed oxidation
5.3.3 Photooxidation
5.3.4 Strategies to prevent oxidation
5.4 Other chemical pathways
5.4.1 β-Elimination reactions
5.4.2 Disulphide exchange reactions
5.5 Conclusion
References

5.1 Introduction

With the advent of recombinant DNA technology and modern synthetic chemistry, peptides and proteins now comprise a significant portion of drugs under clinical development. However, the instability of these macromolecules presents unique challenges to the pharmaceutical scientists who are responsible for developing stable and efficacious formulations or delivery systems for these molecules (Manning *et al.*, 1989; Ahern and Manning, 1992a, 1992b; Wang and Pearlman, 1993; Cleland *et al.*, 1993; Pearlman and Wang, 1996; Sanders and Hendren, 1997; Schöneich *et al.*, 1997).

Unlike small organic molecules, peptides and proteins possess not only a primary structure (sequence of amino acids), but also, in the case of proteins, higher order structures (e.g. secondary, tertiary, and quaternary), which are required for biological activity. Instability of small synthetic peptides involves primarily chemical pathways of degradation. With proteins, the degradation pathways also include chemical transformations,

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 33
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

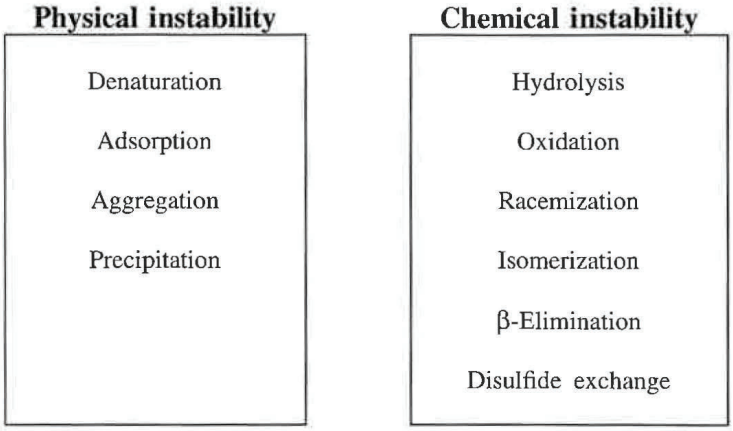


Figure 5.1 Major pathways of peptide and protein degradation.

which are similar to those observed in peptides, as well as physical instability (Manning *et al.*, 1989; Ahern and Manning, 1992a; Cleland *et al.*, 1993; Wang and Pearlman, 1993; Pearlman and Wang, 1996). Chemical instability can be defined as any process involving modification of the protein or peptide by bond formation (e.g. oxidation) or bond cleavage (e.g. deamidation), yielding a new chemical entity (Figure 5.1). In contrast, physical instability does not necessarily involve covalent modifications of the protein. Instead, it generally refers to changes in the higher order structure (secondary and above). This physical instability results in denaturation which can lead to adsorption to surfaces, aggregation and precipitation (Figure 5.1).

The primary objective of this chapter is to address the most prevalent chemical pathways for degradation of peptides and proteins observed during pharmaceutical production, purification, formulation and storage. For additional details about chemical, as well as physical, pathways of peptide and protein degradation, the reader is directed to several reviews (Manning *et al.*, 1989; Cleland *et al.*, 1993; Schöneich *et al.*, 1997) and books (Ahern and Manning, 1992a, 1992b; Wang and Pearlman, 1993; Pearlman and Wang, 1996; Sanders and Hendren, 1997).

5.2 Hydrolytic pathways

The most commonly observed hydrolytic reactions in peptides and proteins involve the side-chain amide groups of asparagine (Asn) and glutamine (Gln), and the peptide bond on the C-terminal side of an aspartic acid (Asp) or a proline (Pro) residue when it is in the penultimate position from the N-terminal end. These hydrolytic reactions are described in detail below.

5.2.1 *Deamidation of Asn and Gln residues*

Deamidation of side-chain amides of Asn and Gln residues in proteins and peptides involves intermolecular attack of water (Figure 5.2, e.g. Asp formation), intramolecular attack by the nitrogen of the N-terminal amine (Figure 5.3, e.g. pyroglutamate formation) or intramolecular attack by the nitrogen of the peptide bond on the C-terminal side (Figure 5.2, e.g. cyclicimide formation) of these amino acid residues (Schöneich *et al.*,

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 34
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

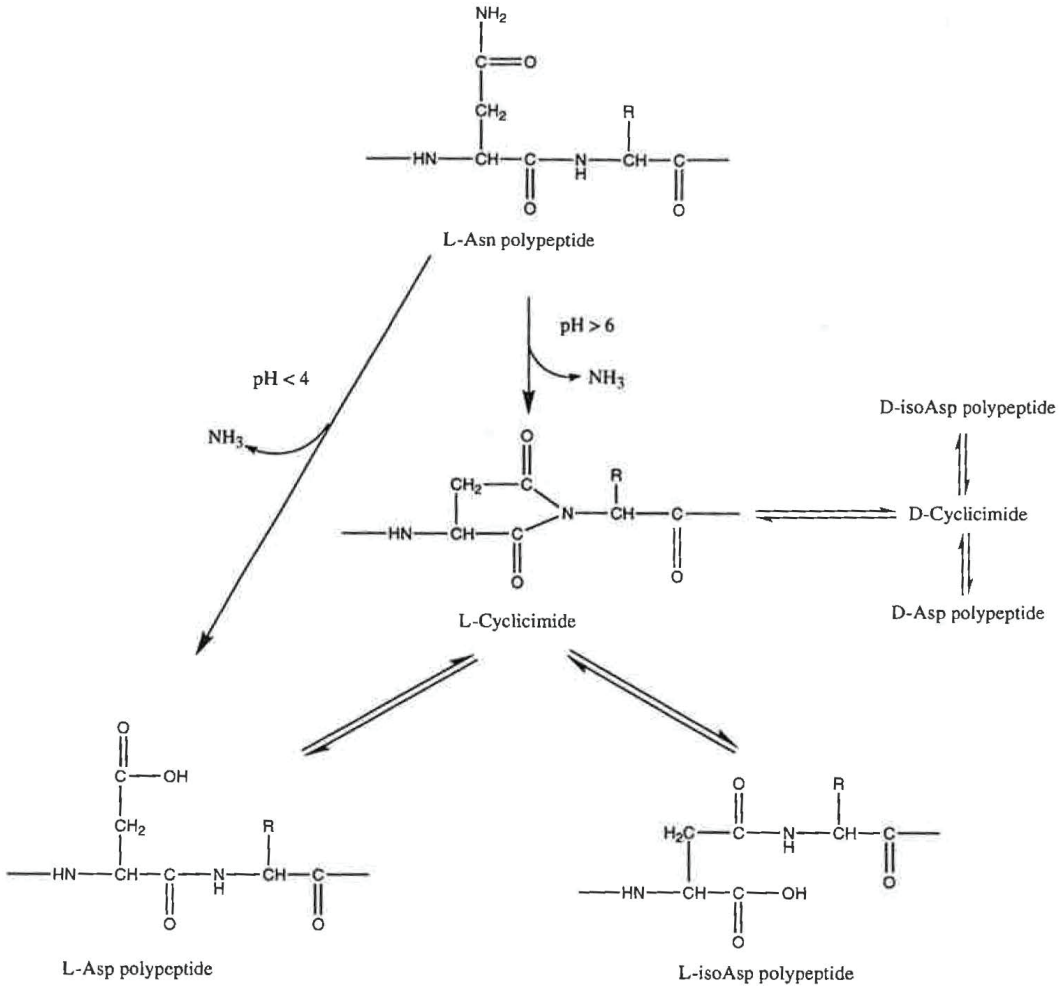


Figure 5.2 Pathways for the deamidation of Asn residues.

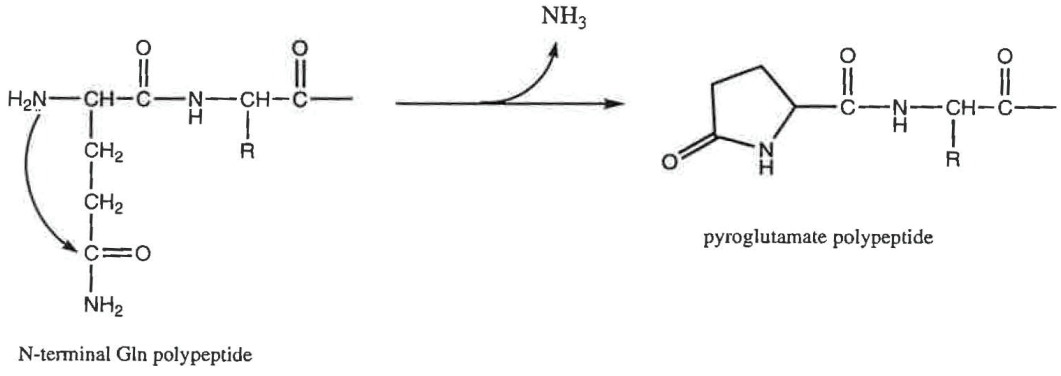


Figure 5.3 Formation of pyroglutamate from N-terminal Gln residue.

1997). Under acidic conditions (pH < 4.0), the predominant reaction with an Asn residue involves direct attack of water on the side-chain amide carbonyl carbon resulting in the formation of an Asp residue (Figure 5.2) (Patel and Borchardt, 1990a, 1990b). In contrast, at neutral or basic conditions (pH > 6.0), the predominant reaction with an Asn residue involves cyclicimide formation followed by attack of water on one of the two carbonyl

carbon atoms to yield Asp and isoaspartic acid (isoAsp) residues (Figure 5.2) (Stephenson and Clarke, 1989; Patel and Borchardt, 1990a). It should be noted that this cyclicimide intermediate can undergo racemization yielding L- and D-Asp- and L- and D-isoAsp-polypeptides as degradation products (Geiger and Clarke, 1987). In general, Asn residues deamidate much more rapidly than Gln residues (Manning *et al.*, 1989; Cleland *et al.*, 1993; Schöneich *et al.*, 1997). However, Gln residues located at the N-terminal end of a polypeptide chain can rapidly deamidate to form a pyroglutamate residue (Figure 5.3).

Many factors, both exogenous (i.e. pH, buffer concentration, buffer species) and endogenous (i.e. primary sequence and secondary and higher ordered structures) can influence the rate of deamidation of Asn residues and the distribution of the degradation products (iso-Asp, Asp). The effects of exogenous factors on the kinetics of deamidation have been extensively studied by our laboratory using an Asn-hexapeptide, Val-Tyr-Pro-Asn-Gly-Ala, residues 22–27 of porcine adrenocorticotropic hormone (ACTH). The most significant exogenous factor determining the rate of deamidation of Asn residues is pH. Using the Asn-hexapeptide described above, we have shown that the degradation is pH-dependent and follows first-order kinetics (Patel and Borchardt, 1990a). Under acidic conditions (pH 1–2), this Asn-hexapeptide undergoes deamidation via direct attack of water on the Asn side-chain carbonyl carbon atom, followed by expulsion of ammonia and formation of an Asp-hexapeptide (Figure 5.2). Interestingly, the formation of a cyclicimide as a minor product was also observed under acidic conditions, although its formation was much slower than the direct hydrolysis reaction. The cyclicimide intermediate was shown experimentally to be highly stable in acidic medium, and under these conditions it does not break down to form the Asp- and isoAsp-hexapeptides. Therefore, the degradation of this Asn-hexapeptide under acidic conditions does not lead to the formation of an isoAsp degradation product. Under basic and neutral conditions (pH 5–12), the Asn-hexapeptide degrades via the formation of a cyclicimide which then undergoes further hydrolysis to produce a mixture of Asp- and isoAsp-hexapeptides in a ratio of approximately 1 to 4, respectively (Figure 5.2). The rate constants of deamidation at basic pH values were greater than the rate constants at acidic pH. The maximal stability of this Asn-hexapeptide was observed at pH 3–4 at 37°C. At all pH values, the ratio of Asp to isoAsp degradation products was independent of buffer and buffer concentrations.

Other exogenous factors reported to influence the rate of deamidation of Asn residues include temperature, buffer species, buffer concentration, ionic strength and formulation excipients (e.g. lactose, mannitol) (Oliyai and Borchardt, 1994b; Oliyai *et al.*, 1994a; Schöneich *et al.*, 1997).

The effect of endogenous factors on the rate of deamidation has been determined by several investigators. Using the model hexapeptides, Val-Tyr-X-Asn-Z-Ala, various laboratories have shown that the major endogenous factor that influences the rate of deamidation is the structural characteristics (i.e. size) of the residue on the C-terminal side (Z) of the Asn residue. For example, determination of the rates of deamidation of the series Val-Tyr-Pro-Asn-Z-Ala, in 0.1 M phosphate buffer (pH 7.4) at 37°C, generated half-lives as follows: Z = Gly, 1.5 days; Z = Ser, 6.8 days; Z = Ala, 20.2 days; Z = Leu, 70 days; and Z = Val, 106 days. As expected, as the steric bulk of the residue on the C-terminal side of Asn residue increased, the rate of deamidation decreased. At pH 1.0, the C-terminal residue (Z) was shown to have no effect on the rate of deamidation (Stephenson and Clarke, 1989; Patel and Borchardt, 1990b). Modifying the residue on the N-terminal side of an Asn residue had little or no effect on the rate of the reaction at acidic, neutral or basic pH values (Patel and Borchardt, 1990b). These observations reinforce the idea that at acidic pH, deamidation occurs by the direct hydrolysis of the side-chain amide of the

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 36
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

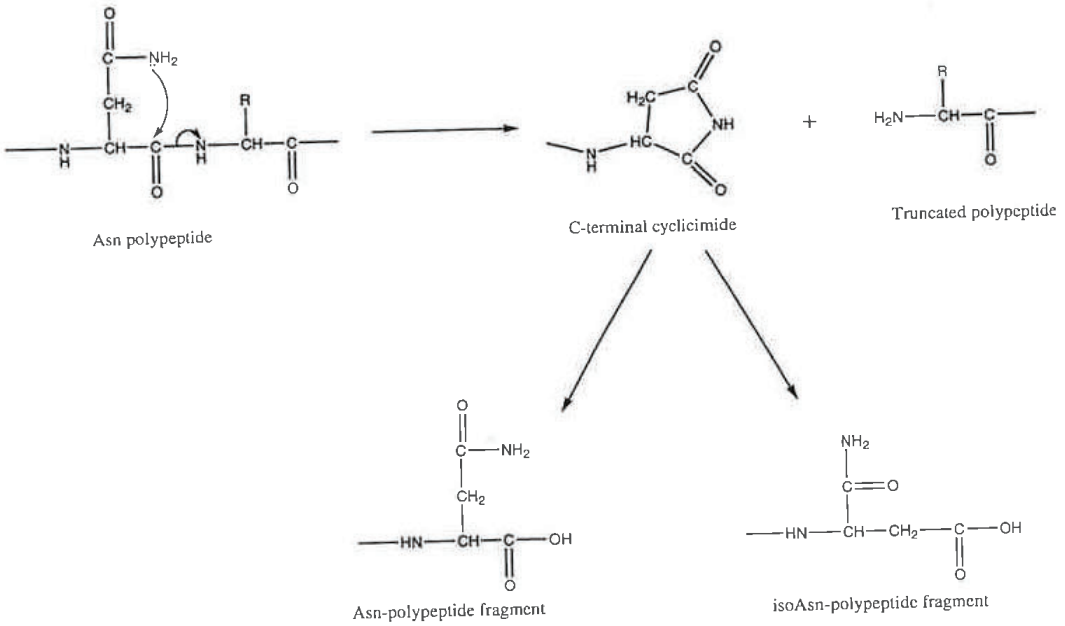


Figure 5.4 Pathways for the spontaneous fragmentation of Asn polypeptides.

Asn residue, while in the neutral-to-basic pH region, Asn residues deamidate via the formation of a cyclicimide intermediate.

Secondary and higher ordered structures can also affect the rates of deamidation of Asn residues in peptides and proteins. Clarke (1987) has shown that native proteins generally exist in conformations in which the nitrogen atom of peptide bond on the C-terminal side of Asn residues cannot approach the side-chain carbonyl carbon atom without large-scale conformational changes. Therefore, certain proteins will only undergo deamidation when they have been denatured and sufficient conformation flexibility exists to allow for the formation of the cyclicimide. Similar observations have been made recently by our laboratory using a series of synthetic growth-hormone-releasing factors (Stevenson *et al.*, 1993) and β-turn peptidomimetics (M. Xie, R.L. Schowen and R.T. Borchardt, unpublished data).

Another pathway of degradation of Asn residues in peptides and proteins involves spontaneous fragmentation (Aswad, 1994). In this process, the side-chain amide nitrogen of the Asn residue attacks the peptide bond carbonyl carbon atom on the C-terminal side, resulting in the cleavage of the peptide bond (Figure 5.4). Of the two fragments produced, one possesses a C-terminal succinimide residue that can undergo further hydrolysis, yielding a mixture of Asn- and isoAsn-peptides. This reaction does not result in the release of ammonia and occurs predominantly, although not exclusively, in peptides in which the Asn residue is followed by a Pro residue.

