

ENABLING TOOLS AND TECHNIQUES FOR ORGANIC SYNTHESIS

A PRACTICAL GUIDE TO EXPERIMENTATION,
AUTOMATION, AND COMPUTATION

EDITED BY **STEPHEN G. NEWMAN**



WILEY

**Enabling Tools and Techniques
for Organic Synthesis**

Enabling Tools and Techniques for Organic Synthesis

A Practical Guide to Experimentation, Automation,
and Computation

Edited by

*Stephen G. Newman
Department of Chemistry & Biomolecular Sciences
University of Ottawa
Ottawa, Canada*

WILEY

Copyright © 2023 by John Wiley & Sons, Inc. All rights reserved.

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 Applied for

ISBN: 9781119855637, ePDF: 9781119855651, epub:9781119855644, oBook:9781119855668

Cover Design: Wiley

Cover Image: © Sergey Tarasov/Adobe Stock Photos

Set in 9.5/12.5pt STIXTwoText by Straive, Pondicherry, India

Contents

List of Contributors xv

Preface xix

1	Biocatalysis 101 – A Chemist’s Guide to Starting Biocatalysis	1
	<i>Pablo Díaz-Kruik, David Lim, and Francesca Paradisi</i>	
	Glossary	1
1.1	Introduction	1
1.1.1	Enzymes – the Green and Sustainable Way of the Future	1
1.1.2	Enzymatic and Organic Catalysis Are Not too Different from Each Other	3
1.1.3	Enzymes 101	4
1.2	When Should I Choose an Enzyme over a Chemical Catalyst?	4
1.3	Key Considerations for Running Biocatalytic Reactions	6
1.3.1	Dispelling Myths	6
1.3.1.1	Enzymes Are Not Safe to Use	7
1.3.1.2	Enzymes Are Not as Readily Available as Chemical Catalysts	7
1.3.1.3	Enzymes Are Seldom Useful Due to Their Limited Substrate Scope	7
1.3.1.4	The Cost of Enzyme Production Is Very High	8
1.3.1.5	Enzymes Are Functionally Unstable Under Organic Conditions	9
1.3.1.6	Sustainability	9
1.3.2	Challenges of Using Enzymes: the Need for Strict Reaction Conditions	9
1.3.2.1	Enzymes from Extremophiles	10
1.3.2.2	Solvents (and Co-solvents)	10
1.3.2.3	Concentration and Ionic Strength of the Buffer	10
1.3.2.4	pH Dependence	11
1.3.2.5	Concentration of Reactants	11
1.3.2.6	Enzyme Concentration	12
1.3.2.7	Enzyme Forms	12
1.3.2.8	Toxicity	13
1.3.3	What Do I Need to Start Biocatalytic Experiments in My Lab?	13
1.3.4	Additional Considerations	14
1.4	Transformations Catalyzed by Enzymes	15
1.4.1	EC – The Enzyme Commission Number	15

1.4.1.1	EC 1 – Oxoreductases	15
1.4.1.2	EC 2 – Transferases	16
1.4.1.3	EC 3 – Hydrolases	16
1.4.1.4	EC 4 – Lyases	17
1.4.1.5	EC 5 – Isomerases	18
1.4.1.6	EC 6 – Ligases	18
1.4.1.7	EC 7 – Translocases	18
1.4.2	Some Applications of Selected Commercially Available Enzymes	19
1.4.2.1	Horseradish Peroxidase	19
1.4.2.2	Lysozyme	19
1.4.2.3	Trypsin	20
1.4.2.4	Candida Lipase B	20
1.4.2.5	Amino Acid Dehydrogenase	20
1.4.2.6	Glycosidases	21
1.4.3	Engineered (Unnatural) Reactions	21
1.5	New Trends and Technologies in Biocatalysis	21
1.5.1	Flow Biocatalysis and New Technologies	21
1.5.1.1	What Is Flow Biocatalysis?	21
1.5.1.2	How Does Flow Biocatalysis Work?	21
1.5.1.3	When Is a Flow Process More Beneficial for a Specific Transformation?	23
1.5.1.4	Should One Implement Every Enzymatic Reaction in Flow?	23
1.5.2	Enzyme Engineering	24
1.5.3	Photobiocatalysis	25
1.6	Flow Chart to Biocatalysis	25
1.7	Case Study: Setting up a Biotransformation	27
1.8	Concluding Remarks	31
	Additional Resources	31
	References	31

2 Introduction to Photochemistry for the Synthetic Chemist 37

Stefano Protti, Davide Ravelli, and Maurizio Fagnoni

Glossary 37

2.1	Introduction	38
2.1.1	Light to Make Your Synthesis Greener	38
2.1.2	A Way to Overcome HOMO/LUMO Interactions	39
2.2	How to Plan a Photochemical Synthesis	45
2.2.1	The Choice of the Solvent	45
2.2.2	Concentration of the Absorbing Species	47
2.2.3	The Reaction Vessel	48
2.2.4	Light Sources	48
2.2.4.1	Low-Pressure Mercury Arcs	49
2.2.4.2	Medium- and High-Pressure Mercury Arcs	50
2.2.4.3	Other Light Sources	50
2.2.5	From Batch to Flow Conditions	52

- 2.2.6 Preparation of the Sample 54
- 2.2.7 Safety Equipment 54
- 2.3 Selected Applications of Photochemical/Photocatalyzed Reactions 55
 - 2.3.1 Reactions Involving the C=C Double Bond 55
 - 2.3.2 Reactions Involving the C=O Double Bond 58
 - 2.3.3 Reactions Involving a Photoinduced Homolysis 60
 - 2.3.4 Reactions Involving Singlet Oxygen 62
 - 2.3.5 Reactions Involving a Photocatalytic Step 62
- 2.4 Conclusions 67
 - Acknowledgment 67
 - References 67

