Biocatalysis in the Pharmaceutical Industry: The Need for Speed

Matthew D. Truppo*

Merck & Co., Inc., 126 East Lincoln Avenue, Rahway, New Jersey 07065, United States

ABSTRACT: The use of biocatalysis in the pharmaceutical industry continues to expand as a result of increased access to enzymes and the ability to engineer those enzymes to meet the demands of industrial processes. However, we are still just scratching the surface of potential biocatalytic applications. The time pressures present in pharmaceutical process development are incompatible with the long lead times required for engineering a suitable biocatalyst. Dramatic increases in the speed of protein engineering are needed to deliver on the ever increasing opportunities for industrial biocatalytic processes.

KEYWORDS: Biocatalysis, protein engineering, sustainability, enzymes

Biocatalysis is the use of enzymes in chemical synthesis. These enzymes can be used as isolated preparations or in whole cell format, prepared either in their native cells or as recombinantly expressed proteins in alternate host cells. The use of biocatalysis has expanded significantly over the last few decades, and it has impacted chemical synthesis in multiple industries including pharmaceuticals, fine chemicals, and food.[1−3] My commentary in this Innovation will focus specifically on the use of biocatalysis in the pharmaceutical industry, but many of the central themes apply across industries.

The Benefits of Biocatalysis. Biocatalysis is an appealing technology for the pharmaceutical industry for several reasons. Enzyme-based catalysis meets the ever increasing demands for highly selective, safe, and sustainable industrial processes. Unlike their chemocatalyst counterparts, biocatalysts have a very large three-dimensional structure that makes multiple points of contact with a substrate of interest, allowing for exquisite selectivity (Figure 1). Through protein engineering,
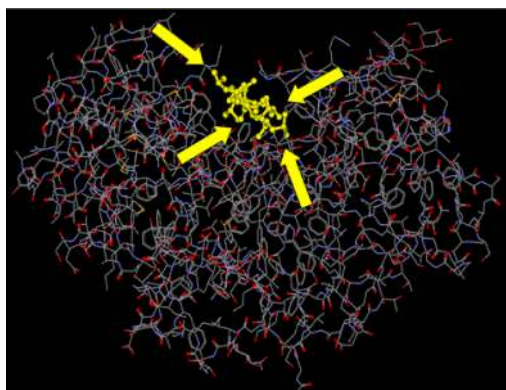
Figure 1. Representation of enzyme structure making multiple points of contact with a substrate for excellent selectivity.

modifications to the protein sequence and therefore structure can be easily made to change the properties of the biocatalyst. The excellent regio- and stereoselectivity of enzyme catalysts along with their ability to work under mild reaction conditions (thus protecting existing functionality within a molecule) enables transformations without the need for multiple protection and deprotection steps within a synthesis.

Additionally, biocatalysis offers both economic and environmental advantages over chemocatalytic methods. Enzymes are produced from inexpensive renewable resources and are themselves biodegradable, fulfilling the central tenants of green chemistry and sustainable development outlined by Graedel:[4]

(1) "using natural resources at rates that do not unacceptably draw down supplies over the long term"

(2) "generating and dissipating residues at rates no higher than can be assimilated readily by the natural environment"

