Synthetic biology: lessons from the history of synthetic organic chemistry

Brian J Yeh & Wendell A Lim

The mid-nineteenth century saw the development of a radical new direction in chemistry: instead of simply analyzing existing molecules, chemists began to synthesize them—including molecules that did not exist in nature. The combination of this new synthetic approach with more traditional analytical approaches revolutionized chemistry, leading to a deep understanding of the fundamental principles of chemical structure and reactivity and to the emergence of the modern pharmaceutical and chemical industries. The history of synthetic chemistry offers a possible roadmap for the development and impact of synthetic biology, a nascent field in which the goal is to build novel biological systems.

In 1828, the German chemist Friedrich Wöhler could hardly contain his excitement as he wrote to his former mentor, Jöns Jakob Berzelius, of a new finding[1,2]: "I must tell you that I can prepare urea without requiring a kidney of an animal, either man or dog." At the beginning of the nineteenth century, the synthesis of this small organic molecule was earth-shattering news. At that time, chemists believed there was a clear distinction between molecules from living beings (referred to as 'organic') and those from nonliving origin ('inorganic'). It was known that organic substances could be easily converted to inorganic compounds through heating or other treatments; however, chemists could not perform the reverse transformation. Surely, a 'vital force' present only in living organisms was required to convert the inorganic into organic. Wöhler's discovery that ammonium cyanate could be converted to urea in the laboratory was a key nail in the coffin of vitalism, and in the next few decades, chemists began to synthesize hundreds of other organic molecules. In a particularly interesting example in 1854, the French chemist Marcellin Berthelot synthesized the fat molecule tristearin from glycerol and stearic acid, a common naturally occurring fatty acid. Taking this a step further, he realized that he could replace stearic acid with similar acids not found in natural fats, thus generating non-natural molecules that had properties similar to those of natural fats[3]. These and other early syntheses demonstrated that chemists could indeed make 'living' molecules as well as new compounds that went beyond those that naturally occurred, thus giving birth to synthetic organic chemistry. It was unclear where this field would lead, and many feared these advances could lead to goals such as the creation of living beings. Today, however, nearly all aspects of our lives are touched by synthetic molecules that mankind has learned to make.

J. Iwasa

Synthetic approaches may transform biology just as they transformed chemistry.

Brian J. Yeh is in the Chemistry and Chemical Biology Graduate Program, University of California, San Francisco, 600–16th Street, San Francisco, California 94158-2517, USA. Wendell A. Lim is in the Department of Cellular and Molecular Pharmacology, University of California, San Francisco, 600–16th Street, San Francisco, California 94158-2517, USA, and is a member of the Cell Propulsion Lab, USF/UCB National Institutes of Health Nanomedicine Development Center, and the US National Science Foundation Synthetic Biology Engineering Research Center.
e-mail: lim@cmp.ucsf.edu

Advances in our ability to build and modify ‘organic’ molecules on increasingly larger scales have continued to push the frontier of our understanding of the physical principles underlying living systems. For example, chemical synthesis of DNA oligonucleotides (first performed by Gobind Khorana) led directly to the elucidation of the genetic code[4]. Bruce Merrifield’s complete synthesis of RNase A demonstrated that chemical structure (primary sequence) is sufficient to confer tertiary structure and the seemingly magical activity of enzymes[5]. More recently, complete chemical synthesis of poliovirus complementary DNA was a vivid demonstration that genetic instructions are sufficient to specify an active biological system[6].

Over the last several years, this line of research has culminated in the emergence of a field known as ‘synthetic biology.’ Whether synthetic biology represents a truly new field is open to debate, but the boldness of the stated goals—to learn how to precisely and reliably engineer and build self-organizing systems that both recapitulate biological function and show new functions—is unquestionably novel. These goals hold promise for harnessing the efficiency and precision of living systems for diverse purposes: microbial factories that manufacture drugs, materials or biofuels[7]; cells that seek and destroy tumors[8]; cells that can carry out rapid tissue repair and regeneration; cells that can direct the assembly of nanomaterials; even living systems that can compute. Synthetic biology, however, is more than a broad set of visionary applications. Much effort is going into developing a common toolkit of well-defined biological parts and devices as well as strategies to link them together into predictable systems[9–12]. These foundational efforts are aimed at one day making engineering cells as reliable and predictable as engineering an electronic device. Indeed, synthetic biology can be viewed as a subdiscipline of bioengineering that is focused on engineering self-organizing cellular devices (as opposed to instruments that interrogate living systems or materials that interface with organisms).

