

Fa Liu

Peptide Science

Chemical Ligation, Lead Generation,
and Therapeutic Advances

WILEY

Peptide Science

Peptide Science

Chemical Ligation, Lead Generation, and
Therapeutic Advances

Edited by

Fa Liu

*Focus-X Therapeutics
Watchung, USA*

WILEY

Copyright 2025 by Fa Liu. All rights reserved. All rights reserved, including rights for text and data mining and training of artificial intelligence technologies or similar technologies.

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.

Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at <http://www.wiley.com/go/permission>.

Trademarks

Wiley and the Wiley logo are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates in the United States and other countries and may not be used without written permission. All other trademarks are the property of their respective owners. John Wiley & Sons, Inc. is not associated with any product or vendor mentioned in this book.

Limit of Liability/Disclaimer of Warranty

While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic formats. For more information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Cataloging-in-Publication Data

Hardback: 9781119824671

ePDF: 9781119824688

epub: 9781119824695

oBook: 9781119824701

Cover Image: Courtesy of Fa Liu

Cover design by: Wiley

Contents

List of contributors *xiii*

Preface *xvii*

Section I New Developments in Peptide Chemistry 1

1	Chemical Peptide Ligations	3
	<i>Yihui Cao and Xuechen Li</i>	
1.1	Introduction	3
1.2	Ligation Driven by <i>Trans</i> -esterification	4
1.2.1	Native Chemical Ligation	4
1.2.1.1	Desulfurization	5
1.2.1.2	Auxiliary-assisted NCL	7
1.2.1.3	Backbone-installed Split Intein-assisted Ligation	7
1.2.2	Diselenide/Selenoester Ligation	8
1.3	Ligation Driven by Imine Capture	8
1.3.1	Serine/Threonine Ligation	8
1.3.2	Cysteine/Penicillamine Ligation	10
1.4	Ligation Initiated by Hydroxylamine	10
1.5	Ligation in Protein Synthesis	12
1.5.1	Sequential Ligation and Chemical Protein Synthesis	12
1.5.1.1	Terminal Latent and Activation	12
1.5.1.2	Ligation Intermediates Facilitate Difficult Peptide Synthesis	14
1.5.1.3	Chemical Protein Synthesis	15
1.5.2	Protein Semi-synthesis	16
1.6	Ligations in Drug Development	17
1.6.1	Ligation in Macrocyclic Peptide Synthesis	17
1.6.2	Mirror-image Phage Display	19
1.7	Summary and Outlook	21

2 Protein and Peptide Ligation Using Peptide Ligases 29*Anna Koijen and Leendert J. van den Bos*

- 2.1 Introduction 29
- 2.2 Peptide Ligases Mechanism of Action 32
- 2.3 Sortase 33
- 2.4 Peptide Asparaginyl Ligases 38
- 2.5 Subtilisin-derived Peptide Ligases 41
- 2.6 Trypsinligase 44
- 2.7 Conclusion 46

3 Modern Methods for Late-stage Peptide Modification 55*Yuxuan Ding and Zachary T. Ball*

- 3.1 Introduction 55
- 3.2 Methionine 57
- 3.3 Aspartate and Glutamate 59
- 3.4 Arginine 61
- 3.5 Tryptophan 62
- 3.6 Functionalization of sp^2 C—H Bonds: Tyr, Phe, and His 66
 - 3.6.1 Tyrosine 66
 - 3.6.2 Phenylalanine 68
 - 3.6.3 Histidine 70
- 3.7 Functionalization of sp^3 C—H Bonds 71
- 3.8 Other X—H Functionalizations 73
- 3.9 Conclusion 74

Section II Peptide Drug Hit Identification 87**4 Generating Phage Libraries Bearing Unnatural Pharmacophores 89***Christin Kossmann, Arunika Ekanayake and Ratmir Derda*

- 4.1 Introduction 89
- 4.2 Biotin Capture Assay to Detect Chemical Modification of Phage Libraries 93
- 4.3 N-terminal Oxidation and Functionalization to Incorporate Unnatural Fragments 106
- 4.4 Macrocyclization of Phage-displayed Peptides Bearing Cysteines 110
- 4.5 Macrocyclization and Installation of Warheads in Phage-displayed Libraries 114
- 4.6 Outlook 120

5	Computational Peptide Design for Diverse Structures and Functions	127
	<i>Stephen Rettie, David Juergens and Gaurav Bhardwaj</i>	
5.1	The Various Types of Peptides	128
5.2	<i>In Silico</i> Peptide Design: From Physics-based Methods to DL	129
5.3	Physics-based Methods	130
5.3.1	Motif-based Design with Physics-based Methods	131
5.3.2	<i>De Novo</i> Design with Physics-based Methods	133
5.3.2.1	Peptide Design with Rosetta Generalized Kinematic Closure	133
5.3.2.2	<i>De Novo</i> Peptide Binder Design with Physics-based Methods	135
5.3.2.3	Molecular Dynamics-based Peptide Design	136
5.4	DL Methods for Peptide Design	136
5.4.1	Structure-based DL Methods for Peptide Design	137
5.4.2	PLMs for Peptide Design	139
5.5	Optimizing for Membrane Permeability and Stability	140
5.6	Remaining Challenges	141
5.7	Conclusion	143

