



Whither Enzymology in the Twenty First Century?

E. N. G. Marsh*

Departments of Chemistry and Biological Chemistry, University of Michigan, Ann Arbor, MI, USA

Keywords: enzyme design, computational modeling, enzyme mechanisms, radical SAM enzymes, microbiome

INTRODUCTION

This “grand challenge” article marks the creation of a new specialty section shared between *Frontiers in Chemistry* and *Frontiers in Molecular Biology* that is dedicated to enzymology and protein chemistry. My aim in this article is to provide a personal perspective on the challenges and opportunities for field of enzymology and the directions in which it could and should move in the coming years. I hope this will challenge the reader to think more deeply about this discipline and its broader impact on science. But before gazing into the future, it is worthwhile briefly reviewing how our field has evolved until now, as it is this perspective that frames our outlook.

Whither Enzymology? was the title of a symposium held in 1990 to honor the late Jeremy Knowles on the occasion of his Sixtieth birthday (Raines, 2008). Knowles was in the vanguard of a pioneering group of enzymologists who first employed the tools of physical organic chemistry to study enzyme reactions. In the 1970's his group was the first to determine the full set of elementary rate constants defining the free energy profile for an enzyme-catalyzed reaction—that of triose phosphate isomerase (Albery and Knowles, 1976). This information, put together with the crystal structure for this enzyme, allowed us to understand enzyme catalysis with the kind of precision previously only possible for simple organic molecules. Subsequently, his lab was among the first to use site-specific mutagenesis to make precisely defined changes to enzyme active sites, which combined with detailed measurements of their effect on catalysis, allowed the relationships between enzyme structure and mechanism to be elucidated.

Studies such as these blew away the “fluffy clouds” obscuring the view of earlier generations of enzymologists, revealing enzymes to be exquisitely evolved, highly complex and incredibly efficient organic catalysts, but nevertheless subject to, and understood through, the basic principles of chemistry—in Knowles' words: “not different, just better” (Knowles, 1991). This view has shaped enzymology for the past 25 years, and by extension our view of cellular biology. It directly underpins the burgeoning discipline of synthetic biology, a field that involves tinkering with enzymes on a scale unimaginable in Knowles' era.

The tools available to enzymologists have advanced immeasurably in the last 25 years, resulting in a paradigm shift in our approach to studying them. Not least, much of the drudgery (and perhaps also some of the fun!) of obtaining enzymes for study has been removed and the accessibility of enzymes has expanded enormously. The advent of genomic sequencing, combined with cheap, commercially available gene synthesis and protein affinity tags, allows one to select essentially any enzyme for study and produce large quantities of pure protein with minimal effort. What would once have taken many person years to accomplish can now be achieved in a few days. (Although some classes of proteins, e.g., integral membrane proteins still remain challenging to express.) Advances in structural biology mean that determining the 3-D structure of most enzymes, if not trivial, is now routine; indeed, thanks to structural proteomics initiatives, the structure is increasingly the first information to emerge for a new enzyme. Moreover, techniques for studying enzymes have become increasingly powerful: instrumentation more sensitive (for some techniques to the level of single molecules) and data analysis more sophisticated and rapid due to exponential increases in computing power.

OPEN ACCESS

Edited by:

T. Joseph Kappock,
Purdue University, USA

Reviewed by:

Erika Anne Taylor,
Wesleyan University, USA
Tobias Jürgen Erb,
ETH Zurich, Switzerland

*Correspondence:

E. N. G. Marsh
nmarsh@umich.edu

Specialty section:

This article was submitted to
Protein Chemistry and Enzymology,
a section of the journal
Frontiers in Chemistry

Received: 08 March 2016

Accepted: 08 April 2016

Published: 22 April 2016

Citation:

Marsh ENG (2016) Whither
Enzymology in the Twenty First
Century? *Front. Chem.* 4:20.
doi: 10.3389/fchem.2016.00020

With such powerful tools and so many enzymes to explore, it seems reasonable some 25+ years on from Knowles' apotheosis, once again to pose the question: whither enzymology? In attempting to answer this question, I propose the three "grand challenges" discussed below. The first two I have chosen because they fundamentally challenge our understanding of enzyme catalysis. The first challenge is to predict enzyme catalytic activity from structure and dynamic information; the second challenge to design enzymes from scratch that approach the activity of naturally evolved enzymes. The third challenge focuses on expanding the frontiers of enzymology by discovering new enzymes that catalyze new reactions.