At this threshold, where our view of biology as something to be observed is transitioning into a view of biology as something that can be engineered, there are many important questions. Why even attempt synthetic biology when our understanding of biological systems is still incomplete? And should we choose to proceed, how should we go about it? Many reviews have compared synthetic biology to electrical engineering, noting that cellular networks, like electronic circuits, are built in a hierarchical fashion from modular components that perform logical computations[9,11–13]. Although this comparison is useful, in some respects a comparison with the development of synthetic organic chemistry may be more appropriate[10]. In this Commentary, we consider the historical role of synthetic approaches in the development of modern organic chemistry in order to extract some lessons that might help guide the development of synthetic biology.

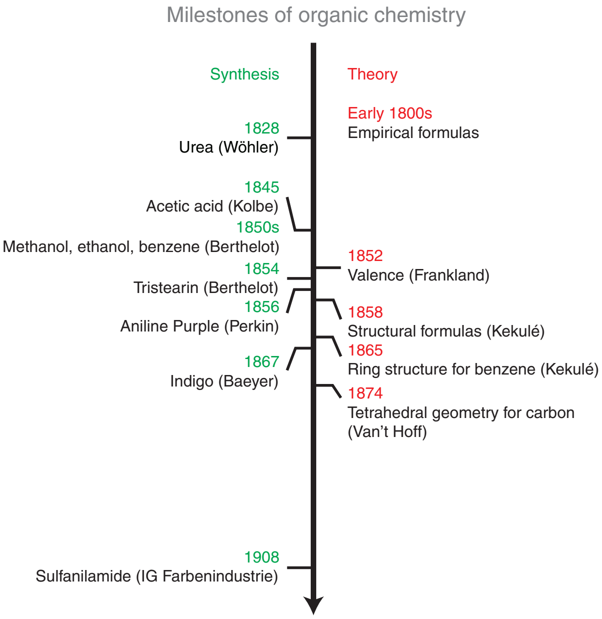

Figure 1 Chemical synthesis and theories of structure emerged concurrently. Significant milestones in chemical synthesis (left of timeline, dates shown in green) and theories of chemical structure (right of timeline, dates shown in red) are shown.

Synthesis and analysis transform chemistry

Before the time of Wöhler and Berthelot, the understanding of even simple molecules was as naïve as our current understanding of complex biological systems. How the composition of organic compounds determined their properties and reactivity was unknown, and the concept of molecules having defined structures was still undeveloped. Ultimately, analytical and synthetic approaches synergized to produce an explosive growth in our knowledge of chemical principles, and fundamental theories of chemical structure developed concurrently with the explosion of synthesis (**Fig. 1**).

A critical early advance at the beginning of the nineteenth century was precise measurement and analysis of compounds as they reacted[3]. For example, by collecting and precisely weighing the carbon dioxide and water that formed upon combustion of organic molecules, it became possible to determine the atomic compositions of these molecules, and therefore their empirical formulas. This careful analysis led to the discovery of isomers—the shocking finding that compounds with very different physical and chemical properties can have identical empirical formulas. Clearly, a more sophisticated understanding of chemical structure would be required.

The nascent field of synthetic chemistry, with its many new reactions and molecules, provided a complementary body of information that contributed to the development of modern theories of chemical structure and reactivity. It was not until after the mid-nineteenth century—after synthesis of small compounds was already becoming routine—that chemists began to develop models to explain bonding between atoms[3]. In 1852, Edward Frankland proposed that each atom had an ability to combine with a fixed number of other atoms. Kekulé used this ‘theory of valence’ to propose structures for many simple organic molecules in 1858, and in the 1860s, Alexander Bulterov pointed out that these structural formulas could explain the majority of isomers. This notion that atoms were held in fixed arrangements was a critical advance. In 1865, Kekulé proposed the structural formula for benzene. Although it had already been synthesized by Berthelot in the 1850s, benzene’s stability could not be adequately explained until Kekulé conceived the ring structure, thereby cementing the usefulness of this paradigm. Later, these structures were extended to three dimensions based on the notion that carbon bonds are arranged in a tetrahedral geometry, an idea proposed by Jacobus Van't Hoff in 1874.