Section III Peptide Drug Lead Generation and Optimization 151

6	Peptide Discovery: Lead Generation	153
	<i>Yvonne Angell, Wendy Hartssock, Hans Melo and Timothy M. Reichart</i>	
6.1	Introduction	153
6.1.1	Inspiration from Nature	164
6.1.2	Artificial Libraries	168
6.1.3	Phage Display	170
6.1.4	mRNA Display	172
6.2	Approved Drugs Discovered Through Phage Display	174
6.2.1	Romiplostim (NPlate)	174
6.2.2	Ecallantide (Kalbitor)	175
6.2.3	Peginesatide (Omontys)	176
6.2.4	Pegcetacoplan (Empaveli)	176
6.3	Approved Drugs Discovered Through mRNA Display	177
6.3.1	Zilucoplan (ZILBRYSQ, Ra Pharma, UBC)	177
6.3.2	Plate-based Array Synthetic Libraries	178
6.3.3	Split and Pool (Tea Bag) Approach	179
6.3.4	Split-and-pool DNA-encoded Libraries Approach (Wendy)	180
6.3.5	Functional Selection for Peptide Hit Identification	180
6.3.6	Bacterial Display	181

- 6.3.7 Computational Approaches 182
- 6.3.8 *De Novo* Peptide Design 183
- 6.3.8.1 Physics-based versus ML-based Approaches 183
- 6.3.9 Physics-based Design: The Process of Designing Cyclic Peptides 183
- 6.4 A Successful Example 185
- 6.4.1 Current Challenges and Future Directions in Computational Peptide Design 185
- 6.5 Conclusion 188

7 A Survey and Guide to the Development of Fatty Acid-derivatized Peptide Therapeutics 203

Michael T. Jacobsen, Christian Poulsen, Katharina L. Kopp, Jørgen Olsen and Nick Cox

- 7.1 Introduction 203
- 7.2 Basics of Fatty Acid Derivatization of Peptides 204
- 7.3 Evolution of GLP-1-based Incretins into Stable, Long-acting Therapeutics 205
- 7.4 Diverse Mechanisms of Fatty Acid Derivatization to Extend Insulin Half-life 210
- 7.5 Development of a Stable, Fatty Acid-derivatized Amylin Analog 214
- 7.6 Stabilization and Fatty Acid Derivatization of Peptide YY (PYY) 217
- 7.7 Guidance for Advancing Fatty Acid-derivatized Peptides 222
 - 7.7.1 Routine Methods for Evaluating Peptide Solubility 223
 - 7.7.2 Concepts and Methods for Biophysical Characterization 225
 - 7.7.2.1 Aggregates: Undesired Multimeric Species 225
 - 7.7.2.2 Self-associates: Better-defined, Acceptable Multimeric Species 226
 - 7.7.3 Assessment of Peptide Immunogenicity 228
 - 7.7.3.1 Impact of Fatty Acid Derivatization on Immunogenicity 233
 - 7.7.3.2 Evaluating Metabolic and Chemical Stability 234
 - 7.7.3.3 Notes on Relevant Chemical Instabilities 236
 - 7.7.5 General Manufacturing Considerations Related to Peptide Optimization 237
 - 7.7.6 Specific Advice for Peptide Fatty Acid Derivatization and Characterization 240
- 7.8 Concluding Remarks 242

Section IV Peptide Drug Case Studies 269**8 Insulin in Its Second Century: History, Evolution, and Aspirations 271**

John P. Mayer, Kishore Thalluri, Ethan Mickelson, Michael H.B. Stowell and Richard D. DiMarchi

- 8.1 Introduction 271
- 8.2 The Evolution of Insulin Therapy 272
 - 8.2.1 Prandial Insulins 273
 - 8.2.1.1 First Generation 274
 - 8.2.1.2 Second Generation 274
 - 8.2.2 Basal Insulins 275
 - 8.2.2.1 First Generation 276
 - 8.2.2.2 Second Generation 276
- 8.3 Structure 278
- 8.4 Synthesis 280
- 8.5 Aspirational Aims 286
 - 8.5.1 Continuous Glucose Monitoring 286
 - 8.5.2 Glucose-responsive Insulins 287
 - 8.5.2.1 Polymeric GRI Systems 287
 - 8.5.2.2 Non-polymeric GRI Systems 289
- 8.6 Concluding Thoughts 290

9 Incretins and Development of Incretin-based Therapeutics 305

Nan Zheng, Xiao-xuan Su, Shengping Zhang and Weijun Shen

- 9.1 Introduction 305
 - 9.1.1 Incretin Hormones 305
 - 9.1.2 Biological Functions of Incretin Hormones 306
 - 9.1.3 Incretin Hormones in Obesity and Type 2 Diabetes 308
- 9.2 Development of Novel GLP-1 and GIP-based Therapeutics 309
 - 9.2.1 GLP-1R Mono-agonists 309
 - 9.2.2 GIPR Mono-agonists 312
 - 9.2.3 GLP-1R/GIPR Dual Agonists 313
- 9.3 Development of Multi-receptor Agonists Targeting Incretin and Other Hormonal Receptors 315
 - 9.3.1 Gut-brain Hormonal Regulation of Glucose and Energy Homeostasis 315
 - 9.3.2 Multi-receptor Agonists Targeting GLP-1R and Other Hormonal Receptors 318