CHALLENGE 1—PREDICTIVE ENZYMOLOGY

Much of our effort continues to focus on understanding what makes enzymes "better" catalysts, with advances in techniques and methodology allowing us to tackle more complex enzymes and peer more deeply into the fundamental principles by which enzymes achieve such enormous rate enhancements (Wolfenden and Snider, 2001). But could enzymes perhaps also be different? For many enzymes we now have very detailed descriptions of their three-dimensional structures, the chemical identities of each intermediate in the reaction, the elementary rate constants connecting each step in the mechanism, the pK_a s of active site residues etc. However, I would argue that we still don't fully understand how enzymes achieve such remarkable rate accelerations of the reactions they catalyze, often of 10^{18} -fold, or even more.

Certainly we can appreciate the importance of contributions made by active site acids and bases (Toney and Kirsch, 1989), electrostatic interactions and transition state binding energy (Wolfenden, 1969) to catalysis. But when we attempt to quantify these various contributions to the reaction profile of a specific enzyme we nearly always come up short of catalytic power—usually by many orders of magnitude. For example, if we were to strictly apply the transition state binding formalism to explain rate enhancements, for many enzymes we would have to postulate transient, non-covalent interactions of remarkable strength—far tighter than even the strongest known stable interactions such as biotin-streptavidin.

Clearly the reductionist approach can only take us so far. What is missing from our understanding is how the large, complex and dynamic structure of the protein contributes to catalysis. Current models focus on the role of low frequency vibrational modes that are postulated promote catalysis by driving the substrates through the transition state to products (Kamerlin and Warshel, 2010; Kohen, 2015). Evidence for such "promoting vibrations" remains rather indirect as much of the vibrational action is thought to lie in the "terahertz gap," a region of the electromagnetic spectrum that is extremely hard to access. Breakthroughs in experimental techniques may open up new opportunities to gain insight into catalysis. But it is worth emphasizing that to properly interpret such data for molecules as complex as enzymes it will still be necessary to understand

the details of the chemical physics underpinning these vibrations and that for that foundational studies on small molecules will be essential.

Equally important are advances in computation that will allow us to accurately simulate enzyme reactions. We have seen great progress in the area of computational simulation over the last few decades, driven by improved methodology and, of course, the exponential increase in computing power. We are now approaching the point where, at least for simple enzymes, reaction trajectories and free energy profiles can be calculated that agree reasonably well with experimentally obtained data (Senn and Thiel, 2009; Van Der Kamp and Mulholland, 2013). One problem is that currently many simulations aim to reproduce prior experimental measurements, and it is unclear how much this prior knowledge influences the choice of parameters used to model the enzyme reaction.

Looking forward, an important goal will be to advance computational simulations to the point where they have predictive power rather simply being descriptive. Ideally, this would mean that starting from the 3-dimensional structure of an enzyme-substrate complex, and possibly including experimental information on protein dynamics, one could simulate the enzyme reaction at a level that accurately and reliably predicts parameters that can be experimentally determined, such as elementary rate constants for the reaction, and ultimately commonly measured kinetic parameters such as k_{cat} and K_M . Among the challenges here are combining simulation techniques to deal with the many timescales that contribute to enzyme catalysis. Whereas chemical bonds are made and broken on the femtosecond timescale, i.e., within one bond vibration, protein conformational changes necessary for substrate binding and product release occur on the millisecond timescale. In between are events that are essential to catalysis such as coordinated protein motions and the formation of chemical intermediates that occur on a range of timescales from picoseconds to milliseconds. If we can fully recapitulate an enzyme reaction *in silico*, we might reasonably claim to have understood how an enzyme works.

