



# Back to the future: Why we need enzymology to build a synthetic metabolism of the future

Tobias J. Erb<sup>1,2</sup>

## Review

Open Access

### Address:

<sup>1</sup>Max-Planck-Institute for Terrestrial Microbiology, Department of Biochemistry & Synthetic Metabolism, Karl-von-Frisch-Str. 10, D-35043 Marburg, Germany and <sup>2</sup>LOEWE Center for Synthetic Microbiology (SYNMIKRO), Marburg, Germany

### Email:

Tobias J. Erb - toerb@mpi-marburg.mpg.de

### Keywords:

enzymes; in vitro biochemistry; metabolic engineering; synthetic biology

*Beilstein J. Org. Chem.* **2019**, *15*, 551–557.

doi:10.3762/bjoc.15.49

Received: 12 October 2018

Accepted: 29 January 2019

Published: 26 February 2019

This article is part of the thematic issue "Enzymes in chemical transformations".

Guest Editor: K. N. Allen

© 2019 Erb; licensee Beilstein-Institut.  
License and terms: see end of document.

## Abstract

Biology is turning from an analytical into a synthetic discipline. This is especially apparent in the field of metabolic engineering, where the concept of synthetic metabolism has been recently developed. Compared to classical metabolic engineering efforts, synthetic metabolism aims at creating novel metabolic networks in a rational fashion from bottom-up. However, while the theoretical design of synthetic metabolic networks has made tremendous progress, the actual realization of such synthetic pathways is still lacking behind. This is mostly because of our limitations in enzyme discovery and engineering to provide the parts required to build synthetic metabolism. Here I discuss the current challenges and limitations in synthetic metabolic engineering and elucidate how modern day enzymology can help to build a synthetic metabolism of the future.

## Introduction

One of the most important and disruptive events in the history of chemistry was its transformation from a purely analytical-descriptive into a synthetic-constructive discipline, which took place more than one hundred years ago [1,2]. Understanding the elemental composition of matter as well as the nature and reactivity of the chemical bond enabled chemists to use their knowledge to create new molecules and materials [3,4]. This development provided humankind with new chemical compounds, such as color dyes, pharmaceuticals,

as well as polymers and plastics. Given its transforming nature, it is beyond any doubt that synthetic chemistry has been one of the key enabling technologies of the 20th century, which has virtually changed the world we are living in. Biology is currently at the verge of a similar transition [5]. Over the last decades, our ability to analyze and manipulate living systems has provided the intellectual as well as technological basis to create biological features that are new to nature.

## Review

### Classical metabolic engineering: Exploiting natural metabolic networks

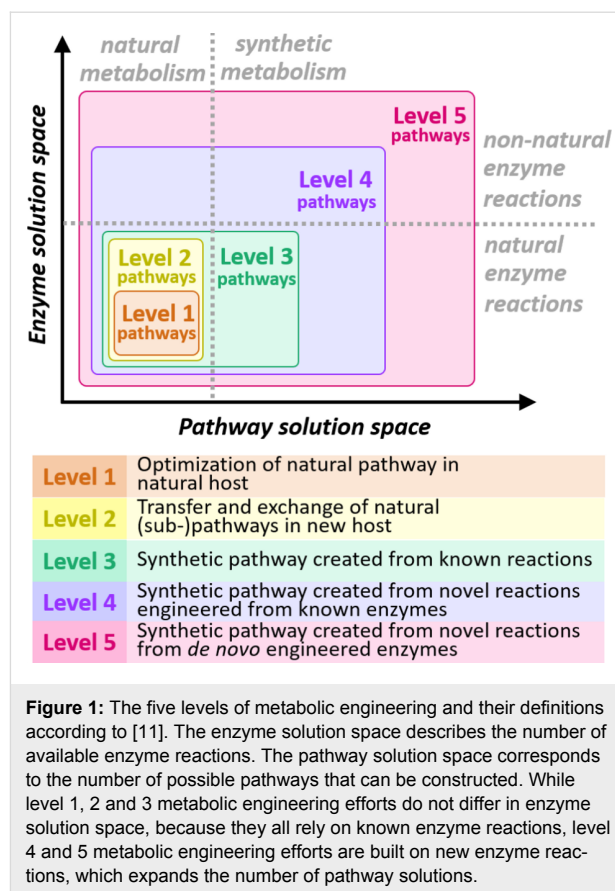
A fundamental feature of living systems is metabolism, which can be defined as the dynamic chemistry that allows life to organize itself in three and four dimensions [6]. The incredible metabolic potential of biology is impressively demonstrated by the more than 2,000 different chemical transformations that can simultaneously take place inside of an *Escherichia coli* cell [7,8], as well as by the more than 200,000 different molecules that have been isolated from different biological systems so far [9]. This diversity has inspired generations of biologists to use living cells as small chemical factories for the production of chemicals.