- 9.3.2.1 GLP-1R/GCGR Dual Agonists 318
- 9.3.2.2 GLP-1R/GIPR/GCGR Triple Agonists 320
- 9.3.2.3 GLP-1R/GLP-2R Dual Agonists 322
- 9.3.2.4 GLP-1 Co-agonist with Amylin and/or Calcitonin 323
- 9.3.2.5 GLP-1R/Y2R Co-agonists 325
- 9.3.2.6 GLP-1/FGF21 Dual Agonists 327
- 9.3.2.7 GLP-1/GDF15 Dual Agonists 328
- 9.4 Summary and Perspective 329

10 Peptide-based Therapeutics for Inflammatory and Autoimmune Diseases 339

Sunay V. Chankeshwara, Werngard Czechtizky and Wu Su

- 10.1 Introduction 339
- 10.2 Literature Search and Methods 341
- 10.3 Representative Anti-inflammatory Peptides in Clinical Trials 342
 - 10.3.1 AMTX-100 342
 - 10.3.2 B27PD 343
 - 10.3.3 EA-230 344
 - 10.3.4 Lupuzor 345
 - 10.3.5 SP16 346
 - 10.3.6 JNJ2113 347
 - 10.3.7 Larazotide 348
 - 10.3.8 LSALT 349
 - 10.3.9 RLS-0071 350
 - 10.3.10 GLP-1 and Antidiabetic GLP-1R Agonists 350
- 10.4 Summary and Discussion 351

11 Radioligand as a Unique Rising Modality for Cancer Diagnosis and Treatment 359

Ming-Jin Jheng, Xiaozhu Wang, Pengyun Li and Fa Liu

- 11.1 Radioligand Introduction 359
- 11.2 Radionuclides and Chelators Developed for Medical Imaging and Cancer Treatment 361
 - 11.2.1 Radionuclides Being Developed as Radioligands 361
 - 11.2.2 Radiometal Chelators Being Utilized in Radioligands 363
- 11.3 Notable Radioligands Developed for Cancer Diagnosis and Treatment 364
 - 11.3.1 SSTR2-targeting Theranostic Radioligands for SSTR2-overexpressed Cancers 364
 - 11.3.2 PSMA-targeting Theranostic Radioligands for Prostate Cancer 367
- 11.4 Emerging Targets 371

- 11.4.1 FAP-targeting Radioligands as Pan-tumor Diagnostic Agents and Potentially Therapeutic Drugs 371
- 11.4.2 Gastrin-releasing Peptide Receptor 374
- 11.4.3 Neurotensin Receptor 1 (NTSR1) 376
- 11.4.4 Glypican 3 (GPC3) 377
- 11.4.5 P-cadherin (CDH3) 378
- 11.5 Emerging Binder Modalities as Radioligand Vector 379
- 11.5.1 Small Molecules 379
- 11.5.2 Antibody and Antibody Fragments 381
- 11.5.3 Engineered Scaffold Protein 382
- 11.5.4 Oligonucleotide Aptamers 383
- 11.6 Perspectives: Promises and Challenges 384

12 Peptide-small Molecule Drug Conjugates for Cancer, Metabolic Diseases, and Beyond 401

Rongjun He, Sumeet Singh and Adam Mezo

- 12.1 Introduction 401
- 12.2 Peptide-drug Conjugates for Cancer 402
- 12.2.1 PEN-221 402
- 12.2.2 BT5528 404
- 12.2.3 G202 406
- 12.2.4 ANG1005 407
- 12.2.5 TH1902 408
- 12.3 Peptide-drug Conjugates for Metabolic Diseases 409
- 12.3.1 Glucagon-like Peptide-1-estrogen 409
- 12.3.2 GLP-1-tesaglitazar 411
- 12.3.3 Glucagon-T3 412
- 12.3.4 NPY1-tesaglitazar 413
- 12.4 Peptide-drug Conjugates for Other Diseases 415
- 12.5 Challenges in Peptide-drug Conjugate Design 415
- 12.6 Summary 416

13 Emerging Landscape of Therapeutic Oligonucleotide Delivery by Cell-penetrating Peptide 425

Xiang Li, Xiulong Shen, Ashweta Sahni and Ziqing Qian

- 13.1 Introduction 425
- 13.2 Cell-penetrating Peptides 426
- 13.2.1 Discovery of CPPs 426
- 13.2.2 Structural Features of CPPs 427
- 13.2.3 Mechanisms of Internalization 428
- 13.2.3.1 Direct Translocation 430

- 13.2.3.2 Endocytosis and Endosomal Escape 430
- 13.2.4 Overview of CPPs for the Delivery of Therapeutic Cargos 431
- 13.3 Applications of CPPs for the Delivery of Therapeutic Oligonucleotides 433
 - 13.3.1 Overview of Therapeutic Oligonucleotides 433
 - 13.3.2 Development of CPPs for PMO Delivery 435
 - 13.3.3 Applications of CPP-PMOs to Modulate Pre-mRNA Splicing 437
 - 13.3.3.1 Duchenne Muscular Dystrophy 438
 - 13.3.3.2 Spinal Muscular Atrophy 440
 - 13.3.3.3 Other Indications 442
 - 13.3.4 Applications of CPP-PMOs to Target mRNA with Aberrant Repeat Expansions 443
 - 13.3.5 Applications of CPP-PMOs to Inhibit mRNA Translation 444
 - 13.3.6 Applications of CPP-PMOs to Modulate RNA Levels 445
 - 13.3.7 Applications of CPP for PNA Delivery 446
- 13.4 Future Perspectives and Summary 447

