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

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

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

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.

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

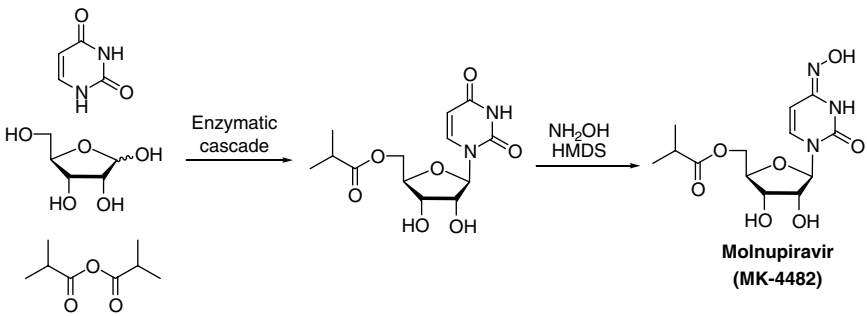

Figure 1.1 A combined enzymatic cascade/hydroxylamination for the synthesis of molnupiravir (MK-4482).

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.

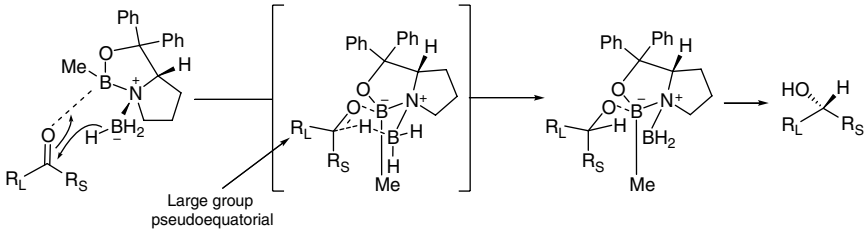

Figure 1.2 The CBS reaction has been used in undergraduate texts as a classic example of where an achiral reactant is stereoselectively transformed in a chiral environment to the corresponding product in high enantiomeric excesses.

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

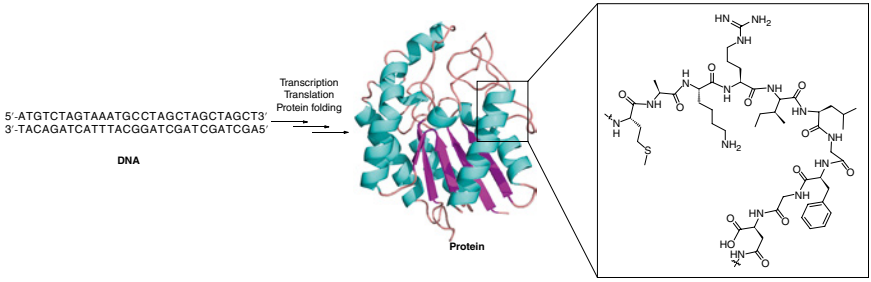

Figure 1.3 Proteins are formed by intracellular machinery that uses genetic information (DNA) to form a polymeric amino acid chain, which is then precisely folded to give a protein.

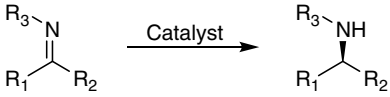

Figure 1.4 Asymmetric reduction of imines to amines in the presence of a chiral catalyst.

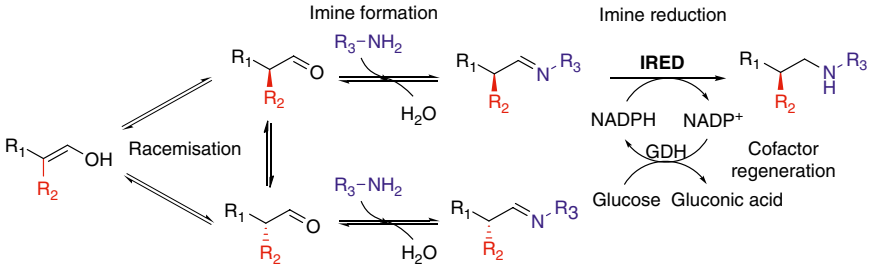

Figure 1.5 Biocatalytic dynamic kinetic resolutions of aldehydes using imine reductases. GDH = glucose dehydrogenase.

and water media. This would be near impossible to achieve with classical reducing agents, such as $NaBH_4$ or $Na(CH_3COO)_3BH$.

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.