3 How to Confidently Become an Electrosynthetic Practitioner 73

Sylvain Charvet, Taline Kerackian, Camille Z. Rubel, and Julien C. Vantourout

- Glossary 73
- Abbreviations 76
- 3.1 Introduction 77
- 3.2 General Definition of Organic Electrosynthesis 78
- 3.3 Why is Organic Electrosynthesis Used? 78
- 3.4 How is Organic Electrosynthesis Performed? 78
- 3.5 Where to Start with Electrosynthesis? 79
 - Selected General Reviews 79
 - Selected General Guides 79
- 3.6 Electrasyn 2.0 80
 - 3.6.1 Machine and Consumables 80
 - 3.6.1.1 Opening the IKA ElectraSyn 2.0 Box 80
 - 3.6.1.2 Cell (Vial and Cap) 81
 - 3.6.1.3 Electrodes 82
 - 3.6.2 Interface 83
 - 3.6.2.1 Hardware 83
 - 3.6.2.2 Menus 84
 - 3.6.3 How to Set Up the Cell 84
 - 3.6.4 How to Start an Experiment 85
 - 3.6.5 During the Reaction 88
 - 3.6.6 After the Reaction 89
- 3.7 Case Study 90
 - 3.7.1 Project Overview 90
 - 3.7.2 Optimization of Parameters 92
 - 3.7.2.1 Designing an Electrochemical Experiment 92
 - 3.7.3 Proof of Concept 94
 - 3.7.3.1 Optimization 94
 - 3.7.3.2 Substrate Scope 102
- 3.8 Conclusion 103
- References 103

4	Flow Chemistry	107
	<i>Yosuke Ashikari and Aiichiro Nagaki</i>	
	Glossary	107
4.1	Introduction	109
4.1.1	What is Flow Microchemistry	109
4.1.1.1	Reaction Time Controllability	110
4.1.1.2	Fast Mixing	111
4.1.1.3	Temperature Controllability	112
4.1.2	Reactions Enabled by Flow Microreactors	112
4.1.2.1	Competitive Sequential Reactions	112
4.1.2.2	Reactions Mediated by Unstable Intermediates	114
4.1.2.3	Reactions Occurring at the Surface: Two-Phase Reactions, Electrochemical Reactions, and Photoreactions	117
4.1.3	Further Applicability of Flow Microsynthesis	118
4.1.3.1	Scalability	118
4.1.3.2	Safety Operation	118
4.2	General Information for Flow Microreactors	118
4.2.1	Tools and Equipment for Flow Chemistry	119
4.2.1.1	Micromixer	119
4.2.1.2	Tube Reactor	120
4.2.1.3	Pump	120
4.2.1.4	Pre-Cooling Tubes	121
4.2.1.5	PTFE Tubes	121
4.2.2	How to Perform Experiments	122
4.2.2.1	Selection of Reaction Conditions	122
4.2.2.2	Preparation of Reagent Solution	125
4.2.2.3	Preparation for Reactions	126
4.2.2.4	Preparation for Reaction Evaluation	128
4.2.2.5	Cleaning Up	129
4.3	Case Studies	129
4.3.1	Competitive Sequential Reaction (General Procedure)	129
4.3.1.1	Preparation	130
4.3.1.2	Experiment	132
4.3.1.3	Screening of Reaction Conditions	133
4.3.1.4	Analysis	134
4.3.1.5	Clean Up	136
4.3.2	Reactions Mediated by Short-Lived Intermediates	136
4.3.3	Reaction Integration	139
4.4	Further Expertise	142
4.4.1	Reaction Integration	142
4.4.2	Chemoselective Reactions	143
4.4.3	Heterogeneous Catalytic Reactions	143
4.5	Summary and Outlook	144
	References	144

5	Reaction Optimization Using Design of Experiments	149
	<i>Laura Forfar and Paul Murray</i>	
	Glossary	149
5.1	Introduction	151
5.1.1	How Do We Experiment and DoE Terminology	151
5.1.2	OVAT vs. DoE	153
5.1.2.1	A Simple Chemical Example	153
5.1.3	A Note on Error, Accuracy, and Precision	156
5.2	When and How Can DoE Be Used?	157
5.3	What Information Can I Get from a DoE and How Is It Obtained?	158
5.3.1	Which Factors Are Important?	159
5.3.2	How Are the Models Generated?	161
5.4	What Types of Design Are Available?	164
5.4.1	Screening Designs	164
5.4.1.1	Fractional Factorial Designs	165
5.4.1.2	Definitive Screening Designs	166
5.4.2	Designs for Optimizing Reactions	167
5.4.3	Response Surface Designs	167
5.5	The DoE Process	169
5.5.1	Aim and Objective	170
5.5.2	Selecting Factors and Ranges	171
5.5.2.1	Factors	171
5.5.2.2	Ranges	173
5.5.3	Selecting Responses	175
5.5.4	Select a Design to Answer the Objective	176
5.5.5	Carry Out Design and Analyze Samples	177
5.5.6	Check Results	178
5.5.7	Model Data	179
5.5.7.1	General Steps for Developing a Model	180
5.5.7.2	Wittig Reaction	181
5.5.7.3	Complementing the Design	187
5.5.8	Validate Predictions	189
5.6	Combining DoE with Other Screening and Optimization Techniques	191
5.7	Software	192
5.8	“I Tried Experimental Design But It Did Not Work”	193
5.9	Conclusion	194
	References	195
6	Introduction to High-Throughput Experimentation (HTE) for the Synthetic Chemist	197
	<i>Stephanie Felten, Michael Shevlin, and Marion H. Emmert</i>	
	Glossary	197
6.1	What Is HTE?	199