When compared to synthetic methodologies that employ precious metals for catalysis, the cost and sustainability benefits are very clear. Metals like rhodium are often employed for asymmetric transformations in chemical synthesis. However, this represents one of the scarcest metals on earth.[5] In addition to the environmental cost of mining for the precious metal, the scarcity and competing demand by other industries (automotive and electronics primarily) leads to large fluctuations in the market price of metals such as rhodium, potentially disrupting supply chains and cost of goods projections (Figure 2). The costs of producing biocatalysts by comparison are stable, predictable, and easily amenable to economic modeling using standard tools.[6,7]
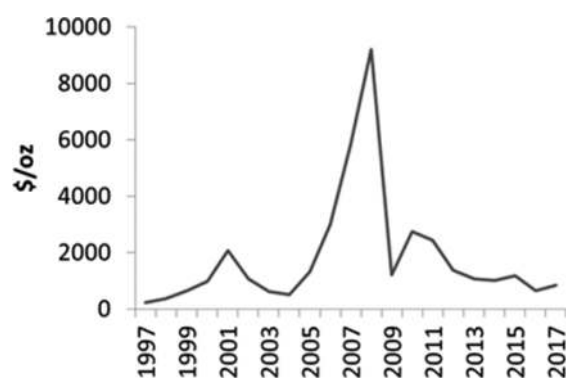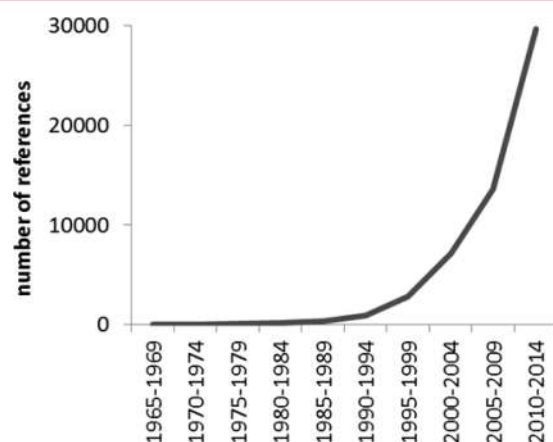
Figure 2. Rhodium price fluctuations over time. Data from publicly available Johnson Matthey price tables.

Despite the clear benefits of biocatalytic processes, the historical number of industrial applications has been modest, with a dramatic increase in the use of biocatalysis only occurring within the last two decades (Figure 3).[8,9] The reason

Figure 3. Number of publications and patents discussing "pharmaceutical biocatalysis" for each 5 year period of the last 50 years. Metrics from Google Scholar.

for this is centered around one key requirement demanded by the pharmaceutical industry... THE NEED FOR SPEED. Simply stated, the ability to identify, obtain, test, and optimize biocatalysts for the synthesis of pharmaceutical intermediates has been too slow to have a significant impact until recently.[10] Two major factors have led to an increase in our ability to develop and implement biocatalytic routes in pharmaceutical syntheses in a relevant time scale. The first is easy access to enzymes, and the second is our ability to engineer those proteins.

Access to Biocatalysts. First we will discuss enzyme access. Enzyme access can be thought of as both, our ability to physically obtain an enzyme and the range of biocatalytic activities available to exploit. About 15 years ago, the industry largely moved away from whole cell biocatalysis to focus almost exclusively on isolated enzyme based processes. Long lead times and significant investment due to complexities with the development of whole cell fermentation processes relegated this work toward the end of the drug development timeline. This meant that the implementation of a whole cell biocatalysis step in a synthesis required supplanting existing chemistry that had been used well into human clinical trials. The switch to isolated enzyme biocatalysis changed that. Stable enzyme

preparations could now be stored in the laboratory fridge and used to quickly deliver pharmaceutical intermediates at any stage of the drug development process. Implementing a biocatalytic step early on in route development is critical to solidifying biocatalysis' role in that synthesis. Ideally, one would initiate biocatalysis efforts early in a discovery program to both accelerate and inspire the development of new chemical matter.[11] This move away from whole cell processes created the need for cofactor recycling.[12] Early on, the vast majority of biotransformations utilized cofactor-independent hydrolases. The lack of need for cofactor recycling and the widespread availability of diverse enzymes with a broad range of substrate specificities made hydrolases uniquely well suited for large scale industrial applications. However, cofactor-dependent enzymes, such as ketoreductases and transaminases exhibited extremely useful synthetic utility, so work was undertaken to develop economical cofactor recycling systems for these enzyme classes.[13] In addition to easy access in the form of isolated enzyme powders, the collective efforts of the field to investigate novel enzyme classes has significantly expanded the repertoire of biocatalytic chemistries available to the synthetic chemist.[14−16] From hydrolytic reactions to reductions, transaminations, oxidations, and halogenations, the toolbox of activities available to synthetic organic chemists has never been greater.[17−21] Our efforts to discover and access new enzyme activities have been enabled by a combination of genomic DNA sequencing and recombinant DNA technology. Advances in sequencing technology have exponentially increased the number of sequences populating publicly available databases. These sequences can be mined using computational tools to search for sequence homology compared to enzymes of known function. Recombinant expression of these enzyme sequences in hosts such as *E. coli* then enables rapid access to putative enzymes with the desired activity.[22,23]