CHALLENGE 2—DESIGNING ENZYMES FROM SCRATCH

In contrast to the over 100 years spent studying natural enzymes, the design of new enzymes is very much a Twenty First century endeavor and one that has been greatly empowered by advances in computation. Enzyme design is important for two reasons. First, the ability to design highly efficient, extremely specific catalysts that function at room temperature and pressure under benign conditions has enormous practical applications; for example in replacing industrial processes that currently require high temperatures and pressures and/or expensive metal catalysts, or in the synthesis of fine chemicals containing complex functionality and many stereocenters. Second, enzyme design provides a rigorous test of our understanding of the principles of enzyme catalysis. If we *really* understand how enzymes work, we should be able to build them from scratch, incorporating the various "components" e.g., a substrate binding pocket, a catalytic

general acid, stabilizing electrostatic interactions, etc., into an existing protein scaffold, or more ambitiously, into a protein specifically designed for the task.

Current *in silico* approaches to enzyme design attempt to achieve this by employing an “inside-out” approach, starting with a transition state for the desired reaction and then surrounding it with a constellation of active residues designed to bind the substrate and provide suitably placed catalytic residues (Mak and Siegel, 2014). Search algorithms are then used to identify a natural protein scaffold that can potentially accommodate the designed active site. The “theozyme” is then synthesized and usually subjected to optimization by several rounds of directed evolution to select enzymes with improved performance. Whether this last step really qualifies as “design” is left to the reader to decide!

So far, enzyme designers have, with good reason, tended to focus on rather simple reactions, for example the Kemp elimination and ester hydrolysis. The best designs exhibit some quite impressive $k_{\text{cat}}/k_{\text{uncat}}$ values of up to $\sim 10^9$ after optimization (Blomberg et al., 2013), with the usual reference state for comparing the uncatalyzed reaction is water at pH 7. However, the average designed enzyme is still around 11 orders of magnitude worse than the “average” natural enzyme before optimization by evolution, improving to 9 orders of magnitude worse after optimization (Mak and Siegel, 2014).

An optimist might argue that, on a log scale, we are halfway to matching natural enzymes where $k_{\text{cat}}/k_{\text{uncat}}$ values of 10^{18} are typical. However, although water is an obvious choice for benchmarking enzyme reactions, it has been suggested that a more appropriate comparison would be a solvent such as acetonitrile, to approximate the dielectric constant of a protein, containing acetate, an effective general base catalyst for the Kemp elimination (Korendovych and Degrado, 2014). This medium would represent a “non-designed” enzyme-like environment that design should improve upon. With this benchmark, the second order rate constant for the acetate-catalyzed Kemp elimination in acetonitrile can be compared with $k_{\text{cat}}/K_{\text{m}}$ for the designed enzyme. It would seem that we ought to be smarter than acetonitrile and acetate – but so far *no* designed Kemp eliminase has beaten this benchmark (!), although subsequent optimization by directed evolution has allowed designed enzymes to surpass this benchmark.

Humbling though the current performance of designed enzymes is, design is an activity where one learns as much, if not more, from carefully analyzing a failure as one does from success. Analysis of k_{cat} and K_{M} for designed enzymes suggests, unsurprisingly, that the challenge lies in catalysis, rather than binding, as K_{M} values are similar to natural enzymes whereas k_{cat} values are much, much smaller. However, this is a field that is still in relative infancy and there are good reasons to be optimistic that designed enzymes will match natural enzymes in the coming decades. For example, current computational design strategies use fairly basic models of transition state theory and don't include features such as protein dynamics and ground state destabilization that we know are important for catalysis. It is worth underscoring that the success of this grand challenge is intimately linked to success of the first challenge. This is especially true for the many multistep reactions that involve

cofactors: if we don't fully understand the mechanism of an enzyme reaction, we are unlikely to succeed in designing such enzymes.