In the past, many efforts centered on manipulating the metabolism of cells to obtain a target molecule. Most of these approaches were based on the concept of metabolic engineering [10]. According to this concept, known pathways and enzymes are manipulated in such a way that a certain molecule can be produced at high purity and yield from a living bacterial cell [7]. In respect to their complexity (Figure 1), these classical metabolic engineering approaches can be classified as level 1 efforts, i.e., the optimization of a natural pathway in a native organism, or level 2 efforts, i.e., the transplantation or reconstruction of a natural pathway in a new host organism [11].

Classical metabolic engineering efforts, however, are limited in a way that they are still bound to existing pathways and reactions, which limit the accessibility of certain compounds, as well as the efficiency with which those compounds can be produced. In an ideal world, the way a target molecule is produced should not be dictated by the serendipity and constraints of evolution, but be accessible through rational design. However, this requires a fundamental understanding of those principles that are necessary for designing, realizing and operating multi-reaction sequences and metabolic networks *de novo*.

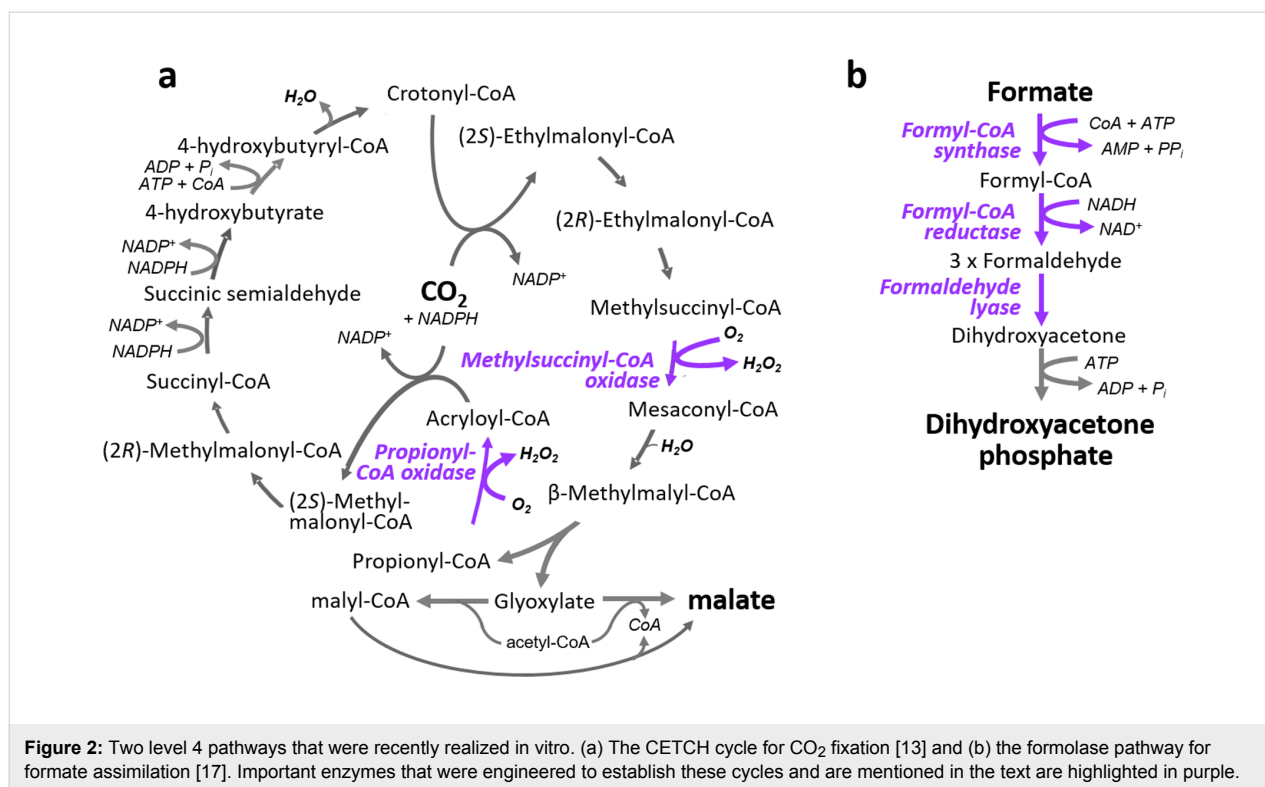
### Metabolic retrosynthesis: Next level metabolic engineering

Recently, the concept of “synthetic metabolism” was developed that aims at overcoming the limitations provided by natural metabolism through the realization of completely novel metabolic networks [11,12]. The novel networks are designed from first principles based on simple physico-chemical considerations, such as kinetics and thermodynamics. For the design, a starting compound and a target molecule are defined and a short, thermodynamically feasible and energetically efficient route connecting the two molecules is identified. While level 3 engineering efforts aim at creating new pathway solutions by mixing and matching known enzymes from different metabolic



pathways, the design efforts in their most advanced form (i.e., level 4 and level 5) do not build on existing enzymes, but only consider plausible chemical transformations and feasible metabolic intermediates [13–16]. In a subsequent realization phase, the corresponding enzymes to realize the theoretical network are identified and/or engineered and a first version of the synthetic network is reconstructed. The network is further optimized or evolved in following rounds in respect to production rate and yield.