Protein Engineering. The second and probably most important factor impacting the speed of implementation for biocatalytic processes is protein engineering.[24] Biocatalysis has moved through three distinct phases over the last century, and I propose that we are starting to move into a fourth phase. These phases have been distinguished by our ability to modify or engineer a protein and its properties to suit our needs. An excellent review by Bornscheuer et al. describes this in detail, but I will summarize some key points here.[2] The first phase consisted of the use of naturally occurring biocatalysts to mediate a desired transformation. The chemistry utilized was dependent on a wild type enzyme's natural proclivity to convert a substrate to the desired product. During the second phase, which took place in the 1980s and 1990s, early protein engineering techniques guided by structural information were used to expand the substrate scope of biocatalysts to non-natural compounds. The third phase accelerated the pace of biocatalyst optimization using directed evolution approaches pioneered by Pim Stemmer and Frances Arnold.[25,26] The rapid generation of enzyme mutants using new molecular biology techniques combined with selective pressure via screening conditions allowed for enzymes to be improved for desired properties at a much more rapid rate, regardless of the existence of a crystal structure. This ubiquitous approach has been used on countless enzymes and is the primary method of improving biocatalysts today, leading to the dramatic uptake in biocatalysis starting in the late 1990s highlighted earlier. The directed evolution approach to enzyme optimization has been

accelerated and commercialized by companies like Codexis, who have combined molecular biology, automated robotics, and integrated software for mutational effect analysis. Indeed it is fair to say that protein engineering via directed evolution is the single greatest enabler of biocatalysis. It allows us to envision an ideal process and create suitable biocatalysts to fit that process rather than rely solely on enzymes provided by nature. Unfortunately, it is also fair to say that protein engineering is the greatest bottleneck to implementing biocatalysis. Despite decades of advancements in the field, the timeline to optimize a biocatalyst for implementation in a pharmaceutical process is still too long.

How Fast Is Fast Enough? One of the most successful and publicized examples of a pharmaceutical process using a highly evolved biocatalyst is the sitagliptin process used in the commercial manufacture of our company's largest product by both volume and sales, Januvia. The directed evolution of the transaminase enzyme used in this process took one year.[18,27] The new process then required a refile with regulatory agencies as the product was already on the market by the time the biocatalytic synthesis was ready for implementation. Our goal should be delivering the best chemistry at product launch to accrue the financial and supply chain benefits that the most efficient process affords. These benefits include reduced cost of goods, decreased inventory requirements, and the ability to more rapidly respond to market demand changes. Implementing the best process at product launch also avoids the resource consuming and costly prospect of refiling a new process with regulatory agencies around the world.

So how much do we need to accelerate protein engineering to meet the needs of the pharmaceutical industry? Simply relying on the sitagliptin example would suggest a doubling of speed is sufficient to meet our needs. However, stopping there would limit the prospects of biocatalysis to single steps in a synthesis across just a few programs in a drug company's pipeline. The recent successes of biocatalytic routes have generated many more "customers", and entire drug portfolios are now assessed for potential biocatalysis impact. Additionally, the field is not limiting itself to a single chiral transformation in a synthesis. Biocatalytic cascades are proving to be very attractive methodologies to rapidly build up molecular complexity from inexpensive starting materials in one pot while driving reaction equilibrium toward the desired products.[28−31] Combining the desire to file the best chemistry at product launch, the increased number of biocatalysis programs in the pipeline, and the possibility for multiple enzymatic steps within a synthesis, we can easily see how a 10× improvement in protein engineering speed is needed to deliver on all of these opportunities.