A critical lesson here is that a complete understanding of chemical principles was not a prerequisite for the emergence of synthetic chemistry. Rather, synthetic and analytical approaches developed in parallel and synergized to shape our modern understanding of chemistry. The synthetic approach—empirically learning how to systematically manipulate and perturb molecules—directly contributed to developing theories of chemical structure and reactivity (**Fig. 2a**). At the simplest level, synthesizing a molecule was often the ultimate proof of the proposed structure. More significantly, the ability to synthesize variants of known molecules allowed a rigorous and systematic exploration of the principles underlying chemical structure and reactivity, thus beginning the field of physical organic chemistry.

Biology: synthesis redux

The convergence of analytical and synthetic approaches seems to be replaying itself in modern biology (**Fig. 2b**). Biology has historically been a field based almost entirely on observation and analysis, and it is currently undergoing exponential growth in the accuracy and throughput of measurement. Modern experimental techniques—genome sequencing, DNA microarrays, molecular structure determination, and high-throughput microscopy coupled with *in vivo* biosensors—represent major analytical advances, giving us extensive parts lists and descriptions of biological systems and their behaviors, including the abundance, subcellular localization and interactions of entire proteomes. These developments are akin to the advances in analytical chemistry of the early nineteenth century. However, the history of organic chemistry suggests that synthesis will be a necessary complement to analysis in order for biologists to truly understand the mechanisms of complex living systems.

In many ways, synthetic biology can be viewed as *in vivo* reconstitution—an intellectual descendant of simple biochemistry. Reconstitution methods (which essentially apply the synthetic philosophy to noncovalent systems) allow us to determine not only what is necessary, but also what is sufficient to build a system that carries out a particular function. The ability to build and systematically modify biological systems will allow one to explore their design principles in much deeper ways. Thus, synthetic biological systems will allow experimentation that is not possible with extant living systems, which are encumbered by eccentric evolutionary histories and constraints. What are the minimal systems that can perform a behavior? How does that behavior precisely vary as individual components, the linkages that connect them into larger networks, and specific biochemical parameters are altered? For example, Suel *et al.* have studied a simple genetic circuit in *Bacillus subtilis* that underlies differentiation, both by quantitatively characterizing the existing circuit and by making new connections and analyzing the changes that result[14,15].

What can we expect to learn? Clearly there will be more than a few simple rules explaining how living systems work. As in physical organic chemistry, however, it is likely that unifying patterns will emerge, and an understanding of the logic underlying biological systems will develop. It may even be possible to construct something analogous to the periodic table for biology that facilitates the systematic understanding of network 'elements' and their properties[16,17].

Diverse and unexpected driving applications

If one important goal of synthetic biology is to create useful systems, then what applications should we be targeting? Again, it is useful to consider the early applications of synthetic organic chemistry. In today's world, many tend to link synthetic chemistry with the production of drugs. Indeed, it was abundantly clear to early chemists that synthetic products could improve human health, but their initial efforts actually led to an industrial explosion in an unexpected direction. August von Hofmann and his student William Perkin postulated that it might be possible to synthesize the highly valuable antimalarial agent quinine from aniline, a cheap product of coal tar[3]. However, in attempting to synthesize quinine from aniline in 1856, Perkin unexpectedly produced a brilliant purple compound—a dye. He soon

