Bioengineering novel *in vitro* metabolic pathways using synthetic biology

Andreas Meyer, Rene Pellaux and Sven Panke

Huge numbers of enzymes have evolved in nature to function in aqueous environments at moderate temperatures and neutral pH. This gives us, in principle, the unique opportunity to construct multistep reaction systems of considerable catalytic complexity *in vitro*. However, this opportunity is rarely exploited beyond research scale, because such systems are difficult to assemble and to operate productively. Recent advances in DNA synthesis, genome engineering, high-throughput analytics, model-based analysis of biochemical systems and (semi-)rational protein engineering suggest that we have all the tools available to rationally design and efficiently operate such systems of enzymes, and finally harvest their potential for preparative syntheses.

Addresses
Bioprocess Laboratory, ETH Zurich, Universitaetsstrasse 6, 8092 Zurich, Switzerland

Corresponding author: Panke, Sven (panke@ipe.mavt.ethz.ch)

Current Opinion in Microbiology 2007, **10**:246–253

This review comes from a themed issue on
Ecology and industrial microbiology
Edited by Victor de Lorenzo and Margaret J McFall-Ngai

Available online 4th June 2007

1369-5274/$ – see front matter

DOI 10.1016/j.mib.2007.05.009

Introduction

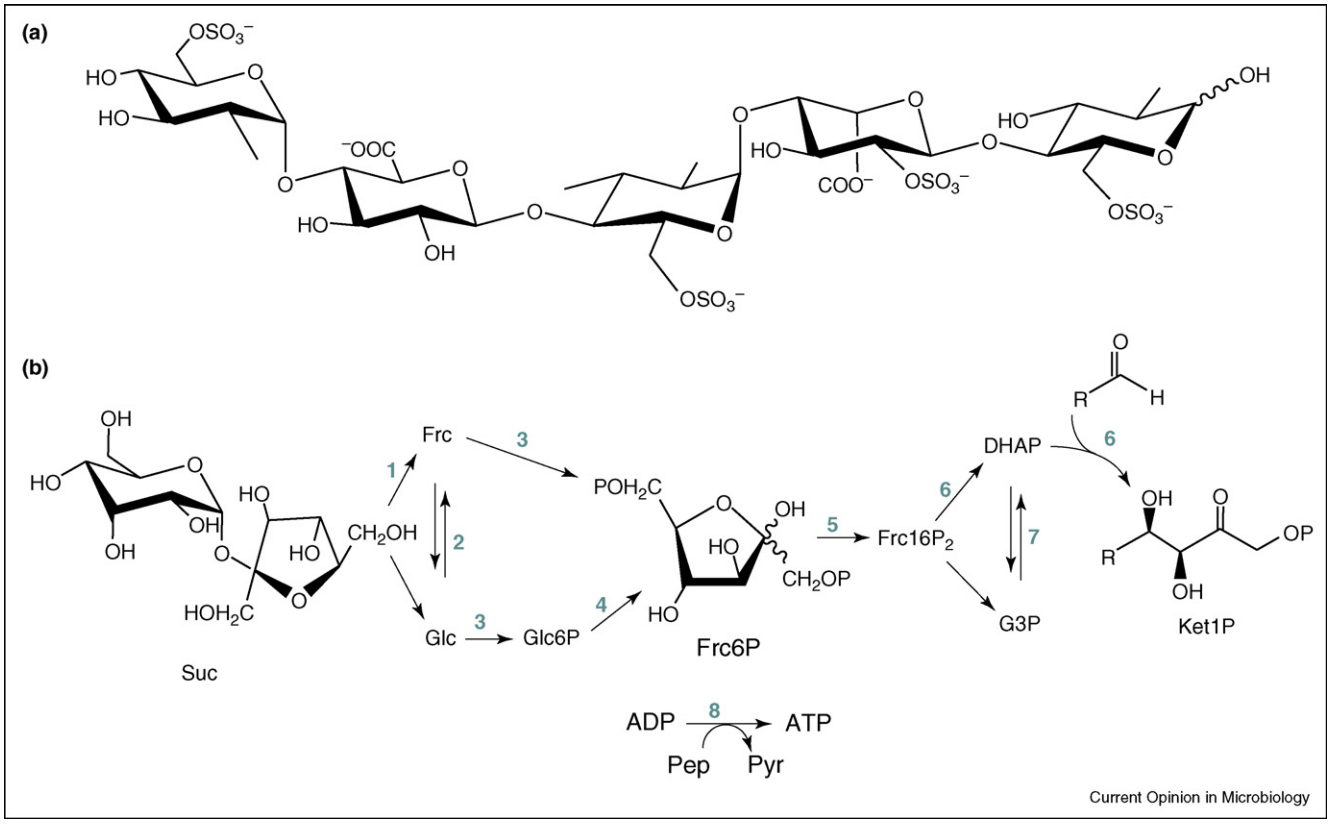
Biotransformations have become an indispensable tool in the fine-chemical industry for catalyzing chemoselectively, regioselectively and enantioselectively challenging reactions [1], where they are typically used as one in a series of chemical transformations, the different steps of which are frequently characterized by drastic changes in the reaction milieu (to temperature, pH and solvent). However, enzymes have a unique advantage in that most of them have been selected to operate under ambient conditions in aqueous environments. This allows, at least in principle, for the simultaneous operation of many different enzymes, thus addressing a much larger catalytic complexity than can be achieved with the typical one-enzyme or two-enzyme biotransformations. Of course, this has been already widely exploited in the manufacturing of chemicals or pharmaceuticals by fermentation, and several excellent examples are available [2,3]. Nevertheless, a large number of pharmaceutically very interesting molecules remain difficult to produce by fermentation, as they cannot cross the membrane, are highly toxic, have a low final concentration and/or yield, or are co-synthesized with many by-products. Prime examples of compounds that require multistep chemical syntheses and that are produced by biotransformations rather than by fermentation come from the group of monosaccharides and oligosaccharides, where multi-enzyme systems have been employed on numerous occasions [4]. Several drugs and vaccines based on saccharide building-blocks are already available (Figure 1a) and more are to follow [5,6], but, in spite of the many elegant enzyme-based systems, chemical routes still prevail; however, because of the polyfunctional nature of these molecules, chemical routes are complex and rarely sufficiently economic.