As an example, several novel level 3 and level 4 pathways for the conversion of CO<sub>2</sub> into organic acids were developed recently [13,14]. These pathways are predicted to be more efficient than the naturally evolved Calvin cycle of photosynthesis, because they require less energy (ATP, redox power and/or photons) and can be supposedly operated at higher catalytic rates compared to natural carbon fixation. Accordingly, the synthetic CO<sub>2</sub>-fixation cycles should be able to convert more carbon dioxide with less energy in a given time and hence succeed natural photosynthesis in volumetric capacity and energetic efficiency. One of these designs, the so-called CETCH cycle (Figure 2a), a synthetic level 4 pathway for the conversion of CO<sub>2</sub> into organic acids, was experimentally realized *in vitro* by combining 17 enzymes (including three engineered ones) from



a total of nine different organisms from all three domains of life [13]. Compared to the first version of the cycle, the system was further improved until version 5.4 by almost a factor of 20, indicating that subsequent system optimization might be as important as initial reconstruction [13].

In a similar fashion, multiple level 3 and level 4 routes for the transformation of the one-carbon compound formate into cellular building blocks were designed that should theoretically outcompete natural formate assimilation pathways [18,19]. Some of the level 3 pathways were recently reconstructed in vivo [20–23] and one of the level 4 solutions – the formolase pathway (Figure 2b) – was demonstrated already in vitro [17]. This pathway relies on three new-to-nature reactions, the most prominent one being the name-giving formolase reaction, which allows the subsequent condensation of three formaldehyde molecules into the three-carbon compound dihydroxyacetone phosphate [24]. In addition to that, several alternative photorespiration, methanol assimilation, as well as glycolytic pathways of levels 3 and 4 were developed that are supposedly more carbon and energy efficient compared to their naturally evolved equivalents [23,25–28].

Yet, while an increasing number of theoretical designs are proposed, the successful experimental realization of many of these designs in the lab is still falling short. This is especially true for pathways of design levels 4 and 5 that feature novel reactions,

for which the corresponding enzymes are unknown (i.e., were not described to date). The realization of these pathways is severely restricted by our limited ability to discover and/or engineer new-to-nature enzymes. Notable exceptions are the formolase pathway and CETCH cycle that required each the establishment of three novel enzymatic reactions for their successful realization. However, other level 4 pathway designs require the establishment of more than ten so-far unknown enzyme reactions, emphasizing the challenge to realize truly synthetic metabolic networks [13].

### The challenge of finding (new) enzymes for synthetic metabolic networks

From above examples it becomes evident that for building completely novel pathways and/or complex reaction cascades, resources are required that provide synthetic biologists with the information to find individual enzymes for a given synthetic metabolic network. More than 116 million proteins were deposited into protein sequence databases, such as UniProtKB [29]. More than 40,000 enzymes were biochemically characterized and the corresponding data is available in specialized enzyme databases, such as BRENDA [30]. This wealth of biological information provides a good starting point to search for enzyme variants that possess a desired catalytic activity.

While existing databases might provide a good resource to find the parts to reconstruct level 3 pathways, this task becomes

more challenging in respect to level 4 and level 5 designs that require new-to-nature reactions. How can these new enzyme reactions be identified or established? One option is the de novo-design of enzymes assisted by computational methods, which have been developed over the last couple of years. When combined with experimental evolution and elaborate screening methods, these efforts have allowed to establish completely novel enzyme reactions from scratch [31-34].

However, even though considerable progress has been made in creating enzymes with the help of computational methods [35], it is a complementary (and equally valid) approach to discover and/or engineer novel reactions from the natural diversity of enzymatic scaffolds [36-39]. One example is formaldehyde lyase (or “formolase”) – the key enzyme of the formolase pathway – that was crafted from a benzaldehyde lyase, which showed initially some side reactivity with formaldehyde [17,24]. Other examples are propionyl-CoA oxidase and methylsuccinyl-CoA oxidase in the CETCH cycles that were engineered from a promiscuous short chain acyl-CoA oxidase and a FAD-dependent methylsuccinyl-CoA dehydrogenase, respectively [13,40,41].

These efforts in exploiting the promiscuity of enzymes to create novel catalysts might profit from new computational methods that succeeded in creating active sites of remarkable promiscuous activities in the scaffold of existing enzymes [42]. Such computationally-created “catalytically diverse active sites” could be further developed towards a new activity through directed evolution. Without any question, screening protein sequence and enzyme databases for suitable candidates is key to advance metabolic retrosynthesis. However, there are still some practical issues in extracting the necessary information from different databases. One particular problem of sequence databases like UniProtKB is the high number of misannotated proteins, which is caused by automatized annotation algorithms that are often based on “simple” sequence similarities [43,44]. In selected enzyme (super)families the annotation error can be as high as 90% [45], which masks or even impedes the identification of novel functions within a given enzyme (super)family. An example are reducing enoyl-CoA carboxylases that were for most of the time annotated as ordinary enoyl-CoA reductases, with which they are phylogenetically related [46,47]. Another example are RubisCO-like proteins [48] that are enolases [49], isomerases [50] and transcarboxylases [51], respectively, which are not capable of fixing CO<sub>2</sub>, but are still found very often misannotated as their CO<sub>2</sub> fixing homologs RubisCO, with which they share a common evolutionary history [52].