Section V Peptide CMC and Regulatory 469

14 Peptides Chemistry, Manufacturing, and Controls 471

Jack Xu

- 14.1 Introduction 471
- 14.2 CMC Information and Format for DS 472
- 14.3 CMC Information and Format for DP 477
- 14.4 Synthetic Peptide CMC Development 483
 - 14.4.1 Peptide Process Development 483
 - 14.4.1.1 Recombinant Production Method 484
 - 14.4.1.2 Chemical Synthesis 484
 - 14.4.2 Phase-appropriate Considerations in CMC Development 491
 - 14.4.2.1 The Graded Nature 491
 - 14.4.2.2 Preclinical and Phase 1 Stage CMC Development 492
 - 14.4.2.3 Late-phase CMC Development 493
 - 14.4.3 Control Strategies for Peptide APIs 494
 - 14.4.3.1 Raw Materials Control 494
 - 14.4.3.2 Manufacturing Process Control 497
- 14.5 Conclusions 504

Index 509

List of Contributors

Yvonne M. Angell

WuXi TIDES
San Francisco, CA, USA

Zachary T. Ball

Department of Chemistry
Rice University
Houston, TX, USA

Gaurav Bhardwaj

Institute for Protein Design
University of Washington
Seattle, Washington, D.C., USA

and

Department of Medicinal Chemistry
University of Washington
Seattle, Washington, D.C., USA

Yihui Cao

Department of Chemistry
University of Hong Kong
Hong Kong SAR, P.R. China

Sunay V. Chankeshwara

Medicinal Chemistry, Research and
Early Development, Respiratory and
Immunology, BioPharmaceuticals R&D
AstraZeneca
Gothenburg, Sweden

Nick Cox

Novo Nordisk, Chemical Biology
Global Research Technologies
Lexington, MA, USA

Werngard Czechtizky

Medicinal Chemistry, Research and
Early Development, Respiratory and
Immunology, BioPharmaceuticals
R&D
AstraZeneca
Gothenburg, Sweden

Ratmir Derda

Department of Chemistry
University of Alberta
Edmonton, AB, Canada

Richard D. DiMarchi

Department of Chemistry
Indiana University
Bloomington, IN, USA

Yuxuan Ding

Department of Chemistry
Rice University
Houston, TX, USA

Arunika Ekanayake

Department of Chemistry
University of Alberta
Edmonton, AB, Canada

Wendy J. Hartsock

Aralez Bio
Berkeley, CA, USA

Rongjun He

Neurocrine Biosciences, Inc.
San Diego, CA, USA

Michael T. Jacobsen

Novo Nordisk, Chemical Biology
Global Research Technologies
Lexington, MA, USA

Ming-Jin Jheng

Full-Life Technologies Limited
Shanghai, China

David Juergens

Institute for Protein Design
University of Washington
Seattle, Washington, D.C., USA

Anna Koijen

EnzyTag BV
Nuth, the Netherlands

Katharina L. Kopp

Novo Nordisk, Functional Assays &
Screening
Global Research Technologies
Måløv, Denmark

Christin Kossmann

Zealand Pharma A/S
Søborg, Denmark

Pengyun Li

Full-Life Technologies Limited
Shanghai, China

Xiang Li

Entrada Therapeutics, Inc.
Boston, MA, USA

Xuechen Li

Department of Chemistry
University of Hong Kong
Hong Kong SAR, P.R. China

and

Laboratory for Marine Drugs and
Bioproducts
Qingdao National Laboratory for
Marine Science and Technology
Qingdao, P.R. China

Fa Liu

Full-Life Technologies Limited
Shanghai, China

John P. Mayer

Department of Molecular, Cellular and
Developmental Biology
University of Colorado
Boulder, CO, USA

Hans Melo

Menten AI
San Francisco, CA, USA

Adam Mezo

Neurocrine Biosciences, Inc.
San Diego, CA, USA

Ethan Mickelson

Department of Bioengineering
University of Washington
Seattle, WA, USA

Jørgen Olsen

Novo Nordisk, Discovery ADME
Global Research Technologies
Måløv, Denmark

Christian Poulsen

Novo Nordisk, Biophysics and
Injectable Formulation
Global Research Technologies
Måløv, Denmark

Ziqing Qian

Entrada Therapeutics, Inc.
Boston, MA, USA

Timothy M. Reichart

Department of Chemistry
Hampden-Sydney College
Hampden-Sydney, VA, USA

Stephen Rettie

Molecular and Cell Biology Program
University of Washington
Seattle, Washington, D.C., USA

and

Institute for Protein Design
University of Washington
Seattle, Washington, D.C., USA

Ashweta Sahni

Entrada Therapeutics, Inc.
Boston, MA, USA

Weijun Shen

Center for Translational Research
(CTR)
Shenzhen Bay Laboratory
Shenzhen, China

Xiulong Shen

Entrada Therapeutics, Inc.
Boston, MA, USA

Sumeet Singh

Neurocrine Biosciences, Inc.
San Diego, CA, USA

Michael H.B. Stowell

Department of Molecular, Cellular and
Developmental Biology
University of Colorado
Boulder, CO, USA