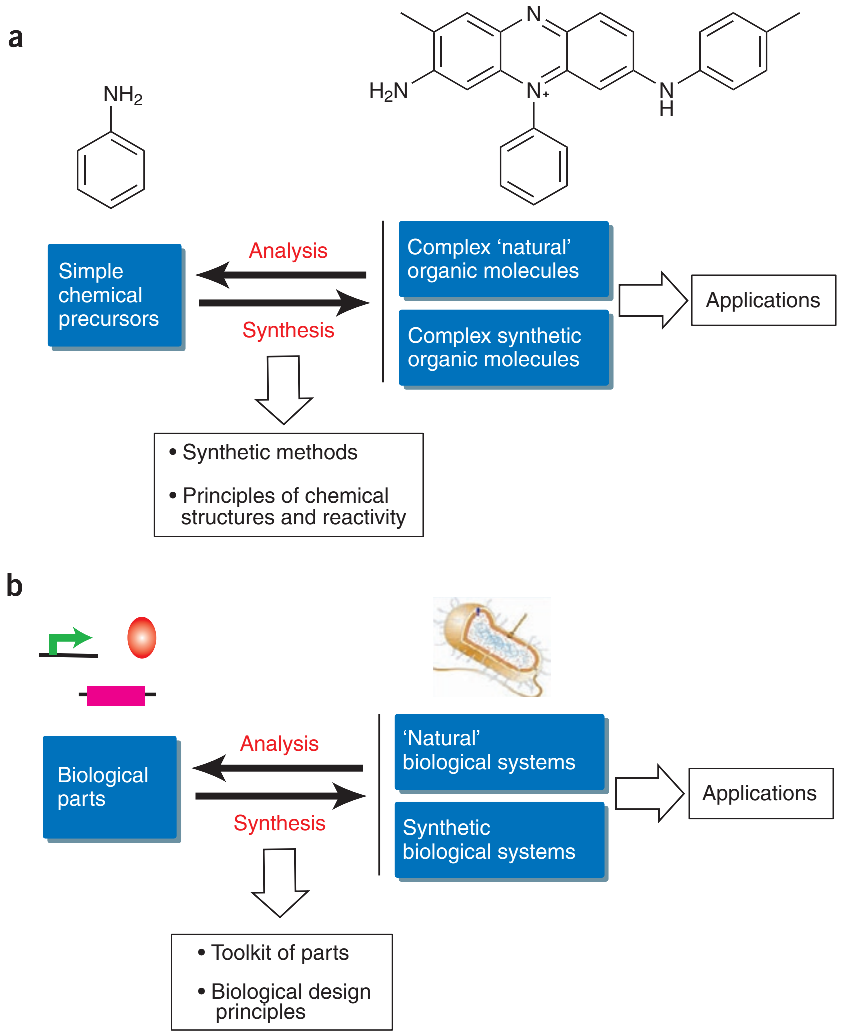

Figure 2 Synthesis and analysis are complementary. (**a**) In organic chemistry, analysis and synthesis were both critical in determining fundamental principles of chemical structure and reactivity. Synthetic molecules have been used for a wide variety of applications. (**b**) Similarly, synthetic approaches will complement analytical methods in elucidating biological principles, and synthetic cellular systems will prove highly useful.

opened a factory to synthesize this molecule, which he called "Aniline Purple," and thus founded the synthetic dye industry. Along with other examples, such as the synthesis of indigo by Adolf von Baeyer in 1867, these advances led to the explosive growth of the German and Swiss dye industry, while simultaneously dismantling the import of indigo and other natural dyes from distant tropical locales. In fact, synthetic indigo remains an important commercial product and is used in today's blue jeans (Levi's was founded in 1873).

Although dyes were the earliest economically important synthetic compounds, the development of the European dye industry would ultimately lead to successes in chemotherapy. Indeed, nearly all of the modern big pharmaceutical companies are in part descended from German or Swiss dye manufacturers. For example, the first effective antibacterials were the sulfa drugs[3,18]. The first of these molecules, sulfanilamide, was synthesized by IG Farbenindustrie in 1908 because of its potential as a dye. In 1932, Gerhard Domagk discovered that sulfanilamide and related compounds have bactericidal activity, and he was aided in his studies by chemists that could make a variety of related compounds.

The lesson here is very clear: synthetic biologists (and their funding agencies) must move forward with an open mind. The progress of synthetic biology cannot be myopically linked to only a few obvious targets; instead, we must be prepared for a variety of potential industrial and therapeutic applications, including unexpected ones that we have not yet foreseen. Most of the successes in synthetic biology so far have been in so-called toy systems; however, we should not underestimate the importance of these achievements in laying a foundation of parts and construction methods that will be useful for diverse applications of synthetic biology. For example, in synthetic chemistry, developing generic protecting-group strategies or classes of reactions to make carbon-carbon bonds allowed the synthesis of diverse organic products with a wide array of applications. Similarly, learning how to link a set of molecules into a generic type of regulatory circuit module, such as an ultrasensitive positive feedback loop, could be useful for a diverse range of synthetic biology applications[9,11,19,20]. In the long term, synthetic biology is likely to have as broad and pervasive a role in our society as the products of both the chemical and electronic revolutions.