- 6.2 Why HTE and What Can It Achieve? 199
 - 6.2.1 Commonly Perceived Barriers to Employing HTE in Synthetic Chemistry 200
 - 6.2.1.1 Cost 200
 - 6.2.1.2 Availability of Dedicated HTE Facilities 200
 - 6.2.1.3 Access to Knowledge and Training 201
 - 6.2.1.4 Perception of HTE as Antithesis of Hypothesis-driven Research 201
 - 6.2.2 Advantages of HTE Workflows vs. Traditional Reaction Setup 203
 - 6.2.2.1 Setup Time per Reaction 203
 - 6.2.2.2 Miniaturization and Efficient Reagent Use 203
 - 6.2.2.3 Multivariable vs. Sequential Optimization 203
 - 6.2.2.4 Visualizing Reactivity Patterns 204
 - 6.2.2.5 Serendipity in Reaction Discovery 204
 - 6.2.2.6 Avoiding Cross-contamination 206
- 6.3 Practical Considerations and Tools for HTE 206
 - 6.3.1 Outline of a Typical HTE Workflow 207
 - 6.3.2 Types of HTE Designs 209
 - 6.3.2.1 HTE for Reaction Discovery 209
 - 6.3.2.2 HTE for Reaction Optimization 210
 - 6.3.3 HTE Design Software: Tools for Building Arrays 211
 - 6.3.4 HTE Reactors and Consumables 214
 - 6.3.4.1 Reaction Blocks 214
 - 6.3.4.2 HTE Vials 214
 - 6.3.4.3 Reaction Blocks with Sealing Top Plate 215
 - 6.3.4.4 Special Reactors for Photochemistry, Electrochemistry, and High-Pressure Reactions 215
 - 6.3.4.5 Reaction Stirring and Temperature Control 217
 - 6.3.4.6 Consumables 219
 - 6.3.5 Considerations for Experimental Setup 220
 - 6.3.5.1 Reaction Atmosphere 220
 - 6.3.5.2 Reagent Preparation and Dispensing 221
 - 6.3.5.3 Storage of Preplated Reagents 223
 - 6.3.5.4 Pipetting 224
 - 6.3.5.5 Solvent Evaporation 225
 - 6.3.6 Analysis of HTE Screens 226
 - 6.3.6.1 Suitable Instrumentation 226
 - 6.3.6.2 Autosampler Configurations 226
 - 6.3.6.3 Analytical Methods 227
 - 6.3.6.4 Internal Standards and Assay Yields 227
 - 6.3.6.5 Data Visualization and Analysis 228
 - 6.3.7 The Role of Automation and Robotics in HTE 229
- 6.4 Section Summary and Outlook 232
- 6.5 Case Study 1: Development of an HTE Platform for Nickel-Catalyzed Suzuki–Miyaura Reactions 233

6.5.1	Motivation	233
6.5.2	Design of Test Reaction and Initial Ligand Screen	233
6.5.3	Second Round of Ligand/Base/Solvent Screens	235
6.5.4	Final Platform Design	237
6.5.5	Validation of Platform Design	237
6.6	Case Study 2: HTE Enabled Reaction Discovery and Optimization of Silyl-Triflate-Mediated C–H Aminoalkylation of Azoles	240
6.6.1	Motivation	240
6.6.2	Reaction Discovery Plate Design	240
6.6.3	Ligand Screen	243
6.6.4	Parallel Optimization of Three Reagents	244
6.6.5	Base Screen	244
6.7	Current Challenges and the Future of HTE	247
6.7.1	Summary and Conclusions	247
6.7.2	Remaining Challenges: The Next Frontiers	248
6.7.2.1	Biphasic Reaction Mixtures	248
6.7.2.2	Flow Chemistry and HTE	248
6.7.2.3	Reaction Profiling	249
6.7.2.4	Building Machine Learning Models to Predict Reactivity	249
6.7.2.5	Addressing Future Challenges	250
	Acknowledgments	250
	Further Recommended Reading	250
	References	250
7	Concepts and Practical Aspects of Computational Chemistry	259
	<i>Martin Breugst</i>	
	Glossary	259
7.1	Introduction	261
7.2	Hardware and Software Requirements for Computational Investigations	264
7.3	Typical Methods in Computational Organic Chemistry	265
7.3.1	General Aspects	265
7.3.2	Molecular Mechanics and Force Fields	266
7.3.3	Wave-Function Methods I – Hartree–Fock Theory	267
7.3.4	Wave-Function Methods II – Post-Hartree–Fock Theory	267
7.3.5	Semiempirical Methods	269
7.3.6	Density Functional Theory	269
7.3.7	Dispersion-Corrected Density Functional Theory	271
7.3.8	Typical Computational Times	272
7.4	Basis Sets Used in Computational Organic Chemistry	273
7.4.1	General Aspects of Basis Sets	273
7.4.2	Introduction to the Mathematical Formalism in Basis Sets	274
7.4.3	Polarization and Diffuse Functions	275
7.4.4	Basis Set Families	276

