Understanding Enzymes: Function, Design, Engineering, and Analysis focuses on the understanding of enzyme function and optimization gained in the past decade, past enzyme function analysis, enzyme engineering, and growing insights from the simulation work and nanotechnology measurement of enzymes in action in vitro or in silico. The book also presents new insights into the mechanistic function and understanding of enzyme reactions, as well as touching upon structural characteristics, including X-ray and nuclear magnetic resonance (NMR) structural methods. A major focus of the book is enzyme molecules' dependency on dynamic and biophysical environmental impacts on their function in ensembles as well as single molecules. A wide range of readers, including academics, professionals, PhD and master's students, industry experts, and chemists, will immensely benefit from this exclusive book.

Allan Svendsen did his master of science in biochemistry and protein chemistry in 1985 from the University of Copenhagen, Denmark. From 1986 to 2008, he worked as a research scientist at Nordisk Gentofte A/S and Novo Nordisk A/S. Later, he joined Novo Nordisk A/S and Novozymes A/S as a science manager and then became senior science manager at Novozymes A/S, where since 2008, he has been science director. His research area is protein engineering in general, but he has been working especially with insulin, proteases, oxidases, amylases, and lipases. His work has been within HPLC analysis, purification, downstream processing, and assay development. Lipase protein engineering and structural computer analysis and design of variants have been a central area for around 26 years.

Dr. Svendsen has worked as a coordinator for a lipase protein engineering project at the Nordic Industrial Foundation. He was the organizer and coordinator of an EU-funded nanotechnology project, “Lipid-Lipase Structure Function Relationship: Novel Methods for Studying Interfacial Enzyme Actions and Relation to Structure of Self-Assembled Lipid Nanostructures,” and the co-organizer of another EU-funded nanotechnology project, “BIO-SCOPE.” He has supervised many master's students and a few PhD students and has acted as censor for a number of PhD and masters theses defenses. He has reviewed manuscripts for various journals, co-edited a special issue of Colloids and Surface B: Biointerfaces (2001), and edited Enzyme Functionality: Design, Engineering and Screening (2002 and 2003 editions) published by Marcel Dekker, USA. He has published more than 180 articles in referred international journals, has more than 90 patents and patent applications, and has made more than 30 presentations at scientific conferences.
Understanding Enzymes
Understanding Enzymes
Function, Design, Engineering, and Analysis

edited by
Allan Svendsen
# Contents

*Introduction*  
xix

## Part I  Enzyme Function

1  A Short Practical Guide to the Quantitative Analysis of Engineered Enzymes  
*Christopher D. Bayer and Florian Hollfelder*

1.1 Introduction  
1.2 Quantifying Reaction Progress  
1.3 Typical Saturation Plots Give Michaelis–Menten Parameters  
1.4 What Can Go Wrong?  
1.5 Dealing with Multiphasic and Pre-Steady-State Kinetics  
1.6 Evaluating Enzymes  

2  Protein Conformational Motions: Enzyme Catalysis  
*Xinyi Huang, C. Tony Liu, and Stephen J. Benkovic*

2.1 Introduction  
2.2 Multidimensional Protein Landscape and the Timescales of Motions  
2.3 Conformational Changes in Enzyme–Substrate Interactions  
2.4 Conformational Changes in Catalysis  
2.4.1 Protein Dynamics of DHFR in the Catalytic Cycle  
2.4.2 Temporally Overlap: Correlation Does Not Mean Causation  
2.4.3 Fast Timescale Conformational Fluctuations  


2.4.4 Effect of Conformational Changes on the Electrostatic Environment 36
2.5 Conservation of Protein Motions in Evolution 38
2.6 Designing Protein Dynamics 39
2.7 Concluding Remarks 40