While extensive research efforts have been devoted to understanding deamidation in aqueous media, little is known about deamidation in the solid state. In the solid state, the kinetics and mechanisms of this reaction are complex, usually involving concomitant existence of other chemical and physical instability processes. Using the hexapeptide Val-Tyr-Pro-Asn-Gly-Ala, Oliyai *et al.* (1994a) have shown that the pH of the lyophilization solution, the moisture level of the lyophilized product, and the storage temperature are important factors in the solid-state stability of this peptide. With a pre-lyophilization solution at pH 3.5, the lyophilized hexapeptide deamidates via direct hydrolysis to produce

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 37
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

the Asp-hexapeptide. The Asp-hexapeptide then undergoes hydrolysis at the Asp-Gly amide bond to generate a small quantity of the tetrapeptide Val-Tyr-Pro-Asp. A significant amount of the cyclicimide was also observed. As the pH, temperature and moisture level were increased, the amount of cyclicimide increased. More recently, our laboratory (M.C. Lai, E.M. Topp, R.L. Schowen, M.J. Hageman and R.T. Borchardt, unpublished data) has determined the effect of polymers (e.g. poly(vinylalcohol), poly(vinylpyrrolidone)) and moisture on the rate of deamidation of the hexapeptide Val-Tyr-Pro-Asn-Gly-Ala. These studies have shown that in solid-state preparations, water can be involved in deamidation of this hexapeptide as either a medium, a reactant or a plasticizer. The plasticizing effect of water obviously affects the mobility of the reactants including the polypeptide. In these studies, interesting differences were also observed in the reactivity of the hexapeptide to deamidation in different polymers, suggesting that prediction of solid-state stability from solution stability data would be problematic.

5.2.2 *Degradation of Asp residues*

It has been established that Asp residues are also potential 'hot spots' in peptides and proteins for chemical degradation. The stability of an Asp residue has been shown to be dependent on its location in the primary sequence and the protein's higher-ordered structure (Ota and Clarke, 1989). Asp residues can degrade by two different pathways depending on the pH of the solution. One pathway involves Asp-to-isoAsp interconversion (Figure 5.5). This pathway proceeds through a cyclicimide intermediate similar to that observed in the deamidation of Asn residues. Depending on the pH, this cyclicimide may or may not degrade further to form isoAsp-products. The cyclicimide intermediate is prone to racemization at the α-carbon, resulting in the potential formation of D,L-cyclicimide-, D,L-Asp- and D,L-isoAsp-containing degradation products (Figure 5.5). The other pathway of Asp degradation involves the hydrolysis of Asp-X and/or X-Asp amide bonds (Figure 5.6). In dilute acid, the rate of hydrolysis of the peptide bonds adjacent to the Asp residue is at least 100 times higher than that of other peptide bonds (Schultz, 1967). The enhanced rate of hydrolysis is attributed to intramolecular catalysis by the side-chain carboxyl group where the X-Asp amide bond cleavage proceeds via a six-membered ring intermediate while the Asp-X bond hydrolysis proceeds via a five-membered ring (Figure 5.6).

Many exogenous and endogenous factors can influence the rate of Asp degradation and the distribution of products. In our laboratory, we have studied the effects of exogenous factors such as pH, buffer species and concentration, and temperature on the kinetics and mechanism of degradation of Asp residues using the model hexapeptide Val-Tyr-Pro-Asp-Gly-Ala (Oliyai and Borchardt, 1993). The rate of the Asp-hexapeptide degradation and the product distribution was shown to be highly dependent on the pH of the solution. Under strongly acidic conditions (pH 1.0–2.0), this hexapeptide degraded predominantly via intramolecular hydrolysis of the Asp-Gly amide bond, producing a tetrapeptide, Val-Try-Pro-Asp, and a dipeptide, Gly-Ala. A small amount of the cyclicimide (Asu-hexapeptide) was also generated which was stable under acidic conditions. As the pH of the solution was increased, the formation of the cyclicimide–hexapeptide became more predominant relative to the hydrolysis of the Asp-Gly amide bond. For example, at pH 4.0, most of the peptide (60%) rearranged to produce the cyclicimide–hexapeptide, which then further hydrolysed to form the isoAsp-hexapeptide, whereas the Asp-Gly peptide bond hydrolysis constituted only approximately 10% of the total degradation products. At

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 38
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

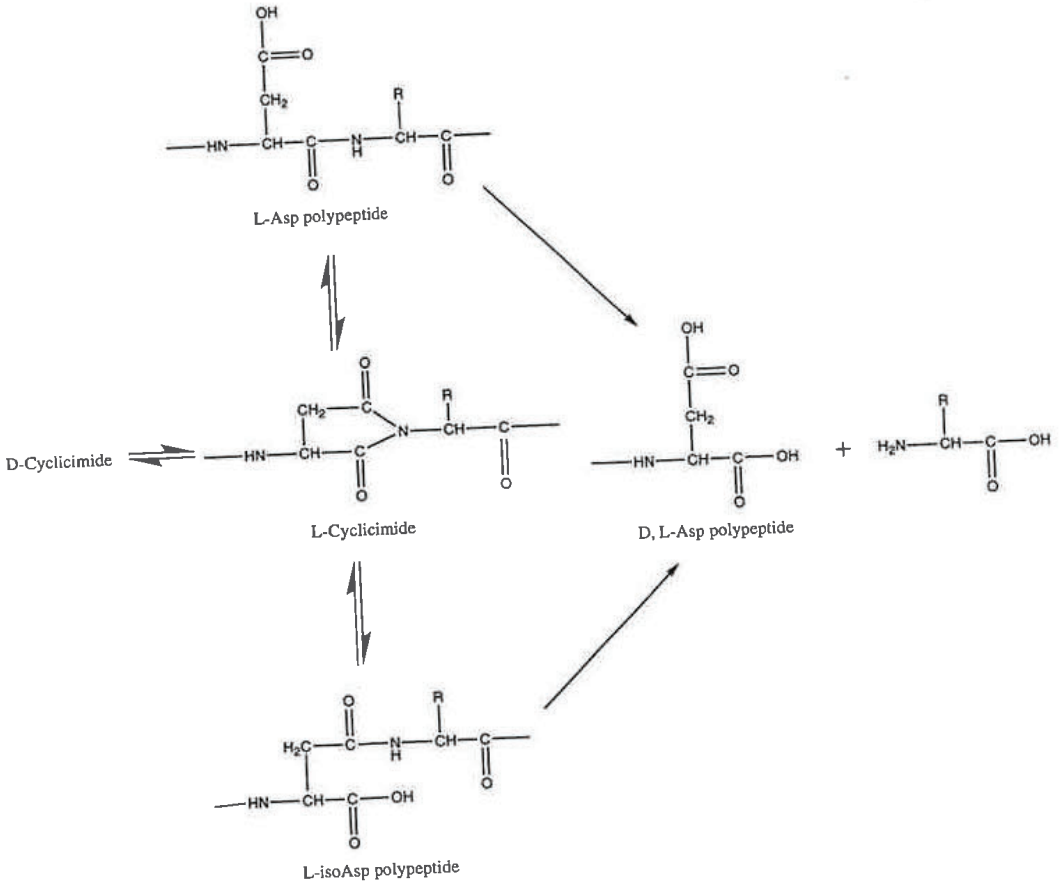


Figure 5.5 Pathways for the degradation of Asp-polypeptides.

pH above 6.0, the contribution from the hydrolysis of the Asp-Gly peptide bond to the overall degradation was negligible. At these pH values, the peptide degraded exclusively to the cyclicimide–hexapeptide, which being extremely unstable in the neutral and basic environment rapidly hydrolysed to the isoAsp- and Asp-hexapeptides. While not determined in this study, the L-cyclicimide hexapeptide should be prone to racemization at the α-carbon, potentially generating the corresponding D-cyclicimide-, D-Asp- and D-isoAsp-hexapeptides (Figure 5.5). Other exogenous factors reported to affect the rates of Asp degradation are temperature (Oliyai and Borchardt, 1993) and dielectric constant of the solvent (Brennan and Clarke, 1993).

The effect of primary sequence on Asp degradation is significant, but not as obvious as that observed for the deamidation of Asn residues. Stephenson and Clarke (1989) have measured the rate of cyclicimide formation at pH 7.4 and 37°C in a series of hexapeptides, Val-Tyr-Pro-Asp-Y-Ala, and observed the following half-lives: Y = Gly, 40.8 days; Y = Ser, 168 days; and Y = Ala, 266 days. Oliyai and Borchardt (1994a) also examined the degradation of a similar series of hexapeptides, Y = Gly, Ser and Val, under acidic (pH 1.1) and alkaline (pH 10.0) conditions. At pH 1.1, the major route of degradation was hydrolysis of the Asp-Y amide bond. They observed that, as the steric bulk of the amino acid Y increased, the rate of hydrolysis decreased, i.e. the Asp-Gly hexapeptide hydrolysed 1.6 times faster than the Asp-Ser peptide and 2.3 times faster than the Asp-Val peptide. A minor amount of cyclicimide–hexapeptide was also observed with the Asp-Gly- and Asp-Ser-hexapeptides but not for the Asp-Val-hexapeptide. At pH 10.0, the

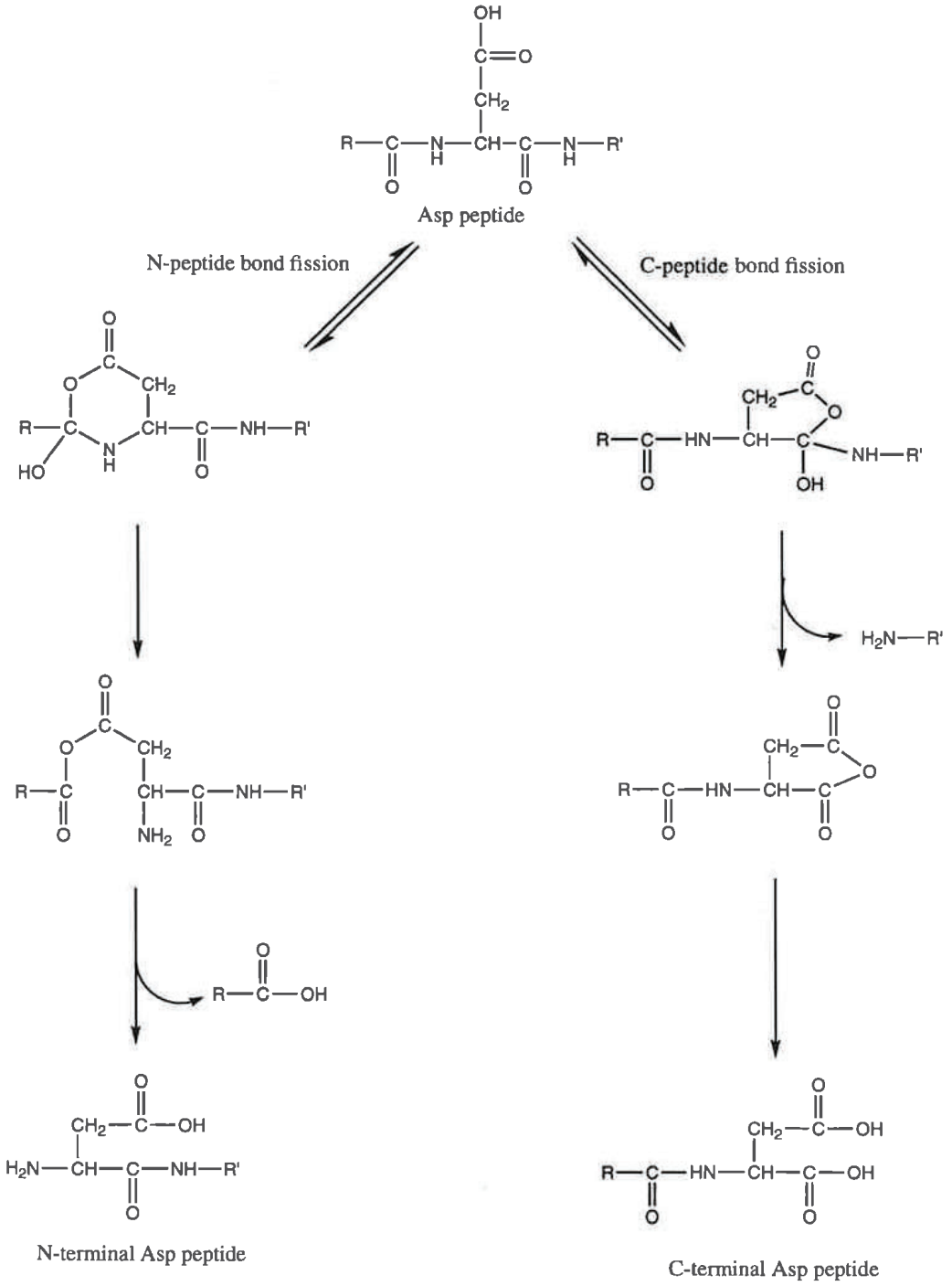


Figure 5.6 Degradation of Asp-peptides in acidic media.

Asp-peptides predominantly degraded to the isoAsp-peptides. As expected, the formation of the cyclicimide was highly sensitive to modifications made in the amino acid located on the C-terminal of Asp. For example, replacing Gly with Ser resulted in a four-fold decrease in the rate of cyclicimide formation. Similarly, substituting Gly with Val further decreased the rate of cyclicimide formation to such an extent that it was not detected.

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 40
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

Furthermore, the degradation of the Asp-Gly- and Asp-Ser-hexapeptides resulted in an isoAsp-to-Asp ratio of 4–5, while the ratio for the Asp-Val-hexapeptide was less than 1. The Asp-Val-hexapeptide also underwent more rapid racemization. In general, the formation of the cyclicimide in an Asn-peptide occurs faster than that of the corresponding Asp-containing peptide (Schöneich *et al.*, 1997).

Degradation of Asp residues in polypeptides in the solid state has been examined by Oliyai *et al.* (1994b) as a function of the pH of the pre-lyophilized solution, moisture level, temperature, and type of bulking agent (e.g. lactose, mannitol) used. The degradation of the Asp-hexapeptide was shown to be dependent on the pH of the pre-lyophilized solution and the moisture content of the lyophilized product. At acidic conditions, the lyophilized peptide degraded predominantly to the cyclicimide–hexapeptide, regardless of the type of excipient present. As the pH of the pre-lyophilized solution was increased, the extent of hydrolysis of the cyclicimide to generate the isoAsp-hexapeptide increased. Thus, at high pH (8.0), the major product of degradation was the isoAsp-hexapeptide. The type of excipient (lactose vs mannitol), amorphous vs crystalline, did not influence the pathway of degradation of the Asp residue, but affected the rate of peptide degradation; i.e. the amorphous formulations were consistently more chemically stable at all pH values, temperatures and moisture levels than were the crystalline formulations.

5.2.3 *Degradation of N-terminal sequences containing penultimate Pro residues via diketopiperazine formation*

Peptides and proteins that possess an N-terminal sequence in which Pro is the penultimate residue undergo non-enzymatic hydrolysis yielding a diketopiperazine (DKP), which arises from the first two amino acids, and a truncated polypeptide. The mechanism of DKP formation involves nucleophilic attack of the N-terminal nitrogen on the carbonyl carbon of the peptide bond between the second and third amino acid residues in the primary sequence. This intramolecular aminolysis reaction occurs readily in aqueous solution and was shown to be catalysed in both acidic and basic conditions (Goolcharran and Borchardt, 1998). Diketopiperazine formation was reported to occur in human growth hormone (Battersby *et al.*, 1994), bradykinin (Straub *et al.*, 1995) and histrelin (Oyler *et al.*, 1991).

Goolcharran and Borchardt (1998) have examined the effect of exogenous (i.e. pH, temperature, buffer species and concentration) and endogenous (i.e. primary sequence) factors on the rate of DKP formation using the dipeptide analogue X-Pro-*p*-nitroaniline (X-Pro-pNA). The kinetics of Phe-Pro-pNA degradation was studied in aqueous buffer solution at 37°C over the pH range 1–10. There are two possible pathways for the degradation of Phe-Pro-pNA; i.e. intramolecular aminolysis yielding the corresponding Phe-Pro-DKP or direct hydrolysis of the Pro-DKP amide bond, which produces the dipeptide Phe-Pro-OH. The Phe-Pro-DKP could then further undergo hydrolysis at either of the two carbonyl groups, producing two different dipeptides, i.e. Phe-Pro-OH and Pro-Phe-OH. In the pH range studied, Phe-Pro-DKP was the only product generated upon degradation of Phe-Pro-pNA. At pH values between 3 and 8, Phe-Pro-DKP was stable. In contrast, at pH values < 3 and > 8, Phe-Pro-DKP undergoes hydrolysis to the dipeptide, Phe-Pro-OH. The rate of disappearance of Phe-Pro-pNA followed pseudo-first-order kinetics and exhibited significant dependence on pH. Buffer catalysis was observed above pH 5.0, but little or no catalysis was observed below pH 4.5. From these data, a pH–rate profile was generated which suggested that the rate of DKP formation depends on the

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 41
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

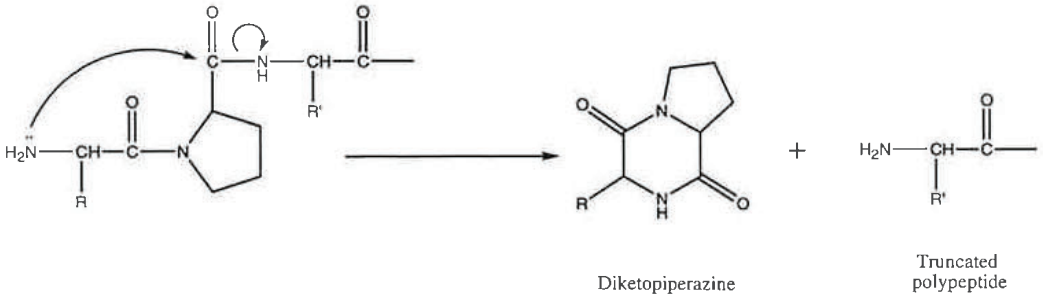


Figure 5.7 Degradation of N-terminal sequences containing penultimate Pro residues via diketopiperazine formation.

degree of ionization of the N-terminal amino group. It is also apparent that the unprotonated N-terminal amino group is much more reactive than the protonated form. At pH < 3, where the reactant exists predominantly in the protonated form, the rate of degradation is independent of pH. From pH 4 to 6, the rate increases with pH, but it is not first-order in hydroxide ion concentration (slope = 0.73). This may be due to the complexity of the mechanism involving different ionic species and buffer catalysis. Between pH 6 and 8 there is a plateau. In this region, the free N-terminal amino group is available for cyclization, but the hydroxide ion concentration is not high enough; thus, degradation occurs by the water-catalysed pathway. In the pH region 9 to 10, the slope of the pH–rate profile is unit positive, indicating specific-hydroxide ion catalysis. A pKa value of 6.1 was kinetically determined for the amino group.