A solution to overcome the problem of misannotation might come from novel computational tools that were developed

recently to analyze the diversity of enzyme (super)families in respect to new functions [53-55]. While these tools have been successfully used to identify and discover new metabolic pathways (Balskus, etc.), they might as well be used to identify interesting candidate enzymes to be screened for new catalytic reactions in metabolic retrosynthesis. Further improvements in homology modeling and virtual docking are expected to increase accuracy and throughput, which will help to map and predict the substrate and reactions catalyzed by an enzyme superfamily and its individual members in the future.

### Enzyme promiscuity: Key and challenge for synthetic metabolism

Another problem is that even in databases that list the experimentally confirmed activity of enzymes, an important aspect is very often not well documented: substrate (and reaction) promiscuity. Yet, this information is essential to identify suitable candidate templates to engineer or evolve a new activity within the backbone of a given enzyme. For example, although the BRENDA database is probably one of the best resources to learn about the detailed catalytic properties of enzymes, it only provides in selected cases detailed information on the activity of a given enzyme with different substrate analogs. Besides providing the necessary information to identify interesting enzyme candidates for level 4 and level 5 pathway construction, more systematic data on enzyme promiscuity would also allow a more holistic view onto the catalytic (and evolutionary) potential of a complete enzyme superfamily [56].

Note that the information on substrate and/or reaction promiscuity is not only important to establish novel enzyme reactions, it is also of very practical information in the actual construction and optimization of synthetic metabolic networks. One problem in realizing metabolic networks from scratch with enzymes that did not evolve in the same physiological context is that the individual enzymes in such mix-and-match networks are prone to feature side reactivities with substrates or products of other enzymes in the synthetic network, most likely because they lack a common evolutionary history that selects for stringent substrate specificity [57]. These unwanted side reactivities are able to compete with the wanted reactions of the synthetic network and can lead to the accumulation of dead-end products, thus decreasing or even inhibiting flux through the whole synthetic network [58]. Consequently, it is important to learn of such unwanted side reactivities before reconstruction of the network to avoid unfruitful interactions and suboptimal functioning of the system.

Again, the CETCH cycle provides a good example, why information on the promiscuity of enzymes is so important for metabolic retrosynthetic efforts. In the first versions of the synthetic

pathway, a promiscuous methylmalyl-CoA lyase caused the accumulation of malyl-CoA from an undesired side reaction of the enzyme with acetyl-CoA, which stalled the cycle. To overcome the problem of unwanted malyl-CoA accumulation, a malyl-CoA thioesterase [59] had to be added to the synthetic network. This enzyme effectively recycles the dead-end metabolite back into two intermediates of the network, malate and free CoA, thus serving as a “proof-reading” enzyme at the periphery of the CETCH cycle to keep the system running. Another problem was posed by the promiscuous activity of propionyl-CoA carboxylase with acetyl-CoA. This problem was solved by replacing the problematic reaction with another enzymatic route. Finally, an initially promiscuous acyl-CoA oxidase was further engineered to increase the catalytic efficiency for the wanted substrate propionyl-CoA compared to the unwanted substrate 4-hydroxybutyryl-CoA by a factor of 50 [13]. Having had known these problematic side reactions beforehand would have probably allowed a more rational design and/or avoided some problems upfront [58].

Yet, it needs to be mentioned that even if complex synthetic metabolic networks can be realized *in vitro*, this does not mean that these metabolic networks can be easily transplanted into living cells. The introduction of new reactions and metabolites into a host cell is expected to create interactions with the native metabolic and regulatory network of the host. Again, promiscuity poses a major challenge. Even though the metabolites and reactions might be completely non-native to the cell, these intermediates might be still drained due to unwanted side reactions or create unwanted metabolic and regulatory effects that negatively affect or even prohibit operation of the synthetic metabolic network inside the host. This problem is exacerbated by the fact that for a well-studied organism like *E. coli*, the function of a large number of enzymes remains still unknown and there are likely to be hundreds if not thousands of unknown reactions and metabolites, often described as catalytic or metabolic “dark matter” [53,60]. Thus, a more detailed understanding of the promiscuity of native enzymes and the interaction of small molecules with the native regulatory network of cells is an important prerequisite to realize synthetic metabolism in the future [61]. In this context, it might also be very interesting to learn, which cellular hosts might be suited best for the transplantation of a given artificial network, or if current approaches to build synthetic cells from the bottom-up might represent a valuable alternative strategy [6].