Today, there are several compelling reasons for the absence of industrial multistep enzyme reactions: first, the assembly of a system of enzymes economically is difficult; second, due to the interactions of substrates, intermediates or products with the enzymes, the dynamic behavior of such a system is complex, frequently not well understood and typically oriented towards cellular homeostasis (maintaining a constant flux despite perturbations) rather than towards unnatural requirements of process productivity; and third, the measures to improve on point two are experimentally very intensive. Nevertheless, such multi-enzyme systems are after all limited in size and complexity (in particular when compared to whole-cell systems), our ability to collect systems of enzymes is improving, the scope and potential of analytical methods are rapidly increasing (driven by advances in metabolomics), and so is our ability to at least semi-rationally design enzyme functionalities. Consequently, we argue that we have all the ingredients that are required for a rational system design or, in other words, to create multi-enzyme systems with synthetic biology (Figure 2) [7••,8].

Assembly of multi-enzyme systems

In a landmark paper for illustrating the potential of multi-enzyme catalysis, Fessner and Walter [9] combined eight enzymes in a one-pot reaction to convert (cheap) sucrose into one of several possible phosphorylated ketoses (Figure 1b). Seven of those enzymes were obtained separately from commercial suppliers, making this set-up unsustainable beyond research-scale. More recently, several highly engineered strains have become available that overexpress several recombinant genes during growth and are then permeabilized or homogenized and used for multi-enzyme catalysis [10,11]. Obviously, the entire cellular background is still present in these

Figure 1



Complex molecules and pathways as targets for synthetic biology. **(a)** Arixtra, an antithrombotic compound currently marketed by GlaxoSmithKline, as an example of an oligosaccharide drug. The compound is currently manufactured chemically in a synthesis consisting of approximately 50 steps. **(b)** A premier example of the synthetic power of systems of enzymes (adapted from [9]): the formation of ketose-1-phosphate compounds from sucrose in an eight-enzyme system. Enzymes: **1**, invertase; **2**, xylose isomerase; **3**, hexokinase; **4**, Glc6P-isomerase; **5**, Frc6P-kinase; **6**, $Frc16P_2$ aldolase; **7**, triose phosphate isomerase; **8**, pyruvate kinase. Enzymes **3** and **5** consume one ATP per conversion, which is regenerated by **8** and the conversion on Pep. Abbreviations: $Frc16P_2$, D-fructose-1,6-bisphosphate; Frc6P, D-fructose-6-phosphate; Frc, D-fructose; Glc, D-glucose; Glc6P, D-glucose-6-phosphate; G3P, D-glyceraldehyde-3-phosphate; Ket1P, ketose-1-phosphate; Pep, phosphoenolpyruvate; Pyr, pyruvate; Suc, D-sucrose.