The Fourth Phase of Biocatalysis. The early phases of biocatalysis progressed with "fits and starts". Pharmaceutical companies often ramped up efforts only to draw down resources years later, due to the inability of the technology to deliver results in the timelines required. I believe the ongoing investment in biocatalysis is here to stay, now that we've achieved escape velocity in the pace of protein engineering. However, the challenge of designing a biocatalyst an order of magnitude faster is still a formidable one. The final enzyme variant is often significantly modified from its parent protein, with 10−20% of the wild type amino acids being substituted for others. Putting this in context, for a protein with 300 amino acids, there are greater than 10^{390} possible sequence combinations that can be generated using the standard 20

amino acids.[32] We can clearly see that this is a daunting numbers game, which cannot be won simply by increasing screening capacity and randomly exploring more sequence space. This highlights our need to enter a fourth phase of biocatalysis in which biocatalysts are engineered through rational directed evolution in a design−make−test cycle combining multiple disciplines in one seamless industrialized workflow (Figure 4). The design phase consists of the use of
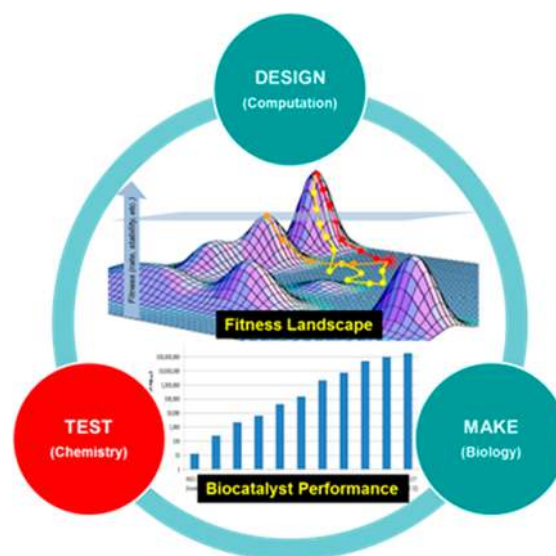

Figure 4. Design−make−test cycle for engineering biocatalysts.

computation and informatics to generate diversity representative of the sequence space we wish to explore to improve the fitness of the enzyme. The make phase leverages the amazing power of biology to generate the physical constructs we want to test. Finally the test phase of the cycle leverages high throughput experimentation to evaluate each construct's performance under the desired experimental conditions. This iterative process allows protein engineers to build up an informed model of how mutations impact a protein's function, marching us ever closer to our target design parameters.

Computational tools are undeniably an important driver of increased speed across multiple aspects of the design cycle including: the development of rationally derived constructs, the statistical deconvolution of mutational effects, and the ever increasing ability to predict the impacts of simultaneous mutations within the protein.[33] The field of *de novo* protein design has led to increasingly interesting results in recent years. However, catalytic efficiencies of these in silico designed biocatalysts are typically much lower than what can be obtained with directed evolution methodologies.[34,35] As a result, we are starting to observe an interesting and useful trend whereby *de novo* designed initial enzyme constructs are further improved using classical directed evolution techniques.[36−38] The future needs to merge these approaches ever closer by leveraging the very large data sets generated across multiple enzyme evolution projects to enrich our predictive models via machine learning algorithms, rather than simple statistical deconvolutions. The field also needs to leverage more information than a single assay for activity or selectivity provides in the test portion of the design cycle. Ideally, high throughput methodologies for complete biophysical characterization of each construct would correlate mutations to a host of parameters: activity, selectivity,

expression, proper folding, stability, etc. Ultimately, these efforts should lead to a decrease in the time it takes to conduct each round of protein engineering and an increase in the performance gain achieved as a result of each round across a variety of metrics (Figure 5).
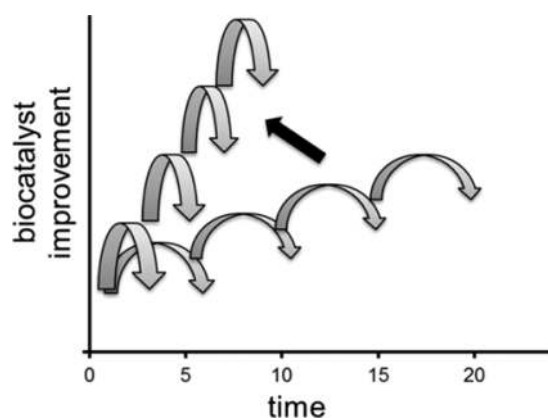
Figure 5. Acceleration of both time and level of improvement in each round of protein engineering.