- 7.4.5 Effective Core Potentials (Pseudopotentials) 278
- 7.4.6 The Basis Set Superposition Error (BSSE) 279
- 7.5 Typical Computational Tasks in Organic Chemistry 279
 - 7.5.1 Preliminary Remarks 279
 - 7.5.2 Single-Point Calculations 281
 - 7.5.3 Geometry Optimizations 281
 - 7.5.4 Frequency Calculations 282
 - 7.5.5 Intrinsic Reaction Coordinate (IRC) Calculations 284
 - 7.5.6 Conformational Analysis 285
- 7.6 Notation of the Model Chemistry 286
- 7.7 The Diels–Alder Reaction as a Tutorial Case Study 286
 - 7.7.1 General Aspects and Requirements 286
 - 7.7.2 Preparing Input Files 288
 - 7.7.3 Conformational Sampling – Generation of Initial Geometries 290
 - 7.7.4 Geometry Optimizations of Starting Materials and Products 291
 - 7.7.5 Locating the Transition States 294
 - 7.7.6 Verifying the Nature of the Transition State 298
- 7.8 More Advanced Aspects 300
 - 7.8.1 General Comments 300
 - 7.8.2 Influence of Solvation 300
 - 7.8.3 Integration Grid 302
 - 7.8.4 Standard States 302
 - 7.8.5 Treating Unpaired Electrons 303
- 7.9 Important and Frequently Used Keywords 304
- 7.10 Practical Considerations 304
- 7.11 Conclusions 306
- References 306

8 NMR Prediction with Computational Chemistry 313

Amy T. Merrill, Wentao Guo, and Dean J. Tantillo

Glossary 313

- 8.1 Introduction 314
- 8.2 Quantum-Chemistry-Based Computational NMR 315
 - 8.2.1 Methods 315
 - 8.2.1.1 Time/Resources for Calculations 316
 - 8.2.1.2 Structural Considerations in Modeling 317
 - 8.2.1.3 Geometry Optimizations 323
 - 8.2.1.4 Calculating Isotropic Shielding Constants 324
 - 8.2.1.5 Common Pitfalls and How to Address Them 328
 - 8.2.1.6 Converting to Chemical Shifts 329
 - 8.2.1.7 Calculating Coupling Constants 330
 - 8.2.2 Confidence Analysis 330
 - 8.2.3 Computer-Aided Automated Approaches 332

- 8.2.3.1 CASE 332
- 8.2.4 A Case Study 336
- 8.2.5 Practicing ^1H and ^{13}C Chemical Shift Prediction 338
- 8.3 Summary and Outlook 339
 - Key References 339
 - References 340

9 Introduction to Programming for the Organic Chemist 347

Jason M. Stevens

- 9.1 Introduction 347
- 9.2 Better Visualizations: Communicating Structure–Data Relationships 351
- 9.3 Text Extraction: Automating Density Functional Theory Calculations 354
- 9.4 Statistical Analysis: Deriving Insight from Historical Data 357
- 9.5 Machine Learning: A Predictive Model for Deoxyfluorination 359
- 9.6 Working with Public Datasets: Identifying Reactivity Cliffs 364
- 9.7 Running Simulations: Process Greenness 367
- 9.8 Application Development: Process Mass Intensity Predictor 371
- 9.9 Machine Learning for Reaction Optimization 374
- 9.10 Executing Robotic Tasks 378
- 9.11 Autonomous Reaction Optimization 381
- 9.12 Conclusion 384
 - References 385

10 Machine Learning for the Optimization of Chemical Reaction Conditions 393

A. Filipa de Almeida and Tiago Rodrigues

Glossary 393

- 10.1 Introduction 394
- 10.2 Prior Art and Alternative Methods for Rational Reaction Optimization 396
- 10.3 Reaction Optimization Using LabMate.ML 400
 - 10.3.1 Step One: Accessing the LabMate.ML Code and Installation 401
 - 10.3.2 Step Two: Initializing the Optimization Routine in LabMate.ML 402
 - 10.3.3 Step Three: Iterative Optimization Routine 404
 - 10.3.4 Examples 406
- 10.4 Primer on Evaluation Guidelines 408
 - 10.4.1 Code and Dataset Availability 408
 - 10.4.2 Retrospective Evaluation 409
 - 10.4.3 Baselines and Comparing Tools 410
 - 10.4.4 Prospective Evaluation 412

- 10.5 Outlook 414
- References 416

- 11 Computer-Assisted Synthesis Planning 423**
Zhengkai Tu, Itai Levin, and Connor W. Coley
- Glossary 423
- 11.1 Introduction to Computer-Aided Synthesis Planning 424
- 11.1.1 Defining the Tasks and Use Cases 424
- 11.1.2 Historical Approaches to Computer-Aided Synthesis Planning 425
- 11.1.3 The Inflection Point of CASP Methods 425
- 11.1.4 Preliminaries on Molecular Representation and Cheminformatics 426
- 11.1.5 Outline of the Rest of the Chapter 428
- 11.2 Approaches and Algorithms for Retrosynthesis 428
- 11.2.1 Data-driven v. Expert-Driven Programs 428
- 11.2.2 Template-Based Approaches 429
- 11.2.3 Template-free Approaches with Graphs and Sequences 431
- 11.2.4 Multistep Planning Algorithms 433
- 11.3 Approaches and Algorithms for Condition Recommendation and Forward Synthesis 436
- 11.3.1 Condition Recommendation Approaches 436
- 11.3.2 Forward Synthesis Approaches 437
- 11.4 Select Examples of Software Tools for CASP 439
- 11.4.1 Open-Source Tools 439
- 11.4.1.1 ASKCOS 439
- 11.4.1.2 AiZynthFinder 440
- 11.4.1.3 Retro* 442
- 11.4.2 Closed-Source Tools 443
- 11.4.3 CASP Tools for Enzymatic Catalysis 446
- 11.4.4 Practical Considerations for CASP Programs 446
- 11.4.4.1 Traceability to Literature Precedent 447
- 11.4.4.2 How to Use CASP: Command Line Versus Graphical User Interface 447
- 11.4.4.3 Data Privacy 448
- 11.4.4.4 Customization Ability 448
- 11.5 Case Studies 448
- 11.5.1 Segler et al.'s Data-driven Program and A/B Testing Success 449
- 11.5.2 MIT's ASKCOS Program and Robotic Synthesis Demonstration 449
- 11.5.3 Grzybowski's Chematica/Synthia Program's Experimental Validations and Acquisition 450
- 11.6 Conclusion 451
- Key References 453
- References 453