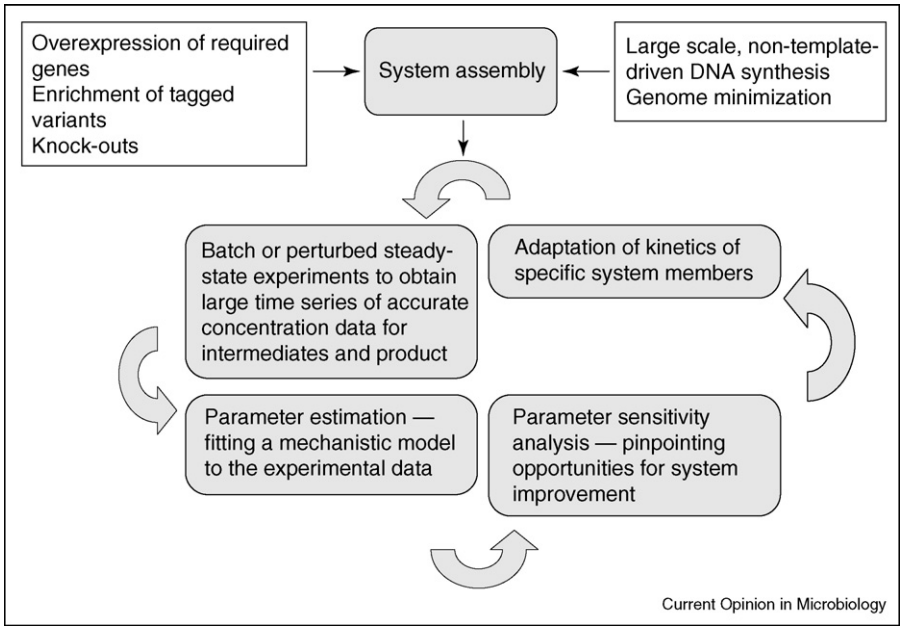
settings and can potentially interfere with the yield of the reaction. However, a certain 'insulation' of the enzymatic network is achieved through the overexpression of the target genes relative to the background. In many cases, given the current comprehensive genome-scale knowledge on model organisms, it might be possible to eliminate genes with interfering gene products from the host organism [12•]. Alternatively, multiple genes have been manipulated to provide the entire desired set of enzymes with one common tag-sequence, so that with only one step, the enzyme set can be enriched from those products with disturbing activities [13]. Further extensions of such systems, for example to include cheap and efficient energy regeneration schemes, as have been developed for cell-free translation systems [14], will advance the 'synthetic' character of these systems (as opposed to the 'modifying' character of biotransformations) even further.

Even though impressive pioneering work in this direction has already been performed, we argue that two current developments will drastically facilitate such endeavors. First, the assembly of large collections of tailored genes will be hugely facilitated by recent advances in non-template-driven design of large stretches of DNA [15••], which should enable easy assembly of the multiple genes for such artificial systems of enzymes. Second, the undesired interference of background with the synthetic enzyme network should become much less of a problem when using hosts with thoroughly minimized genomes based, for example, on the current *Escherichia coli* minimal genome project [16••]. These developments should help to remove the essential obstacles from designing multi-step biotransformations.

Data generation

Once the network is assembled *in vitro*, we need to quantitatively understand its dynamics in order to overcome, for example, the constraints of homeostasis. In other words, we need a mechanistic model, the parameters of which are connected to clear physical

Figure 2



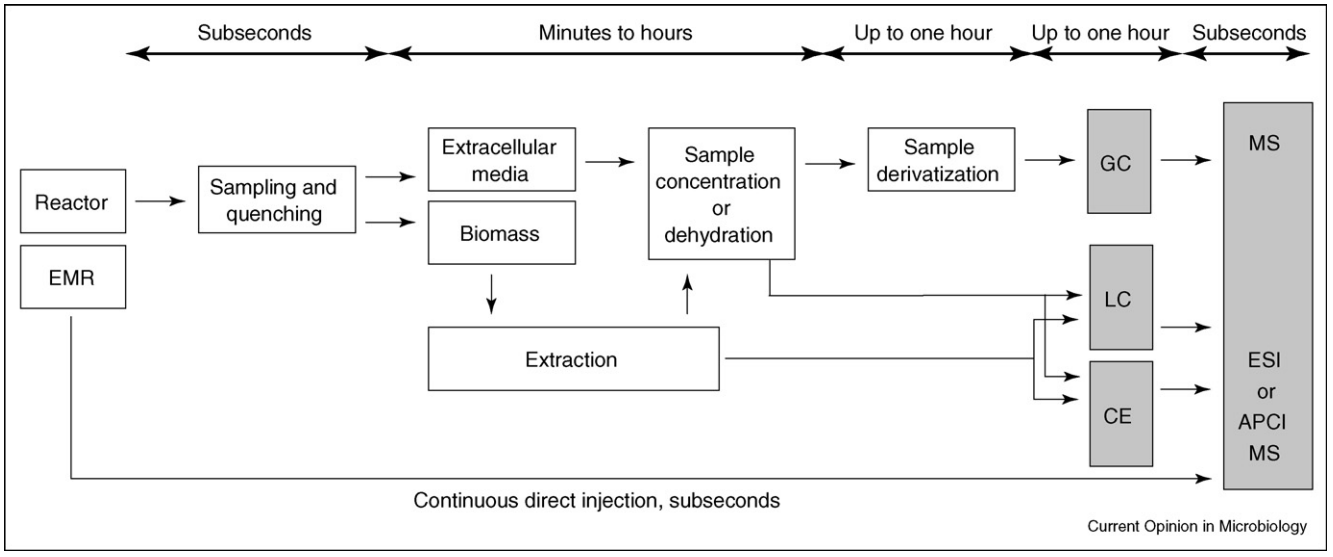
The engineering cycle for the rational design of systems of enzymes.