Modification of the amino acid residue located on the N-terminal side of Pro was shown to have a major influence on the rate of DKP formation (Goolcharran and Borchardt, 1998). For example, with the series of dipeptide analogues X-Pro-pNA, the half-lives of DKP formation in 0.05 M phosphate buffer (pH 7.0) at 37°C were reported as follows: X = Gly, 5.1 days; X = Ala, 1.1 days; X = Val, 2.5 days; X = Phe, 0.5 days; X = β-cyclohexylalanine, 0.8 days; and X = Arg, 0.7 days. The effect of alkyl and aryl substituents in intramolecular reactions has been well documented, with the results suggesting that an increase in the bulkiness of the substituent increases the rate of cyclization (Borchardt and Cohen, 1972). This does not seem to be true for the series studied above, as the Ala analogue cycles twice as fast as the bulkier Val analogue. Thus, simple steric bulk cannot be used to explain the effect of the N-terminal residue on the rate of DKP formation. The effect may be due to contributions from several different factors, including the ability of the X-Pro peptide bond to undergo *cis–trans* isomerization.

5.3 Oxidation pathways

Oxidation is one of the most commonly observed chemical degradation pathways of peptides and proteins (Stadtman, 1992; Li *et al.*, 1995; Schöneich *et al.*, 1997). Among the amino acids most susceptible to oxidative modification are those that contain a sulphur (e.g. Met and Cys) or an aromatic ring (e.g. His, Tyr, and Trp) (Stadtman, 1992; Li *et al.*, 1995; Schöneich *et al.*, 1997). There are a number of mechanisms that may result in oxidative modification of amino acids (Stadtman, 1993). Despite the variations in conditions used to induce oxidation, the underlying principle involves the activation of oxygen to generate a few key reactive oxygen species, which ultimately then react with

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 42
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

the protein or peptide. In the following text, brief overviews describing the various characteristics of autooxidation, metal-catalysed oxidation, and photooxidation are provided.

5.3.1 Autooxidation

Autooxidation refers to the direct reaction between ground state molecular oxygen and a compound in the absence of any catalytic processes (Schöneich *et al.*, 1997). True autooxidation is of minor importance as a degradation pathway of proteins and peptides for the following reason. Ground state oxygen exists in the triplet state, i.e. there are two unpaired electrons of parallel spin direction in each of the two antibonding orbitals of oxygen. The majority of organic molecules differ from oxygen in that they are in the singlet ground state, i.e. their valence electrons are in paired antiparallel electron spins. Therefore, the direct reaction of oxygen in the triplet state with an organic molecule in the singlet state is a spin-restricted reaction and no significant bond formation is expected (Schöneich *et al.*, 1997). The triplet state configuration of oxygen permits the acceptance of a single electron but not a pair of electrons of antiparallel spin directions. Therefore, any significant oxidation reaction between ground state oxygen and an organic molecule in the singlet state is expected to occur via an outer-sphere electron transfer (Schöneich *et al.*, 1997).

The improbability of such an autooxidation reaction was shown by Schöneich *et al.* (1997) for a hypothetical formulation that contained a standard protein of molecular weight 22 000 in an air-saturated and metal- and peroxide-free aqueous buffer. The rates of autooxidation of the most labile amino acid residues, Cys and Met, were calculated for a pseudo-first-order process. The approximate half-life for the autooxidation of one deprotonated Cys residue was calculated to be 32 090 days. Moreover, the approximate half-life of the Met residue was calculated to be greater than 8.8×10^7 years. Given these calculations for a hypothetical protein formulation, it is apparent that true autooxidation of peptides and proteins is a very slow process that poses no significant threat to the stability of these compounds. However, many so-called cases of autooxidation are observed for peptides and proteins. In fact, it is likely that these cases of oxidation are metal-catalysed or light-induced.

5.3.2 Metal-catalysed oxidation

In the presence of a transition metal ion (i.e. Fe(III) and Cu(II)), oxygen, and an electron donor (reducing agent), metal-catalysed oxidation of proteins and peptides can occur at significant rates (Stadtman, 1993). Pharmaceutical processes such as protein/peptide synthesis, purification, storage of bulk drug, and storage of the dosage form can provide conditions leading to metal-catalysed oxidation reactions (Manning *et al.*, 1989; Cleland *et al.*, 1993). Metals such as Fe(III) and Cu(II) are commonly encountered as contaminants that may originate from reagents and, particularly, from buffers used in processing and formulation of proteins and peptides (Li *et al.*, 1995). Normally, transition metal ions in their oxidized states (i.e. Cu(II) and Fe(III)) do not readily react with oxygen to generate a more reactive oxygen species. An electron donor or reducing agent known as a prooxidant is needed to reduce the transition metal ion; this reduced form may then interact with oxygen to generate reactive oxygen species. The prooxidant in pharmaceutical formulations may arise from contaminants in buffers or may have been added to the

formulation, ironically, as an antioxidant (i.e. ascorbic acid) (Schöneich *et al.*, 1997). In addition, the side-chains of certain amino acid residues such as Trp, Cys, and Tyr found in proteins and peptides can act as prooxidants (Timmins *et al.*, 1982; Schöneich *et al.*, 1997). Equation 5.1 illustrates the prooxidant donating an electron to reduce the transition metal ion Fe(III).

$$\text{Fe(III)} + \text{prooxidant} \rightarrow \text{Fe(II)} + \text{oxidized prooxidant}^{-} \tag{5.1}$$

At this point, it is important to note that the transition metal ion being reduced may be ligated or complexed to appropriate ligands (Stadtman, 1993). Further reaction of the ligated or complexed reduced metal with molecular oxygen O_2 to generate reactive oxygen species occurs at the metal binding site (Stadtman, 1993). The generated reactive oxygen species reacts preferentially with an amino acid residue at the metal binding site before diffusion into the bulk solution. This feature makes metal-catalysed oxidation a site-specific process. That is, given a certain protein, one or at most only a few amino acid residues located at the metal-binding site are specific targets (Stadtman, 1993). As illustrated in equations 5.2 and 5.3, the reduced form of the transition metal ion, complexed by appropriate ligands, may donate an electron to reduce oxygen to form superoxide radical. In these equations, C_x represents a complex that may include a number of species in solution (i.e. amino acid residue(s), buffer, reducing agent, etc.). The pertinent reacting species of the complex (C_x) are illustrated individually.

$$O_2 + C_xFe(II) \rightarrow C_xFe(II)-O-O^{\cdot} \rightarrow C_xFe(III)-O-O^{-} \tag{5.2}$$

$$C_xFe(III)-O-O^{-} \rightarrow C_xFe(III) + O_2^{\cdot -} \tag{5.3}$$

Superoxide anion radical may then react with protons to generate hydrogen peroxide

$$2O_2^{\cdot -} + 2H^{+} \rightarrow H_2O_2 \tag{5.4}$$

It is also possible that hydrogen peroxide is directly generated

$$O_2 + C_xFe(II) \rightarrow C_xFe(II)-O-O^{\cdot} \tag{5.5}$$

$$C_xFe(II)-O-O^{\cdot} + C_xFe(II) + 2H^{+} \rightarrow 2C_xFe(III) + H_2O_2 \tag{5.6}$$

Hydrogen peroxide can then oxidize susceptible amino acids such as Cys or Met directly or it can undergo various other reactions to form a number of intermediate complexes with the reduced and oxidized forms of the metal. These intermediate complexes have been proposed to oxidize amino acids (Yamazaki and Piette, 1991; Stadtman, 1990). A reactive oxygen species commonly proposed to be the damaging species in metal-catalysed oxidation is the hydroxyl radical or an equivalent bound/complexed form of it. The hydroxyl radical may arise from a Fenton-type reaction where there is a one-electron reduction of hydrogen peroxide

$$C_xFe(II) + H_2O_2 \rightarrow C_xFe(III) + {}^{\cdot}OH + {}^{-}OH \tag{5.7}$$

The hydroxyl radical or an equivalent bound/complexed form may then react immediately with the amino acid at the metal-binding site before diffusion into bulk solution. It is important to note that other reactive oxygen species may be involved in metal-catalysed oxidation. At this time, only speculation about the exact nature of the reactive oxygen species has been possible because of the site-specific nature of the reaction. The amino acids most susceptible to metal-catalysed oxidation are His, Arg, Lys, Pro, Met, and Cys (Stadtman, 1993). Table 5.1 provides a list of the common degradation products arising from metal-catalysed reaction of these amino acids. As a specific example, the

Table 5.1 Common degradation products formed from metal-catalysed oxidation of amino acid residues

Oxidized residue	Products formed
His	2-oxo-His, aspartic acid, asparagine
Lys	2-amino-adipicsemialdehyde
Cys	–S–S–disulphide crosslinks
Pro	Glutamic acid, glutamic semialdehyde, 2-pyrrolidone, *cis/trans*-4-hydroxyproline
Arg	Glutamic semialdehyde
Met	Met-sulphoxide

reader is referred to a recent paper by Khossravi and Borchardt (1998) dealing with His oxidation by ascorbate/Cu(II)/O_2. Proposed mechanisms of interaction of the reactive oxygen species with susceptible amino acid residues to generate the degradation product can be found in literature references cited in a review by Stadtman (1993).

In metal-catalysed oxidation, the location of the amino acid in the protein, i.e. whether it is exposed or buried, is not a significant factor in its oxidation potential. Amino acids are most susceptible to metal-catalysed oxidation because their side-chains aid in the formation of favourable metal-binding sites in proteins, making them specific targets for oxidation. In contrast to the site-specific nature of metal-catalysed oxidation reactions, oxidation induced by light (photooxidation) or contaminants is of a non-site-specific nature. In such a mechanism, the surface area exposure of the susceptible amino acid becomes an issue in its oxidation potential (Shechter *et al.*, 1975). In other words, the oxidation potential of the amino acid residue depends on whether it is exposed or buried within the protein. An example of a non-site-specific oxidation by contaminants is the oxidation of Met by hydrogen peroxide. In the presence of hydrogen peroxide, Met in peptides and proteins is oxidized to Met sulphoxide (Nguyen *et al.*, 1993). Peroxide contaminants in formulations have been know to arise from surfactants such as polysorbate 80 (Hora *et al.*, 1991) and polyethylene glycols (MacGinity *et al.*, 1975).

5.3.3 Photooxidation

Some pharmaceutical proteins are photosensitive and, therefore, undergo oxidative modification when exposed to light. Light-induced oxidation of protein pharmaceuticals may occur during protein processing and storage (Li *et al.*, 1995). The detailed mechanism of photooxidation of protein pharmaceuticals is complex and beyond the scope of this chapter. Therefore, a simple introduction to the conditions that initiate photooxidation reactions is provided. Photooxidation is initiated when a compound absorbs a certain wavelength of light, which provides energy to raise the molecule to an excited state (Halliwell and Gutteridge, 1989). The excited molecule can then transfer the energy to oxygen, converting it to singlet oxygen, while returning to ground state. Alternatively, the excited molecule can react directly with other molecules. This compound, which in essence initiates photooxidation reactions, is often referred to as a photosensitizer (Halliwell and Gutteridge, 1989). Efficient photosensitizers commonly have certain structural characteristics, such as aromatic systems and/or high conjugation with appropriate chromophores, which favour

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 45
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

Table 5.2 Common degradation products formed from reaction of singlet oxygen with amino acid residues

Oxidized residue	Products formed
His	2,5-endoperoxide, aspartic acid, asparagine
Trp	N-formylkynurenine, kynurenine
Cys	–S–S–disulphide crosslinks, sulphonic acid
Met	Met-sulphoxide

the absorbance of light (Halliwell and Gutteridge, 1989). Some compounds that have acted to initiate light-induced reactions are phenothiazines (used as tranquillizers), tetracycline antibiotics, various porphyrins (free and bound to proteins), riboflavin, and bilirubin (Halliwell and Gutteridge, 1989). One of the reactive oxygen species commonly involved in photooxidation reactions is singlet oxygen. The electron configuration of singlet oxygen allows this reactive oxygen species to cause oxidative damage through direct interaction with molecules. The electrons in the highest occupied orbital of singlet oxygen are in antiparallel spin directions, in contrast to triplet oxygen, which has electrons in parallel spin directions. Therefore, singlet oxygen is not spin-restricted in its reaction with an organic molecule. The interaction of singlet oxygen with amino acid residues is investigated in laboratories by initiation of photooxidation reactions using efficient photosensitizer dyes such as methylene blue, rose bengal, and acridine orange (Foote, 1968; Halliwell and Gutteridge, 1989). The amino acids found to be most susceptible to these photooxidation reactions are His, Trp, Met, and Cys (Spikes and Straight, 1967; Halliwell and Gutteridge, 1989). A list of the most common degradation products reported to arise from the reaction of singlet oxygen with these amino acids is provided in Table 5.2 (Spikes and Straight, 1967; Halliwell and Gutteridge, 1989). Further information on the mechanism of photooxidation reactions can be found in references provided in a book by Halliwell and Gutteridge (1989) and a review by Foote (1968).

5.3.4 *Strategies to prevent oxidation*

There are several strategies for stabilization of compounds against oxidation. The three main strategies of stabilization covered here are intrinsic, physical, and chemical.

An intrinsic method of stabilization utilizes site-directed mutagenesis. Stabilization of the protein towards oxidation is achieved by directly replacing the oxidation-labile amino acid. It is important to understand that a major shortcoming of this method of stabilization is that the structural and functional characteristics of the protein may be altered by replacement of the amino acid residue(s).

The physical method relies on the physical state of the formulation (solid vs liquid) as a basis for stabilization. Oxidation may occur in the solid-lyophilized or liquid-solution states of the protein. However, oxidation is usually slower in the solid state (Kenarke and Richards, 1966). Therefore, stabilization of the protein towards oxidation may be achieved by choosing a lyophilized solid-state formulation during storage.

Chemical additives may be utilized as another method of stabilization. Two of the most common are antioxidants and chelating agents. Antioxidant additives tend to end free

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 46
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

radical reactions by reacting with radicals to terminate the chain. However, care should be taken that the antioxidant (i.e. thiol, ascorbic acid) does not function as a prooxidant in the presence of trace metals to initiate metal-catalysed oxidation reactions. Chelating agents are used to remove metals such as iron and copper from pharmaceutical formulations, and usually tend to protect proteins from metal-catalysed reactions. However, chelates have been shown to induce rather than prevent oxidation in certain situations (Zhao *et al.*, 1996). For example, Fe(III)–EDTA chelate has been observed to induce peptide oxidation (Zhao *et al.*, 1996). Overall, it is important that special care is taken in choosing the right means of stabilization of proteins and peptides against oxidation.

5.4 Other chemical pathways

Several other chemical reactions have been shown to be important degradation pathways for peptides and proteins. These include β-elimination reactions, which can also lead to racemization, and disulphide exchange reactions, which can result in incorrect pairing of disulphide bonds and, consequently, affect the three-dimensional structure of the protein. Each of these types of reaction is discussed below.

5.4.1 *β-Elimination reactions*

β-Elimination is another pathway of peptide/protein degradation. Amino acid residues that undergo β-elimination reactions include Cys, Ser, Thr, Phe, and Lys (Manning *et al.*, 1989). β-Elimination reactions result from the abstraction of a proton from the α-carbon of an amino acid residue in a polypeptide chain resulting in the formation of a carbanion intermediate. Addition of a proton to the opposite face of the molecule can lead to racemization. Alternatively, the carbanion intermediate can undergo further reaction to form a dehydroalanine residue (R = H) (Figure 5.8). β-Elimination reactions have been observed in a number of proteins, i.e. lysozyme (Nashef *et al.*, 1977) and bovine pancreatic ribonuclease A (Zale and Klibanov, 1986).

5.4.2 *Disulphide exchange reactions*

Disulphide bonds are of great importance to the structural stability of many proteins, as they are the most frequently encountered covalent crosslinks in proteins (Kosen, 1992). The disulphide bond may serve to join two independent polypeptide chains in an intermolecular fashion or may form between two Cys residues of one polypeptide chain in an intramolecular fashion. Thus, the interchange and/or cleavage of disulphide bonds can lead to an altered protein three-dimensional structure, which may result in the loss of biological activity (Kosen, 1992).

The stability of the disulphide bond is threatened by the presence of catalytic amounts of thiols. The thiolate ion is a reactive species that acts as a nucleophile attacking the sulphur atom of the disulphide resulting in disulphide exchange at neutral-to-alkali pH values

$$R^*S^- + R''\text{-S-S-}R'' \rightleftharpoons R^*\text{-S-S-}R'' + R''S^- \quad (5.8)$$

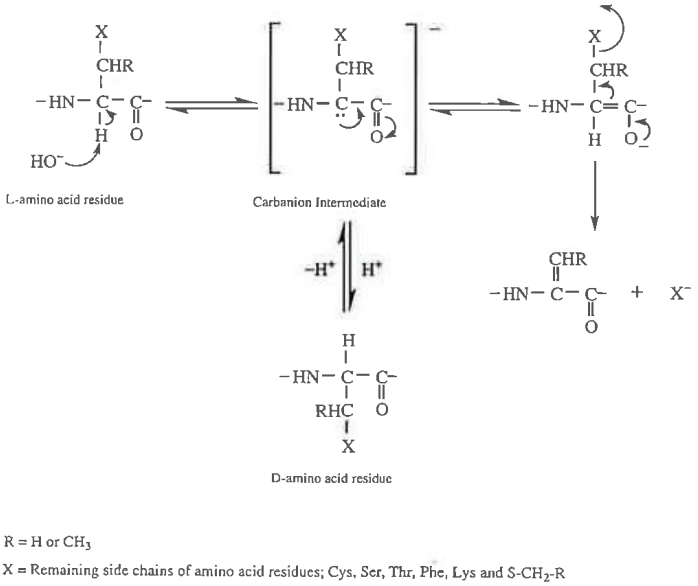


Figure 5.8 Mechanism of β-elimination reactions of amino acid residues at alkali pH.