Wu Su

Medicinal Chemistry, Research and
Early Development, Respiratory and
Immunology, BioPharmaceuticals
R&D
AstraZeneca
Gothenburg, Sweden

Xiao-xuan Su

Center for Translational Research
(CTR)
Shenzhen Bay Laboratory
Shenzhen, China

Kishore Thalluri

Department of Chemistry
Indiana University
Bloomington, IN, USA

Leendert van den Bos

EnzyTag BV
Nuth, the Netherlands

Xiaozhu Wang

Full-Life Technologies Limited
Shanghai, China

Jack Xu

Elix tide LLC
San Marcos, CA, USA

Shengping Zhang

Center for Translational Research
(CTR)
Shenzhen Bay Laboratory
Shenzhen, China

Nan Zheng

Center for Translational Research
(CTR)
Shenzhen Bay Laboratory
Shenzhen, China

Preface

With great excitement, the past 10 years have witnessed significant advancements in the broad peptide science field. These include the development of the fundamentally novel peptide ligation and protein modification chemistries, the re-engineering of automated peptide synthesizers enabling fast and high-quality protein synthesis, and numerous milestone achievements in peptide drug discovery and development. On the chemistry front, novel chemical methods have resulted in promising progress in enhancing the permeability of peptides to reach the targets inside cells or behind the blood-brain barrier (BBB). Relatedly but differently, new formulation approaches have enabled the oral administration of peptide drugs, where an incredible milestone was marked by the 2019 approval of Rybelsus[®], the oral tablet formulation of semaglutide, a 31-mer glucagon-like peptide 1 (GLP-1) analog. On the therapeutics front, the huge success of peptide drugs in metabolic diseases, such as semaglutide, tirzepatide, liraglutide, and multiple modern insulin analogs, has now been expanded into oncology and other therapeutic areas. Two notable recent successes are the approval of Lutathera[®] in 2018 and Pluvicto[®] in 2022. Lutathera[®] is a peptide-radionuclide conjugate consisting of a somatostatin receptor agonist as the homing vector and a beta-emitter Lutetium-177 as the warhead to treat gastroenteropancreatic neuroendocrine tumors. Pluvicto[®] shares the same mechanism of action as Lutathera[®], while targeting prostate-specific membrane antigen for the treatment of metastatic castration-resistant prostate cancer. Further in the metabolic disease areas, the 2022 approval of tirzepatide (Mounjaro[®]), a dual-acting peptide that can activate both GLP-1 receptor and gastric inhibitory polypeptide receptor, opens the era of clinical application of unimolecular polypharmacy where numerous promising multi-acting peptides are racing in the late development phases. Finally, it needs to be noted that there are many other equally exciting successes that were also developed within the same time frame. However, these are not mentioned above, only due to the limited space in this Preface.

These important advancements in general peptide chemistry and broad peptide therapeutics areas prompted us to assemble this book: *Peptide Science: Chemical*

Ligation, Lead Generation, and Therapeutic Advances. The motivation was to provide the readers with a collective set of quality reviews and case studies that cover the representative developments within the past 10 years or so in a logical and systematic fashion. With a total of five sections, it is intended to offer a highlight of modern peptide chemistries and comprehensive coverage of peptide therapeutics, including discovery, development, and regulatory.

Section I, *New Developments in Peptide Chemistry*, highlights recent developments in chemical and enzymatic peptide ligation technologies and modern peptide and protein modification chemistries. Please note that this section is not arranged to provide lecture chapters on the fundamentals of peptide chemistry, but rather focuses on selected advancements.

Section II, *Peptide Drug Hit Identification*, aims to review the approaches and technologies that have been frequently practiced for peptide hit identification, an important early step of the peptide drug discovery process. The classical source is naturally existing peptides, which are still one of the most reliable starting places as these peptides are already well-optimized via evolution to effectively execute their endogenous biological tasks. In addition to the nature source, artificial libraries, including phage display library and its chemically modified versions, and mRNA peptide display libraries, have become increasingly important as quality hit sources. More recently, computational-based approaches have also been rapidly developed, and impressive outcomes on peptide hit generation have been frequently achieved.

Section III, *Peptide Drug Lead Generation and Optimization*, includes two chapters focusing on lead generation and lead optimization, respectively. There is some overlapping between Chapter 6 and the chapters in Section II. The intention is for Chapter 6 to provide a thorough review of the hit and lead generation while Section II allows readers to dive deeply into the individual specific approaches. Chapter 7, contributed by a group of Novo Nordisk scientists, serves as a great reference for peptide therapeutics lead optimization with a focus on long-acting peptide hormone analogs, particularly the fatty acid derivatization approach.

Section IV, *Peptide Drug Case Studies*, dives deeply into the established peptide therapeutics and emerging developments in selected therapeutic areas. It consists of six chapters covering several therapeutic areas, including diabetes, obesity, NASH, cardiovascular diseases, cancer, inflammation, and autoimmune. It also involves various peptide drug molecular formats such as insulin (a mini protein with two chains and three pairs of disulfide bonds), incretin (a linear peptide with 30–40 residues), peptide-radionuclide conjugate (radioligand therapy, also called RDC), peptide-cytotoxin conjugate, and finally, peptide-oligonucleotides conjugates.

Section V, *Peptide CMC and Regulatory*, discusses peptide CMC processes and the related regulatory requirements.