3 Enzymology Meets Nanotechnology: Single-Molecule Methods for Observing Enzyme Kinetics in Real Time 47
Kerstin G. Blank, Anna A. Wasiel, and Alan E. Rowan
3.1 Introduction 48
3.2 Single-Turnover Detection 53
3.2.1 Fluorescent Reporter Systems 53
3.2.2 Measurement Setup 56
3.2.3 Data Analysis 57
3.3 Single-Enzyme Kinetics 60
3.3.1 Candida antarctica Lipase B 63
3.3.2 Thermomyces lanuginosus Lipase 67
3.3.3 α-Chymotrypsin 73
3.3.4 Nitrite Reductase 78
3.3.5 Summary 84
3.4 New Developments Facilitated by Nanotechnology 88
3.4.1 Nano-optical Approaches 89
3.4.2 Nano-electronic Approaches 96
3.4.3 Nanomechanical Approaches 103
3.4.4 Summary 108
3.5 Conclusion 110

4 Interfacial Enzyme Function Visualized Using Neutron, X-Ray, and Light-Scattering Methods 125
Hanna Wacklin and Tommy Nylander
4.1 Phospholipase A2: An Interfacially Activated Enzyme 126
4.1.1 Neutron Reflection 129
4.1.2 Ellipsometry 130
4.1.3 Activity of Naja mossambica mossambica PLA2 130
4.1.4 Fate of the Reaction Products 133
4.1.5 The Lag Phase and Activation of Pancreatic PLA2 135
4.1.6 Distribution of Products during the Lag Phase 138
4.1.7 Hydrolysis of DPPC by Pancreatic PLA₂ 139
4.1.8 Role of the Reaction Products in PLA₂ Activation 141
4.1.9 Effect of pH and Activation by Me-β-cyclodextrin 144
4.2 Other Lipolytic Enzyme Reactions on Surfaces 150
4.2.1 Triacylglycerol Lipases and the Role of Lipid Liquid Crystalline Nanostructures 150
4.3 Cellulase Enzymes 154
4.4 Conclusion 158

5 Folding Dynamics and Structural Basis of the Enzyme Mechanism of Ubiquitin C-Terminal Hydroylases 167

5.1 Introduction 169
5.1.1 UCH-L1 171
5.1.1.1 Genetic association between UCH-L1 and neurodegenerative diseases 171
5.1.1.2 UCH-L1 in oncogenesis 175
5.1.2 Molecular Insights into the Pathogenesis Associated with UCH-L1 175
5.1.3 UCHL3 177
5.1.4 UCHL5 178
5.1.5 BAP1 179
5.2 UCH Structures 180
5.3 Folding Dynamics and Kinetics 183
5.4 Substrate Recognition 184
5.5 Enzyme Mechanism 186
5.6 Conclusion 189

6 Stabilization of Enzymes by Metal Binding: Structures of Two Alkalophilic Bacillus Subtilases and Analysis of the Second Metal-Binding Site of the Subtilase Family 203

6.1 Introduction: Subtilases and Metal Binding 203
6.1.1 Calcium-Binding Sites in Bacillus: Proposal for a Standard Nomenclature 209
6.1.2 The Weak Metal-Binding Site 214

6.2 Two New Structures of Subtilases with Altered Calcium Sites 216
   6.2.1 Proteinase SubTY 216
      6.2.1.1 The overall fold 216
      6.2.1.2 The active site 216
      6.2.1.3 SubTY calcium and sodium sites 218
      6.2.1.4 SubTY disulfide bridge 219
   6.2.2 SubHal 220
      6.2.2.1 The unliganded form of SubHal 220
      6.2.2.2 The SubHal:CI2A complex 221
      6.2.2.3 Termini, surface, and pH stability of SubHal 221
      6.2.2.4 The two crystallographically independent SubHal:CI2A complexes 223
      6.2.2.5 The calcium sites in SubHal 224
      6.2.2.6 The active site of SubHal 226
   6.2.3 Enzymatic Activity of SubTY and SubHal 228
   6.2.4 Comparison of SubTY and SubHal with Other Subtilases 228
   6.2.5 The SubHal C-domain Compared to the Eukaryotic PCs, Furin and Kexin 232
      6.2.5.1 Active site comparison 233
      6.2.5.2 The specificity pockets 234
      6.2.5.3 Inhibitor CI2A binding 234
   6.2.6 Activity Profiles 236
   6.2.7 Comparison of Metal Binding at the Strong and Weak Sites in the S8 Family 236
   6.2.8 The Ca-II and Na-II Metal-Binding Sites 237