Enzyme Immobilization. While the primary driver of biocatalysis' success has been protein engineering, enzyme immobilization can play an important role in the adoption of biocatalytic processes. Most synthetic chemistry is carried out in organic solvents. Enzyme immobilization enables the use of biocatalysts in organic solvent systems.[39,40] The ability to run the enzyme catalyzed transformation in solvents compatible with upstream and downstream chemistry eliminates many of the isolation and solvent switching steps required with conventional aqueous biocatalysis, providing for significantly more efficient processes with lower cost and less waste. This facilitates reaction telescoping, facile enzyme recovery and reuse, easier rejection of protein related impurities from the final active pharmaceutical ingredient, and continuous processing methodologies. While many examples of enzyme immobilization exist, a general methodology for the rapid and inexpensive immobilization of biocatalysts still eludes us. Solving this problem will require a cross functional effort between chemistry, molecular biology, material science, and chemical engineering.

The Future Is Bright. The impact of biocatalysis continues to grow within the pharmaceutical industry. Rapid access to enzymes with a variety of natural and even unnatural catalytic activities has expanded the synthetic organic chemist's toolbox.[41−43] Fully leveraging opportunities for lower cost and more sustainable biocatalytic processes requires a dramatic increase in the speed of protein engineering. The only way to achieve this is with a multidisciplinary approach, combining high throughput experimentation, advanced molecular biology techniques, and computation tools. Finally, just as a holistic view is necessary to accelerate protein engineering, the design of biocatalytic syntheses as a whole is a global optimization problem. Cascade biocatalysis and the ability to run in organic solvents with immobilized enzymes gets us another step closer to the ultimate goal. We need to think of "biocatalysis" as just "catalysis". Biocatalysis is simply another set of reactions within the world of organic chemistry that allows us to develop the ultimate synthesis of a molecule.

AUTHOR INFORMATION

Corresponding Author

*E-mail: matthew.truppo@merck.com.

ORCID ●

Matthew D. Truppo: 0000-0003-4917-7838

Notes

The author declares no competing financial interest.

REFERENCES

(1) Turner, N. J.; Truppo, M. D. Biocatalysis enters a new era. *Curr. Opin. Chem. Biol.* **2013**, *17*, 212−214.

(2) Bornscheuer, U. T.; Huisman, G. W.; Kazlauskas, R. J.; Lutz, S.; Moore, J. C.; Robins, K. Engineering the third wave of biocatalysis. *Nature* **2012**, *485*, 185−194.

(3) Reetz, M. T. *J. Am. Chem. Soc.* **2013**, *135*, 12480−12496.

(4) Graedel, T. E. In *Handbook of Green Chemistry and Technology*; Clark, J., Macquarrie, D. J., Eds.; Wiley: New York, 2002; pp 56−61.

(5) Alonso, E.; Field, F. R.; Kirchain, R. E. A case study of the availability of platinum group metals for electronics manufacturers. *IEEE* **2008**, 1.

(6) Liu, G.; Zhang, J.; Bao, J. Cost evaluation of cellulose enzyme for industrial − scale cellulosic ethanol production based on rigorous Aspen Plus modeling. *Bioprocess Biosyst. Eng.* **2016**, *39*, 133−140.

(7) Klein-Marcuschamer, D.; Oleskowicz-Popiel, P.; Simmons, B. A.; Blanch, H. W. The challenge of enzyme cost in the production of lignocellulosic biofuels. *Biotechnol. Bioeng.* **2012**, *109*, 1083−1087.

(8) Choi, J.-M.; Han, S.-S.; Kim, H.-S. Industrial applications of enzyme biocatalysis: current status and future prospects. *Biotechnol. Adv.* **2015**, *33*, 1443−1454.

(9) Schmid, A.; Dordick, J. S.; Hauer, B.; Kiener, A.; Wubbolts, M.; Witholt, B. Industrial biocatalysis today and tomorrow. *Nature* **2001**, *409*, 258−268.