CHALLENGE 3—PROSPECTING FOR NEW ENZYMES

In the process of understanding the fundamental biochemistry underpinning cellular metabolism and physiology we have discovered a remarkably diverse array of enzyme-catalyzed reactions. However, it is likely that we have only scratched the surface, and that there are many more fascinating enzymes waiting to be discovered. Much enzymological research has focused on human/mammalian metabolism, unsurprisingly given the importance of understanding enzymes, and the mechanisms by which they can be inhibited, for the design of effective drugs. But even in this intensely studied area, the functions of 20–25% of human protein-encoding genes remain completely unknown, and a good fraction of these are likely to encode new enzymes.

Compared with plants and microbes, animals are rather limited range of molecules they synthesize, relying on diet to obtain half the amino acids needed to build proteins and for the precursors to many essential enzyme cofactors (as vitamins). Therefore, the opportunities to discover new and useful enzymes are vastly greater among plants and microbes. These organisms synthesize an astonishing array of molecules, ranging from very simple organic molecules to natural products of remarkable complexity and a wide variety of polymeric materials. Microbes also colonize some of the least hospitable environments—from naturally occurring sulfurous hot springs to highly polluted sites where they can utilize some of the most persistent organic pollutants as carbon sources.

Where life exists under extreme conditions, there are certainly likely to be interesting and unusual enzymes waiting to be discovered, which, furthermore, may have important uses. For examples, consider viral reverse transcriptase (Gallo, 1971), bacterial restriction enzymes (Meselson and Yuan, 1968) and archaeal thermostable polymerases (Chien et al., 1976), which were all discovered long before their importance in revolutionizing molecular biology and medical research would be appreciated. CO_2 fixation represents another highly timely example, as this step is generally considered to be the bottleneck in carbon assimilation by plants. Although the Calvin cycle is the best known and studied, six alternative microbial pathways for the fixation of CO_2 have been identified so far, four of them in the last decade (Erb, 2011). The search for novel microbial pathways has resulted in the isolation of novel enzymes such as crotonyl-CoA carboxylase/reductase (Erb et al., 2009), the most efficient CO_2 -fixing enzyme described so far.

The importance of microbes, literally, much closer to our hearts has also recently been appreciated. The human microbiota represents a symbiotic collection of organisms that have recently been estimated to be present in similar numbers to our own cells (Sender et al., 2016), most of which reside in our gut. There is increasing evidence that these organisms play an

important role in human health; they may, for example, process dietary molecules to increase their nutritional value, or render them less toxic, through metabolic pathways that we lack. For example there is recent evidence that changes in gut microbiota caused by the consumption of artificial sweeteners may induce glucose intolerance (Suez et al., 2014). The enzymology of the microbiome remains largely unexplored, but some of these undiscovered enzymes may prove vital for human health.

Finally, as an illustrative example of enzyme discovery, consider the radical SAM enzyme family. These enzymes catalyze a remarkably wide range of chemical transformations that are initiated by the one-electron reduction of S-adenosylmethionine by an iron-sulfur cluster. This generates a reactive adenosyl radical (and methionine) that subsequently abstracts a hydrogen atom from the substrate to initiate the catalytic cycle (Wang et al., 2014). Originally identified in studies on lysine fermentation in Clostridial bacteria (Chirpich et al., 1970), these enzymes, which are very oxygen sensitive, were for a long time thought to be isolated curiosities that were restricted to a few anaerobic fermentation pathways. However, it is now known that radical SAM enzymes play vital roles in the biosynthesis of many vitamins, in the modification of tRNA and rRNA, in post-translational modifications of proteins and in the biosynthesis of various secondary metabolites. Although almost all of the enzymes studied so far are microbial, it is very likely, despite their oxygen sensitivity, that radical SAM enzymes are also present in animals, including humans.

