Published in final edited form as:

Curr Opin Chem Biol. 2009 February ; 13(1): 1–2. doi:10.1016/j.cbpa.2009.02.022.

Engineering enzymes by “intelligent” design

Romas Kazlauskas and Stefan Lutz

Editorial Overview

The current research challenges in biocatalysis and biotransformation evolve around enzyme discovery, design and optimization. How can we find or create enzymes that catalyze important synthetic reactions, even reactions that may not exist in Nature? How can we optimize these enzymes for non-biological target substrates and industrial reaction conditions? How can we effectively assemble multiple enzymes into multi-step pathways? The common strategies to find answers to these questions include bioinformatics, directed evolution, and rational design, yet there are many different flavors within each approach and even more ways of combining these tools. The six articles in this special issue display a menu of different current approaches.

Undoubtedly the biggest success in biocatalysis in 2008 was the successful design of two catalytically active enzymes by the Baker group reported in *Science* and *Nature*. One designed enzyme catalyzes an aldol addition, an important carbon-carbon bond forming reaction; the other performs a Kemp elimination, notable because it is a reaction not found in Nature. Several of the reviews in this issue touch on this work, highlighting the many ways on how these results have impacted the field. These two success stories also emphasize one of the remaining challenges - increasing the efficiency of catalysts. Enzyme design and redesign is possible but often yields inefficient catalysts not ready for industrial applications. It is not currently clear what is missing from these designs to increase the efficiency.

Directed evolution has emerged as a highly effective process to engineer enzymes with new properties. Although the idea is straightforward - make enzyme variants, screen for the improved ones and repeat - the precise strategies are an active area of research. Is it best to do a few cycles of large libraries with multiple mutations or should one focus on more cycles of smaller libraries with fewer mutations? Tracewell and Arnold argue for the second approach: accumulating beneficial mutations one at a time or as their title states “Climbing fitness peaks one amino acid at a time.” They reason that most enzyme engineering solutions are not unique; there are numerous different variants that will have the right properties and also numerous paths one could take to find each one of these. The one-beneficial-mutation-at-a-time path is easy to follow. Each single mutation brings the enzyme closer to the desired property and the combination of single mutations achieves the desired property. The authors also acknowledge pitfalls in this strategy: one may reach a dead end when the protein is not stable enough to accept new mutations; one will miss mutations that are only beneficial cooperatively but not individually, and one cannot make major changes in protein architecture.

Divergent evolution has created families of enzymes with the same protein fold, but different catalytic activities. If nature did it, can researchers also do it in the laboratory? Gerlt and Babbitt

review this emerging research area and conclude yes, we can -- but not (yet) as well as nature did. The article offers a thorough review of recent efforts to change the catalytic activity of enzymes. From the discussion of the various strategies, one common theme for identifying conserved residues for a particular reaction type is the sequence comparison of enzymes catalyzing the same reaction. In addition, x-ray structures and mechanistic insight are essential as guiding principles. Together, the information can identify the critical few amino acid substitutions that form the basis for a particular enzymatic reaction, even when the members of the superfamily differ by hundreds of amino acid substitutions.

In their article, Schwaneberg and coworkers remind us of the staggering numbers associated with protein sequence space, as well as recent efforts to more effectively negotiate this multidimensional landscape through the development of new computational tools and experimental protocols. Individually, as well as in combination with existing methods, these new tools take the field one step closer to generating high-quality libraries, small in size and rich in beneficial mutations. Beyond addressing important fundamental questions concerning the boundaries of laboratory evolution, set by protein fitness and mutational thresholds, the review introduces a series of new computer algorithms to guide experimentalists in their efforts to balance maximal library diversity while minimal screening efforts. In addition, a predictive framework for evaluating the randomness of existing mutagenesis protocols can lend bench researchers a helpful hand in choosing the most suitable method for their project. Translating these *in silico* data into experimental results, a number of new strategies for creating focussed libraries sum up the review.