6.3 Conclusion: Implications for Structural Studies of Enzymes 248

6.4 Materials and Methods 249
   6.4.1 SubTY 249
      6.4.1.1 Protein production and purification 249
      6.4.1.2 Purification of the SubTY:CI2A (1:1) complex 250
      6.4.1.3 Crystallization 250
      6.4.1.4 Structure determination 251
6.4.2 SubHal

6.4.2.1 Protein production and purification 251
6.4.2.2 Purification of the SubHal:Cl2A (1:1) complex 252
6.4.2.3 Crystallization 252
6.4.2.4 Structure determination 253

6.4.3 Protease Assays 256
6.4.4 pH Stability 257
6.4.5 Data Deposition 257

7 Structure and Functional Roles of Surface Binding Sites in Amylolytic Enzymes 267
Darrell Cockburn and Birte Svensson

7.1 Introduction 267
7.2 Identification of SBSs: X-Ray Crystallography 271
7.3 Bioinformatics of SBS Enzymes 273
7.4 Binding Site Isolation 275
7.5 Protection of Binding Sites from Chemical Labeling 277
7.6 Nuclear Magnetic Resonance 277
7.7 Binding Assays 278
7.8 Activity Assays 282
7.9 Future Prospects 283
7.10 Conclusion 286

8 Interfacial Enzymes and Their Interactions with Surfaces: Molecular Simulation Studies 297
Nathalie Willems, Mickaël Lelimousin, Heidi Koldsø, and Mark S. P. Sansom

8.1 Introduction 297
8.2 Enzyme Interactions at Interfaces 299
8.3 Molecular Dynamic Simulations of Biomolecular Systems 301
8.4 Lipases 303
8.4.1 Atomistic MD Studies of Lipase Interactions with Interfaces 304
8.4.2 The Role of Water in Lipase Catalysis at Interfaces 307
8.5 Coarse-Grained MD Studies of Interfacial Enzymes: Orientation and Interactions
   8.5.1 Phospholipase A2 309
   8.5.2 PTEN 310
8.6 Conclusions 311

PART II ENZYME DESIGN

9 Sequence, Structure, Function: What We Learn from Analyzing Protein Families 321
   Michael Widmann and Jürgen Pleiss
   9.1 Introduction 321
   9.2 Detection of Inconsistencies Utilizing a Standard Numbering Scheme 323
   9.3 Identification of Functionally Relevant Positions 327
   9.4 The Modular Structure of Thiamine Diphosphate–Dependent Decarboxylases 330
   9.5 Stereoselectivity-Determining Positions: The S-Pocket Concept in Thiamine Diphosphate–Dependent Decarboxylases 333
   9.6 Regioselectivity-Determining Positions: Design of Smart Cytochrome P450 Monoxygenase Libraries 336
   9.8 Conclusion 341

10 Bioinformatic Analysis of Protein Families to Select Function-Related Variable Positions 351
   Dmitry Suplatov, Evgeny Kirilin, and Vytas Švedas
   10.1 Introduction 352
   10.2 Bioinformatic Analysis of Evolutionary Information to Identify Function-Related Variable Positions 359
      10.2.1 Problem Definition 359
      10.2.2 Scoring Schemes in the Variable Position Selection: High-Entropy, Subfamily-Specific, and Co-Evolving Positions 361
      10.2.3 Association of the Variable Positions with Functional Subfamilies 366
Contents

10.2.4 How to Select Functionally Important Positions as Hotspots for Further Evaluation: Implementation of Statistical Analysis 366

10.3 The Bioinformatic Analysis of Diverse Protein Superfamilies 369
  10.3.1 Bioinformatic Challenges at Studying Enzymes 369
  10.3.2 Zebra: A New Algorithm to Select Functionally Important Subfamily-Specific Positions from Sequence and Structural Data 370