What I hoped to assemble while missing in this book is the alternative formulation approach, such as oral tablet formulation of peptides, the methods enabling targeting intracellular receptors and targets beyond BBB, and another chapter in Section V to cover more regulatory side of peptide drug development.

In any case, I expect this book can serve as a valuable resource for the entire peptide science community, from junior researchers to university professors and seasoned industrial veterans, whoever is interested in the highlight of recent developments in peptide chemistry and a deep dive into the discovery and development of peptide therapeutics.

I thank all chapter authors for their contributions. The time they devoted to this book is highly appreciated. I also thank all Wiley editors and friends involved in this project for their long-lasting support. Finally, I thank my daughter Doris (fifth grade), who designed the cover image for the book.

Section I

New Developments in Peptide Chemistry

1

Chemical Peptide Ligations

Yihui Cao and Xuechen Li

Department of Chemistry, the University of Hong Kong, Hong Kong SAR, P.R. China

1.1 Introduction

Proteins play crucial roles in basic physiological processes and are responsible for a variety of biochemical functions, including signaling transduction, energy utilization, and immune response. Correlating protein structure with function has always been a charming topic among researchers. Although recombinant expression from bacteria or cell lines is a convenient means to produce proteins, it is still difficult to control specific post-translational modifications such as glycosylations, incorporate any uncanonical amino acid, or introduce unnatural reporters such as fluorescent tags, using the natural cellular machinery [1]. Herein, chemical protein synthesis that assembles protein sequence through atom-by-atom control provides a solution for generating site-specific natural or unnatural modification(s), and mirror-image proteins.

The solid-phase peptide synthesis (SPPS) by Merrifield provides an efficient peptide synthesis approach [2]. Utilization of SPPS methodology, along with the condensation of protected peptide fragments, has significantly expanded the range of polypeptide lengths that can be achieved via chemical synthesis. However, SPPS is limited by the peptide length. Due to the statistical reasons for linear stepwise coupling, each step during SPPS is incomplete, which causes byproducts to accumulate with peptide chain elongation. The peptide length from SPPS mostly remains within 50 amino acids to maintain good synthesis quality. Besides, the limited solubility of protected peptides in organic solvents hampers the ability of this method to meet the increasing synthesis demand for

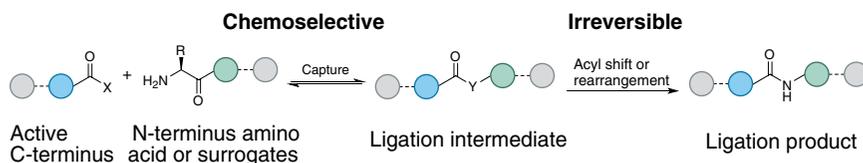


Figure 1.1 Generic chemical peptide ligation.

complex protein structures [3]. Consequently, novel synthetic approaches that can be conducted in aqueous buffers for handling unprotected peptide segments are strongly demanded.

The concept of peptide ligation, which allows condensation of unprotected peptide segments, was first proposed in the 1980s [4,5]. It involves a weakly activated peptide C-terminus to chemoselectively react with the N-terminus of the second peptide, resulting in a ligation intermediate that links the two fragments together, followed by an irreversible rearrangement step to form a natural peptide linkage (Figure 1.1). In this chapter, the driving forces of ligation, chemoselectivity details, and their applications in protein synthesis are discussed.

1.2 Ligation Driven by *Trans*-esterification

1.2.1 Native Chemical Ligation

Native chemical ligation (NCL) (Figure 1.2a), developed by Kent et al. in 1994, is the most widely applied ligation method [6]. NCL requires one peptide with a C-terminal thioester and the second peptide with an N-terminal cysteine (Cys). Ligation occurs when two fragments are mixed in a neutral or slightly basic aqueous buffer. The thiol group of N-terminal Cys undergoes reversible *trans*-thioesterification, replacing the thioester at the C-terminal of the first fragment. After that, a rapid [1,4] S-to-N acyl transfer converts the thioester intermediate into a native Xaa-Cys peptide (Xaa represents any amino acid) and generates the desired ligation product. Chemoselectivity can be considered to originate from “soft base-soft acid” interaction between the thioester (soft acid) and the thiol group (soft base) from free Cys. Other nucleophiles present on unprotected peptides, such as amines, are “hard” bases that do not have the same reactivity as the thiol group. Next, the irreversible and rapid acyl transfer drives the equilibrium. Even though the internal Cys could be involved in the reversible *trans*-thioesterification, the reaction equilibrium cannot move forward and does not produce a stable product. The ligating C-terminal residue and the thioester type highly affect the NCL reaction rate. For instance, β -branched amino acids (Val, Ile, Thr) significantly decrease ligation rates due to their bulky side chains [7]. Additionally, because of carbonyl oxygen interference,