Besides the two computational *de novo* enzyme design coups of the Baker lab, the review by Damborsky examines a number of new and exciting computational approaches for remodeling active sites, ligand access tunnels, and protein-protein interfaces, as well as for generating more stable enzymes. Despite impressive changes and improvements in function, the author does concede that even the best of these designs is no match for native biocatalysts in regard to performance as several parameters relevant to enzyme function are poorly understood and/or are neglected in current computational models. Furthermore, the article offers an alternative method to the one-at-a-time strategy suggested by Tracewell and Arnold, which suggests the creation of small libraries, consisting of maybe a few hundred members with up to 10 mutations. Functional evaluation of these library members for a desired property is combined with machine-learning algorithms for data analysis, which assigns a net effect (positive, neutral, negative) to each tested amino acid position. This latter strategy was first implemented by Fox and coworkers at Codexis and has, upon combination of multiple positive substitutions in a single enzyme, resulted in significant improvements in the targeted function.

Beyond the studies of individual enzymes, the exploration and tailoring of entire biosynthetic pathways continues to excite and challenge scientists and engineers alike. In their review, Bernhardt and O'Conner recapitulate recent advances in the field, focusing on multi-enzyme complexes such as polyketide synthases (PKSs) and nonribosomal peptide synthetases (NRPSs), as well as biocatalysts involved in alkaloids and isoprenoids biosynthesis. Besides crucial new insights from structural studies of NRPSs and new discoveries regarding the mechanism of iterative PKSs, the authors report on advances in controlling the stereochemistry of the polyketide backbone in linear PKSs through mutations in the keto and enol reductase domains. Separate studies show that product derivatization does not have to stop with the polyketide skeleton; directed evolution of glycosyltransferases has yielded new biocatalysts for product glycosylation with a broad range of natural and unnatural sugars. Reflecting the future trend in protein engineering away from large, random libraries towards small, targeted libraries, the review closes with a discussion of three impressive examples of semi-rationally designed synthases with altered substrate and product specificities.

Finally, Tao and Xu take a look at biocatalysis from an applied standpoint. Citing five examples of high-value pharmaceutical intermediates, the authors illustrate the potential benefits upon integration of biocatalytic steps. While this integration might require a complete process redesign, potential benefits might include shorter reaction sequences, energy conservation, and waste reduction. In the cases of Zocor® and Lipitor®, new chemienzymatic routes enabled a drastic reduction in the number of reaction steps while the new bio-processes for Lyrica®, Paxil® and Keppra® significantly boost product yields by capitalizing on the enzymatic resolution of enantiomers in combination with racemization of the “wrong” enantiomer. Equally important, the examples discussed in the review demonstrate that new and improved processes do not necessarily depend on fancy new enzymes. Instead, well established biocatalysts such as subtilisin Carlsberg can be candidates of choice as much as enzymes isolated from metagenomic libraries or engineered by laboratory evolution techniques.

Biographies

Romas Kazlauskas studied chemistry at the Massachusetts Institute of Technology (Ph.D.) and Harvard University (postdoc with George Whitesides). He worked at General Electric Company (1985–88) and McGill University, Montreal, Canada (1988–2003) and is currently a professor in Biochemistry, Molecular Biology and Biophysics at the University of Minnesota, Twin Cities. He is an expert in enantioselective organic synthesis using enzymes.

Stefan Lutz holds a B. S. degree from the Zurich University of Applied Sciences (Switzerland), and a M.S. degree from the University of Teesside (UK). He obtained his Ph.D. from the University of Florida and spent three years as a postdoctoral researcher with Stephen Benkovic at Pennsylvania State University under a fellowship of the Swiss National Science Foundation. Since 2002 he has been a Chemistry professor at Emory University in Atlanta, Georgia (USA). The research in the Lutz laboratory focuses on the structure-function relationship of proteins through combinatorial protein engineering.