The rate of reaction is greatest at alkali pH values, since the thiolate ion will predominate at higher pH values. An example of a protein that undergoes thiol disulphide exchange reactions is lysozyme. A more thorough discussion of the conditions leading to this reaction may be found in references in Kosen (1992).

Disulphide exchange reactions may also occur in acidic media. Benesch and Benesch (1958) proposed a mechanism of disulphide exchange in acidic media. The reaction intermediate is the sulphenium cation (R^*S^+), which may be formed by heterolytic cleavage of a protonated disulphide bond

$$R^*\text{-S-S-}R^* + H^+ \rightleftharpoons [R^*\text{-S-SH}^+\text{-}R^*] \rightleftharpoons R^*S^+ + R^*SH \quad (5.9)$$

The electrophilic R^*S^+ may then react with the sulphur atom of a disulphide bond leading to disulphide exchange

$$R^*S^+ + R'\text{-S-S-}R' \rightleftharpoons R^*\text{-S-S-}R' + R'\text{-}S^+ \quad (5.10)$$

Stabilization of the reaction at acidic pH values may be achieved by the addition of an external thiol that can scavenge R^*S^+ (Benesch and Benesch, 1958). However, the reaction mechanism that occurs at neutral-to-alkali pH values (equation 5.8) is catalysed by the addition of an external thiol. Inhibition of thiol disulphide exchange reactions at neutral pH can be achieved by the addition of efficient thiol scavengers such as *p*-mercuribenzoate and N-ethylmaleimide (Zale and Klibanov, 1986).

5.5 Conclusion

The chemical instability of proteins and peptides poses a problem in their development as pharmaceuticals. A better understanding of the underlying mechanisms of instability of these complex molecules is essential to provide knowledge such that optimum stability conditions can be achieved in the pharmaceutical development processes.

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 48
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

References

AHERN, T.J. and MANNING, M.C. (eds), 1992a, *Stability of Protein Pharmaceuticals. Part A: Chemical and Physical Pathways of Protein Degradation*, New York: Plenum Press.

(eds), 1992b, *Stability of Protein Pharmaceuticals. Part B: In Vivo Pathways of Degradation and Strategies for Protein Stabilization*, New York: Plenum Press.

ASWAD, D.W., 1984, Stoichiometric methylation of porcine adrenocorticotropin by protein carboxyl methyltransferase requires deamidation of asparagine 25, *Journal of Biological Chemistry*, **259**, 10714–10721.

(ed.), 1994, *Deamidation and Isoaspartate Formation in Peptides and Proteins*, Baco Raton: CRC Press.

BATTERSBY, J.E., HANCOCK, W.S., CONNOVA-DAVIS, E., OESWEIN, J. and O'CONNOR, B., 1994, Diketopiperazine formation and N-terminal degradation in recombinant human growth hormone, *International Journal of Peptide and Protein Research*, **44**, 215–222.

BENESCH, R.E. and BENESCH, R., 1958, The mechanism of disulfide interchange in acidic solution; role of sulfenium ions, *Journal of the American Chemical Society*, **80**, 1666–1669.

BORCHARDT, R.T. and COHEN, L.A., 1972, Stereopopulation control. II. Rate enhancement of intramolecular nucleophilic displacement, *Journal of the American Chemical Society*, **94**, 9166–9174.

BRENNAN, T.V. and CLARKE, S., 1993, Spontaneous degradation of polypeptides at aspartyl and asparaginyl residues: effects of the solvent dielectric, *Protein Science*, **2**, 331–338.

CLARKE, S., 1987, Propensity of spontaneous succinimide formation from aspartyl and asparaginyl residues in cellular proteins, *International Journal of Peptide and Protein Research*, **30**, 808–821.

CLELAND, J.L., POWELL, M.F. and SHIRE, S., 1993, The development of stable protein formulations: a close look at protein aggregation, deamidation, and oxidation, *Critical Reviews in Therapeutic Drug Carrier Systems*, **10**, 307–377.

FOOTE, C.S., 1968, Mechanisms of photosensitized oxidation, *Science*, **162**, 963–970.

GEIGER, T. and CLARKE, S., 1987, Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides, *Journal of Biological Chemistry*, **262**, 785–794.

GOOLCHARRAN, C. and BORCHARDT, R.T., 1998, Formation of diketopiperazine in model peptides, *Journal of Pharmaceutical Sciences*, **87**, 283–288.

HALLIWELL, B. and GUTTERIDGE, J.M., 1989, The chemistry of oxygen radicals and other oxygen-derived species. *Free Radicals in Biology and Medicine*, 2nd edn, pp. 22–85, Oxford: Clarendon Press.

HORA, M.S., RANA, R.K., WILCOX, C.L., HIRTZER, P., WOLFE, S.N. and THOMPSON, J.W., 1991, Development of a lyophilized formulation of interleukin-2, *Developments in Biological Standardization*, **74**, 295–306.

KENARKE, U.W. and RICHARDS, F.M., 1966, The histidyl residues in ribonuclease-S. Photooxidation in solution and in single crystals; the iodination of histidine-12*, *Journal of Biological Chemistry*, **241**, 3197–3206.

KHOSSRAVI, M. and BORCHARDT, R.T., 1998, Chemical pathways of peptide degradation. IX. Metal-catalyzed oxidation of histidine in model peptides, *Pharmaceutical Research*, **15**, 1096–1102.

KOSEN, A.P., 1992, Disulfide bonds in proteins. *Stability of Protein Pharmaceuticals. Part A: Chemical and Physical Pathways of Protein Degradation*, edited by T.J. AHERN and M.C. MANNING, pp. 31–59, New York and London: Plenum Press.

LI, S., SCHÖNEICH, C. and BORCHARDT, R.T., 1995, Chemical instability of protein pharmaceuticals: mechanisms of oxidation and strategies for stabilization, *Biotechnology and Bioengineering*, **48**, 490–500.

MACGINITY, J.W., HILL, J.A. and LA VIA, A.L., 1975, Influence of peroxide impurities in polyethylene glycols on drug stability, *Journal of Pharmaceutical Sciences*, **64**, 356–359.

MANNING, M.C., PATEL, K. and BORCHARDT, R.T., 1989, Stability of protein pharmaceuticals, *Pharmaceutical Research*, **6**, 903–918.

NASHEF, A.S., OSUGA, D.T., LEE, H.S., AHMED, A.I., WHITAKER, J.R. and FEENEY, R.E., 1977, Effects of alkali on proteins. Disulfides and their products, *Journal of Agricultural and Food Chemistry*, **25**, 245–251.

NGUYEN, T.H., BURNIER, J. and MENG, W., 1993, The kinetics of relaxin oxidation by hydrogen peroxide, *Pharmaceutical Research*, **10**, 1563–1571.

OLIYAI, C. and BORCHARDT, R.T., 1993, Chemical pathways of peptide degradation. IV. Pathways, kinetics, and mechanism of degradation of an aspartyl residue in a model hexapeptide, *Pharmaceutical Research*, **10**, 95–110.

1994a, Chemical pathways of peptide degradation. VI. Effect of the primary sequence on the pathways of degradation of aspartyl residues in model hexapeptides, *Pharmaceutical Research*, **11**, 751–758.

1994b, Chemical instability of proteins in solution and lyophilized formulation. *Protein Formulation and Delivery*, edited by J. CLELAND and R. LANGER, pp. 46–48, Washington, DC: American Chemical Society.

OLIYAI, C., PATEL, J., CARR, L. and BORCHARDT, R.T., 1994a, Solid-state chemical instability of an asparaginyl residue in a model hexapeptide, *Journal of Parenteral Science & Technology*, **48**, 167–173.

1994b, Chemical pathways of peptide degradation. VII. Solid state chemical instability of an aspartyl residue in a model hexapeptide, *Pharmaceutical Research*, **11**, 901–908.

OTA, I.M. and CLARKE, S., 1989, Calcium affects the spontaneous degradation of aspartyl/asparaginyl residues in calmodulin, *Biochemistry*, **28**, 4020–4027.

OYLER, A.R., NALDI, R.E., LLOYD, J.R., GRADEN, D.A. and SHAW, C.J., 1991, Characterization of the solution degradation products of histrelin, a gonadotropin releasing hormone (LH/RH) agonist, *Journal of Pharmaceutical Sciences*, **80**, 271–275.

PATEL, K. and BORCHARDT, R.T., 1990a, Chemical pathways of peptide degradation. II. Kinetics of deamidation of an asparaginyl residue in a model hexapeptide, *Pharmaceutical Research*, **7**, 703–711.

1990b, Chemical pathways of peptide degradation. III. Effects of primary sequence on the pathways of deamidation of asparaginyl residues in hexapeptides, *Pharmaceutical Research*, **7**, 787–793.

PEARLMAN, R. and WANG Y.J., 1996, *Formulation, Characterization, and Stability of Protein Drugs. Case Histories*, New York: Plenum Press.

SANDERS, L.M. and HENDREN R.W., 1997, *Protein Delivery: Physical Systems*, New York: Plenum Press.

SCHÖNEICH, C., HAGEMAN, M.J. and BORCHARDT, R.T., 1997, Stability of peptides and proteins. *Controlled Drug Delivery Challenges and Strategies*, edited by K. PARK, pp. 205–228, Washington, DC: American Chemical Society.

SHECHTER, Y., BURSTEIN, Y. and PATCHORNIK, A., 1975, Selective oxidation of methionine residues in proteins, *Biochemistry*, **14**, 4497–4503.

SCHULTZ, J., 1967, Cleavage at aspartic acid, *Methods in Enzymology*, **11**, 255–263.

SPIKES, J.D. and STRAIGHT R., 1967, Sensitized photochemical processes in biological systems, *Annual Review of Physical Chemistry*, **18**, 409–436.

STADTMAN, E.R., 1990, Metal ion-catalyzed oxidation of proteins: biochemical mechanism and biological consequences, *Free Radical Biology & Medicine*, **9**, 315–325.

1992, Protein oxidation and aging. *Science*, **257**, 1220–1224.

1993, Oxidation of free amino acids and amino acid residues in proteins by radiolysis and by metal-catalyzed reactions, *Annual Review of Biochemistry*, **62**, 797–821.

STEPHENSON, R.C. and CLARKE, S., 1989, Succinimide formation from aspartyl and asparaginyl peptides as a model for the spontaneous degradation of proteins, *Journal of Biological Chemistry*, **264**, 6164–6170.

STEVENSON, C.L., FRIEDMAN, A.R., KUBIAK, T.M., DONLAN, M.E. and BORCHARDT, R.T., 1993, Effect of secondary structure on the rate of deamidation of several growth hormone releasing factor analogs, *International Journal of Peptide and Protein Research*, **42**, 497–503.

STRAUB, J.A., AKIYAMA, A., PARMER, P. and MUSSO, G.F., 1995, Chemical pathways of the bradykinin analog, RMP-7, *Pharmaceutical Research*, **12**, 305–308.

TIMMINS, P., JACKSON, I.M. and WANG, Y.J., 1982, Factors affecting captopril stability in aqueous solution, *International Journal of Pharmaceutics*, **11**, 329–336.

WANG, Y.J. and PEARLMAN, R., 1993, *Stability and Characterization of Protein and Peptide Drugs. Case Histories*, New York: Plenum Press.

YAMAZAKI, I. and PIETTE, L.H., 1991, EPR spin trapping study on the oxidizing species formed in the reaction of the ferrous ion with hydrogen peroxide, *Journal of the American Chemical Society*, **113**, 7588–7593.

ZALE, S.E. and KLIBANOV, A.M., 1986, Why does ribonuclease irreversibly inactivate at high temperatures? *Biochemistry*, **25**, 5432–5444.

ZHAO, F., YANG, J. and SCHÖNEICH, C., 1996, Effects of polyaminocarboxylate metal chelators on iron-thiolate induced oxidation of methionine- and histidine-containing peptides. *Pharmaceutical Research*, **13**, 931–938.

6

Physical Stability of Proteins

JENS BRANGE

Brange Consult, Klampenborg, Denmark

6.1 Introduction
6.2 Protein structure
6.2.1 Stabilizing interactions
6.2.2 Role of water in structure and stability
6.3 Protein destabilization (denaturation)
6.3.1 Unfolding intermediates (molten globule)
6.3.2 Temperature-induced changes
6.3.3 Influence of pH
6.3.4 Influence of pressure
6.4 Aggregation and precipitation
6.4.1 Mechanisms of aggregation
6.4.2 Precipitation and fibrillation phenomena
6.4.3 Factors influencing aggregation and precipitation
6.5 Surface adsorption
6.6 Solid phase stability
6.6.1 Lyophilization-induced aggregation
6.7 Stabilization of protein drugs
6.7.1 Stabilization strategies
References

6.1 Introduction

Maintenance of the structural integrity of a therapeutic protein is essential for its efficacy in relation to physiological and pharmacological activity. Therefore, a major challenge confronting the pharmaceutical scientist working with protein production and formulation is the instability of the protein drug during processing, handling, storage and use. The structural changes taking place in a protein can be broadly classified as involving physical or chemical alterations of the molecule. In practice, however, these two distinct routes of transformation of the protein structure have a mutual influence, as changes in conformation affect the susceptibility of the protein to chemical reactions, and vice versa. Chemical instability involves covalent modification in the amino acid sequence (primary structure), i.e. bond formation or cleavage, resulting in a new chemical entity. Physical instability refers to any change of the folded state that does not include bond cleavage or

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 52
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

formation, i.e. changes in the spatial, three-dimensional conformation (secondary, tertiary and quaternary structure) of the protein. Once partially or wholly unfolded, the protein can undergo further changes by aggregation with other protein molecules, by adsorption to surfaces or by covalent modifications due to chemical reactions. When aggregation leads to macroscopic ensembles, the process is termed precipitation.

For successful application of proteins as pharmaceuticals, it is essential to understand the stability issues relevant to their production, formulation and use. The processes that affect physical stability include denaturation, i.e. loss of the higher-order, globular structure that a protein adopts upon folding to its native structure, aggregation, precipitation, and adsorption to surfaces. The term *aggregation* is here used to describe a non-native 'self-association' of protein molecules resulting in aggregates that remain in solution and are not visible to the naked eye. *Precipitation*, on the other hand, is a macroscopic process producing a visible change of the protein solution in the form of an increase in viscosity or clouding of the solution. Native, folded proteins may precipitate under certain conditions, most notably salting-out and isoelectric precipitation. Such precipitates, however, are readily distinguished from non-native aggregates by their solubility in normal solvents. Both covalent (polymerization) and noncovalent (aggregation) reaction pathways can be responsible for processes leading to formation of insoluble precipitates, and means to differentiate such mechanistically different processes have been suggested (Costantino *et al.*, 1994b). It is emphasized that this chapter will focus only on the aspects of non-covalent transformation of proteins.

This chapter first gives an overview of the structural details of protein architecture and the different forces contributing to its stability, and then gives a general description of factors favouring protein destabilization which can lead to aggregation and adsorption phenomena. Finally, different approaches for physical stabilization of protein drugs are reviewed. Protein stability during downstream processing has recently been reviewed by Hejnaes *et al.* (1998), and will not be dealt with in depth here.

6.2 Protein structure

The relationship between protein structure and stability is one of the important issues of modern protein science. By understanding the native state of a protein and the mechanism by which it is stabilized, the pharmaceutical scientist can more successfully predict the formulation that will result in a stable dosage form. Therefore, in order to discuss protein stability, a brief description of the characteristics and principles of protein structure and its hierarchical order is appropriate.

The apparently complex structure of proteins is the product of an evolution that probably optimized it for function and not necessarily for stability. One of the most important features of protein structure is that many proteins with quite different amino acid sequences have a similar general architecture, meaning that the overall three-dimensional structures are determined not by all the details of amino acid sequences but by some of their key characteristics. The transformation of the linear sequence (primary structure) from a disordered unfolded state into a compact and functional state is governed by a set of relatively simple principles (Chotia, 1984). Local folding of the polypeptide chain into β-strands and helical segments (secondary structure) is stabilized by a hydrogen bonding network, whereas the packing of α-helices and β-strands (or β-sheets) into a compact globular conformation (tertiary structure) mainly is stabilized by van der Waals' interactions and by the hydrophobic effect (resulting from interaction between non-polar amino

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 53
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

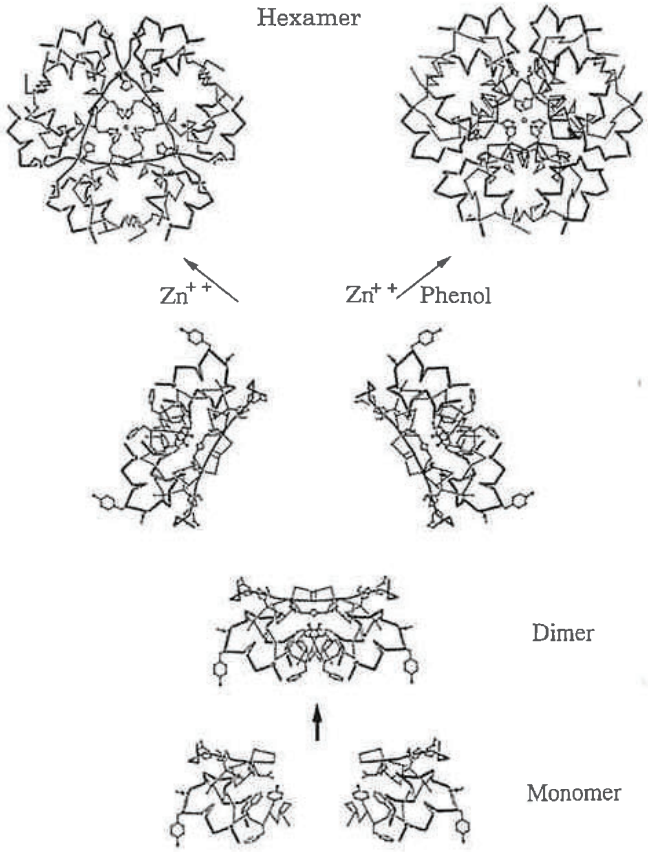


Figure 6.1 Self-association of insulin from the monomer to the dimer, and of three dimers into a hexamer in the presence of zinc ions. If phenol is also present, the gathering of the three dimers is accompanied by a conformational transformation of the B-chain N-terminal from an extended chain into an α-helix.

acid side-chains being more favourable than those they have with water – see section 6.4). Often the fully folded protein molecules will self-associate with one another and form dimers or higher-order subunit associates (quaternary structure). Subunit association is highly specific as it requires the subunits to be in their proper conformation before they coalesce to form the native quaternary structure. Close packing, i.e. optimum complementarity of the subunit interfaces, is the prerequisite of stable self-association, and binding of ligands, often cross-linking the subunits, is further able to stabilize the structure (Figure 6.1).