10.4 Subfamily-Specific Positions as a Tool for Enzyme Engineering 375

10.5 Conclusion 377

11 Decoding Life Secrets in Sequences by Chemicals 387
  Zizhang Zhang

  11.1 Introduction 388
  11.2 Linking an Enzyme’s Activity to Its Sequence 389
  11.3 Refining the Sequence Space to a Specific Function by Directed Evolution 395
  11.4 Linking Chemistry to -Oomics with High-Throughput Screening Methods 398
  11.5 Finding Large Sequence Space of a Specific Function from Microbial Diversity 400
  11.6 Linking Sequences to Substromes at the Molecular Level 404
    11.6.1 Biocatalytic Study of EHs 405
    11.6.2 Pharmacological Study of EHs 407
    11.6.3 Mechanistic Study of EHs 407
    11.6.4 What We Have Learned from the Studies of EH 410
    11.6.5 Technologies with Potentials in Genochemistry Approach 410
  11.7 Correlating with Computational Methods 410
  11.8 Problems That Genochemistry Can Potentially Tackle 413
  11.9 Conclusion 414
12 Role of Tunnels and Gates in Enzymatic Catalysis 421
Sérgio M. Marques, Jan Brezovsky, and Jiri Damborsky

12.1 Introduction 421
12.2 Protein Tunnels 423
  12.2.1 Structural Basis and Function 423
  12.2.2 Identification Methods 427
  12.2.3 Molecular Engineering 429
12.3 Protein Gates 431
  12.3.1 Structural Basis and Function 431
  12.3.2 Identification Methods 437
  12.3.3 Molecular Engineering 440
12.4 Conclusions 442

13 Molecular Descriptors for the Structural Analysis of Enzyme Active Sites 465
Valerio Ferrario, Lydia Siragusa, Cynthia Ebert, Gabriele Cruciani, and Lucia Gardossia

13.1 Introduction: Molecular Descriptors for Investigation of Enzyme Catalysis 465
13.2 Molecular Descriptors Based on Molecular Interaction Fields 467
13.3 Multivariate Statistical Analysis for Processing and Interpretation of Molecular Descriptors 472
13.4 Grind Descriptors for the Study of Substrate Specificity 475
13.5 VolSurf Descriptors for the Modeling of Substrate Specificity 477
13.6 Differential MIF$_3$ Descriptors for the Study of Enantioselectivity 479
13.7 Hybrid MIF$_3$ Descriptors for the Computation of Entropic Contribution to Enantiodiscrimination 481
13.8 Analysis of Enzyme Active Sites for Rational Enzyme Engineering 484
13.9 BioGPS Descriptors for in silico Rational Design and Screening of Enzymes 489
13.10 Conclusions 495
14 Hydration Effects on Enzyme Properties in Nonaqueous Media Analyzed by MD Simulations  
*Diana Louisa, António M. Baptista, and Cláudio M. Soares*

14.1 Enzyme Reactions in Nonaqueous Solvents 502
14.2 Classes of Nonaqueous Solvents 503
14.3 The Role of Water in Nonaqueous Biocatalysis 504
14.4 Effect of Water Content on Enzyme Structure and Dynamics 504
14.5 Effect of Water Content on Enzyme Selectivity 507
14.6 Hydration Mechanisms of Enzymes in Polar and Nonpolar Solvents 508
14.7 Enzyme Behavior as a Function of Water Activity 510
14.8 Hydration Effects on Enzyme Reactions in Ionic Liquids 512
14.9 Hydration Effects on Enzyme Reactions in Supercritical Fluids 514
14.10 Conclusions 516

15 Understanding Esterase and Amidase Reaction Specificities by Molecular Modeling  
*Per-Olof Syrén*

15.1 Introduction 523
15.2 Fundamental Catalytic Concepts 525
  15.2.1 Fundamental Chemistry of Amides and Esters 525
  15.2.2 Esterases and Amidases and Their Metabolic Significance 525
  15.2.3 Fundamental Chemical Aspects of Amidase and Esterase Catalysis 526
  15.2.4 Impact of Stereoelectronic Effects on the Enzymatic Reaction Mechanism 529
15.3 Molecular Modeling of Fundamental Catalytic Concepts 529
  15.3.1 QM Calculations on Amidases and Esterases 529
  15.3.2 MD Simulations on Amidases and Esterases 535
  15.3.3 QM/MM Simulations on Amidases and Esterases 539
15.4 Outlook and Implications for Enzyme Design 544
15.5 Additional Comments 546