Goal-oriented biology

Focusing on specific goals and applications in the context of biological systems is a feature that clearly distinguishes synthetic biology from the traditional discipline of biology. This philosophy may seem strange to many academic biologists, but it is actually quite appropriate. A defining feature of evolution is the constant selection of organisms that achieve the best performance (fitness) in a particular niche. Understanding how to build biological systems to achieve well-defined performance specifications will force us to understand biology at a far more quantitative level. For example, efforts to engineer *Escherichia coli* to synthesize isoprenoids have shown that it is not sufficient to simply express the right collection of enzymes to create a new pathway. Rather, accumulation of metabolic intermediates can greatly limit cell growth and overall flux through the pathway[21]. Thus, precisely balancing intermediate synthetic steps is critical for maximizing yield of the final product.

Synthetic biology will also require sociological reorganization of how biologists work on problems. Achieving success will require close integration of interdisciplinary teams of scientists focused on clear, practical goals. Many synthetic biology centers organized around target goals have emerged over the last few years, such as the US National Science Foundation Synthetic Biology Engineering Research Center, the US National Institutes of Health Nanomedicine Development Centers, and the US Department of Energy Joint Bioenergy Institute.

The present and future of synthetic biology

Just as in the early days of synthetic chemistry, it will be necessary to develop both an intellectual framework and diverse tools for synthetic biology; engineering a cell with specific quantitative specifications will be challenging and will require development on many fronts. First, we must understand the different languages or currencies of biology. Early progress in synthetic biology was limited to transcriptional networks[12,20], but more recently, attention has turned to engineering networks based on phosphorylation[22], GTPases[23] and RNA interference[24]. Second, we must maximize the efficiency of the cellular "chasses" that will perform the desired functions; therefore there are efforts to construct cells that only have the minimal requirements for replication[25]. In addition, researchers are engineering orthogonal components that only interact with each other, and not existing cellular molecules. These components, which include novel ribosome-mRNA pairs[26] and protein-protein interactions[27], should lead to more reliable and predictable behavior when used in cellular engineering. Ultimately, engineering cells may require the ability to rewrite entire genomes. By successfully replacing the genome of one bacterial cell with that of another species, Lartigue *et al.* have shown that this will indeed be possible[28].

It is important to remember that early synthetic chemists did not always know what to expect in their reactions, only that something interesting could happen[18]. Most importantly, they were prepared to follow up these experiments to understand what did happen. Similarly, in these early days of synthetic biology, it will be very difficult to predict the behavior of novel biological systems. Therefore, directed evolution and combinatorial methods will be useful[29]; analyzing libraries of synthetic circuits and systems will maximize the probability of obtaining the targeted biological behavior. Furthermore, systematically varying many parameters will produce structure-activity relationships at the biological network level that will improve future designs.

Developing educational initiatives will continue to be an important emphasis in synthetic biology. Perkin was a teenager when he initially synthesized Aniline Purple—during Easter vacation in a small laboratory in his home[3]. Similarly, it is young scientists who are likely to truly view biology from this new perspective, and who will shape the yet unforeseen 'killer applications' of synthetic biology. The innovative spirit of Perkin lives on today in the growing number of undergraduates and high school students participating in the International Genetically Engineered Machine (iGEM) competition. Every summer, teams of these students compete to design and build new synthetic biological systems. In 2007, 56 teams from North America, Europe and Asia have registered for the fourth year of the competition (http://www.igem2007.com).

Like all technologies, synthetic organic chemistry also introduced its own set of problems. In addition to the plethora of beneficial drugs and polymers that have significantly increased our standard of living, harmful or 'dual-use' compounds, including explosives and chemical weapons, have also been created. There is no question that synthetic biological systems will also bring a mixed array of potential applications. Synthetic biologists are not ignoring this possibility; discussions of biosafety and security have been a major component of synthetic biology conferences (http://syntheticbiology.org/SB2.0/Biosecurity_resolutions.html). Initial attention has focused on preventing commercial DNA synthesis companies from supplying pathogenic or otherwise dangerous sequences[30].

Conclusions

In the coming decades, we are likely to see a revolution in biology akin to the revolution in chemistry that occurred in the latter half of the nineteenth century. The development of increasingly sophisticated methods to alter and build biological systems will provide essential synthetic tools that will synergize with analytical methods, which together will ultimately lead to a far deeper understanding of the physical principles underlying the behavior and design of cellular systems. The applications of synthetic biology will be highly varied, and progress and innovation is likely to come from unexpected areas. We might also expect that understanding the engineering principles of biological systems will have a transformative effect on other fields of science. For example, man-made materials, even at the nanoscale, are currently templated or built using directed assembly—exactly the opposite of how biological molecules create structure and function. Biological molecules are self-assembling systems that can adapt to change, show robust homeostasis, and can self-repair. There may come a day when man-made materials also have these properties. Universities, industry, governmental agencies and scientists will have to work together in an open-minded and responsible way to foster productive growth of this exciting field.