The distribution of hydrophobic and hydrophilic residues along a sequence is a major determinant of the overall folding of a protein (Miranker and Dobson, 1996). Furthermore, the size, shape and conformation of side-chains determine which particular class of packing occurs and modulate its exact geometry (Chotia, 1984). The protein assumes the most thermodynamically favourable state in a given solvent environment, and close packing strongly restricts the number of possible structures, ensuring a unique organization of the residues for a given protein (Levitt *et al.*, 1997).

Proteins are closely packed both in the solid state and in solution. In fact, they are more tightly packed than almost any other organic matter, as the overall packing efficiency of atoms in the protein core is greater than in crystals of organic molecules. Whereas organic liquids have packing density of less than 0.44, the density of globular proteins is

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 54
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

close to 0.75, which is roughly what is expected for close-packing of hard spheres (Privalov and Gill, 1988). At the surface proteins are packed less tightly than in the core, whereas packing at protein–protein interfaces is roughly comparable to that in the protein interior (Levitt *et al.*, 1997).

6.2.1 Stabilizing interactions

For the effective and reliable functioning of a protein, the stability of its structure must be great enough for the protein to maintain its native conformation within a certain range of conditions. On the other hand, the structure must not be so rigid as to preclude smaller conformational changes, considered an integral part of protein function.

The stability of the native state of a protein is the difference between the free energy of the folded and the unfolded protein. Due to a delicate balance of stabilizing and destabilizing interactions, proteins are only marginally stable. The large and opposing forces include joint actions of several different energetic and entropic effects, which are dominated by the rigid parts of protein as the flexible, solvent-exposed parts contribute little (Matthews, 1993). The intermolecular and intramolecular forces contributing to the stability of the native structure include van der Waals' interactions, hydrogen bonding and other electrostatic interactions, and, not to be forgotten, the so-called hydrophobic effect (see also section 6.2.2).

It is generally agreed that hydrophobic interaction, representing a combined action of the water-ordering effect by the non-polar groups (the hydrophobic effect) and the van der Waals' interaction between these groups, is the major factor in stabilizing the folded structure of globular proteins. However, recent studies suggest that hydrogen bonding, weak electrostatic interactions in the form of charge–dipole or dipole–dipole interactions, also are major contributors to stability (Creighton, 1991). At least for some proteins, hydrogen bonding and the hydrophobic effect make comparable contributions to the conformational stability (Shirley *et al.*, 1992). In addition, strong electrostatic interactions between oppositely charged groups (salt bridges), weak electrostatic interactions between α-helix dipoles, and between the electrons surrounding the nuclei of atoms of aromatic side-chains contribute to some extent to protein stability (Burley and Petsko, 1989).

Ligands such as cofactors, substrates, specific organic molecules or ions may stabilize or destabilize native structures of proteins, depending on whether the ligands bind with highest affinity for the native or the denatured protein. Metal ions can dramatically increase protein stability when they bind to the flexible fragments of the polypeptide chain, in partiular to the bends (Mozhaev *et al.*, 1988).

The stability of proteins at near-physiological temperatures is not high. The values of free energy of protein stability for most globular proteins are only 5 to 15 kcal/mol (Mozhaev *et al.*, 1988). Amino acids with non-polar side-chains occupy about one half of the total volume of protein molecules (Mozhaev *et al.*, 1988). The burial of non-polar amino acid residues in the interior of a protein represents a very well-fitting three-dimensional jigsaw puzzle, and is a major contributor to stability as 1.3 kcal/mol in stability is gained for each CH_2 group buried by folding (Pace, 1992).

The content of hydrogen bonds in various proteins is almost the same and amounts to 0.75 per amino acid residue. Most intramolecular hydrogen bonds contribute 1.3 kcal/mol to the stability of globular protein structures (Shirley *et al.*, 1992). In a typical protein only one-third of the charged groups are involved in salt bridges, but when present they make a significant contribution to protein stability: up to 5 kcal/mol when they are localized

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 55
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

inside the globule, and 1–2 kcal/mol when they are on its surface (Creighton, 1991; Mozhaev *et al.*, 1988). The helix–dipole interactions contribute only about 0.6 kcal/mol to stability (Matthews, 1993).

There is considerable experimental evidence for the stabilizing role of carbohydrate moieties in glycoproteins, whereas only a few cases have been observed in which deglycosylation increases protein stability (Mozhaev and Martinek, 1990).

Protein–protein interactions stabilize the native structure (or the denatured – see section 6.4) by burying hydrophobic clusters localized on the surface of the protein molecule whereby the thermodynamically unfavourable contact of these clusters with water is eliminated. In addition to the shielding of the hydrophobic regions on the surface of the protein, the rigidity of the protein molecule is changed as a result of its multipoint interaction with another molecule. The free energy of stabilization is usually several kcal/mol per protein–protein contact (Mozhaev *et al.*, 1988). Hydrophobic effect is a dominant driving force also in protein–protein association but, in general, hydrophobic amino acids are more frequent in the interior of a globular protein than in the interior of the protein–protein interfaces, and the contribution of the hydrophobic effect to protein self-association is not as strong as to protein folding (Tsai *et al.*, 1997).

The higher-order structure of proteins suggests what forces can be used in stabilization, and the more a formulator knows about the structure of the protein to be formulated, the better the likelihood of stabilizing the protein drug. By understanding the interactions, the formulation chemist can choose excipients to contribute to the thermodynamic stabilization of the protein. It is also important to recognize the potentially denaturing forces that may be applied to the molecule during processing.

6.2.2 *Role of water in structure and stability*

Solvent–protein interactions strongly influence the conformation of the protein molecule. An understanding of proteins requires knowledge of protein hydration as the protein–water interactions play an important role in the folding, stability, and function of proteins. There is a layer of non-covalently bound water molecules on a protein surface that differs essentially from bulk water. The number of bound water molecules is 300–500 for proteins of molecular weight 15–25 kDa, or approximately a monolayer of water, and several states of bound water are distinguished corresponding to the hydration of different types of functional groups (Mozhaev and Martinek, 1990). The layer of bound water associated with a protein in aqueous solution tends to accumulate ionized water molecules oppositely charged to the ionized groups on the surface of the protein molecule.

The role of water in protein structure and stability has been reviewed by Zaks (1992). Protein-bound water molecules can be divided into two main populations: water with hindered rotational freedom that is strongly bound by two or more hydrogen bonds, and weakly bound water with less restriction of rotational freedom. When the number of rotationally free molecules grows, flexibility of the protein structure increases. Hydrophobic effect is, as already mentioned, considered the most important factor contributing to protein stability and such interactions are derived predominantly from an unfavourable entropy decrease of water in the vicinity of non-polar groups. The solvation of a hydrophobic compound decreases its entropy due to increase in ordering of the water molecules surrounding the non-polar groups. To avoid this unfavourable entropy change, non-polar amino acid side-chains are driven into the hydrophobic core, minimizing interaction with the aqueous phase.

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 56
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

In addition to forming a hydration layer around the protein, the presence of water molecules contributes to van der Waals' interactions, salt bridges and, especially, to the hydrogen bonding network within the protein molecule. The molecular events that occur upon binding of water to a protein are well characterized for lysozyme and involve three stages (Zaks, 1992): up to 7% (w/w) of water, the hydration process is dominated by the interaction of water with charged groups; between 7% and 25%, the remaining polar sites are saturated; and finally, the weakly interacting parts of the protein become hydrated and full monolayer coverage of the lysozyme molecule occurs at 38% water corresponding to 300 water molecules per protein molecule.

Organic solvents added to water lower its dielectric constant and reduce polar interactions between the protein and surrounding water molecules, resulting in a decrease in the solubility of the protein. A protein can retain its correct conformation on transfer from a hydrophilic aqueous to a hydrophobic organic solvent system when a layer of bound or 'essential' water remains associated with the folded protein. Dehydration does not seem to disturb the native protein conformation but restricts the mobility. This is facilitated by the ionization of acidic and basic groups of the protein and requires less than 7% water. Lack of protein flexibility associated with dehydration can be reversed by the addition of small amounts of other compounds, such as formamide and glycerol, that have a high potential to form hydrogen bonds (Zaks, 1992). When water is removed from the vicinity of the protein, the protein becomes inert with respect to interactions with hydrophobic solvents, and in dry organic medium (0.02% water) proteins may display a remarkably high stability. Increased resistance to cold denaturation (see section 6.3.2) is another remarkable phenomenon of protein behaviour at low water activity (Zaks, 1992).

6.3 Protein destabilization (denaturation)

To assure protein stability in both liquid and solid formulations, the processes that cause physical destabilization of protein must be understood. Proteins are, as already mentioned, marginally stable and hence readily denatured by various stresses encountered in solution or solid states. Denaturation refers to an alteration of the global fold of a protein molecule, i.e. a disruption of the higher-order structure, such as tertiary and often also the secondary structure. An extreme description of this state is that only random non-covalent interactions are present although conformational preferences do exist. This situation is often referred to as a 'random coil' configuration. However, the random coil model is frequently not an adequate description of denatured proteins in solution. Unfolded states can vary considerably in the extent of residual structure, but some feature that relates to the structure of the native protein will remain and some hydrophobic interactions may persist even under extreme denaturation conditions (Dobson, 1992). Thus, the denatured state can be defined as any non-native conformation and the terms 'denaturation' and 'unfolding' will be used interchangeably in the following, and should be considered equivalent.

Caused by a variety of conditions, the denaturation process is in theory always reversible as the protein molecule, even if aggregated and precipitated, can be returned to its native state after dissolution by 'normalizing' the conditions. Depending on the requirements for reversal of the process, it can be envisioned as reversible or irreversible. However, in pharmaceutical practice, if the unfolded protein cannot easily recover its native state by refolding, for instance by lowering the temperature, then denaturation is considered to be irreversible (Manning *et al.*, 1989).

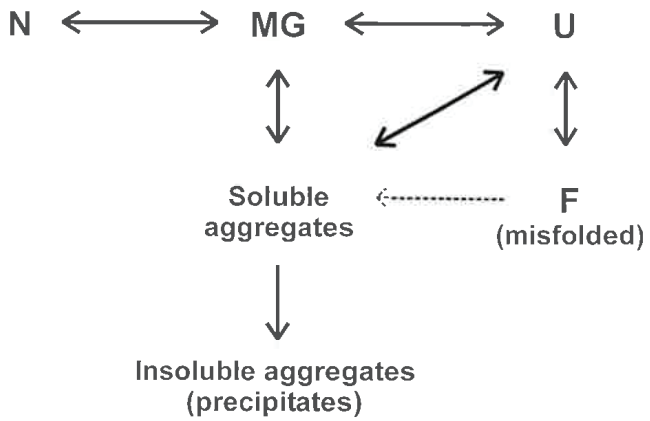


Figure 6.2 Unfolding, aggregation and precipitation. The transition from the native state (N) to the molten globule state (MG) and from MG to the unfolded state (U) are reactions of the first order, i.e. independent of protein concentration, whereas the aggregation reactions are second- or higher-order reactions. In the U state the denatured protein sometimes folds into misfolded (or off-pathway) structures (F) which also may form aggregates.

Proteins often undergo denaturation–renaturation cycles during extraction and purification. Conformational transition of globular proteins into the denatured state, as well as the back-transition into the native state, is a fast process as the protein assumes the most thermodynamically favourable state in a given solvent. Thus both α-helix and β-structure formation show half-lives in the microsecond time range (Jaenicke, 1991). The denaturation of small globular proteins involves a two-state model in which the native state, N, exists in equilibrium with the denatured state D. The denatured state may be an ensemble average of several denatured conformations which frequently exhibit significant amounts of persistent residual structure (Shortle, 1993).

Proteins can be unfolded by various factors, typically by raising or lowering temperature, by extremes of pH, or by addition of organic solvents, surface-active agents or chaotropic denaturants, whose effect is to solvate hydrocarbons more readily. Depending on how denaturation is brought about, the characteristics of the denatured state can differ greatly. For example, the organic solvents and detergents produce a denatured state that still contains some type of highly ordered structure (Shirley *et al.*, 1992).

6.3.1 Unfolding intermediates (molten globule)

In recent years it has become increasingly clear that denaturation of some proteins, especially larger ones, proceeds through a definable thermodynamic state, often native-like, referred to as an unfolding intermediate. There seems to be considerable similarity between folding and unfolding events, and therefore studies of protein folding are providing invaluable insight into the nature of protein stability (Dobson, 1992). A large number of folding studies have during the past decade provided increasing experimental evidence that proteins can exist in states that are intermediate between a fully folded and a highly unfolded state, often referred to as compact denatured states or molten globules (MGs) (Figure 6.2). Many proteins do not show one-step folding but have a distinct, populated intermediate (I state or molten globule state) between the fully unfolded (U state) and the

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 58
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

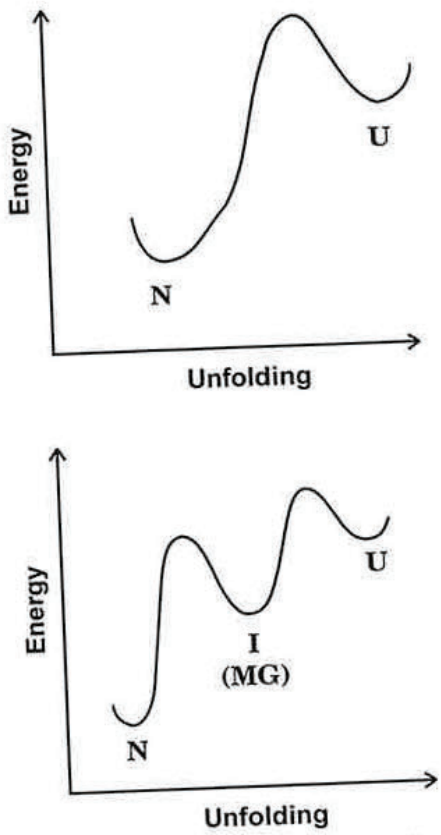


Figure 6.3 Simplified representation of the free-energy profiles for protein unfolding as a result of increasing temperature or increasing concentration of denaturant. Top: One-step unfolding of the native protein (N) into an unfolded structure (U). Bottom: Unfolding via a molten globule intermediate (I).

native state of proteins (N state). The intermediate states are considerably stabilized relative to the U state in the presence of disulphide bonds, and the stability of the intermediate state increases with molecular weight of the protein (Clarke and Waltho, 1997). The MG concept is now widely accepted, and some researchers regard the MG state as a third thermodynamic state of protein molecules (Ptitsyn and Uversky, 1994). This view, however, has recently been challenged (Pfeil, 1998). The free-energy profiles for one-step unfolding and for unfolding via an MG state are shown in Figure 6.3.

The MG intermediates found for different proteins appear to have some common features:

1 they are less compact than the native state, but with Stokes radius close to that of the native protein

2 they contain extensive secondary structure

3 they have loose and disordered tertiary contacts without tight side-chain packing.

The passage of the native state to the MG state involves the breakage of only a small fraction of all intramolecular hydrogen bonds, a fraction similar to that involved in tertiary structure hydrogen bonds (Xie *et al.*, 1991). The more flexible nature of the MG state permits some internal non-polar groups to become exposed to water, making the

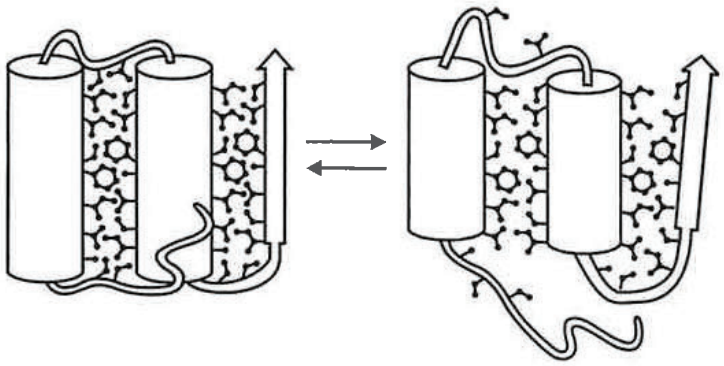


Figure 6.4 Schematic model of the native (left) and the molten globule (right) states of a hypothetical protein molecule. Only hydrophobic side-chains are shown.

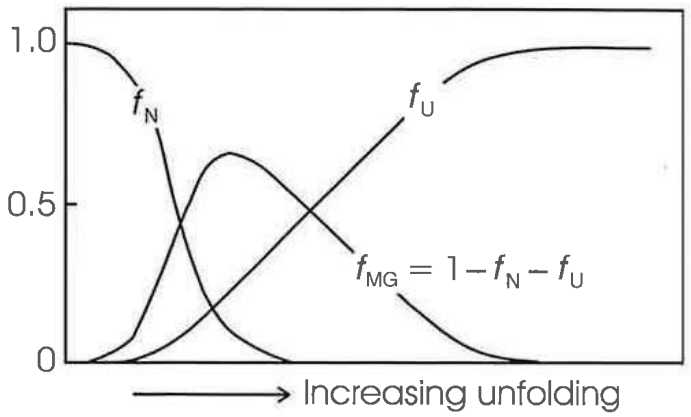


Figure 6.5 Relative population of native (N), molten globule (MG) and unfolded (U) states during unfolding of a hypothetical protein via an intermediate state.

surface of this state more hydrophobic than for the native state, as illustrated in Figure 6.4 (Ptitsyn, 1995), and frequently proteins tend to aggregate in the MG state as measured by equilibrium sedimentation or by light scattering.