### Linking enzymology and synthetic biology

In summary, synthetic biology can develop its full potential, if it becomes able to harness the diversity of the millions of different enzyme variants and homologs that naturally exist. While such information is collected and made available by many enzy-

mology and biochemistry laboratories worldwide in a community effort, it is not provided in an optimal way so that it can be used for the synthetic metabolism community. How could this apparent gap be bridged?

First, it will be necessary to collect enzymatic data in a more standardized fashion. As a matter of fact, standardization has been an important driver in the development of synthetic biology. This is probably best demonstrated by the BioBrick System [62] and the multitude of standardized genetic elements that are available for the assembly of complex genetic networks. The STREND standard [63,64] might provide a good blueprint, how enzyme data could be organized and reported by enzymologists in the future so that the synthetic biologist could better compare and evaluate different enzymes in respect to their suitability for a given pathway.

Second it will be important not only to investigate a given enzyme in respect to its native reaction, but also study its (potential) side reactivities more systematically. For every new enzyme characterized, it would be helpful if the enzymologist tested at least a small set of substrate and/or cofactor analogs. Even though a detailed kinetic data would not necessarily be required, the fact that a certain side reactivity exists in the scaffold of a given enzyme would already be a highly useful and relevant information for the synthetic biologist. On the one hand, this information could be used to identify a target enzyme for further engineering to develop the side reactivity as main activity [38,39]. On the other hand, this data would allow the synthetic biologist to anticipate potentially unwanted side reactions in the metabolic network by a given part and take corresponding countermeasures [58].

Third and lastly, there cannot enough enzymes be described. The discovery of new enzymes as well as the characterization of homologs of known enzymes needs to be continued and eventually even intensified. Only these efforts will allow to build an exhaustive library of enzyme parts for level 3, level 4 and level 5 metabolic engineering. At the same time the methods of (re-)engineering and the *de-novo* design of enzymes need to be further developed. This will allow to further develop and improve catalytic activities in enzymes and create new enzyme reactions that cannot be found naturally. Altogether, these activities will expand the limits of natural metabolism and pave the way for synthetic metabolic networks. Enzymology is far from being an old-fashioned business, its most fruitful era might just have begun.

### Acknowledgements

The author wants to thank Carsten Kettner for organizing the 2017 Beilstein Enzymology Symposium that served as inspira-

tion for this review, as well as the two anonymous reviewers for their valuable suggestions during the review process. Research of the author on the design and realization of synthetic metabolism is supported by the Max Planck Society, the European Research Council (ERC “Syborg”), as well as the BMBF (“FormatPlant”).