Our exploration of radical SAM enzymes was greatly aided by the identification of a characteristic sequence motif associated with the SAM-coordinating iron-sulfur cluster, which is unique to this family of enzymes. To date there are ~ 50,000 putative radical SAM sequences cataloged. A recent bioinformatics study was able to classify these into 50 different subgroups associated with a known or putative enzyme function. Yet this analysis also identified subgroups for which no known function exists and other putative radical SAM sequences that were unique. Furthermore, it is quite likely that even within the known functional subgroups that there are enzymes catalyzing unknown reactions. This suggests that we have much more to learn about the reactions catalyzed by radical SAM enzymes, and the same is undoubtedly true for many other enzyme families.

REFERENCES

- Albery, W. J., and Knowles, J. R. (1976). Evolution of enzyme function and development of catalytic activity. *Biochemistry* 15, 5631–5640. doi: 10.1021/bi00670a032
- Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C., and Sarma, V. R. (1965). Structure of hen egg-white lysozyme – a 3-dimensional fourier synthesis at 2 Å resolution. *Nature* 206, 757–761. doi: 10.1038/206757a0
- Blomberg, R., Kries, H., Pinkas, D. M., Mittl, P. R. E., Gruetter, M. G., Privett, H. K., et al. (2013). Precision is essential for efficient catalysis in an evolved Kemp eliminase. *Nature* 503, 418–421. doi: 10.1038/nature12623
- Chien, A., Edgar, D. B., and Trela, J. M. (1976). Deoxyribonucleic acid polymerase from extreme thermophile *Thermus aquaticus*. *J. Bacteriol.* 127, 1550–1557.

CONCLUDING REMARKS—THE GREATER CHALLENGE

In conclusion, enzymology will not wither any time soon! We still have a lot to learn about enzymes as catalysts. But, going forward, what we do learn will be profoundly useful to society.

In closing, however, I want to mention one overarching challenge. For any discipline to thrive going forward it needs to recruit a diverse group of bright new minds. Cultivating the next generation of enzymologists requires engaging students at the undergraduate level, or even earlier. In contrast to the advances in the lab, enzymology in the classroom seems to have stood still. Comparing the chapters on proteins and enzymes in the biochemistry textbook I used as an undergraduate 30 years ago and the book I teach from today, the same three examples: hemoglobin, lysozyme and chymotrypsin, still dominate the pages. Of course, these were among the very first proteins to have their structures determined (Blake et al., 1965; Perutz, 1965; Matthews et al., 1967) and at one time these examples really did represent the cutting edge of our understanding of enzymes. However, other areas of biochemistry, for example transcription, translation and the role of RNA biology, appear to have done a much better job in remaining current and its unclear why enzymology seems to have been left behind. Updating the way we teach enzymology to reflect the discipline in the Twenty First century and, more broadly, recruiting people into science from a wider, more diverse cross-section of society is a challenge equal in importance to any of the scientific challenges discussed above.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and approved it for publication.

ACKNOWLEDGMENTS

Research in enzymology in the author's laboratory is funded by grants from the National Institutes of Health (GM 93088), the National Science Foundation (CBET 1336636) and the Defense Threat Reduction Agency (HDIRTA 1-11-1-0019).

- Chirpich, T. P., Zappia, V., Costilow, R. N., and Barker, H. A. (1970). Lysine 2,3-aminomutase – purification and properties of a pyridoxal phosphate and S-adenosylmethionine activated enzyme. *J. Biol. Chem.* 245, 1778–1789.
- Erb, T. J. (2011). Carboxylases in natural and synthetic microbial pathways. *App. Environ. Microbiol.* 77, 8466–8477. doi: 10.1128/AEM.05702-11
- Erb, T. J., Brecht, V., Fuchs, G., Muller, M., and Alber, B. E. (2009). Carboxylation mechanism and stereochemistry of crotonyl-CoA carboxylase/reductase, a carboxylating enoyl-thioester reductase. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8871–8876. doi: 10.1073/pnas.0903939106
- Gallo, R. C. (1971). Reverse transcriptase, DNA polymerase of oncogenic RNA viruses. *Nature* 234, 194–198. doi: 10.1038/234194a0
- Kamerlin, S. C. L., and Warshel, A. (2010). At the dawn of the 21st century: Is dynamics the missing link for understanding enzyme catalysis? *Proteins Struct. Funct. Bioinform.* 78, 1339–1375. doi: 10.1002/prot.22654