interpretations of the causes of specific observations. These parameters can then become the target for system modification (see below). Ideally, model building for a system of enzymes can benefit from a large body of research carried out on purified proteins over past decades, which should result in the availability of many kinetic parameters. However, for whole pathways, such data sets are only complete for a small number of model systems, such as glycolysis [17–19]. Furthermore, kinetic parameters have usually been determined in an artificial situation — using purified enzymes, only one or a few substrates at a time — and it is not always clear how parameters change when we look at a complex chemical situation. Finally, whereas parameters such as in Michaelis–Menten kinetics or inhibition constants might be easily transferred to more complex systems, the maximum reaction rate remains typically elusive, as there is usually no quick method to determine the amount of enzyme in a system. Consequently, model building on the basis of literature data will only take us part of the way.

The alternative is to estimate kinetic parameters from adequately large and correct sets of experimental data, such as time series of concentrations of substrate, intermediates and product. Traditionally, this is a rather difficult task. Using yeast glycolysis as a model system, a kinetic model (more specifically, a model containing only the reactions from hexokinase to pyruvate kinase) based on the algebraic expressions available from decades of biochemical research, contains between 38 and 45 parameters [17,18], which are all part of highly nonlinear equations. This emphasizes the requirement for large, accurate data sets with a resolution in time that matches the dynamics of the enzyme system under study. Recent advances in mass spectrometry (MS) [20], however, make such an endeavor feasible (Figure 3). Various separation techniques (e.g. gas and liquid chromatography and capillary electrophoresis) coupled to one of a diverse set of MS systems are available to record large sets of metabolites in complex matrix-loaded mixtures in a quantitative fashion [21–23], and have been exploited to investigate the response of intracellular metabolite concentrations to system perturbations. Typically, steady or quasi-steady state cultures are challenged with a (glucose) pulse and the translation of the pulse through the relevant sequence of reactions is followed. First, this approach requires rapid (frequently sub-second) and efficient sampling and quenching of intracellular metabolism immediately after the stimulus in order to capture the behavior of the system at the moment of stimulation, and various effective approaches have been followed to achieve this (discussed in [24]). Next, an efficient and high-throughput-compatible quantitative analysis is required, which has already been established for intermediates of central carbon metabolism, and also for aromatic amino acid formation pathways [25–28] and is now increasingly applied to answer physiologically and biotechnologically relevant questions about intracellular metabolite dynamics [29,30].

Luckily, the analysis of cell-free enzyme systems is much less demanding with respect to experimental set-up and

Figure 3



The main steps involved in quantitative MS-based analysis of intracellular metabolites and typical time ranges required for individual steps. Continuous direct injection has so far only been used for qualitative purposes, not quantitative, and requires MS/MS technology.

shows some other distinct properties: first, it is much easier to efficiently stop reactions after (even short) time periods because the enzyme system can be placed, for example, in enzyme membrane reactors that are continuously operated and where the outgoing fluid has been effectively separated from the catalysts; second, the system of metabolites is significantly smaller, which should help in shortening the still considerable analysis time for MS analysis (Figure 3); third, cell-free enzyme systems can be perturbed in a much larger variety of ways than can living systems — cell-free approaches can vary enzyme concentrations, feed intermediates or combinations of substrates and intermediates, vary pulse type and concentration, and interrogate the same system several times — and finally, cell-matrix effects can be reduced by diafiltration, for example, in the membrane reactor. In summary, a single experiment can be made much more specific and informative, which will help in the next step of the process. Consequently, it appears reasonable to assume that the generation of sufficiently large and accurate data sets for model identification (see below) is possible.