List of Contributors

Yosuke Ashikari

Department of Chemistry
Faculty of Science
Hokkaido University
Sapporo, Japan

Martin Breugst

Institut für Chemie
Technische Universität Chemnitz
Chemnitz, Germany

Sylvain Charvet

Institut de Chimie et Biochimie
Moléculaires et Supramoléculaires
(ICMBS, UMR 5246 du CNRS)
Université Lyon
Villeurbanne, France

Connor W. Coley

Department of Chemical Engineering
Massachusetts Institute of Technology
Cambridge, MA, USA

and

Department of Electrical
Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, MA, USA

Pablo Díaz-Kruik

Department of Chemistry
Biochemistry and Pharmaceutical
Sciences
Bern, Switzerland

Marion H. Emmert

Process Research & Development
Merck & Co., Inc.
Rahway, NJ, USA

Maurizio Fagnoni

PhotoGreen Laboratory
Department of Chemistry
University of Pavia
Pavia, Italy

Stephanie Felten

Process Research & Development
Merck & Co., Inc.
Rahway, NJ, USA

A. Filipa de Almeida

Instituto de Investigação do
Medicamento (iMed)
Faculdade de Farmácia
Universidade de Lisboa
Lisbon, Portugal

Laura Forfar

Paul Murray Catalysis Consulting Ltd
Yate, Bristol, UK

Wentao Guo

Department of Chemistry
University of California
Davis, CA, USA

Taline Kerackian

Institut de Chimie et Biochimie
Moléculaires et Supramoléculaires
(ICMBS, UMR 5246 du CNRS)
Université Lyon
Villeurbanne, France

Itai Levin

Synthetic Biology Center
Department of Biological Engineering
Massachusetts Institute of Technology
Cambridge, MA, USA

and

Department of Chemical Engineering
Massachusetts Institute of Technology
Cambridge, MA, USA

David Lim

Department of Chemistry
Biochemistry and Pharmaceutical
Sciences
Bern, Switzerland

Amy T. Merrill

Department of Chemistry
University of California
Davis, CA, USA

Paul Murray

Paul Murray Catalysis Consulting Ltd
Yate, Bristol, UK

Aiichiro Nagaki

Department of Chemistry
Faculty of Science
Hokkaido University
Sapporo, Japan

Francesca Paradisi

Department of Chemistry
Biochemistry and Pharmaceutical
Sciences
Bern, Switzerland

Stefano Protti

PhotoGreen Laboratory
Department of Chemistry
University of Pavia
Pavia, Italy

Davide Ravelli

PhotoGreen Laboratory
Department of Chemistry
University of Pavia
Pavia, Italy

Tiago Rodrigues

Instituto de Investigação do
Medicamento (iMed)
Faculdade de Farmácia
Universidade de Lisboa
Lisbon, Portugal

Camille Z. Rubel

Institut de Chimie et Biochimie
Moléculaires et Supramoléculaires
(ICMBS, UMR 5246 du CNRS)
Université Lyon
Villeurbanne, France
and

Department of Chemistry
The Scripps Research Institute
La Jolla, CA, USA

Michael Shevlin

Process Research & Development
Merck & Co., Inc.
Rahway, NJ, USA

Jason M. Stevens

Bristol Myers Squibb
Summit, NJ, USA

Dean J. Tantillo

Department of Chemistry
University of California
Davis, CA, USA

Zhengkai Tu

Computational Science and
Engineering
Massachusetts Institute of Technology
Cambridge, MA, USA

Julien C. Vantourout

Institut de Chimie et Biochimie
Moléculaires et Supramoléculaires
(ICMBS, UMR 5246 du CNRS)
Université Lyon
Villeurbanne, France

Preface

At the undergraduate level, the organic chemistry curriculum at most universities is similar. Professors emphasize the fundamental concepts necessary to understand how, when, and why organic molecules interact, while lab instructors familiarize students with important hands-on aspects of carrying out experiments. Bachelor's students can expect to finish their studies with an idea of how molecules behave, how they are made, and how technologies such as NMR and IR can be used for their characterization. Those that enter graduate school are often surprised at the breadth of powerful technologies that make advanced organic chemistry the discipline it is today. Instead of a typical stirred round-bottomed flask, many reactions are better done using photochemical, electrochemical, or flow reactors. Computational chemistry, once reserved for dedicated experts, can now be used by organic chemists to help predict outcomes, understand selectivity, and decipher reaction mechanisms. Automation technology can be used to generate large amounts of data with limited amounts of material, and data processing software can be used to extract subtle trends.