- Knowles, J. R. (1991). Enzyme catalysis – not different, just better. *Nature* 350, 121–124. doi: 10.1038/350121a0
- Kohen, A. (2015). Role of dynamics in enzyme catalysis: substantial versus semantic controversies. *Acc. Chem. Res.* 48, 466–473. doi: 10.1021/ar500322s
- Korendovych, I. V., and Degrado, W. F. (2014). Catalytic efficiency of designed catalytic proteins. *Curr. Opin. Struct. Biol.* 27, 113–121. doi: 10.1016/j.sbi.2014.06.006
- Mak, W. S., and Siegel, J. B. (2014). Computational enzyme design: transitioning from catalytic proteins to enzymes. *Curr. Opin. Struct. Biol.* 27, 87–94. doi: 10.1016/j.sbi.2014.05.010
- Matthews, B. W., Sigler, P. B., Henderso,R., and Blow, D. M. (1967). 3-Dimensional structure of tosyl-alpha-chymotrypsin. *Nature* 214, 652–656. doi: 10.1038/214652a0
- Meselson, M., and Yuan, R. (1968). DNA restriction enzyme from *E. coli*. *Nature* 217, 1110–1112. doi: 10.1038/2171110a0
- Perutz, M. F. (1965). Structure and function of haemoglobin. A tentative atomic model of hoerse oxyhaemoglobin. *J. Mol. Biol.* 13, 646. doi: 10.1016/S0022-2836(65)80133-4
- Raines, R. T. (2008). Jeremy R. Knowles (1935–2006). *ACS Chem. Biol.* 3, 262–264. doi: 10.1021/cb800099n
- Sender, R., Fuchs, S., and Milo, R. (2016). Revised estimates for the number of human and bacteria cells in the body. *bioRxiv*. doi: 10.1101/036103
- Senn, H. M., and Thiel, W. (2009). QM/MM Methods for biomolecular systems. *Angew. Chem. Intl. Edn.* 48, 1198–1229. doi: 10.1002/anie.200802019.
- Suez, J., Korem, T., Zeevi, D., Zilberman-Schapira, G., Thaiss, C. A., Maza, O., et al. (2014). Artificial sweeteners induce glucose intolerance by altering the gut microbiota. *Nature* 514, 181–184. doi: 10.1038/nature13793
- Toney, M. D., and Kirsch, J. F. (1989). Direct Bronsted analysis of the restoration of activity to a mutant enzyme by exogenous amines. *Science* 243, 1485–1488. doi: 10.1126/science.2538921
- Van Der Kamp, M. W., and Mulholland, A. J. (2013). Combined quantum mechanics/molecular mechanics (QM/MM) methods in computational enzymology. *Biochemistry* 52, 2708–2728. doi: 10.1021/bi400215w
- Wang, J., Woldring, R. P., Roman-Melendez, G. D., McClain, A. M., Alzua, B. R., and Marsh, E. N. G. (2014). Recent advances in radical SAM enzymology: new structures and mechanisms. *ACS Chem. Biol.* 9, 1929–1938. doi: 10.1021/cb5004674
- Wolfenden, R. (1969). Transition state analogues for enzyme catalysis. *Nature* 223, 704–706. doi: 10.1038/223704a0
- Wolfenden, R., and Snider, M. J. (2001). The depth of chemical time and the power of enzymes as catalysts. *Acc. Chem. Res.* 34, 938–945. doi: 10.1021/ar000058i

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Marsh. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.