## References

- Jaffe, B. *Crucibles: The Story of Chemistry from Ancient Alchemy to Nuclear Fission*; Dover Publications: New York, 1976.
- Wöhler, F. *Ann. Phys. (Berlin, Ger.)* **1828**, *88*, 253–256. doi:10.1002/andp.18280870206
- Asimov, I. A. *Short History of Chemistry*; Anchor Books: Garden City, New York, 1967.
- Nicolaou, K. C.; Vourloumis, D.; Winssinger, N.; Baran, P. S. *Angew. Chem., Int. Ed.* **2000**, *39*, 44–122. doi:10.1002/(sici)1521-3773(2000103)39:1<44::aid-anie44>3.0.co;2-|
- Yeh, B. J.; Lim, W. A. *Nat. Chem. Biol.* **2007**, *3*, 521–525. doi:10.1038/nchembio0907-521
- Schwille, P.; Spatz, J.; Landfester, K.; Bodenschatz, E.; Herminghaus, S.; Sourjik, V.; Erb, T. J.; Bastiaens, P.; Lipowsky, R.; Hyman, A.; Dabrock, P.; Baret, J.-C.; Vidakovic-Koch, T.; Bieling, P.; Dimova, R.; Mutschler, H.; Robinson, T.; Tang, T.-Y. D.; Wegner, S.; Sundmacher, K. *Angew. Chem., Int. Ed.* **2018**, *57*, 13382–13392. doi:10.1002/anie.201802288
- Weaver, D. S.; Keseler, I. M.; Mackie, A.; Paulsen, I. T.; Karp, P. D. *BMC Syst. Biol.* **2014**, *8*, 79. doi:10.1186/1752-0509-8-79
- Orth, J. D.; Conrad, T. M.; Na, J.; Lerman, J. A.; Nam, H.; Feist, A. M.; Palsson, B. O. *Mol. Syst. Biol.* **2011**, *7*, 535. doi:10.1038/msb.2011.65
- Wink, M. *Phytochemistry* **2003**, *64*, 3–19. doi:10.1016/s0031-9422(03)00300-5
- Woolston, B. M.; Edgar, S.; Stephanopoulos, G. *Annu. Rev. Chem. Biomol. Eng.* **2013**, *4*, 259–288. doi:10.1146/annurev-chembioeng-061312-103312
- Erb, T. J.; Jones, P. R.; Bar-Even, A. *Curr. Opin. Chem. Biol.* **2017**, *37*, 56–62. doi:10.1016/j.cbpa.2016.12.023
- Bilgin, T.; Wagner, A. *PLoS One* **2012**, *7*, e39903. doi:10.1371/journal.pone.0039903
- Schwander, T.; Schada von Borzyskowski, L.; Burgener, S.; Cortina, N. S.; Erb, T. J. *Science* **2016**, *354*, 900–904. doi:10.1126/science.aah5237
- Bar-Even, A.; Noor, E.; Lewis, N. E.; Milo, R. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 8889–8894. doi:10.1073/pnas.0907176107
- Carbonell, P.; Parutto, P.; Baudier, C.; Junot, C.; Faulon, J.-L. *ACS Synth. Biol.* **2014**, *3*, 565–577. doi:10.1021/sb4001273
- Rodrigo, G.; Carrera, J.; Prather, K. J.; Jaramillo, A. *Bioinformatics* **2008**, *24*, 2554–2556. doi:10.1093/bioinformatics/btn471
- Yu, H.; Liao, J. C. *Nat. Commun.* **2018**, *9*, No. 3992. doi:10.1038/s41467-018-06496-4
- Bar-Even, A. *Biochemistry* **2016**, *55*, 3851–3863. doi:10.1021/acs.biochem.6b00495
- Bar-Even, A.; Noor, E.; Flamholz, A.; Milo, R. *Biochim. Biophys. Acta, Bioenerg.* **2013**, *1827*, 1039–1047. doi:10.1016/j.bbabi.2012.10.013
- Yishai, O.; Bouzon, M.; Döring, V.; Bar-Even, A. *ACS Synth. Biol.* **2018**, *7*, 2023–2028. doi:10.1021/acssynbio.8b00131
- Bang, J.; Lee, S. Y. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, E9271–E9279. doi:10.1073/pnas.1810386115
- Tashiro, Y.; Hirano, S.; Matsun, M. M.; Atsumi, S.; Kondo, A. *Metab. Eng.* **2018**, *47*, 211–218. doi:10.1016/j.ymben.2018.03.015