Due to their prominence in the recent literature, trainees and established chemists alike would benefit from gaining expertise with these technologies to be best prepared for solving the diverse synthetic challenges that come their way. However, the barrier to learning techniques without formal instruction can be high. Even if one is fortunate enough to have access to advanced training, the expert instructor may not necessarily curate the course to the needs and the background of a synthetic chemist. The primary literature and recent textbooks have similar limitations – while there is no shortage of resources, experts generally write to other experts, and the interested organic chemist may have critical gaps in their understanding and struggle with subdiscipline-specific jargon.

The goal of this text is to help fill this gap by providing synthetic chemists with a user-friendly starting point to initiate their journey in developing new skills and knowledge. In each of the 11 chapters, experts communicate basic information about an impactful technology in a manner accessible to a classically trained

synthetic chemist. Chapters also includes a glossary of common terminology, a general introduction to the technology of interest, case-study examples of how it may useful to synthetic chemists, a practical discussion about steps one may take to put knowledge into practice, and references to recommended further reading. The book seeks to be a go-to resource for organic chemists at or above the graduate level that wish to expand the breadth of tools they can use to perform, analyze, and interpret chemistry experiments. After completion, the reader will be armed with the practical knowledge needed to comprehend the literature, to assess the strengths and limitations of each technique, and to begin applying modern tools to solve synthetic challenges. This will make it useful as a general resource for graduate students looking to expand their expertise, for instructors of graduate-level courses on advanced techniques for organic synthesis, and for industrial scientists seeking a beginner-friendly way to expand their knowledge.

The book is organized into four subsections. Chapters 1–4 describe different enabling technologies for performing chemical experiments – biocatalysis, photochemistry, electrochemistry, and flow chemistry. While none of these topics are fundamentally new, their power as a tool for organic synthesis is becoming increasingly evident. These chapters will help the reader overcome the technical barrier hindering them from comfortably replicating experiments and designing their own. Chapters 5 and 6 focus on improved approaches to select, carry out, and analyze experiments. Specifically, Chapter 5 describes a statistical approach to experimentation that can be used to understand and optimize chemical reactions. This Design of Experiments (DoE) technique is commonly employed by practicing scientists in many fields but is seldom taught to chemists. Chapter 6 describes techniques that researchers can use to get more data using less time and fewer resources. This high-throughput experimentation (HTE) approach shows the reader how to carry out reactions in parallel and how the collected data can be interpreted to gain insights that might otherwise be missed. Chapters 7 and 8 introduce the reader to computational chemistry tools that enable molecules and reactions to be modeled *in silico*, providing predictions and mechanistic insight to supplement experimentation. Chapter 7 provides a general overview of the most common computational tasks that an organic chemist may want to carry out and walks the reader through a beginner-friendly case study wherein the reactants, transition states, and products of a Diels–Alder reaction are calculated. Chapter 8 builds upon the general knowledge given in the previous chapter and describes how computational chemistry can be used to predict the NMR spectrum of organic molecules. The goal of this chapter is to put this powerful technique into the hands of experimental chemists, which should be achievable after familiarizing themselves with the simplified approach detailed throughout. Chapters 9–11 provide the reader with an introduction to programming and machine learning. Computers already play a critical role in the daily life of a synthetic chemist, and

a little bit of familiarity with modern techniques can go a long way. Chapter 9 provides a blueprint for understanding how and why a chemist may go about familiarizing themselves with programming. Chapter 10 describes a deep dive case study for using machine learning to facilitate reaction optimization, providing a step-by-step guide that a beginner may follow to use the tool and to gain confidence in harnessing other published algorithms. Chapter 11 explains how computers can facilitate the planning of multistep synthesis by suggesting synthetic routes and reaction conditions. Helpful discussions on the current tools available, how they work, and their associated strengths and weaknesses are also described.

This project was only possible due to an immense amount of work by the authors who generously agreed to share their knowledge and meet the formidable task of communicating with a general audience. I am also indebted to the many students and postdoctoral fellows at the University of Ottawa that served as reviewers to help ensure that the content serves as a welcoming and beginner-friendly introduction to these topics that are becoming increasingly important to the modern synthetic chemists. I hope the readers agree that this goal has been met and that this marks the beginning of their journey to being a more well-rounded scientist capable of tackling diverse problems that come their way.

Stephen G. Newman
Ottawa

July 2023

1

Biocatalysis 101 – A Chemist’s Guide to Starting Biocatalysis

Pablo Díaz-Kruik, David Lim, and Francesca Paradisi

Department of Chemistry, Biochemistry and Pharmaceutical Sciences, Bern, Switzerland

Glossary

API Active pharmaceutical ingredient

BRENDA A comprehensive enzyme information system

CALB Lipase B from *Candida antarctica*

Cofactor A non-protein chemical compound or metallic ion that is required for an enzyme’s role as a catalyst

DKR Dynamic kinetic resolution

IRED Imine reductase

NAD⁺/NADH Nicotinamide adenine dinucleotide

Protein expression Biological process where the protein is synthesized inside a cell

Recombinant DNA DNA scaffold that contains the protein sequence of interest

TRIS Tris(hydroxymethyl)aminomethane

1.1 Introduction

1.1.1 Enzymes – the Green and Sustainable Way of the Future

Recent efforts by chemists to actively reduce toxic waste production and minimize costs have led to the discovery of many green and sustainable technologies. Not surprisingly, the use of enzymes, Nature’s catalysts, has seen a major resurgence in academic and industrial interest over the past decade – not only for their sustainability and natural activities but for engineering them to perform novel transformations beyond capabilities observed in a synthetic organic lab [1, 2].