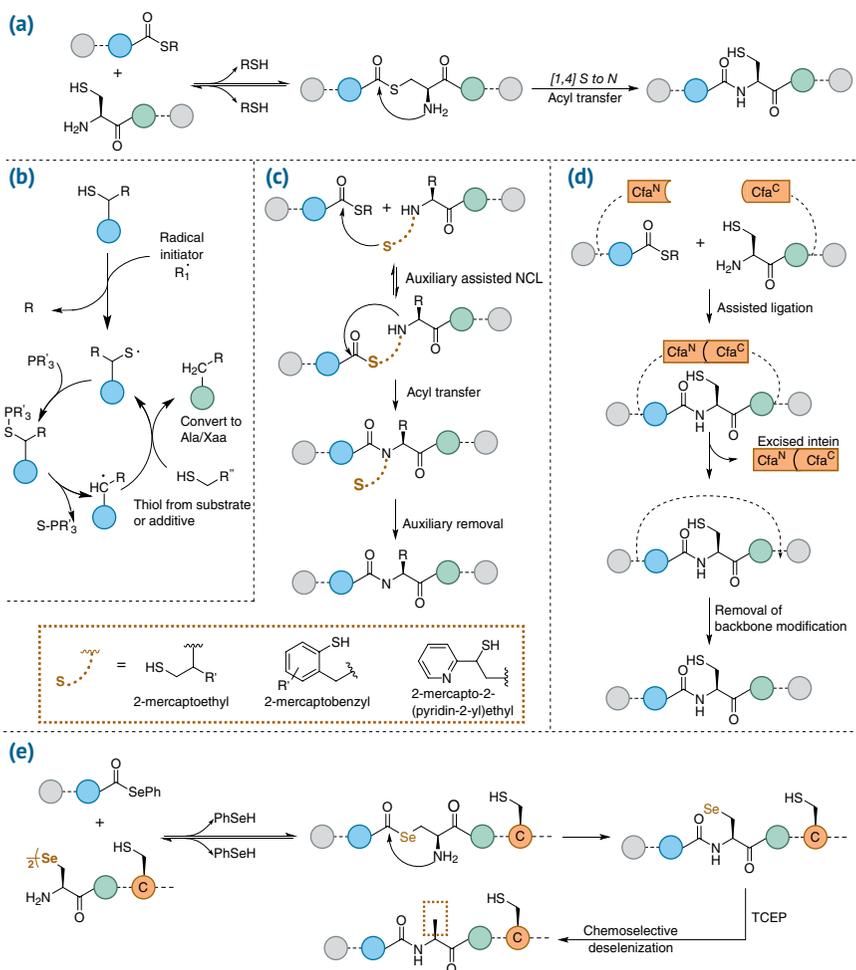


Figure 1.2 Ligation through *trans*-esterification.

proline carbonyl is less electrophilic, which restricts *trans*-esterification [6]. For the formation of intermolecular thioesters, 4-mercaptophenylacetic acid (MPAA) is widely used as an additive due to its low pKa, good water solubility in NCL buffer, and odorless nature. Other thiols with higher pKa values, such as trifluoroethanethiol (TFET), can be also used [8].

1.2.1.1 Desulfurization

Even though NCL is a powerful technique, the demand for one N-terminal Cys restricts its potential applications in protein chemical synthesis due to its low

natural abundance (1.8%). The situation has changed after the invention of post-ligation desulfurization. Desulfurization, first reported in 2001 by Dawson, took advantage of metal-catalyzed reduction under a hydrogen atmosphere [9]. Nevertheless, the requirement of excess metal could occasionally cause side reactions and result in low yields. In addition, utilization of metal catalysts potentially induced epimerization of secondary alcohols and caused reduction of thiols and thioesters [10]. Later, the establishment of free radical-based desulfurization by Danishefsky et al. provided a milder and more reliable means for chemoselective peptide desulfurization (Figure 1.2b) [10]. Radical-induced desulfurization requires a radical initiator, phosphine compound, and hydrogen source. The reaction is initiated by water-soluble radical initiator 2,2 α' -azobis[2-(2-imidazolin-2-yl)propane]dihydrochloride (VA-044) at 37 °C. The initiator radical attacks the Cys on the peptide, generating a thiyl radical. Subsequently, the thiyl radical rapidly reacts with tris(2-carboxyethyl)phosphine (TCEP) to generate the phosphonium radical. After that, the phosphine sulfide TCEP=S is cleaved from the complex, driving the alkyl radical formation. The desulfurized product is eventually generated after the application of hydrogen atom transfer (HAT) from thiol additives such as *tert*-butylthiol (tBuSH) or glutathione (GSH). Other thiyl radicals generated continue the radical chain reaction until its full conversion.

Based on the radical desulfurization strategy, photon-induced radical initiation has been developed. Pyane's group designed a flow chemistry system allowing NCL and UV-induced desulfurization to take place sequentially [11]. Later, visible light-induced metal complexes [12] and peroxide [13] radical initiator were reported, expanding the scope of radical initiator. Lately, Li et al. applied a novel radical generator, tetra-organylborate, to peptide desulfurization [14]. This strategy significantly increases the desulfurization reaction rate. It can be accomplished through a simple add-and-done procedure to finish within 30 seconds. Sodium tetraethylborate effectively serves as a radical initiator in the presence of atmospheric oxygen, inducing peptide desulfurization. Besides, the byproduct triethylborane from the initiation step serves as a hydrogen donor, demonstrating comparable efficiency to thiol additives such as tBuSH. Therefore, odorous thiol additives are not necessary for this strategy. In addition, the expeditious production of radicals can surpass the inhibitory impact of traces of MPAA residue (<100 μ M), which allows NCL and desulfurization to perform in one pot. Furthermore, the mild conditions are compatible with some reductive functional groups, such as serotonylated substrates.