- Siegel, J. B.; Smith, A. L.; Poust, S.; Wargacki, A. J.; Bar-Even, A.; Louw, C.; Shen, B. W.; Eiben, C. B.; Tran, H. M.; Noor, E.; Gallaher, J. L.; Bale, J.; Yoshikuni, Y.; Gelb, M. H.; Keasling, J. D.; Stoddard, B. L.; Lidstrom, M. E.; Baker, D. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 3704–3709. doi:10.1073/pnas.1500545112
- Poust, S.; Piety, J.; Bar-Even, A.; Louw, C.; Baker, D.; Keasling, J. D.; Siegel, J. B. *ChemBioChem* **2015**, *16*, 1950–1954. doi:10.1002/cbic.201500228
- Trudeau, D. L.; Edlich-Muth, C.; Zarzycki, J.; Scheffen, M.; Goldsmith, M.; Khersonsky, O.; Avizemer, Z.; Fleishman, S. J.; Cotton, C. A. R.; Erb, T. J.; Tawfik, D. S.; Bar-Even, A. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, E11455–E11464. doi:10.1073/pnas.1812605115
- Bogorad, I. W.; Lin, T.-S.; Liao, J. C. *Nature* **2013**, *502*, 693–697. doi:10.1038/nature12575
- Erb, T. J.; Zarzycki, J. *Curr. Opin. Chem. Biol.* **2016**, *34*, 72–79. doi:10.1016/j.cbpa.2016.06.026
- Shih, P. M.; Zarzycki, J.; Niyogi, K. K.; Kerfeld, C. A. *J. Biol. Chem.* **2014**, *289*, 9493–9500. doi:10.1074/jbc.c113.543132
- UniProt Consortium. *Nucleic Acids Res.* **2018**, *46*, 2699. doi:10.1093/nar/gky092
- Schomburg, I.; Jeske, L.; Ulbrich, M.; Placzek, S.; Chang, A.; Schomburg, D. *J. Biotechnol.* **2017**, *261*, 194–206. doi:10.1016/j.jbiotec.2017.04.020
- Giger, L.; Caner, S.; Obexer, R.; Kast, P.; Baker, D.; Ban, N.; Hilvert, D. *Nat. Chem. Biol.* **2013**, *9*, 494–498. doi:10.1038/nchembio.1276
- Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Rothlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas, C. F., III; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D. *Science* **2008**, *319*, 1387–1391. doi:10.1126/science.1152692
- Obexer, R.; Godina, A.; Garrabou, X.; Mittl, P. R. E.; Baker, D.; Griffiths, A. D.; Hilvert, D. *Nat. Chem.* **2017**, *9*, 50–56. doi:10.1038/nchem.2596
- Khersonsky, O.; Kiss, G.; Röthlisberger, D.; Dym, O.; Albeck, S.; Houk, K. N.; Baker, D.; Tawfik, D. S. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 10358–10363. doi:10.1073/pnas.1121063109
- Kries, H.; Blomberg, R.; Hilvert, D. *Curr. Opin. Chem. Biol.* **2013**, *17*, 221–228. doi:10.1016/j.cbpa.2013.02.012
- Hammer, S. C.; Kubik, G.; Watkins, E.; Huang, S.; Minges, H.; Arnold, F. H. *Science* **2017**, *358*, 215–218. doi:10.1126/science.aao1482
- Arnold, F. H. *Angew. Chem., Int. Ed.* **2018**, *57*, 4143–4148. doi:10.1002/anie.201708408
- Khersonsky, O.; Malitsky, S.; Rogachev, I.; Tawfik, D. S. *Biochemistry* **2011**, *50*, 2683–2690. doi:10.1021/bi101763c
- Bornscheuer, U. T.; Kazlauskas, R. J. *Angew. Chem., Int. Ed.* **2004**, *43*, 6032–6040. doi:10.1002/anie.200460416
- Burgener, S.; Schwander, T.; Romero, E.; Fraaije, M.; Erb, T. *Molecules* **2017**, *23*, 68. doi:10.3390/molecules23010068
- Schwander, T.; McLean, R.; Zarzycki, J.; Erb, T. *J. Biol. Chem.* **2018**, *293*, 1702–1712. doi:10.1074/jbc.ra117.000764
- Khersonsky, O.; Lipsh, R.; Avizemer, Z.; Ashani, Y.; Goldsmith, M.; Leader, H.; Dym, O.; Rogotner, S.; Trudeau, D. L.; Prilusky, J.; Amengual-Rigo, P.; Guallar, V.; Tawfik, D. S.; Fleishman, S. J. *Mol. Cell* **2018**, *72*, 178–186.e5. doi:10.1016/j.molcel.2018.08.033