Enabling Tools and Techniques for Organic Synthesis: A Practical Guide to Experimentation, Automation, and Computation, First Edition. Edited by Stephen G. Newman.

© 2023 John Wiley & Sons, Inc. Published 2023 by John Wiley & Sons, Inc.

The attractiveness of using enzymes for transformations stems from their exquisite regio- and stereoselectivities – something that traditional chemists still struggle to achieve in the lab – that enzymes often execute effortlessly. Moreover, we have seen the emergence of multienzyme cascades for the synthesis of active pharmaceutical ingredients (APIs). A recent landmark example involves the synthesis of molnupiravir (MK-4482), an orally dosed ribonucleoside analogue and inhibitor of influenza viruses, which has demonstrated activity against COVID-19 when administered in animal models [3, 4]. In this work, McIntosh et al. developed a scalable three-step route toward MK-4482 [5]. Using a cascade of five enzymes, MK-4482 could be accessed from 5-isobutyrylribose (Figure 1.1).

To the uninitiated, entering the world of enzyme-catalyzed chemical transformations can be incredibly daunting, especially when one is not equipped with a foundational understanding of what an enzyme is and how these macromolecules work. However, you may be surprised to hear that enzymology and chemistry are not too different from each other at all! With an undergraduate chemistry background, a chemist can easily harness the power of enzymes to perform desired transformations – a *fact* that we aim to convince you of over the next few pages.

However, while this chapter aims to illustrate the power of enzymes for novel and sustainable transformations, we do not want to inadvertently imply the use of these macromolecules is the be-all-end-all solution – sometimes the use of traditional organic synthesis to access target molecules is the more logical solution. Therefore, when an enzyme *might* be used is a weighted question often involving the combination of various intricate factors, including efficiency and cost.

Over the following sections, we will do our best to educate you on these factors so that you can begin making an informed decision on this matter. We also aim to

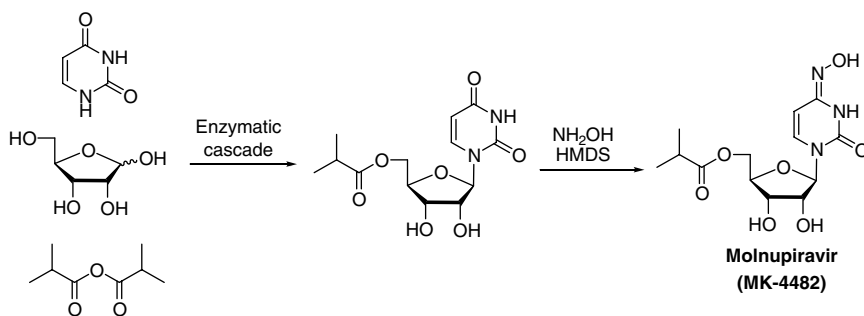


Figure 1.1 A combined enzymatic cascade/hydroxylamination for the synthesis of molnupiravir (MK-4482).

convince the reader that the use of enzymes is not limited to biologists and biochemists but also readily available for use by synthetic chemists. With the following breakdown of important considerations to make when using an enzyme, we hope to instill confidence in the reader that a biological catalyst is not too dissimilar to a chemical catalyst and can be readily obtainable from common suppliers.

We will also dispel common misconceptions and myths surrounding the use of enzymes and then give an overview of several classes of reactions that can be performed with enzymes, including recent developments into more exotic transformations such as photobiocatalysis.

This chapter will then conclude with a snippet into recent trends and technologies that have harnessed the use of enzymes in novel ways. We hope that the information gained from reading this chapter will provide a strong foundation for the reader to develop confidence in the use of enzymes and begin their venture into the world of biocatalysis.

1.1.2 Enzymatic and Organic Catalysis Are Not too Different from Each Other

A seasoned chemist may be quite familiar with several stereoselective reactions whereby stereocontrol is dictated by the chiral environment of the reaction. For example, one model for the Corey–Bakshi–Shibata (CBS) reduction involves coordination of the respective carbonyl to the CBS catalyst in a specific spatial orientation, leading to stereoselective reduction of the carbonyl to the corresponding alcohol (Figure 1.2) [6]. Enzymes utilize a very similar concept to this reaction – the catalyst (enzyme) places a reactant (substrate) in a chiral environment (the active site), whereby stereoselectivity is dictated by the local reactive environment, leading to a selective reaction outcome. In the next subsection, we will look at how an enzyme achieves these feats.

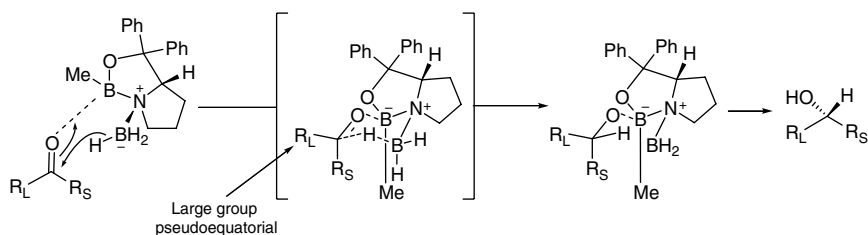


Figure 1.2 The CBS reaction has been used in undergraduate texts as a classic example of where an achiral reactant is stereoselectively transformed in a chiral environment to the corresponding product in high enantiomeric excesses.