Model of the enzyme system

This next step in the engineering cycle requires the condensation of the experimental data into the mechanistic model, which consists of a set of differential equations (the mass balances of substrate, intermediates and product[s]) and a set of non-linear algebraic expressions that define the various reaction rates and contain the parameters (Hill coefficients, Michaelis–Menten and inhibition constants, and rate constants). These parameters have to be determined ('estimated') on the basis of the dynamic concentration data sets (see above). However, the large numbers of parameters in even relatively small systems of enzymes and the highly non-linear nature of the system of equations makes the value estimation far from trivial. Specifically, the identification of the globally best estimate for the parameters (rather than one of a series of possible other estimates, which are only the local best in a subset of the possible parameter space) is particularly challenging. Mathematically, parameter estimation can be formulated as the minimization of a cost function that punishes deviation between experimental data and model predictions when the parameters are varied [31]. This estimation problem can be most efficiently solved with stochastic global optimization procedures, in particular with evolution strategies (ES), although at large computational cost [32]. This effort can be reduced with hybrid methods, where the ES method is used to identify a solution in the vicinity of the global minimum and another algorithm to identify the exact location of the minimum [33]. Such powerful parameter estimation algorithms are available from software repositories (e.g. see [34–36]) and can efficiently help in completing an enzyme system model.

Another fundamental problem in parameter estimation is the treatment of errors in the model structure. When addressing realistically complex systems, it is very likely that, at the beginning, several influences are not correctly implemented in the model, leading to poor parameter estimations. Recently, two methods have been applied to biochemical systems to address this problem systemati-

cally. The first supports the parameter estimation process with the automated generation of families consisting of similar models, which are subsequently screened for feasibility by evaluating stoichiometric constraints. For the suitable models, the parameters need to be estimated separately and evaluated [37]. Alternatively, grey box modeling can be applied, in which the structure of the deviations of experimental data and model prediction can be used to direct the modeler to the required changes and even to uncover relations that involve unmeasured variables [38]. Taken together, these advances should tremendously facilitate the establishment of suitable models that describe the dynamics of the enzyme system sufficiently correctly in terms of accuracy and mechanism.

Mathematical optimization of the network

Once the model has been established and validated, it needs to be evaluated with respect to the question of which changes need to be made in order to improve the performance of the corresponding enzyme. This is usually done by carrying out a sensitivity analysis on the estimated parameters, that is, to determine the extent to which an interesting output variable (such as the concentration of the final product of the enzyme network) varies with a varying parameter (such as an inhibition constant in one of the enzymes of the network). The most sensitive parameters are then selected as the most promising targets for manipulating the network. It is well known from metabolic control analysis that the control of pathway performance is frequently distributed over an entire pathway, so optimization of pathway performance in some cases might need to involve the manipulation of several enzymes, whereas in others the manipulation of single enzymes might be enough [39]. However, in many cases it will be even more relevant to optimize a reaction network, subject to constraints, such as a requirement to maintain low concentrations of specific pathway intermediates (e.g. to suppress side reactions). Such 'conditioned' improvements have recently been investigated for overproduction of serine in *E. coli*, based on a mechanistic model of *E. coli* central metabolism [19]. In the first approach, the overall enzyme level was required to remain constant and intermediate metabolite concentration changes (relative to a reference state) were required to remain small on average. The resulting optimization problem was then solved with a non-linear optimization algorithm [40]. The second approach linearized the mechanistic model around a reference state and then applied mixed-integer linear programming to identify the optimal combination of knockouts and enzyme modulations under similar constraints as above [41]. Both optimizations predicted that an increase in the phosphotransferase system, the phosphofructokinase, and the glyceraldehyde phosphate dehydrogenase would be most beneficial, suggesting that the rules of pointing out optimal points for interfering with the network components are well understood.