Whether two-state or three-state unfolding is observed may depend solely on the stability of the MG state (Baldwin, 1996), and for some small proteins accumulation of intermediates is not observed (Schindler *et al.*, 1995). Recent work has emphasized the complexity of protein folding but has also shed light on folding pathways, the role of intermediates in folding, and the general principles by which proteins adopt their native folds. Many small proteins fold without detectable intermediates, but there are many examples of other proteins whose folding routes pass through partially folded MG-like intermediates (Figure 6.5). However, the presence or absence of intermediates for any given protein is sensitive to folding conditions such as temperature and pH (Pande *et al.*, 1998).

Side-chain interactions (particularly those involving hydrophobic groups) are relatively weak in the intermediate state but form the same pattern as in the native structure, i.e. the I state appears to be neither randomly collapsed nor appreciably misfolded (Clarke and Waltho, 1997). The existence of partially unfolded, stable intermediate conformations has been observed for pharmaceutically relevant proteins, including human growth hormone (Bam *et al.*, 1996) and insulin (Millican and Brems, 1994).

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 60
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

6.3.2 *Temperature-induced changes*

The protein native state is stable in a limited temperature range. When exposed to sufficiently elevated temperatures, all proteins eventually lose native structure. With increasing temperature, hydrogen bonding becomes progressively weaker while the hydrophobic interactions are strengthened with increasing temperature, at least up to 60°C (Volkin and Middaugh, 1992). At physiological temperatures hydrophobic interactions provide the major contribution to stability of the folded state (Franks *et al.*, 1988) but the stability of proteins at this temperature is not high (approximately 12 kcal/mol).

Upon an increase in temperature, heat ultimately disrupts these non-covalent forces. At elevated temperatures, proteins become more flexible, leading to partial unfolding, and their collision frequency increases, resulting in a propensity to form aggregated states. Thermally induced denaturation is classified as either reversible or irreversible depending on whether native structure is recovered following return to ambient temperature.

When small proteins are reversibly denatured by heat, a two-state transition is commonly observed where only the folded or the unfolded states are present. The temperature at which 50% of the molecules are unfolded is called the thermal melting temperature (T_m). Reversible thermal denaturation is dependent on the precise solution conditions. Thus, T_m can increase by 20–25 K when the pH of the solution is changed from 3 to 7, or as a function of the presence and type of neutral salt (Volkin and Middaugh, 1992). One must exercise caution in extrapolating stability data obtained under accelerated storage conditions to less extreme temperature conditions. Such extrapolation is valid only if the kinetics clearly exhibit Arrhenius behaviour (Cleland *et al.*, 1993).

Whereas heat denaturation is driven principally by the loss of polar contacts and the gain of conformational entropy of the chain, cold denaturation is explained by the weakening of the hydrophobic effect upon cooling (Dill *et al.*, 1989; Privalov and Gill, 1988). For most proteins the temperature of cold denaturation is far below the freezing point of aqueous solutions (Privalov *et al.*, 1986). At the temperature of maximal stability (T_s), generally between –10 and 35°C (Pace and Laurents, 1989), the entropy of protein unfolding is zero and stabilization of the native structure is achieved only by the enthalpy factor (Figure 6.6).

As a practical matter, the freezing of a protein solution may also lead to denaturation. As water molecules crystallize, solutes like salts and buffers are concentrated, causing potentially dramatic shifts in pH and ionic strength (Volkin and Middaugh, 1992).

6.3.3 *Influence of pH*

The pH of a protein solution is one of the most important factors determining the structural state of the protein. For example, at low pH decreased stability can be observed as a result of unfavourable electrostatic interactions introduced by the increase in the positive charge on the protein. A similar argument accounts for the decrease in stability observed in highly basic solutions due to repulsion of negative charges. Not surprisingly, most proteins have maximal thermodynamic stability at or near the isoelectric point where the net charge is zero.

Many proteins form an extensively unfolded denatured state at acid pH (2–3) in the absence of high concentrations of anions. If anion concentration is increased (or pH further decreased), and electrostatic repulsion thereby neutralized, the structure becomes much more compact, a state which earlier was called the A state but now more often is

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 61
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

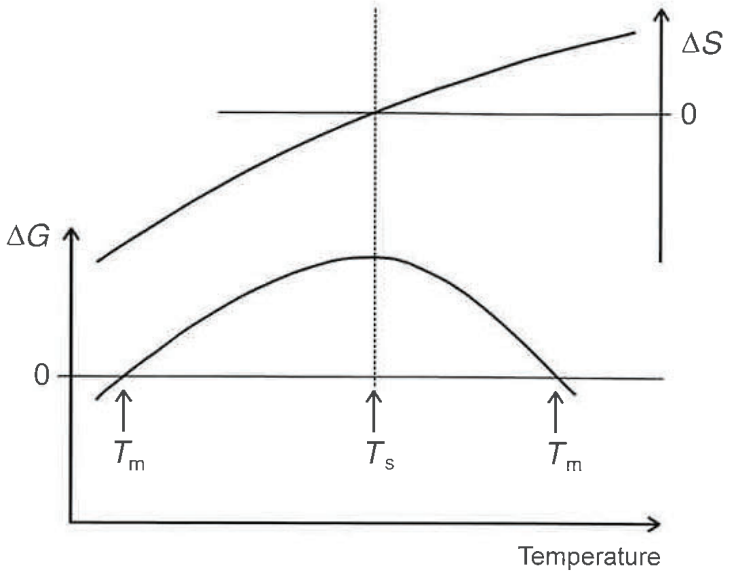


Figure 6.6 Schematic representation of the changes of entropy (ΔS) and Gibbs' free energy (ΔG) of protein unfolding as a result of change in temperature. T_m is the temperature at which the free energy difference between the native and the unfolded state is equal to zero, and represents the temperatures of cold and heat denaturation. T_s is the temperature at which the stability of the native state is maximal and the ΔS is zero.

termed the molten globule state (Shortle, 1993). The effectiveness of an anion in inducing the conformational transition is dependent on the anion species, and the major factor responsible for the transition is electrostatic binding to the positively charged sites of the protein (Goto *et al.*, 1990).

6.3.4 *Influence of pressure*

Proteins are not very sensitive to hydrostatic pressure, and only at extremely large values of pressure do they exhibit the changes which are very similar to those observed in temperature- and pH-induced denaturation. Whereas moderate pressures of 0.5–2 kbar (500–2000 atm) may induce dissociation of oligomeric proteins, denaturation of the protein monomer requires pressures from 4 to 8 kbar (Mozhaev *et al.*, 1996). Hydrophobic contacts are disfavoured by pressure, and electrostatic interactions become much weaker at elevated pressures. In contrast, hydrogen bonding is stabilized by high pressures (Mozhaev *et al.*, 1996). High pressure can lower the freezing point of water by as much as 20 K; this has been used to facilitate the study of cold denaturation of proteins (Nash and Jonas, 1997).

6.4 Aggregation and precipitation

The formation of soluble or insoluble protein ensembles can be due to different phenomena involving either covalent or non-covalent interactions, as follows.

Chemical reactions (polymerization) between different protein molecules result in the formation of covalently linked protein dimers or polymers. Because the origin of protein

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 62
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

aggregates or precipitates is frequently not known, the term 'aggregation' is unfortunately often used in protein literature to describe this process, which should rather be termed 'polymerization'. In practice, however, protein ensembles/precipitates often are mixtures of covalently and non-covalently linked protein molecules (Costantino *et al.*, 1994a, 1994b). Chemical transformation of proteins is dealt with in Chapter 5.

There are two basic types of protein non-covalent interaction, which for clarity here are termed *association* if the process involves protein molecules with native structure, and *aggregation* when denatured protein molecules are involved.

Self-association of the native protein happens as a result of changes in solvent environment (solvent composition, pH, ionic strength, protein concentration, etc.) analogous to the conditions facilitating protein crystallization, isoelectric precipitation, salting out, etc. Thus, if a protein is placed in a high concentration of salt, the surface charges on the protein become masked such that charge–charge repulsion between different native protein molecules does not occur. A similar phenomenon of reduced charge repulsion can occur when the pH of the solution approaches the isoelectric point of the protein. In both cases the surface charge neutralization can result in association of protein molecules with native tertiary structure, and, if the protein concentration is high enough, precipitation of the protein oligomers occurs. In such associates the basic structure of the protein is native-like if not fully native, and the association process is easily reversed by reducing the protein concentration or, if salt and pH changes do not irreversibly alter the conformation of the protein, by adjusting the pH or reducing the salt concentration. Self-association of proteins has often, with insulin as a good example, developed evolutionarily as a requirement for biosynthesis and for ensuring *in vivo* storage stability.

Aggregation of denatured protein molecules was once assumed to arise from ensembles of completely unfolded states of the protein, but recent observations suggest that aggregation is often likely to arise from partially unfolded intermediates. The characteristics and properties of such intermediates may be significantly different from the folded native state as well as the completely denatured protein, and aggregation will be favoured by factors and conditions that favour population of these partly unfolded intermediates (Fink, 1998).

Thus, the term 'association' is here used to characterize ordered, native quaternary structure formation, and the term 'aggregation' is reserved for those cases where a protein forms a regular or non-regular oligomeric structure of non-natively structured proteins. Aggregation may occur during pharmaceutical processing and use of protein drugs by exposure to hydrophobic surfaces, change in temperature (heating or cooling), sonification, or shear stress during filtration or agitation.

6.4.1 Mechanisms of aggregation

Protein aggregation can occur from a conformational intermediate produced during folding or denaturation, or from more extensively unfolded (fully denatured) protein molecules. Protein aggregation is most commonly described as a two-step process but if the protein in its native state is oligomeric, an additional initial process of dissociation into monomeric species may apply at high protein concentration where the protein self-associates (Figure 6.7). The first step involves more or less unfolding of the protein molecules whereby buried, hydrophobic residues are exposed to the aqueous solvent. The unfolding is primarily caused by the solvent environment (pH, salt, co-solvents, etc.), temperature, or surface interactions. In the second step, the hydrophobic residues of the

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 63
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

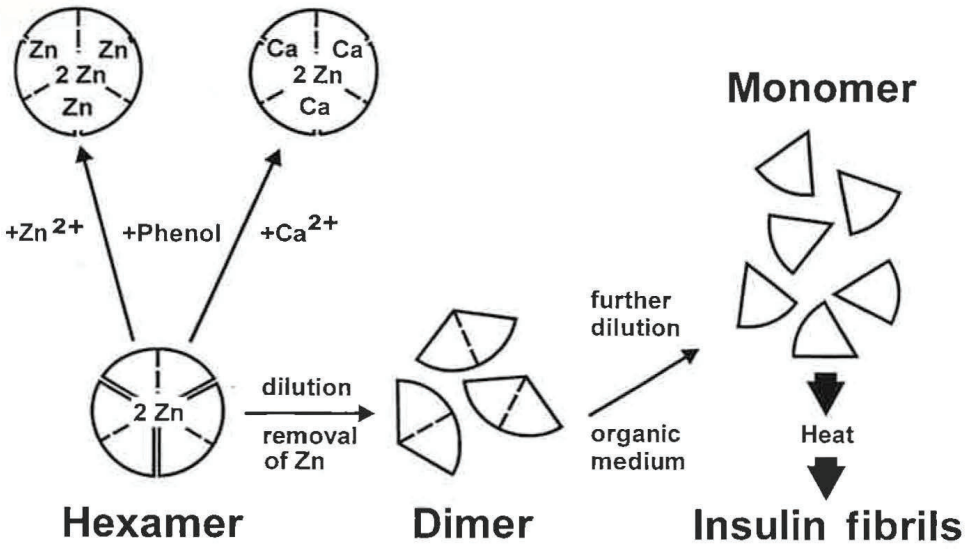


Figure 6.7 Schematic representation of the dissociation of the insulin hexamer as caused by dilution or addition of organic medium. Dissociation into the monomeric species is required for insulin fibril formation, and stabilization of the hexameric state by addition of certain ligands therefore counteracts insulin fibril formation (Brange *et al.*, 1997a).

unfolded protein molecules, now on the surface of the intermediates (or the fully unfolded protein monomers), make intermolecular hydrophobic interactions in order to minimize the unfavourable exposure of the hydrophobic amino acid residues to the solvent. The initial stages of aggregation are quite specific in the sense that they involve the interaction of specific structural subunits of one molecule with 'matching' hydrophobic surface areas of structural subunits of a neighbouring molecule. Two sites can be sufficient, in which case the aggregation most likely propagates in a linear fashion forming long fibres. In any case, the process leads to larger aggregates and eventually their size will exceed the solubility limit.

It has been emphasized that intermediates are more prone to aggregation than the unfolded state, simply because hydrophobicity in the unfolded state is scattered randomly in many small regions, whereas in the intermediates there will be large patches of contiguous surface hydrophobicity with a much stronger propensity for aggregation (Fink, 1998). Some of these surfaces are those that also interact in an intramolecular or intermolecular manner to form the native conformation and associates, respectively (Fink, 1998, Brange *et al.*, 1997b). The hydrophobic interactions in aggregates have often been considered non-specific. However, recent results have provided strong evidence that aggregation occurs by highly specific intermolecular interactions, as a mixture of folding intermediates from two different proteins did not coaggregate with each other, but only with themselves (Speed *et al.*, 1998). The nature of the intermediate state preceding aggregation is the key to understanding the aggregation pathway, but this state has been hypothesized to exist as anything from a small, transient expansion of the native state to a MG structure to a fully unfolded molecule. The sequence of states leading from the unfolded to the native state also allow for the possibility of off-pathway (misfolded) intermediates (see Figure 6.2) (Pande *et al.*, 1998).

Energy transfer – heat, radiation, ultrasound, etc. – is normally required to unfold the protein molecule. Conformational flexibility increases with temperature, and the higher the energy input, induced by, for example, temperature or ultrasound, the more unfolding

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 64
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

and subsequent aggregation there will be. Most proteins aggregate under denaturing conditions at concentrations of 0.01–10 mg/ml. There is a kinetic competition between refolding to the native state and higher-order aggregation reactions. Therefore, at high protein concentration aggregation dominates over folding and eventually leads to irreversible formation of insoluble protein aggregates (Kiefhaber *et al.*, 1991; Volkin and Middaugh, 1992). However, as concentration also influences the equilibrium between native protein oligomers and monomers, the relationship between aggregation and protein concentration is not simple. Thus, an increase of insulin concentration under conditions that favour self-association of the hormone into hexamers actually stabilizes insulin against formation of insoluble aggregates (Brange *et al.*, 1997a). A kinetic model that explains quantitatively the formation of protein aggregates from non-associating proteins has been presented by Kiefhaber *et al.* (1991). With appropriate assumptions with regard to relative magnitude of the various reaction rates, the complex aggregation pathways of self-associating proteins have also been modelled (Lencki *et al.*, 1992).

A common feature of thermally-induced protein aggregation is the formation of an intermolecular hydrogen-bonded, antiparallel β-sheet structure (Dong *et al.*, 1995); often the aggregation occurs in a linear fashion eventually resulting in the formation of long fibres. Aggregation is often considered to be an irreversible process, but as long as only a few molecules are forming the individual aggregates, the aggregation process may be reversible. Initially the oligomeric aggregates will be soluble, but once the aggregates reach a certain size and become insoluble, the process is not reversible without dramatic changes in solvent environment.

6.4.2 *Precipitation and fibrillation phenomena*

Precipitation of a protein from solution is a widely used procedure during downstream processing. Solubility of the protein is here lowered by changing the pH to the isoelectric point of the protein (iso-precipitation), by adding high concentrations of salt (salting out) or by inclusion of organic solute such as ethanol, acetone or polyethylene glycol (Hejnaes *et al.*, 1998). Addition of certain metal ions such as zinc or calcium can for some proteins, such as insulin, further reduce solubility (Brange *et al.*, 1987).

However, precipitation during manufacture or storage of a pharmaceutical preparation of a protein is an undesired event. Such precipitation, initially seen as increasing turbidity or viscosity of the solution, can occur for various reasons relating to either chemical (polymerization, often due to disulphide exchange reactions – see Chapter 5) or physical changes of the protein. Sometimes a protein precipitate can be the result of both covalent and non-covalent aggregation pathways (Costantino *et al.*, 1994a, 1994b), and often, when protein aggregation and precipitation is reported in literature, the nature of the precipitate is not known.

Protein precipitation from a solution involves three reactions: formation of stable aggregates of a few molecules (nucleation), development of these centres into larger aggregates or fibres (growth), and floccule formation from soluble aggregates (precipitation). The rate of nucleus formation is normally low and the time period during which stable nuclei build up is called the lag phase. The length of the lag phase depends on the protein concentration and other experimental conditions, but if preformed aggregates are added to the solution, even in small amounts, the lag phase is completely eliminated (Figure 6.8).