PART III ENZYME DIVERSITY

16 Toward New Nonnatural TIM-Barrel Enzymes Using Computational Design and Directed Evolution Approaches 561
   Mirja Krause and Rik K. Wierenga
   16.1 Introduction 562
   16.2 General Aspects of Protein Engineering 566
      16.2.1 Library Creation Methods 569
      16.2.2 Structure-Based Library Design 572
      16.2.3 Optimal Libraries for Directed Evolution Methods 574
      16.2.4 Data-Driven Design (Semirational Design) 578
      16.2.5 Protein Engineering by Selection and Screening Methods 579
   16.3 Directed Evolution Studies with TIM-Barrel Enzymes 584
      16.3.1 Protein Engineering Studies of TIM-Barrel Proteins 586
      16.3.2 The Kemp Eliminases 590
   16.4 Concluding Remarks 596

17 Handling the Numbers Problem in Directed Evolution 613
   Carlos G. Acevedo-Rocha and Manfred T. Reetz
   17.1 Introduction 614
   17.2 Saturation Mutagenesis in Directed Evolution 617
   17.3 Statistical Analyses 620
      17.3.1 Conventional Statistics Based on the Patrick and Firth Algorithm 620
      17.3.2 Statistics Based on the Nov Algorithm 624
   17.4 How to Group and Randomize Amino Acid Positions 626
   17.5 Fitness Landscapes 628
      17.5.1 Fujiyama vs. Badlands Fitness Landscapes 628
      17.5.2 Fitness-Pathway Landscapes and How to Escape from Local Minima 630
   17.6 Conclusions and Perspectives 636
18 Hints from Nature: Metagenomics in Enzyme Engineering 643
   Esther Gabor, Birgit Heinze, and Jürgen Eck
   18.1 Metagenomics and the Ideal Enzyme 644
   18.2 Molecular Microdiversity 647
   18.3 Metagenomic Enzyme Chimera 650
   18.4 Outlook 653

19 A Functional and Structural Assessment of Circularly
   Permutated Bacillus circulans Xylanase and Candida
   antarctica Lipase B 657
   Stephan Reitinger and Ying Yu
   19.1 Introduction 657
   19.2 Naturally Occurring Circular Permutations:
      Selected Examples 658
   19.3 Circular Permutation of Bacillus circulans
      Xylanase 661
   19.4 Circular Permutation on Candida antarctica
      Lipase B 669
   19.5 Conclusion 674

20 Ancestral Reconstruction of Enzymes 683
   Satoshi Akanuma and Akihiko Yamagishi
   20.1 Introduction 683
   20.2 Reconstruction of an Ancestral Protein Sequence
      20.2.1 Overview 684
      20.2.2 Methods for Ancestral Sequence
         Reconstruction 684
      20.2.3 Early Works 686
   20.3 The Commonote 687
      20.3.1 The Last Universal Common Ancestor, the
         Commonote 687
      20.3.2 Theoretical Studies on the Environmental
         Temperature of the Commonote 688
      20.3.3 Reconstruction of an Ancestral Nucleoside
         Diphosphate Kinase 689
      20.3.4 Estimation of the Environmental
         Temperature of the Commonote 692
   20.4 Application to Designing Thermally Stable Proteins 693
20.4.1 Design of Thermally Stable Proteins 693
20.4.2 Case Studies to Create Thermally Stable Enzymes by Introducing Ancestral Residues as Amino Acid Substitutions 694
20.4.3 Reconstruction of Thermally Stable, Ancestral DNA Gyrase Using a Small Set of Homologous Amino Acid Sequences 696