1.1.3 Enzymes 101

Enzymes are known to accelerate reactions by more than 10^{17} -fold [7]. How an enzyme achieves these colossal rate increases under aqueous conditions requires an understanding of the active site architecture in great molecular detail.

There are 20 essential amino acids found in nature. Enzymes are formed by cellular machinery, which stitch together combinations of these amino acids in a genetically pre-defined sequence, making one very long polymer. This polymer is folded to give a precise three-dimensional structure (Figure 1.3). The active site is defined as the region of the enzyme where substrates bind and undergo catalysis. The catalytic cycle begins with the binding of the substrate in the active site. This process precisely positions all molecules involved in the catalysis (metals, solvents, cofactors, etc.) in their respective orientations ready to achieve regio- and stereoselectivity. Subsequent activation of the substrate initiates the reaction, generating a transition state, which is stabilized by interactions with the active site residues of the enzyme. Following effective conversion of the substrate, the product is then released from the active site of the enzyme, completing one turnover and returning the catalyst back to its original state.

1.2 When Should I Choose an Enzyme over a Chemical Catalyst?

The choice of using a chemical catalyst over a biochemical solution needs to be assessed on a case-by-case basis, often involving a detailed cost–benefit analysis. For example, chemical asymmetric imine reduction often requires the use of expensive precious metals, such as Ir, Rh, Ru, and Pd (Figure 1.4) [8]. While recent methods have moved toward Earth-abundant solutions, such as employing iron or nickel, all these still require decoration with expensive chiral ligands that cannot be recycled [8], making the overall synthesis very environmentally and economically demanding.

In contrast, imine reductases (IREDs) can perform stereoselective reductions without the use of expensive metals and can be performed under aqueous conditions mitigating the need for organic solvents. Since the initial report of IREDs in 2010 [9], many advancements have been made to use these enzymes for novel synthetic transformations [9, 10]. In fact, Matzel et al. published an elegant procedure for performing biocatalytic dynamic kinetic resolutions (DKRs) of aldehydes using IREDs (Figure 1.5). This method exploits the stereo-preference of the enzyme for either the *R*- or *S*-chiral center [11].

The use of enzymes in this case showcases the re-opening of the chemical window, enabling unprecedented reaction conditions, merging asymmetric reduction

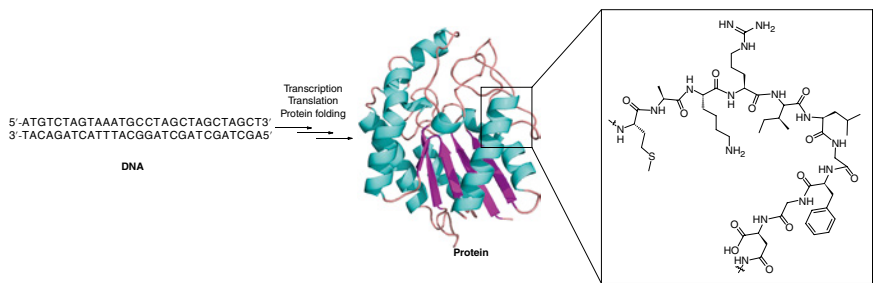


Figure 1.3 Proteins are formed by intracellular machinery that uses genetic information (DNA) to form a polymeric amino acid chain, which is then precisely folded to give a protein.

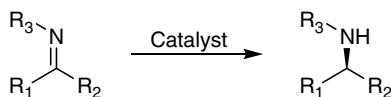


Figure 1.4 Asymmetric reduction of imines to amines in the presence of a chiral catalyst.

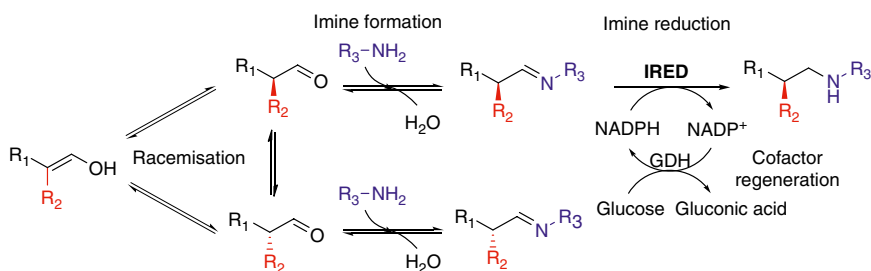


Figure 1.5 Biocatalytic dynamic kinetic resolutions of aldehydes using imine reductases. GDH = glucose dehydrogenase.

and water media. This would be near impossible to achieve with classical reducing agents, such as NaBH_4 or $\text{Na}(\text{CH}_3\text{COO})_3\text{BH}$.

Next, we move the reader on to build an understanding of the variables associated with using a biocatalyst and here, they will gain foundational knowledge on how to gauge whether a chemical or biochemical solution is appropriate for solving a target problem.

1.3 Key Considerations for Running Biocatalytic Reactions

With a primary level of appreciation of the power of enzymes, we can now continue our journey by addressing the variables that define a biocatalytic reaction. We will also illustrate how these variables change depending on the system that is being applied for the reaction. Before we advance in that direction, we will first begin by dissipating common myths surrounding the use of enzymes.

1.3.1 Dispelling Myths

The uptake of biocatalysis in academic and industrial applications has increased significantly in recent years [12, 13]. Despite the positive perception of the technology, the breadth of applications remains rather modest. A factor that contributes to this lack of progress may be associated with the perceived notion of the limitations of biocatalysts – their availability, cost, ease of use, substrate scope, and operational stability. We aim to address these factors in the following subsections.