(10) Pollard, D. J.; Woodley, J. M. Biocatalysis for pharmaceutical intermediates: the future is now. *Trends Biotechnol.* **2006**, *25*, 66−73.

(11) Molinaro, C.; Bulger, P. G.; Lee, E. E.; Kosjek, B.; Lau, S.; Gauvreau, D.; Howard, M. E.; Wallace, D. J.; O'Shea, P. D. CRTH2 Antagonist MK-7246: A Synthetic Evolution from Discovery through Development. *J. Org. Chem.* **2012**, *77*, 2299−2309.

(12) Wachtmeister, J.; Rother, D. Recent advances in whole cell biocatalysis techniques bridging from investigative to industrial scale. *Curr. Opin. Biotechnol.* **2016**, *42*, 169−177.

(13) Truppo, M. D. Cofactor recycling for enzyme catalyzed processes. *Comprehensive Chirality*; Carreira, E. M., Yamamoto, H., Eds.; Elsevier, 2012; pp 46−70.

(14) Faber, K.; Fessner, W.-D.; Turner, N. J. *Biocatalysis in Organic Synthesis*; Thieme, 2015; Vol. 3.

(15) Liese, A.; Pesci, L. Enzyme classification and nomenclature and biocatalytic retrosynthesis. In *Science of Synthesis Biocatalysis in Organic Synthesis*; Thieme, 2015; pp 41−74. DOI: 10.1055/sos-SD-214-00028.

(16) Prier, C. K.; Arnold, F. H. Chemomimetic biocatalysis: exploiting the synthetic potential of cofactor-dependent enzymes to create new catalysts. *J. Am. Chem. Soc.* **2015**, *137*, 13992−14006.

(17) Turner, N.; Carr, R. In *Biocatalysis in the Pharmaceutical and Biotechnology Industries*; Patel, R., Ed.; CRC Press: Boca Raton, FL, 2007; pp 743−755.

(18) Savile, C.; Janey, J.; Mundorff, E.; Moore, J.; Tam, S.; Jarvis, W.; Colbeck, J.; Krebber, A.; Fleitz, F.; Brands, J.; Devine, P.; Huisman, G.; Hughes, G. *Science* **2010**, *329*, 305−309.

(19) Kroutil, W.; Velikogne, S.; Schrittwieser, J. H. Biocatalytic imine reduction and reductive amination of ketones. *Adv. Synth. Catal.* **2015**, *357*, 1655−1685.

(20) Weichold, V.; Milbredt, D.; van Pee, K.-H. Specific enzymatic halogenation-from the discovery of halogenated enzymes to their applications in vitro and in vivo. *Angew. Chem., Int. Ed.* **2015**, *55*, 6374−6389.

(21) Huisman, G. W.; Collier, S. J. On the development of new biocatalytic processes for practical pharmaceutical synthesis. *Curr. Opin. Chem. Biol.* **2013**, *17*, 284−292.

(22) Davids, T.; Schmidt, M.; Bottcher, D.; Bornscheuer, U. T. Strategies for the discovery and engineering of enzymes for biocatalysis. *Curr. Opin. Chem. Biol.* **2013**, *17*, 215−220.

(23) Jacques, P.; Bechet, M.; Bigan, M.; Caly, D.; Chataigne, G.; Coutte, F.; Flahaut, C.; Heuson, E.; Leclere, V.; Lecouturier, D.; Phalip, V.; Ravallec, R.; Dhulster, P.; Froidevaux, R. High-throughput strategies for the discover and engineering of enzymes for biocatalysis. *Bioprocess Biosyst. Eng.* **2017**, *40*, 161−180.

(24) Woodley, J. M. Protein engineering of enzymes for process applications. *Curr. Opin. Chem. Biol.* **2013**, *17*, 310−316.

(25) Arnold, F. H. The nature of chemical innovation: new enzymes by evolution. *Q. Rev. Biophys.* **2015**, *48*, 404−410.

(26) Denard, C. A.; Ren, H.; Zhao, H. Improving and repurposing biocatalysts via directed evolution. *Curr. Opin. Chem. Biol.* **2015**, *25*, 55−64.