ACKNOWLEDGMENTS

We thank A. Chau, M. Cohen, N. Helman, B. Rhau, J. Taunton and A. Watters for their comments on this Commentary. This work was supported by the US National Institutes of Health Roadmap, the US National Science Foundation, the Packard Foundation and the Rogers Family Foundation.

COMPETING INTERESTS STATEMENT

The authors declare no competing financial interests.

1. Jaffe, B. *Crucibles: The Story of Chemistry from Ancient Alchemy to Nuclear Fission* (Dover Publications, New York, 1976).
2. Wöhler, F. *Poggendorff's Ann. Phys.* **12**, 253–256 (1828).
3. Asimov, I. *A Short History of Chemistry* (Anchor Books, Garden City, New York, USA, 1965).
4. Khorana, H.G. *Fed. Proc.* **24**, 1473–1487 (1965).
5. Gutte, B. & Merrifield, R.B. *J. Biol. Chem.* **246**, 1922–1941 (1971).
6. Cello, J., Paul, A.V. & Wimmer, E. *Science* **297**, 1016–1018 (2002).
7. Khosla, C. & Keasling, J.D. *Nat. Rev. Drug Discov.* **2**, 1019–1025 (2003).
8. Anderson, J.C., Clarke, E.J., Arkin, A.P. & Voigt, C.A. *J. Mol. Biol.* **355**, 619–627 (2006).
9. Andrianantoandro, E., Basu, S., Karig, D.K. & Weiss, R. *Mol. Syst. Biol.* **2**, 2006.0028 (2006).
10. Benner, S.A. & Sismour, A.M. *Nat. Rev. Genet.* **6**, 533–543 (2005).
11. Endy, D. *Nature* **438**, 449–453 (2005).
12. Voigt, C.A. *Curr. Opin. Biotechnol.* **17**, 548–557 (2006).
13. Arkin, A.P. & Fletcher, D.A. *Genome Biol.* **7**, 114 (2006).
14. Suel, G.M., Garcia-Ojalvo, J., Liberman, L.M. & Elowitz, M.B. *Nature* **440**, 545–550 (2006).
15. Suel, G.M., Kulkarni, R.P., Dworkin, J., Garcia-Ojalvo, J. & Elowitz, M.B. *Science* **315**, 1716–1719 (2007).
16. Alon, U. *Nat. Rev. Genet.* **8**, 450–461 (2007).
17. Alon, U. *Nature* **446**, 497 (2007).
18. Hoffmann, R. *The Same and Not the Same* (Columbia University Press, New York, 1995).
19. Becskei, A., Seraphin, B. & Serrano, L. *EMBO J.* **20**, 2528–2535 (2001).
20. Kaern, M., Blake, W.J. & Collins, J.J. *Annu. Rev. Biomed. Eng.* **5**, 179–206 (2003).
21. Pitera, D.J., Paddon, C.J., Newman, J.D. & Keasling, J.D. *Metab. Eng.* **9**, 193–207 (2007).
22. Park, S.H., Zarrinpar, A. & Lim, W.A. *Science* **299**, 1061–1064 (2003).
23. Yeh, B.J., Rutigliano, R.J., Deb, A., Bar-Sagi, D. & Lim, W.A. *Nature* **447**, 596–600 (2007).
24. Rinaudo, K. *et al. Nat. Biotechnol.* **25**, 795–801 (2007).
25. Forster, A.C. & Church, G.M. *Mol. Syst. Biol.* **2**, 45 (2006).
26. Rackham, O. & Chin, J.W. *Biochem. Soc. Trans.* **34**, 328–329 (2006).
27. Kortemme, T. & Baker, D. *Curr. Opin. Chem. Biol.* **8**, 91–97 (2004).
28. Lartigue, C. *et al. Science* (in the press).
29. Haseltine, E.L. & Arnold, F.H. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 1–19 (2007).
30. Bugl, H. *et al. Nat. Biotechnol.* **25**, 627–629 (2007).