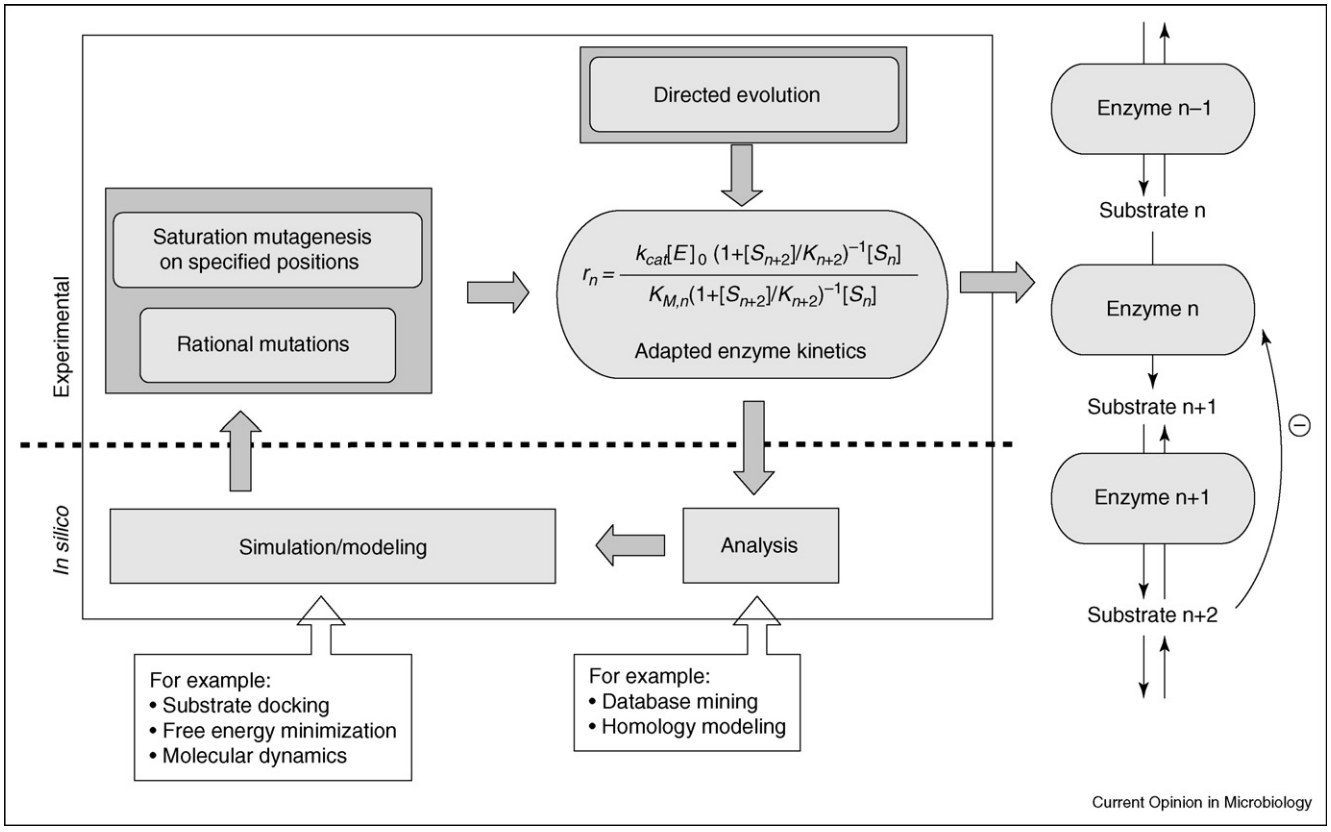
Real optimization of the network

The optimization of the model (see above) should provide clear suggestions on which property of which enzyme in the network should be modified. Changes in enzyme concentrations can be relatively easily accommodated, for example, by (reinforced) gene overexpression. For manipulations of the other parameters, two principal roads can be followed: evolutionary strategies or (semi-)rational design.

For a directed evolution approach to parameter manipulation to work, a suitable assay must be developed. For example, when screening for enzymes with decreased Michaelis–Menten or increased inhibition constant, the reaction is performed at substrate concentrations lower than the Michaelis–Menten constant of the wild-type protein or in the presence of the inhibitory compound, respectively (see e.g. [42–44]). A typical library for such a purpose contains 10^4–10^6 individual members [45] and consequently its (possibly repeated) screening is a laborious affair, although one with a high chance of success. However, taking into account the possibility that comprehensive re-engineering of a system of enzymes might require the modification of several members, a directed evolution strategy seems distinctively less attractive. On the other hand, many of the compounds that are particularly interesting for systems of enzymes, would exploit (parts of) the central carbon metabolism of, for example, *E. coli*. The corresponding enzymes have been very well investigated, so it might make sense to also consider rational design strategies.

(Semi-)rational concepts of protein design are characterized as iterative processes of *in silico* and *in vivo* methods (Figure 4) [46]. In general, the goal is the elucidation of sequence-structure-function relationships of a protein and the prediction of the effects of specific amino acid substitutions by computational tools, for example, by calculation of van der Waals forces, hydrogen bonds, hydrophobic interactions or salt bridges, within an amino acid chain or between two binding partners. This enables the calculation of the influence of mutations on protein stability, for example, with the FoldX force field [47], or predicting the effects of amino acid substitutions on the affinity of protein complexes [48]. Because these calculations are based on the structure and not on the amino acid sequence of the corresponding protein, the availability of a high-resolution image is a prerequisite, which might limit the application of such methods. An alternative approach to get access to sequence-structure-function relationships is based on the comparison of specific folds within protein families [49]. Several members of the α/β hydrolase-fold family, a versatile group of proteins including important biocatalysts, could be engineered using functional information from such comparative studies [50]. Even though this method is today restricted to a set of well-characterized protein families, novel high-