With the effective desulfurization technique established, the development of thiolated amino acids has drawn wide attention. The first thiolated amino acid to be used in NCL-desulfurization was β -thiolated phenylalanine [15]. NCL-desulfurization via both the nickel method and the free radical VA-044 method has been successfully applied to it [16,17]. Remarkably, commercially

available penicillamine (Pen) has been used as the valine surrogate to expand the ligation site to one of the most abundant amino acids (6.8%) [18]. So far, a number of research groups have contributed to developing 13 Fmoc-SPPS-compatible thiolated amino acids, including β -thiolated, γ -thiolated, and δ -thiolated amino acids [1,19,20]. Although desulfurization may be difficult to some steric hindrance residue [18], NCL surrogate has been significantly expanded.

1.2.1.2 Auxiliary-assisted NCL

In parallel with the development of desulfurization, another strategy to overcome the low Cys abundance problem is the auxiliary-assisted NCL (Figure 1.2c). It introduces a thiol-containing handle at or near the peptide N-terminus, which mediates *trans*-thioesterification and *N*-to-*S* acyl transfer in a similar manner as Cys. After ligation, the auxiliary thiol handle can be removed. For glycopeptides, the branching sugar can serve as the handle through a thiol substitution. Its ligation site can be 1–6 amino acids away from the sugar-substituted amino acid [21,22]. In such cases, the acyl transfer via a large ring transition state significantly affects the reaction rate and efficiency, as compared to NCL. On the other hand, a variety of N^α substitutions, including N^α -2-mer-captoethyl type and N^α -2-mercaptobenzyl type [23], have been developed. They are designed to allow for acyl transfer through a five- or six-member ring for a rapid transfer rate. However, the secondary N-terminal amine carrying an auxiliary has more steric hindrance than a typical N-terminus. Thus, the *N*-to-*S* acyl transfer rate is much slower than that of NCL and becomes a rate-determining step. Due to the low ligation rate, the significant hydrolysis of the peptide thioesters and C-terminal residues epimerization problem may occur. Although this strategy could be applied to any ligation site in theory, a less hindered ligation site such as Gly-Gly or Gly-Ala is preferred. Recently, another auxiliary, 2-mercapto-2-(pyridine-2-yl) ethyl (MPyE), has been reported to have a significant improvement in the acyl transfer efficiency [24]. Its pyridine structure on auxiliary handle serves as an internal base catalyst that accepts hydrogen from the amine group, providing an extra driving force for acyl transfer.

1.2.1.3 Backbone-installed Split Intein-assisted Ligation

NCLs usually require millimolar (mM) concentration to proceed. Otherwise, their conversion yields could be dramatically low. However, such concentration cannot be achieved for low-solubility peptides, especially membrane protein fragments, due to their high aggregation potency. Recently, Liu's group developed backbone-installed split intein-assisted ligation (BISIAL) to overcome this problem (Figure 1.2d) [25]. The technique combines NCL and the consensus-fast (Cfa) split intein, which has a unique sequence allowing condensation and further engaging fusion of N- and C-extein without the presence of transpeptidases under denaturing

conditions [26,27]. In this strategy, some short peptide sequences were designed as “exteins” (Exts) and installed on the desired peptide segment as a branch of removable backbone modification (RBM). Then, Cfas were linked to the extein containing peptides through NCL. BISIAL occurred once the Cfa^N-containing peptide and Cfa^C-containing peptide are mixed together. The reaction maintained a high reaction rate (within 2 hours) and high conversion yield at low concentration (50 μM). After NCL, Cfas was automatically split, and only connected Exts remained on the ligated product through RMB. Finally, RBMs were cleaved by acidic cleavage cocktails, providing a native peptide structure. Moreover, BISAL showed high efficiency on a variety of C-terminal amino acids, even those with steric hindrance, including Val (88% yield) and Pro (79% yield).

1.2.2 Diselenide/Selenoester Ligation

The 21st amino acid selenocysteine (Sec) was also investigated for peptide ligation with a similar mechanism and chemoselectivity with NCL (Figure 1.2e) [28]. Compared with Cys, Sec is more prone to undergo air oxidation. Thus, reducing reagents are required under ligation conditions to maintain the existence of selenolate monomer [29]. Due to Sec's lower pKa [1], the ligation can occur under slightly lower pH, under which less hydrolysis happens on the thioester. In the meantime, peptide selenoesters exhibit fast *trans*-selenoesterification even under high diluted concentrations due to its liability. Moreover, the removal of a selenol group is milder and more accessible than desulfurization due to a weaker C–Se bond. Deselenization does not require a radical initiator, and it can be achieved with a reducing agent TCEP and hydrogen donor present [30]. The production of a phosphine selenide, which is highly thermodynamically favorable, further facilitates the cleavage of the C–Se bond. The produced β-carbon-centered radical can then abstract a hydrogen atom to form the alanine residue [1]. In-chain native Cys residues do not participate in the deselenization conditions, so specific Cys protection is not necessarily required. Similar to NCL, researchers have contributed to the synthesis of selenylated amino acids, thereby synthesizing five β- or γ-substituted selenylated amino acids.

1.3 Ligation Driven by Imine Capture

1.3.1 Serine/Threonine Ligation

Chemoselective imine formation between the N-terminus and a C-terminus aldehyde has been also used to realize chemoselective condensation. Kemp et al. first demonstrated the possibility of applying imine capture and *O*-to-*N* acyl transfer to the amide bond formation with benzaldehyde [31]. Later, unprotected peptides with