Precipitation of non-covalent aggregates happens when the aggregate nuclei have grown into primary particles, typically with a diameter of 100 to 200 nm corresponding

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 65
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

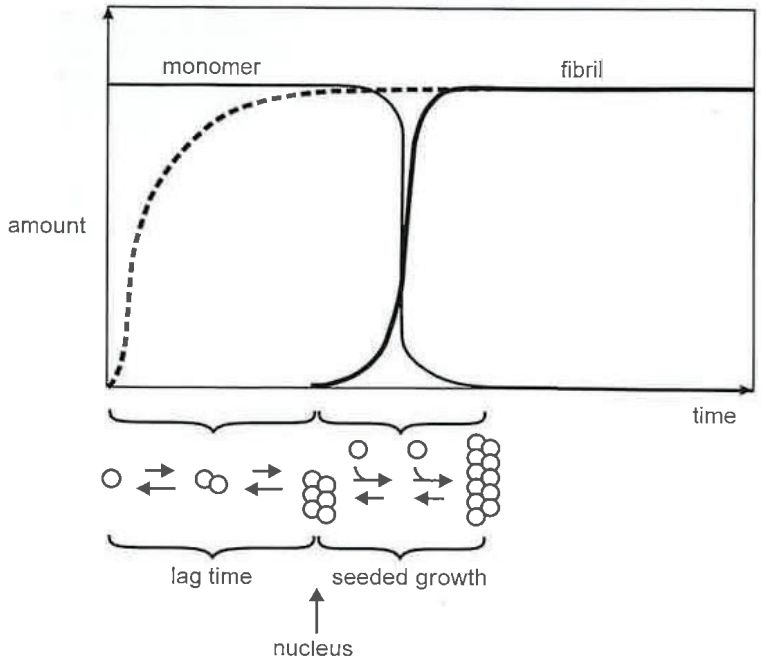


Figure 6.8 Idealized diagram of the formation of protein aggregates (fibrils) as a function of time. The lag time represents the time until a stable nucleus is formed. If fibrils or fibril nuclei are added initially, the lag time disappears and seeded growth starts immediately as represented by the broken curve.

to about 1000 protein molecules (Glatz, 1992). The size and number of primary particles formed depend on the total amount of protein that comes out of solution and the competition between the growth of existing particles and the nucleation of new particles. In the electron microscope, aggregates are often seen as long fibres with diameters ranging from 5 to 150 nm and lengths up to several μm, as in the case of insulin (Brange *et al.*, 1997a). Depending on the exact experimental conditions, the fibres can take different forms (Figure 6.9). The subsequent flocculation of primary particles results in formation of precipitates large enough (10 to 20 μm) to be observed with optical microscopy (Glatz, 1992).

The propensity of insulin, under the influence of heat and exposure to hydrophobic surfaces, to undergo conformational changes resulting in successive, linear aggregation (fibrillation), and formation of a viscous gel or insoluble precipitates, has been one of the most widely studied phenomena in relation to aggregation and precipitation of proteins (for a review see Brange *et al.*, 1997a). The classical method to induce insulin fibrillation has been to heat a concentrated acid solution. The formation of the insulin fibres can be observed as an increase in viscosity or precipitate formation. Whereas nucleation seems to require temperature above normal, the subsequent growth can proceed at ambient or even lower temperature and, depending on the conditions, the growth leads to long fibres resulting in a thixotropic gel or to shorter fibres with a tendency to arrange radially to spherolites with precipitation as the consequence (Waugh *et al.*, 1953; Waugh, 1957). The growth of insulin fibrils has been found to be a function of the surface area of the fibril population and the concentration of insulin in solution (Waugh, 1954), and thus is a highly cooperative process that can remove the insulin quantitatively from solution in the fibrous form. An increase in ionic strength increases the rate of growth and, in particular,

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 66
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

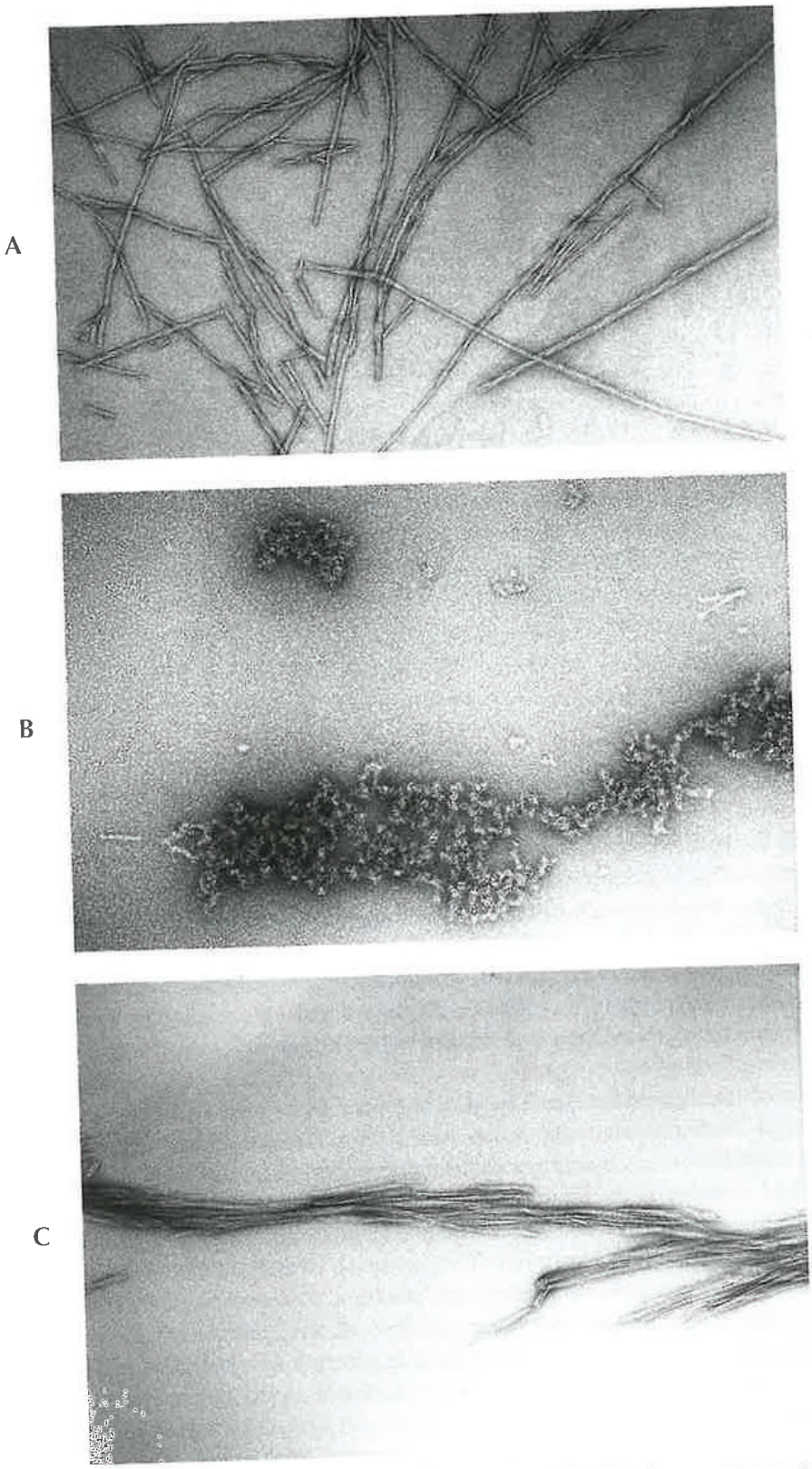


Figure 6.9 Transmission electron micrographs of bovine insulin aggregates (fibrils) formed under various conditions. A: 2% insulin in 0.1 N HCl, pH 1.7, treated at 80°C for 30 min. B: 0.4% insulin, 0.2% phenol, 0.02 M NaCl, pH 7.5, treated by agitation at 37°C for 3 days. C: 2% insulin in 8.3 M acetic acid, 0.1 N NaCl, pH 1.7, treated at 80°C for 2.5 h.

nucleation (Waugh, 1957), whereas high concentrations of organic acids, urea and phenol suppress the nucleation reaction while allowing the growth reaction to proceed (Waugh *et al.*, 1953).

Human insulin does not differ from porcine insulin in fibrillation tendency, whereas bovine insulin, with only two amino acid residues differing in the sequence, surprisingly is significantly more prone to form fibres than the other two species of insulin (Brange *et al.*, 1997a). Apparently the hydrophobicity of the A8 residue (Ala in bovine and Thr in porcine and human insulin), which is on the surface of the molecule, plays a role in the fibrillation process by directing the interactions of insulin with hydrophobic surfaces. With the presence of certain types of insulin-like substances the tendency to fibrillation decreases (Brange *et al.*, 1982). Higher molecular contaminants such as covalent insulin dimer and proinsulin were active in counteracting fibrillation when present in amounts corresponding to a few per cent of the total protein material, whereas more simple derivatives such as desamidoinsulins were ineffective even at 5–10% concentration. Presumably the covalent dimer or proinsulin is able to block further linear aggregation by providing only one of the two complementary interacting surfaces in the insulin fibril.

6.4.3 Factors influencing aggregation and precipitation

Environmental conditions that favour a population of partially unfolded intermediates are likely to lead to aggregation. The major factors that determine the extent and rate of *in vitro* protein aggregation are, besides the amino acid sequence of the protein, pH, temperature, ion strength, protein concentration, and the presence of co-solutes (including chaotropes) or ligands that interact selectively with either the native or non-native conformations of the protein. The propensity for a given protein to aggregate may be determined by the lifetime of partially unfolded intermediates (Fink, 1998).

Effect of excipients and cosolvents

Excipients, often referred to as co-solvents because they affect the solvation of the protein, affect the physical stability of proteins by preferential hydration. The driving force for preferential hydration varies for different types of co-solvent compounds. There are two basic mechanisms by which co-solvents are preferentially excluded from the surface of proteins (Timasheff, 1992):

1 the interactions with protein are determined strictly by the properties of the solvent, as the protein remains essentially inert and only presents a surface

2 the interactions (attractive or repulsive) are to a major extent determined by the chemical nature of the protein surface.

The first mechanism comprises steric exclusion of the co-solvent from the surface of the protein (accounting for the preferential exclusion of polyethylene glycol (PEG) from the surface of native proteins) and perturbation of the surface tension of water by the co-solvent, resulting in a depletion of the additive in the surface layer. Lowering of the surface tension appears to be the most prevalent mechanism by which co-solvents are excluded from proteins. Thus, the preferential hydration caused by addition of carbohydrates is related to this effect (Arakawa and Timasheff, 1982).

The second mechanism comprises repulsion from the charges on the protein surface and the solvophobic effect, which is similar in nature to the hydrophobic effect. The

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 68
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

solvophobic effect causes co-solvent molecules to migrate away from non-polar regions of the protein and into the bulk solvent because the contacts between these non-polar regions and the water–cosolvent mixture are entropically even more unfavourable than contact between these regions and water. Glycerol and a number of other polyols belong to this latter category although the effect of these compounds is complex. The preferential hydration caused by addition of glycerol is related both to the high tendency of glycerol to repel hydrophobic regions and to its ability to enter the water lattice and strengthen solvent structure (Gekko and Timasheff, 1981a, 1981b). On the other hand, glycerol also has affinity for polar regions of proteins, mainly due to hydrogen bond formation. The net effect is a balance between exclusion and binding, with preferential exclusion generally dominating relative to binding (Timasheff, 1992).

6.5 Surface adsorption

The presence of an air–water or solid–water interface has an important influence on the normal forces stabilizing protein higher-ordered structure in solution. Partial unfolding of the protein can occur at the interface which leads to adsorption to the surface through interaction of hydrophobic amino acid residues and a hydrophobic surface, or through binding of polar amino acid residues to charged surfaces. Therefore, proteins can be adsorbed both to non-polar solid (or air) surfaces and to surfaces with ion-exchange properties (such as glass), although greater adsorption normally occurs at hydrophobic than at hydrophilic interfaces.

Normally a protein in solution reaches a surface by diffusion, and therefore the rate of adsorption is a function of the protein concentration. A characteristic property of adsorption is the saturation effect. The amount of protein on the surface increases steeply with time, followed by a levelling-off to a plateau value which normally falls within the range expected for a close-packed monolayer of protein corresponding to 0.1 to 0.5 $\mu g/cm^2$. Sometimes, however, proteins do not display a well-defined plateau but instead protein concentration continues to rise, although more slowly than initially (Horbett, 1992), probably as a result of protein aggregation at the surface. Eventually, the local protein concentration at the interface may be 1000 times higher than the initial protein concentration in the solution. This means that many partially unfolded protein molecules are in close proximity to each other, often resulting in extensive aggregation of the protein.

Like many solid surfaces, gas interfaces are hydrophobic and can cause protein denaturation and aggregation. At air–water interfaces a large fraction of the protein may be adsorbed to the interface when the solution is agitated, and the turbulence creates additional interface area. In a shaken solution there is a continual creation of new interface for protein to adsorb to, and the large effective surface area can lead to massive protein adsorption and subsequent aggregation. The protein denaturation caused by agitation is the result of both the interfacial and the shear forces on the protein (Cleland *et al.*, 1993).

Binding to hydrophobic surfaces virtually always leads to unfolding of the protein, and exposure of its non-polar interior would be expected to enhance hydrophobic surface interactions. Thus, the more hydrophobic the surface, the greater is the destabilizing effect, probably as a result of stabilization of the denatured form of the protein by the hydrophobic surface. In contrast, when proteins bind to polar surfaces the change in protein structure is small (if any), and the interaction between the surface and the protein may be due to more specific polar contacts between the surface and the protein (Steadman *et al.*, 1992).

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 69
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

Insulin aggregation at solid–water interfaces has been studied by several groups (Brange *et al.*, 1982; Sluzky *et al.*, 1992; Feingold *et al.*, 1984; Chawla *et al.*, 1985). In systematic studies on the influence of different materials, Sluzky *et al.* (1992) showed the fibrillation tendency to increase with the surface area of hydrophobic material, and fibrillation tended to proceed faster with increasing hydrophobicity of the material. They also provided evidence that aggregation caused by agitation is the result of both the interfacial and shear forces on the protein. However, because an insulin solution can be rotated in a glass vial (without any air present) for several months (Brange *et al.*, 1982) or agitated with glass beads for many days (Sluzky *et al.*, 1992) without any fibrillation, the shear forces probably play a minor role compared with the surface interactions. Insulin fibrillation at the agitated air–water interface increases with increasing rate of agitation (Sluzky *et al.*, 1992). The insulin monomer, with its exposed hydrophobic surfaces normally buried by dimer and hexamer formation, has potentially higher binding affinity for a hydrophobic interface than the dimer and hexamer. However, at high insulin concentration the equilibrium is shifted towards the hexamer, which may populate the interface and reduce access for the monomeric species.

6.6 Solid phase stability

Most commercially available therapeutic proteins are available as a lyophilized powder which may undergo physical changes during processing and storage. The stability of a protein in the solid state is highly dependent on the moisture content of the solid and there is an optimum level of residual moisture required to maintain stability (Banga, 1995). Most often moisture induced aggregation is due to covalent reactions but also noncovalent forces seem sometimes to be involved (Costantino *et al.*, 1994a, 1994b). The aggregation reactions and means to prevent them are discussed in detail in Chapter 9.

6.6.1 Lyophilization-induced aggregation

When a protein is not sufficiently stable in aqueous solution, lyophilization is the method most commonly used to ensure stability during long-term storage. While lyophilization may improve storage stability, the process in itself can cause damage of the protein as both freezing and drying stress the protein. During freezing the concentration of all solutes increases dramatically in the non-ice phase. This localized high salt or solute concentration may result in denaturation of the protein, and subsequently, after rehydration, lead to aggregate formation, especially if residual moisture is present. The degree of aggregation noted after rehydration correlates directly with the degree to which the protein is unfolded in the dried state (Dong *et al.*, 1995). Therefore special precautions must be taken when proteins are lyophilized, and it is often necessary to add solutes (cryoprotectants and lyoprotectants) that stabilize against both freezing and drying stresses to ensure protection of the protein. This is discussed in depth in Chapter 9.

6.7 Stabilization of protein drugs

Methods of protein stabilization have been improving steadily in recent years. The various general approaches by which increased physical stability can be achieved are briefly reviewed in the following. A more detailed presentation can be found in Chapters 7

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 70
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

(solution stability) and 9 (solid-state stability). Stabilization of proteins with respect to chemical degradation reactions is discussed in Chapter 5.

The key factors for providing a rational basis for protein stabilization are: the nature of the species responsible for aggregation, the detailed mechanism that leads to aggregation and the underlying kinetic scheme, the structure of the aggregates, the specificity of the intermolecular interaction, and how environmental conditions affect the rate and the amount of aggregation.

6.7.1 Stabilization strategies

From the previous sections, it is obvious that rational strategies to improve the practical physical stability of protein drugs include the following main approaches:

1 stabilize the native structure of the protein (i.e. promote self-association, improve native hydrophobic interaction by adding co-solvents or by genetic engineering
2 prevent aggregation of unfolded structures (polyethylene glycols, cyclodextrins?)
3 avoid or block unwanted hydrophobic interfaces (surfactants, avoid head space in container)
4 reduce shear forces (avoid head space in container).

A protein molecule can be stabilized by adding co-solvents or particular ions and salts to the solution; it can also be stabilized by a covalent modification such as glycosylation, or by mutation of its amino acid residues (semisynthesis or genetic engineering). When choosing additives for improving long-term stabilities of proteins in solution or in the solid state, the various factors must be considered on a protein-by-protein basis, as reviewed by Arakawa and Timasheff (1982) and Carpenter *et al.* (1997). Often a combination of stabilizers, working by different mechanisms, may be required for obtaining the desirable stability.

Additives

Solvent environment dictates the ability of the protein to maintain its native state and avoid denaturation from physical forces. It has long been recognized that the addition of a number of polyhydric alcohols, carbohydrates, and some amino acids to aqueous solutions of proteins leads to their stabilization. The mechanism of protein structure stabilization by these compounds was elucidated by Timasheff and coworkers (for reviews see Timasheff, 1992, 1993), who found that the polyols are preferentially excluded from the immediate domain of the protein, which in their presence is surrounded by an environment relatively enriched in water (preferential hydration of the protein). Since this exclusion is thermodynamically unfavourable, the system tends to reduce it by reducing the area of protein–solvent interaction.