20.5 Conclusion 697

PART IV ENZYME SCREENING AND ANALYSIS

21 High-Throughput Screening or Selection Methods for Evolutionary Enzyme Engineering 707
Shuobo Shi, Hongfang Zhang, Ee Lui Ang, and Huimin Zhao
21.1 Introduction 708
21.2 Selection 710
21.2.1 Solid-Medium-Based Selection 717
21.2.2 Liquid-Medium-Based Selection 719
21.2.3 Display-Based Selection 722
21.3 Screening 724
21.3.1 Chromatography- and Mass-Spectrometry-Based Screening 725
21.3.2 Solid-Medium-Based Screening 726
21.3.3 Microtiter-Plate-Based Screening 727
21.3.4 Yeast Two-/Three-Hybrid System 729
21.3.5 FACS-Based Screening 729
21.3.6 Microfluidics-Based Screening 732
21.4 Conclusions and Prospects 734

22 Nanoscale Enzyme Screening Technologies 745
Helen Webb-Thomasen and Andreas H. Kunding
22.1 Introduction 745
22.2 Approaches to Nanocompartmentalization of Enzymes 746
22.2.1 Liposomes 747
22.2.1.1 Addressability 747
22.2.1.2 Reagent exchange 749
Contents

22.2.2 Polymersomes and VirusLike Particles 751
22.2.3 Water-in-Oil Emulsion Droplets 752
   22.2.3.1 Addressability 755
   22.2.3.2 Reagent exchange 755
22.3 Microfabricated Chip Devices for Enzyme Compartmentalization and Screening 756
   22.3.1 Microfluidic-Generated Emulsion Droplets 757
   22.3.2 Microfabricated Arrays 762
      22.3.2.1 Optical fiber microarrays 762
      22.3.2.2 Elastomeric microarrays 763
      22.3.2.3 Surface tension microarrays 765
22.4 Conclusion and Current Challenges 767
22.5 Future Improvements 769

23 Computational Enzyme Engineering: Activity Screening Using Quantum Chemistry 777
   Martin R. Hediger
   23.1 Motivation 778
   23.2 Introduction 779
   23.3 Methods 780
      23.3.1 Calculation Engines 780
      23.3.2 Molecular Modeling 782
      23.3.3 Software 786
   23.4 Applications 786
      23.4.1 Overview 786
      23.4.2 Engineering Candida antarctica Lipase B 787
      23.4.3 Engineering Bacillus circulans Xylanase 793
   23.5 Conclusions 800

24 In Silico Screening of Enzyme Variants by Molecular Dynamics Simulation 805
   Hein J. Wijma
   24.1 Potential Applications of MD Simulations For Improving Enzymes 805
   24.2 Molecular Dynamics vs. Other in silico Methods 809
   24.3 Improving Catalytic Activity by MD Screening 812
      24.3.1 Transition-State Simulation 812
      24.3.2 High-Energy Intermediate Simulation 814


24.3.3 Substrate Simulation with Near-Attack Conformations 815
24.3.4 Substrate Simulation with Monitoring of H Bonds 817
24.4 Predicting and Improving Binding Affinity 818
24.5 MD Screening to Improve Enzyme Stability 819
24.6 Improving Correlation between MD and Experiment 822
24.6.1 Force Field Inaccuracies 822
24.6.2 Sampling Concerns 823
24.6.3 Other Concerns 824
24.7 Outlook and Further Possibilities 825

25 Kinetic Stability of Variant Enzymes 835
Jose M. Sanchez-Ruiz
25.1 Kinetics vs. Thermodynamics in Protein Stability 835
25.2 Mutation Effects on Kinetic Stability: A Description Based on the Transition State for Irreversible Denaturation 838
25.3 Kinetic Stability Linked to the Breakup of Interactions in the Transition State: Pro-dependent Proteases 841
25.4 Kinetic Stability Linked to Substantially Unfolded Transition States: Thioredoxin and Phytase Enzymes 842
25.5 Role of Solvation Barriers in Kinetic Stability: Lipases and Triose Phosphate Isomerases 848
25.6 Concluding Remarks 852

Index 859
Introduction

More than three decades ago, the hope emerged that protein engineering would be able to predict protein and enzyme function on the basis of X-ray crystal structures. The expectations were that we should be able to create goal-oriented functions in the enzyme of interest. A large effort was made to obtain the structures of enzymes of great importance for understanding biological processes and enzymes of general commercial interest in many industries. A large variety of structures of enzymes from many biological pathways, as well as enzymes of commercial interest, have been solved, including carbohydrate-acting enzymes, proteolytic enzymes, and lipolytic enzymes, and have helped tremendously in understanding the structure–function relationships. They have also revealed how much we still need to learn in order to manipulate genes to make enzymes react in a desired way.