43. Zallot, R.; Harrison, K.; Kolaczowski, B.; de Crécy-Lagard, V. *Life* **2016**, *6*, 39. doi:10.3390/life6030039
44. Radivojac, P.; Clark, W. T.; Oron, T. R.; Schnoes, A. M.; Wittkop, T.; Sokolov, A.; Graim, K.; Funk, C.; Verspoor, K.; Ben-Hur, A.; Pandey, G.; Yunes, J. M.; Talwalkar, A. S.; Repo, S.; Souza, M. L.; Piovesan, D.; Casadio, R.; Wang, Z.; Cheng, J.; Fang, H.; Gough, J.; Koskinen, P.; Törönen, P.; Nokso-Koivisto, J.; Holm, L.; Cozzetto, D.; Buchan, D. W. A.; Bryson, K.; Jones, D. T.; Limaye, B.; Inamdar, H.; Datta, A.; Manjari, S. K.; Joshi, R.; Chitale, M.; Kihara, D.; Lisewski, A. M.; Erdin, S.; Venner, E.; Lichtarge, O.; Rentzsch, R.; Yang, H.; Romero, A. E.; Bhat, P.; Paccanaro, A.; Hamp, T.; Kašner, R.; Seemayer, S.; Vicedo, E.; Schaefer, C.; Achten, D.; Auer, F.; Boehm, A.; Braun, T.; Hecht, M.; Heron, M.; Hönigschmid, P.; Hopf, T. A.; Kaufmann, S.; Kiening, M.; Krompass, D.; Landerer, C.; Mahlich, Y.; Roos, M.; Björne, J.; Salakoski, T.; Wong, A.; Shatkay, H.; Gatzmann, F.; Sommer, I.; Wass, M. N.; Sternberg, M. J. E.; Škunca, N.; Supek, F.; Bošnjak, M.; Panov, P.; Džeroski, S.; Šmuc, T.; Kourmpetis, Y. A. I.; van Dijk, A. D. J.; ter Braak, C. J. F.; Zhou, Y.; Gong, Q.; Dong, X.; Tian, W.; Falda, M.; Fontana, P.; Lavezzo, E.; Di Camillo, B.; Toppo, S.; Lan, L.; Djuric, N.; Guo, Y.; Vucetic, S.; Bairoch, A.; Linial, M.; Babbitt, P. C.; Brenner, S. E.; Orengo, C.; Rost, B.; Mooney, S. D.; Friedberg, I. *Nat. Methods* **2013**, *10*, 221–227. doi:10.1038/nmeth.2340
45. Schnoes, A. M.; Brown, S. D.; Dodevski, I.; Babbitt, P. C. *PLoS Comput. Biol.* **2009**, *5*, e1000605. doi:10.1371/journal.pcbi.1000605
46. Erb, T. J.; Berg, I. A.; Brecht, V.; Muller, M.; Fuchs, G.; Alber, B. E. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 10631–10636. doi:10.1073/pnas.0702791104
47. Erb, T. J.; Brecht, V.; Fuchs, G.; Muller, M.; Alber, B. E. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 8871–8876. doi:10.1073/pnas.0903939106
48. Hanson, T. E.; Tabita, F. R. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 4397–4402. doi:10.1073/pnas.081610398
49. Ashida, H.; Saito, Y.; Kojima, C.; Kobayashi, K.; Ogasawara, N.; Yokota, A. *Science* **2003**, *302*, 286–290. doi:10.1126/science.1086997
50. Erb, T. J.; Evans, B. S.; Cho, K.; Warlick, B. P.; Sriram, J.; Wood, B. M.; Imker, H. J.; Sweedler, J. V.; Tabita, F. R.; Gerlt, J. A. *Nat. Chem. Biol.* **2012**, *8*, 926–932. doi:10.1038/nchembio.1087
51. Carter, M. S.; Zhang, X.; Huang, H.; Bouvier, J. T.; Francisco, B. S.; Vetting, M. W.; Al-Obaidi, N.; Bonanno, J. B.; Ghosh, A.; Zallot, R. G.; Andersen, H. M.; Almo, S. C.; Gerlt, J. A. *Nat. Chem. Biol.* **2018**, *14*, 696–705. doi:10.1038/s41589-018-0067-7
52. Erb, T. J.; Zarzycki, J. *Curr. Opin. Biotechnol.* **2018**, *49*, 100–107. doi:10.1016/j.copbio.2017.07.017
53. Zallot, R.; Oberg, N. O.; Gerlt, J. A. *Curr. Opin. Chem. Biol.* **2018**, *47*, 77–85. doi:10.1016/j.cbpa.2018.09.009
54. Atkinson, H. J.; Morris, J. H.; Ferrin, T. E.; Babbitt, P. C. *PLoS One* **2009**, *4*, e4345. doi:10.1371/journal.pone.0004345
55. Gerlt, J. A.; Bouvier, J. T.; Davidson, D. B.; Imker, H. J.; Sadkhin, B.; Slater, D. R.; Whalen, K. L. *Biochim. Biophys. Acta, Proteins Proteomics* **2015**, *1854*, 1019–1037. doi:10.1016/j.bbapap.2015.04.015
56. Vögeli, B.; Erb, T. J. *Curr. Opin. Chem. Biol.* **2018**, *47*, 94–100. doi:10.1016/j.cbpa.2018.09.013
57. Barelier, S.; Cummings, J. A.; Rauwerdink, A. M.; Hitchcock, D. S.; Farelli, J. D.; Almo, S. C.; Raushel, F. M.; Allen, K. N.; Shoichet, B. K. *J. Am. Chem. Soc.* **2014**, *136*, 7374–7382. doi:10.1021/ja501354q
58. Sun, J.; Jeffries, J. G.; Henry, C. S.; Bruner, S. D.; Hanson, A. D. *Metab. Eng.* **2017**, *44*, 150–159. doi:10.1016/j.ymben.2017.10.006
59. Erb, T. J.; Frerichs-Revermann, L.; Fuchs, G.; Alber, B. E. *J. Bacteriol.* **2010**, *192*, 1249–1258. doi:10.1128/jb.01267-09
60. Ellens, K. W.; Christian, N.; Singh, C.; Satagopam, V. P.; May, P.; Linster, C. L. *Nucleic Acids Res.* **2017**, *45*, 11495–11514. doi:10.1093/nar/gkx937
61. Donati, S.; Sander, T.; Link, H. *Wiley Interdiscip. Rev.: Syst. Biol. Med.* **2018**, *10*, e1396. doi:10.1002/wsbm.1396
62. Shetty, R. P.; Endy, D.; Knight, T. F., Jr. *J. Biol. Eng.* **2008**, *2*, 5. doi:10.1186/1754-1611-2-5
63. Gardossi, L.; Poulsen, P. B.; Ballesteros, A.; Hult, K.; Švedas, V. K.; Vasić-Rački, Đ.; Carrea, G.; Magnusson, A.; Schmid, A.; Wohlgemuth, R.; Halling, P. J. *Trends Biotechnol.* **2010**, *28*, 171–180. doi:10.1016/j.tibtech.2010.01.001
64. Swainston, N.; Baici, A.; Bakker, B. M.; Cornish-Bowden, A.; Fitzpatrick, P. F.; Halling, P.; Leyh, T. S.; O'Donovan, C.; Raushel, F. M.; Reschel, U.; Rohwer, J. M.; Schnell, S.; Schomburg, D.; Tipton, K. F.; Tsai, M.-D.; Westerhoff, H. V.; Wittig, U.; Wohlgemuth, R.; Kettner, C. *FEBS J.* **2018**, *285*, 2193–2204. doi:10.1111/febs.14427

## License and Terms

This is an Open Access article under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>). Please note that the reuse, redistribution and reproduction in particular requires that the authors and source are credited.

The license is subject to the *Beilstein Journal of Organic Chemistry* terms and conditions: (<https://www.beilstein-journals.org/bjoc>)

The definitive version of this article is the electronic one which can be found at: [doi:10.3762/bjoc.15.49](https://doi.org/10.3762/bjoc.15.49)