Figure 4



The kinetic properties of enzyme n as an element in a system of enzymes need to be adapted to increase the overall efficiency of the system. The example enzyme is uncompetitively inhibited by substrate n+2 (S refers to substrate, E_0 to the concentration of enzyme n, K_m is the Michaelis–Menten constant, K_{n+2} is the inhibition constant for inhibition by substrate n+2, and k_{cat} is the reaction rate constant). By using either directed evolution or a (semi-)rational approach that drastically limits the required number of mutants to screen, the required adaptation is implemented in an iterative approach.

throughput methods in the fields of genomics and proteomics are poised to change this. Interactions of proteins with small molecules — as is required for the application of systems of enzymes in the manufacturing of chemicals — can also be modified on the basis of molecular dynamics studies and docking simulations with substrates and transition intermediates [51,52]. Such computational profiling was even used to change the type of reaction catalyzed by a protein [53]. However, a major drawback is that most docking tools consider the protein as a rigid structure to reduce the complexity of their calculations, although it has been shown that enzymes might undergo substantial conformational changes upon substrate or cofactor binding.

The ultimate goal of protein engineering, and still an audacious vision, is the *de novo* design of enzymes with specific activities. A prerequisite for this is the ability to generate stable protein folds and catalytically active centers. Remarkably, both processes have been successfully implemented using computational design strategies. A globular protein containing two α-helices and five β-sheets, and thus representing a fold that has not been observed in any structurally elucidated protein, was designed completely *in silico* and subsequently experimentally verified [54]; and in another experiment, a triose phosphate isomerase activity was introduced into the ribose-binding protein [55••]. Regarding the *de novo* design of proteins, these results represent major milestones. Nevertheless, the intricacy of the calculations impedes the methods to become generally applicable so far. In summary, there are considerable advances in the field of (semi-)rational enzyme engineering, and in particular when assembling systems of enzymes from well-known central metabolism enzymes, it appears reasonable to assume that the effort required to manipulate multiple enzymes can be kept within reasonable limits. Nevertheless, it appears that, from all the separate points discussed so far, enzyme manipulation might turn out to be the limiting step in system design.

Conclusions

We argue that the works discussed above indicate that the central elements required for a synthetic biology

approach to designing efficient systems of enzymes are already or are rapidly being established. However, in order to be truly successful, the different elements (system assembly, analysis, modeling and manipulation) have to be considered in an integrated fashion: the measurement of dynamic responses to system perturbations enables estimation of the model parameters, which thus enables model analysis, leading to identification of points of interference, which then drives (semi-)rational enzyme manipulation. The completion of this cycle will be an example for a true engineering design approach in biotechnology and exemplify the power of synthetic biology in the area of fine chemistry.

Acknowledgements

The authors acknowledge support from the NEST-program of the European Union (project EUROBIOSYN).

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Straathof AJJ, Panke S, Schmid A: **The production of fine chemicals by biotransformations**. *Curr Opin Biotechnol* 2002, **13**:548-556.

2. Ro DK, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J *et al.*: **Production of the antimalarial drug precursor artemisinic acid in engineered yeast**. *Nature* 2006, **440**:940-943.

3. Menzella HG, Reid R, Carney JR, Chandran SS, Reisinger SJ, Patel KG, Hopwood DA, Santi DV: **Combinatorial polyketide biosynthesis by *de novo* design and rearrangement of modular polyketide synthase genes**. *Nat Biotechnol* 2005, **23**:1171-1176.

4. Endo T, Koizumi S: **Large-scale production of oligosaccharides using engineered bacteria**. *Curr Opin Struct Biol* 2000, **10**:536-541.

5. Werz DB, Seeberger PH: **Carbohydrates as the next frontier in pharmaceutical research**. *Chem Eur J* 2005, **11**:3194-3206.

6. Koeller KM, Wong CH: **Emerging themes in medicinal glycoscience**. *Nat Biotechnol* 2000, **18**:835-841.

7. •• Endy D: **Foundations for engineering biology**. *Nature* 2005, **438**:449-453.
A conceptual paper illustrating the foundations of the synthetic biology approach.

8. Heinemann M, Panke S: **Synthetic biology — putting engineering into biology**. *Bioinformatics* 2006, **22**:2790-2799.

9. Fessner W-D, Walter C: **''Artificial metabolisms'' for the asymmetric one-pot synthesis of branched-chain saccharides**. *Angew Chem Int Ed Engl* 1992, **31**:614-616.

10. Chen X, Liu J, Zhang J, Zhang W, Kowal P, Wang PG: **Reassembled biosynthetic pathway for large-scale carbohydrate synthesis: α-Gal epitope producing ''superbug''**. *ChemBioChem* 2002, **3**:47-53.