(27) Desai, A. Sitagliptin manufacture: a compelling tale of green chemistry, process intensification, and industrial asymmetric catalysis. *Angew. Chem., Int. Ed.* **2011**, *50*, 1974−1976.

(28) Sigrist, R.; da Costa, B. Z.; Marsaioli, A. J.; de Oliveira, L. G. Nature-inspired enzymatic cascades to build valuable compounds. *Biotechnol. Adv.* **2015**, *33*, 394−411.

(29) Fernandez-Lucas, J. Multienzymatic synthesis of nucleic acid derivatives: a general perspective. *Appl. Microbiol. Biotechnol.* **2015**, *99*, 4615−4627.

(30) Oroz-Guinea, I.; Garcia-Junceda, E. Enzyme catalyzed tandem reactions. *Curr. Opin. Chem. Biol.* **2013**, *17*, 236−249.

(31) Otte, K. B.; Hauer, B. Enzyme engineering in the context of novel pathways and products. *Curr. Opin. Biotechnol.* **2015**, *35*, 16−22.

(32) Currin, A. C.; Swainston, N.; Day, P. J.; Kell, D. B. Synthetic biology for the directed evolution of protein biocatalysts: navigating sequence space intelligently. *Chem. Soc. Rev.* **2015**, *44*, 1172−1239.

(33) Romero-Rivera, A.; Garcia-Borras, M.; Osuna, S. Computational tools for the evaluation of laboratory-engineered biocatalysts. *Chem. Commun.* **2017**, *53*, 284−297.

(34) Kries, H.; Blomberg, R.; Hilvert, D. De novo enzymes by computational design. *Curr. Opin. Chem. Biol.* **2013**, *17*, 221−228.

(35) Wijma, H. J.; Floor, R. J.; Bjelic, S.; Marrink, S. J.; Baker, D.; Janssen, D. B. Enantioselective enzymes by computational design and in silico screening. *Angew. Chem., Int. Ed.* **2015**, *54*, 3726−3730.

(36) Obexer, R.; Pott, M.; Zeymer, C.; Griffiths, A. D.; Hilvert, D. Efficient laboratory evolution of computationally designed enzymes with low starting activities using fluorescence-activated droplet sorting. *Protein Eng., Des. Sel.* **2016**, *29*, 355−366.

(37) Garrabou, X.; Verez, R.; Hilvert, D. Enantiocomplementary synthesis of nitroketones using designed and evolved carboligases. *J. Am. Chem. Soc.* **2017**, *139*, 103−106.

(38) Osuna, S.; Jimenez-Oses, G.; Noey, E. L.; Houk, K. N. Molecular dynamics explorations of active site structure in designed and evolved enzymes. *Acc. Chem. Res.* **2015**, *48*, 1080−1089.

(39) Li, H.; Moncecchi, J.; Truppo, M. D. Development of an immobilized ketoreductase for Enzymatic (*R*)-1-(3,5-bis (trifluoromethyl)phenyl)ethanol production. *Org. Process Res. Dev.* **2015**, *19*, 695−700.

(40) Truppo, M. D.; Strotman, H.; Hughes, G. Development of an immobilized transaminase capable of operating in organic solvent. *ChemCatChem* **2012**, *4*, 1071−1074.

(41) McIntosh, J. A.; Farwell, C. C.; Arnold, F. H. Expanding P450 catalytic reaction space through evolution and engineering. *Curr. Opin. Chem. Biol.* **2014**, *19*, 126−134.

(42) Windle, C. L.; Muller, M.; Nelson, A.; Berry, A. Engineering aldolases as biocatalysts. *Curr. Opin. Chem. Biol.* **2014**, *19*, 25−33.

(43) Tinoco, A.; Steck, V.; Tyagi, V.; Fasan, R. Highly diastereo- and enantioselective synthesis of trifluoromethyl-substituted cyclopropanes via myoglobin-catalyzed transfer of trifluoromethylcarbene. *J. Am. Chem. Soc.* **2017**, *139*, 5293.