Today, there are at least two major focuses on gaining benefit from and knowledge about enzyme function: (1) data analysis and (2) a more detailed understanding. Much learning cannot be said to be statistically feasible, but I hope the scientific society will still accept a few examples as feasible hypotheses to investigate further. With the increasing knowledge on enzyme function, with input from atomistic mobility and hydrogen bonding, the shifting electrostatics situation due to mobility and changes in relative coordinated atoms and macroscopic dependencies on enzyme environment changes leaves us with a very complex multidimensional space for how enzymes work. This makes it nearly experimentally unfeasible to have enough statistics on all the possible impact characteristics, as theoretically needed, making it difficult to draw sound, comprehensive, and significant conclusions. Commonly, even very large data sets will reveal single conclusions but are incorrectly
drawn since the number of data sets for each parameter alone is too few to make findings statistically significant. The data analysis will definitely add to a more detailed understanding and to suggestions for function. Some chapters touch upon data-driven discovery, but most of the chapters are focused on hypothesis-driven research testing one specific enzyme in a specific environment and with few parameters, giving exciting insights into the complexity of enzyme nature.

During my work in developing enzymes for technical use and work on the enzyme–substrate interaction, it has been tempting to combine the information from quantum mechanical calculations of the energetics in the catalytic reaction, and the overall molecular mobility using standard force fields, as well as electrostatics calculations and docking in order to inform on three important topics of enzyme function, namely (1) the initial substrate binding to the enzyme, (2) the important local fitting to accommodate the correct spatial state that can contain the reactive state as seen by molecular dynamics mobility and hydrogen bonding patterns, and (3) the reactive state energetics as measured by quantum mechanical calculations. This overall reaction could be stated in a formula as shown below:

\[ \text{Enzyme function} = f(\text{overall binding}) + f(\text{local fluctuations and interactions}) + f(\text{reactive energy}) \]

Or in other words, enzyme function is a function of three major key factors: (1) the overall fitting of the substrate for binding with the correct orientation for the more detailed local interactions in the nearer active site surroundings, (2) the necessary hydrogen bonding and electrostatic interactions to secure the correct arrangements for the catalysis reaction to take place, and (3) the quantum mechanical energy in the catalysis reaction. Seen from molecular dynamics simulations some hydrogen bonds are only present at a certain time during the simulation, indicating that activity only occurs when the structure is in a certain subdomain structure containing the important hydrogen bonds. If certain hydrogen bonds are in place at the same time the reaction can occur. If one of the three stated factors is not fulfilled at the same time, then no reaction occurs. Examples of important hydrogen bonds are presented in Chapter 15.
Introduction

In Chapter 10 on sequences and design the combination of sequence alignment information, docking, and molecular simulation of variant molecules to extract more combinatorial information is discussed.

This book focuses on the current understanding obtained in the past 10–15 years to the present. In the 1980s focus was on making 3D structures and understanding and analyzing proteins. In the 1990s focus was on diversity methods and screening methods, whereas in the 2000s the focus has been on bioinformatics and simulation methods and statistical methods, as well as ultrahigh-throughput methods with revised views on proteins. Today we hope the analysis of large data will help find the desired results. Many new technologies have brought new insights into enzyme function, with emphasis on single-molecule behavior and molecular mobility and electrostatics, as well as enzymes working on large substrates and complex substrates. Focus on the mobility impact on substrate interaction can be found in Chapter 2.