Polyethylene glycol (PEG) has unique properties with respect to its interaction with proteins. It is excluded non-specifically from native globular proteins caused by the excluded volume effect. But because PEG also has a strong non-polar character and therefore interacts with newly exposed non-polar residues, it is attracted to the protein in its unfolded state (Timasheff, 1993). Thus, PEG binds to the unfolded form of the protein and thereby stabilizes denatured protein and prevents its aggregation, but once folding has taken place, PEG drives the equilibrium towards the native globular state as it is then

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 71
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

excluded from the structure (Lee and Lee, 1987; Banga, 1995). The dual properties of PEG make it difficult to predict its possible effect on a given protein, and the recipe for stabilization must be dictated by each protein individually (Timasheff, 1992).

Cyclodextrins, with a ring structure of six (α-), seven (β-), or eight (χ-cyclodextrin) glucopyranose units, have been proposed as stabilizing excipients for proteins because of their capability to form inclusion complexes due to the hydrophobic nature of the inner core of these carbohydrates (Brewster *et al.*, 1991). The cavity, with a size of 5–8 Å (Banga, 1995), has a particular affinity for aromatic amino acids and may be able to encapsulate them (Cleland *et al.*, 1993). β-cyclodextrin in a concentration of 1.5% has been shown to reduce agitation-induced aggregation of insulin to some degree (Banga and Mitra, 1998).

The major rationale for using surfactants as stabilizers is their ability to bind to hydrophobic interfaces that are potential denaturation sites, i.e. container surfaces and air–liquid interfaces. However, surfactants are also able to bind to exposed hydrophobic domains on proteins whereby aggregation may be reduced (Cleland *et al.*, 1993). The concentration of surfactant must be sufficiently high to cover the surface as a monolayer and the optimum concentration is most likely to be its critical micelle concentration (Banga, 1995). Higher concentrations of surfactants may lead to denaturation of the protein. Non-ionic surfactants such as Tween may act like molecular *in vitro* chaperones in facilitating protein refolding, thus preventing aggregation while not altering the native conformation (Bam *et al.*, 1996).

Genetic engineering

Manipulation of physical stability by protein engineering has been performed on many proteins, and examples of stabilization have been achieved for a number of proteins, in particular the enzymes subtilisin and lysozyme. Stability can be changed by mutating amino acids involved in stabilizing interactions in the folded protein, destabilizing interactions in the unfolded form, or both. Suitable substitutions are those which will increase the hydrophobicity of the protein core and will increase the strength of helical packing (Fersht and Serrano, 1993). Introduction of new S–S bonds in the molecule can either increase or decrease stability (Mozhaev *et al.*, 1988).

Stability of a protein substituted at the interior of the molecule tended to increase linearly with increasing hydrophobicity of the substituted residue, unless the volume of the substituted residue was over a certain limit (Yutani *et al.*, 1987). Stability of some protein molecules has been improved to the level of those from extreme thermophilic organisms. This was achieved by multiple mutations and by utilizing the knowledge gained from protein structures in the extreme thermophiles (Lee and Vasmatzis, 1997).

In order to raise the temperature of thermal denaturation by 10 K it is necessary to add 6 kJ/mol of stabilizing energy, achieved by formation of only one additional ionic or hydrogen bond or a few new hydrophobic contacts (Mozhaev and Martinek, 1990).

References

ARAKAWA, T. and TIMASHEFF, S.N., 1982, Stabilization of protein structure by sugars, *Biochemistry*, **21**, 6536–6544.

BALDWIN, R.L., 1996, On-pathway versus off-pathway folding intermediate, *Folding & Design*, **1**, R1–R8.

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 72
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

BAM, N.B., CLELAND, J.L. and RANDOLPH, T.W., 1996, Molten globule intermediate of recombinant human growth hormone: stabilization with surfactants. *Biotechnol. Prog.*, **12**, 801–809.

BANGA, A.K., 1995, *Therapeutic Peptides and Proteins. Formulation, Processing, and Delivery Systems*, Lancaster, PA: Technomic Publishing Co.

BANGA, A.K. and MITRA, R., 1998, Minimization of shaking-induced formation of insoluble aggregates of insulin by cyclodextrins, *J. Drug Targeting*, **1**, 341–345.

BRANGE, J., HAVELUND, S., HANSEN, P., LANGKJAER, L., SØRENSEN, E. and HILDEBRANDT, P., 1982, Formulation of physically stable neutral insulin solutions for continuous infusion by delivery systems. In: GUERIGUIAN, J.L., BRANSOME, E.D. and OUTSCHOORN, A.S. (eds) *Hormone Drugs*, pp. 96–105, Rockville, MA: United States Pharmacopeial Convention.

BRANGE, J., SKELBAEK-PEDERSEN, B., LANGKJAER, L., DAMGAARD, U., EGE, H., HAVELUND, S., HEDING, L.G., JØRGENSEN, K.H., LYKKEBERG, J., MARKUSSEN, J., PINGEL, M. and RASMUSSEN, E., 1987, *Galenics of Insulin: the Physico-chemical and Pharmaceutical Aspects of Insulin and Insulin Preparations*, Berlin: Springer-Verlag.

BRANGE, J., ANDERSEN, L., LAURSEN, E.D., MEYN, G. and RASMUSSEN, E., 1997a, Toward understanding insulin fibrillation, *J. Pharm. Sci.*, **86**, 517–525.

BRANGE, J., DODSON, G.G., EDWARDS, D.J., HOLDEN, P.H. and WHITTINGHAM, J.L., 1997b, A model of insulin fibrils derived from the x-ray crystal structure of a monomeric insulin (despentapeptide insulin), *Proteins*, **27**, 507–516.

BREWSTER, M.E., HORA, M.S., SIMPKINS, J.W. and BODOR, N., 1991, Use of 2-hydroxypropyl-β–cyclodextrin as a solubilizing and stabilizing excipient for protein drugs, *Pharm. Res.*, **8**, 792–795.

BURLEY, S.K. and PETSKO, G.A., 1989, Electrostatic interactions in aromatic oligopeptides contribute to protein stability, *Tibtech*, **7**, 354–359.

CARPENTER, J.F., PIKAL, M.J., CHANG, B.S. and RANDOLPH, T.W., 1997, Rational design of stable lyophilized protein formulations: some practical advice, *Pharm. Res.*, **14**, 969–975.

CHAWLA, A.S., HINBERG, I., BLAIS, P. and JOHNSON, D., 1985, Aggregation of insulin, containing surfactants, in contact with different materials, *Diabetes*, **34**, 420–424.

CHOTIA, C., 1984, Principles that determine the structure of proteins, *Ann. Rev. Biochem.*, **53**, 537–572.

CLARKE, A.R. and WALTHO, J.P., 1997, Protein folding pathways and intermediates. *Curr. Opin. Biotechnol.*, **8**, 400–410.

CLELAND, J.L., POWELL, M.F. and SHIRE, S.J., 1993, The development of stable protein formulations: a close look at protein aggregation, deamidation, and oxidation, *Crit. Rev. Therapeutic Drug Carrier Systems*, **10**, 307–377.

COSTANTINO, H.R., LANGER, R. and KLIBANOV, A.M., 1994a, Moisture-induced aggregation of lyophilized insulin, *Pharm. Res.*, **11**, 21–29.

1994b, Solid-phase aggregation of proteins under pharmaceutically relevant conditions, *J. Pharm. Sci.*, **83**, 1662–1669.

CREIGHTON, T.E., 1991, Stability of folded conformations, *Curr. Opin. Struct. Biol.*, **1**, 5–16.

DILL, K.A., ALONSO, D.O.V. and HUTCHINSON, K., 1989, Thermal stabilities of globular proteins, *Biochemistry*, **28**, 5439–5449.

DOBSON, C.M., 1992, Unfolded proteins, compact states and molten globules, *Curr. Opin. Struct. Biol.*, **2**, 6–12.

DONG, A., PRESTRELSKI, S.J., ALLISON, S.D. and CARPENTER, J.F., 1995, Infrared spectroscopic studies of lyophilization- and temperature-induced protein aggregation, *J. Pharm. Sci.*, **84**, 415–424.

FEINGOLD, V., JENKINS, A.B. and KRAEGEN E.W., 1984, Effect of contact material on vibration-induced insulin aggregation, *Diabetologia*, **27**, 373–378.

FERSHT, A.R. and SERRANO, L., 1993, Principle of protein stability derived from protein engineering experiments, *Curr. Opin. Struct. Biol.*, **3**, 75–83.

FINK, A.L., 1998, Protein aggregation: folding aggregates, inclusion bodies and amyloid, *Folding & Design*, **3**, R9–R23.

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 73
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

FRANKS, F., HATLEY, R.H.M. and FRIEDMAN, H.L., 1988, The thermodynamics of protein stability. Cold destabilization as a general phenomenon, *Biophys. Chem.*, **31**, 307–315.

GEKKO, K. and TIMASHEFF, S.N., 1981a, Mechanism of protein stabilization by glycerol: preferential hydration in glycerol–water mixtures, *Biochemistry*, **20**, 4667–4676.

1981b, Thermodynamic and kinetic examination of protein stabilization by glycerol, *Biochemistry*, **20**, 4677–4686.

GLATZ, C.E., 1992, Modeling of aggregation–precipitation phenomena. In: AHERN, T.J. and MANNING, M.C. (eds) *Stability of Protein Pharmaceuticals, Part A: Chemical and Physical Pathways of Protein Degradation*, pp. 135–166, New York: Plenum Press.

GOTO, Y., TAKAHASHI, N. and FINK, A.L., 1990, Mechanism of acid-induced folding of proteins, *Biochemistry*, **29**, 3480–3488.

HEJNAES, K., MATTHIESEN, F. and SKRIVER, L., 1998, Protein stability in downstream processing. In: SUBRAMANIAN, G. (ed.) *Bioseparation and Bioprocessing*, Vol. 2, pp. 31–65, Weinheim: Wiley–VCH Verlag.

HORBETT, T.A., 1992, Adsorption of proteins and peptides at interfaces. In: AHERN, T.J. and MANNING, M.C. (eds) *Stability of Protein Pharmaceuticals, Part A: Chemical and Physical Pathways of Protein Degradation*, pp. 195–214, New York: Plenum Press.

JAENICKE, R., 1991, Protein folding: local structures, domains, subunits, and assemblies, *Biochemistry*, **30**, 3147–3461.

KIEFHABER, T., RUDOLPH, R., KOHLER, H.-H. and BUCHNER, J., 1991, Protein aggregation *in vitro* and *in vivo*: a quantitative model of the kinetic competition between folding and aggregation, *Biotechnology*, **9**, 825–829.

LEE, B. and VASMATZIS, G., 1997, Stabilization of protein structures, *Curr. Opin. Biotechnol.*, **8**, 423–428.

LEE, L.L.Y. and LEE, J.C., 1987, Thermal stability of proteins in the presence of poly(ethylene glycols), *Biochemistry*, **26**, 7813–7819.

LENCKI, R.W., ARUL, J. and NEUFELD, R.J., 1992, Effect of subunit dissociation, denaturation, aggregation, coagulation, and decomposition on enzyme inactivation kinetics: I. First-order behavior, *Biotechnol. Bioeng.*, **40**, 1421–1426.

LEVITT, M., GERSTEIN, M., HUANG, E., SUBBIAH, S. and TSAI, J., 1997, Protein folding: the endgame, *Annu. Rev. Biochem.*, **66**, 549–579.

MANNING, M.C., PATEL, K. and BORCHARDT, R.T., 1989, Stability of protein pharmaceuticals, *Pharm. Res.*, **6**, 903–917.

MATTHEWS, B.W., 1993, Structural and genetic analysis of protein stability, *Annu. Rev. Biochem.*, **62**, 139–160.

MILLICAN, R.L. and BREMS, D.N., 1994, Equilibrium intermediates in the denaturation of human insulin and two monomeric insulin analogs, *Biochemistry*, **33**, 1116–1124.

MIRANKER, A.D. and DOBSON, C.M., 1996, Collapse and cooperativity in protein folding, *Curr. Opin. Struct. Biol.*, **6**, 31–42.

MOZHAEV, V.V. and MARTINEK, K., 1990, Structure–stability relationships in proteins: a guide to approaches to stabilizing enzymes, *Adv. Drug Del. Rev.*, **4**, 387–419.

MOZHAEV, V.V., BEREZIN, I.V. and MARTINEK, K., 1988, Structure–stability relationship in proteins: fundamental tasks and strategy for the development of stabilized enzyme catalysts for biotechnology, *CRC Crit. Rev. Biochem.*, **23**, 235–281.

MOZHAEV, V.V., HEREMANS, K., FRANK, J., MASSON, P. and BALNY, C., 1996, High pressure effects on protein structure and function, *Proteins*, **24**, 81–91.

NASH, D.P. and JONAS, J., 1997, Structure of the pressure-assisted cold denatured state of ubiquitin, *Biochem. Biophys. Res. Commun.* **238**, 289–291.

PACE, C.N., 1992, Contribution of the hydrophobic effect to globular protein stability. *J. Mol. Biol.*, **226**, 29–35.

PACE, C.N. and LAURENTS, D.V., 1989, A new method for determining the heat capacity change for protein folding, *Biochemistry*, **28**, 2520–2525.

PANDE, V.S., GROSBERG, A.U., TANAKA, T. and ROKHSAR, D.S., 1998, Pathways for protein folding: is a new view needed? *Curr. Opin. Struct. Biol.*, **8**, 68–79.

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 74
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462

PFEIL, W., 1998, Is the molten globule a third thermodynamic state of protein? The example of α-lactalbumin, *Proteins*, **30**, 43–48.

PRIVALOV, P.L. and GILL, S.J., 1988, Stability of protein structure and hydrophobic interaction, *Adv. Protein Chem.*, **39**, 191–234.

PRIVALOV, P.L., GRIKO, Y.V. and VENYAMINOV, S.Y., 1986, Cold denaturation of myoglobin, *J. Mol. Biol.*, **190**, 487–498.

PTITSYN, O.B., 1995, Molten globule and protein folding, *Adv. Protein Chem.*, **47**, 83–229.

PTITSYN, O.B. and UVERSKY, V.N., 1994, The molten globule is a third thermodynamical state of protein molecules, *FEBS Lett.*, **341**, 15–18.

SCHINDLER, T., HERRLER, M., MARAHIEL, M.A. and SCHMID, F.X., 1995, Extremely rapid protein folding in the absence of intermediates, *Nature Struct. Biol.*, **2**, 663–673.

SHIRLEY, B.A., STANSSENS, P., HAHN, U. and PACE, C.N., 1992, Contribution of hydrogen bonding to the conformational stability of ribonuclease T1, *Biochemistry*, **31**, 725–732.

SHORTLE, D., 1993, Denatured states of proteins and their roles in folding and stability, *Curr. Opin. Struct. Biol.*, **3**, 66–74.

SLUZKY, V., KLIBANOV, A.M. and LANGER, R., 1992, Mechanism of insulin aggregation and stabilization in agitated aqueous solutions, *Biotechnol. Bioeng.*, **40**, 895–903.

SPEED, M.A., WANG, D.I.C. and KING, J., 1998, Specific aggregation of partially folded polypeptide chains: the molecular basis of inclusion body composition, *Nature Biotechnol.*, **14**, 1283–1287.

STEADMAN, B.L., THOMPSON, K.C., MIDDAUGH, C.R., MATSUNO, K., VRONA, S., LAWSON, E.Q. and LEWIS, R.V., 1992, The effect of surface adsorption on the thermal stability of proteins, *Biotechnol. Bioeng.*, **40**, 8–15.

TIMASHEFF, S.N., 1992, Stabilization of protein structure by solvent additives. In: AHERN, T.J. and MANNING, M.C. (eds) *Stability of Protein Pharmaceuticals. Part B: In Vivo Pathways of Degradation and Strategies for Protein Stabilization*, pp. 265–285, New York: Plenum Press.

1993, The control of protein stability and association by weak interactions with water: how do solvents affect these processes? *Annu. Rev. Biophys. Biomol. Struct.*, **22**, 67–97.

TSAI, C.-J., LIN, S.L., WOLFSON, H.J. and NUSSINOV, R., 1997, Studies of protein–protein interfaces: a statistical analysis of the hydrophobic effect, *Protein Sci.*, **6**, 53–64.

VOLKIN, D.B. and MIDDAUGH, C.R., 1992, The effect of temperature on protein structure. In: AHERN, T.J. and MANNING, M.C. (eds) *Stability of Protein Pharmaceuticals, Part A: Chemical and Physical Pathways of Protein Degradation*, pp. 215–247, New York: Plenum Press.

WAUGH, D.F., 1954, Protein–protein interactions, *Adv. Protein Chem.*, **9**, 325–437.

1957, A mechanism for the formation of fibrils from protein molecules. *J. Cell. Comp. Physiol.*, **49** (suppl. 1), 145–164.

WAUGH, D.F., WILHELMSON, D.F., COMMERFORD, S.L. and SACKLER, M.L., 1953, Studies on the nucleation and growth reactions of selected types of insulin fibrils, *J. Am. Chem. Soc.*, **75**, 2592–2600.

XIE, D., BHAKUNI, V. and FREIRE, E., 1991, Calorimetric determination of the energetics of the molten globule intermediate in protein folding: apo-α-lactalbumin, *Biochemistry*, **30**, 10673–10678.

YUTANI, K., OGASAHARA, K., TSUJITA, T. and SUGINO, Y., 1987, Dependence of conformational stability on hydrophobicity of the amino acid residue in a series of variant proteins substituted at a unique position of tryptophan synthase α subunit, *Proc. Natl Acad. Sci. USA*, **84**, 4441–4444.

ZAKS, A., 1992, Protein–water interactions. Role in protein structure and stability. In: AHERN, T.J. and MANNING, M.C. (eds) *Stability of Protein Pharmaceuticals, Part A: Chemical and Physical Pathways of Protein Degradation*, pp. 249–271, New York: Plenum Press.

Regeneron Pharmaceuticals, Inc. Exhibit 2015 Page 75
Celltrion, Inc. v. Regeneron Pharmaceuticals, Inc. IPR2023-00462