The book is divided into four major sections: enzyme function (Chapters 1–8), enzyme design (Chapters 9–15), enzyme diversity (Chapters 16–20), and enzyme screening and analysis (Chapters 21–25). The enzyme function part addresses the enzyme kinetics on simple substrates in Chapter 1, as well as the more complex interaction on larger substrates in Chapters 4, 7, and 8. Also structural aspects are addressed in Chapter 6, NMR structures in Chapter 5, and further dynamic aspects in Chapters 2 and 3. The enzyme design part is focused on the sequence-derived design methods in Chapters 9, 10, and 11, as well as in Chapter 20, and 3D structural methods. The 3D structural design/understanding is mainly discussed in Chapters 12–15. The design area is also covered partly under enzyme diversity, especially in Chapter 16, which has a review of both diversity methods and some design ideas. Further under enzyme diversity are handled metagenomics, circular permutations, and ancestral reconstruction in Chapters 18, 19, and 20, respectively, as well as the number issues in directed evolution in Chapter 17. The enzyme screening and analysis part includes both in silico screening in Chapters 23 and 24 and wet chemistry screening methods in Chapters 21 and 22, as well as an example of analysis of enzyme variants in Chapter 25.
The computer simulations reveal great insight into the function of enzymes and can help in designing new functionalities and activities. The predictive power is still not precise, but we can use the simulations to screen for potential variants of interest, which then need testing for the desired function. Decades ago, one specific predicted variant was selected for testing—today it is commonly understood that a certain number of the, say, top 10 or 100 candidates could potentially be of interest. The speed of computers today allows for this kind of suggestions and sometimes also a reasonable simplification is used for making the screening possible. Chapters 23 and 24 address these possibilities. Also Chapter 16 touches upon the in silico design possibilities.

It is now more than a decade ago that enzyme promiscuity became a major field of interest. The versatility of enzymes and their activities are more open today than ever and the general EC classification system is seldom fully explanatory today. A few chapters touch upon the promiscuity—not from a specificity issue but rather a reaction mechanistic view; see Chapters 15 and 23.

Other screening methods in the wet chemistry part are being developed, and while screening has come out of the first decade in protein engineering, the limitations are getting more visible and the possibilities better utilized. A few chapters address the methodologies (Chapters 16, 17, 21, and 22)—micronanotechnology has gone into the screening area and possibilities for very high numbers have become a reality. Smart techniques to secure the picking of hits are important and an interesting method is mentioned in Chapter 22.

In an earlier book I edited, Enzyme Engineering: Function, Design, Variant Generation and Screening, the focus was more on the variant generation and screening part and less on the function and design part. In this book the main focus is on enzyme function and design and less on variant generation and screening methods. This reflects the fact that many new insights into the more complex enzyme function have emerged during the past many years. Massive quantities of information on variants of enzymes and the multiple states of the structures as well as single-molecule insight have added to the colligative understanding of enzyme function.

The production of many mutations has, besides a lot of data, also resulted in the realization of how little we still understand about
enzyme function. Therefore, this has been emphasized in the first eight chapters with examples from the versatility of factors influencing enzyme activity and enzyme–substrate interaction. Around 20 years ago the main enzyme understanding was based on simple kinetics and soluble substrate interactions. In industry, we are aware that the main enzyme function often occurs under conditions other than the simple substrate–enzyme interaction theory, very well described with mathematical equations. Chapter 3 (on single-enzyme function) and Chapter 2 (on enzyme motions) emphasize the rather complicated behavior of the enzymatic function, which continues to open new depths of understanding. Examples of these complicated behaviors are presented in Chapter 4 on surface-active enzymes and Chapter 7 on the carbohydrate-hydrolyzing enzyme family.

During the work on writing the book chapters representing important directions in enzyme research on enzyme function, design, engineering, and analysis, recent aspects have been published, including enzymes’ use of the energy coming from the catalyzed chemical reaction itself, which adds to the chapters on mobility of the enzymes. Also the importance of electrostatics and the impact on enzyme function has not been directly addressed in the chapters but is clearly a major part of some of the added chapters and has been established as an important factor in enzyme function and catalysis. Clearly, more combinations of these factors mentioned in the chapters and above are needed in the future to further understand the full functional space of enzymes and thus understand how to address improvements by protein engineering.