11. Koizumi S, Endo T, Tabata K, Ozaki A: **Large-scale production of UDP-galactose and globotriose by coupling metabolically engineered bacteria**. *Nat Biotechnol* 1998, **16**:847-850.

12. • Michel-Reydellet N, Calhoun K, Swartz J: **Amino acid stabilization for cell-free protein synthesis by modification of the *Escherichia coli* genome**. *Metab Eng* 2004, **6**:197-203.
This paper gives a straightforward example how strain design can beneficially influence cell-free application.

13. Nahalka J, Liu ZY, Chen X, Wang PG: **Superbeads: immobilization in ''sweet'' chemistry**. *Chem Eur J* 2003, **9**:373-377.

14. Calhoun KA, Swartz JR: **An economical method for cell-free protein synthesis using glucose and nucleoside monophosphates**. *Biotechnol Prog* 2005, **21**:1146-1153.

15. •• Tian J, Gong H, Sheng N, Zhou X, Gulari E, Gao X, Church G: **Accurate multiplex gene synthesis from programmable DNA microchips**. *Nature* 2004, **432**:1050-1054.
This paper describes a breakthrough in large-scale, low error, non-template-driven DNA synthesis.

16. •• Posfai G, Plunkett G, Feher T, Frisch D, Keil GM, Umenhoffer K, Kolisnychenko V, Stahl B, Sharma SS, de Arruda M *et al.*: **Emergent properties of reduced-genome *Escherichia coli***. *Science* 2006, **312**:1044-1046.
The authors describe the current state of the genome minimization project for *E. coli*, which has already lost up to 15% of genomic sequence, including all transposable elements.

17. Teusink B, Passarge J, Reijenga CA, Esgalhado E, van der Weijden CC, Schepper M, Walsh MC, Bakker BM, van Dam K, Westerhoff HV *et al.*: **Can yeast glycolysis be understood in terms of *in vitro* kinetics of the constituent enzymes? Testing biochemistry**. *Eur J Biochem* 2000, **267**:5313-5329.

18. Rizzi M, Baltes M, Theobald U, Reuss M: ***In vivo* analysis of metabolic dynamics in *Saccharomyces cerevisiae*: II. Mathematical model**. *Biotechnol Bioeng* 1997, **55**:592-608.

19. Chassagnole C, Noisommit-Rizzi N, Schmid JW, Mauch K, Reuss M: **Dynamic modeling of the central carbon metabolism of *Escherichia coli***. *Biotechnol Bioeng* 2002, **79**:53-73.

20. Villas-Boas SG, Mas S, Akesson M, Smedsgaard J, Nielsen J: **Mass spectrometry in metabolome analysis**. *Mass Spectrom Rev* 2005, **24**:613-646.

21. Bajad SU, Lu WY, Kimball EH, Yuan J, Peterson C, Rabinowitz JD: **Separation and quantitation of water soluble cellular metabolites by hydrophilic interaction chromatography-tandem mass spectrometry**. *J Chromatogr A* 2006, **1125**:76-88.

22. Roessner U, Wagner C, Kopka J, Trethewey RN, Willmitzer L: **Simultaneous analysis of metabolites in potato tuber by gas chromatography-mass spectrometry**. *Plant J* 2000, **23**:131-142.

23. Soga T, Ueno Y, Naraoka H, Ohashi Y, Tomita M, Nishioka T: **Simultaneous determination of anionic intermediates for *Bacillus subtilis* metabolic pathways by capillary electrophoresis electrospray ionization mass spectrometry**. *Anal Chem* 2002, **74**:2233-2239.

24. Oldiges M, Takors R: **Applying *metabolic profiling* techniques for stimulus-response experiments: chances and pitfalls**. *Adv Biochem Eng Biotechnol* 2005, **92**:173-196.

25. Buchholz A, Takors R, Wandrey C: **Quantification of intracellular metabolites in *Escherichia coli* K12 using liquid chromatographic-electrospray ionization tandem mass spectrometric techniques**. *Anal Biochem* 2001, **295**:129-137.

26. Oldiges M, Kunze M, Degenring D, Sprenger GA, Takors R: **Stimulation, monitoring, and analysis of pathway dynamics by metabolic profiling in the aromatic amino acid pathway**. *Biotechnol Prog* 2004, **20**:1623-1633.

27. van Dam JC, Eman MR, Frank J, Lange HC, van Dedem GWK, Heijnen SJ: **Analysis of glycolytic intermediates in *Saccharomyces cerevisiae* using anion exchange chromatography and electrospray ionization with tandem mass spectrometric detection**. *Anal Chim Acta* 2002, **460**:209-218.

28. Wu L, Mashego MR, van Dam JC, Proell AM, Vinke JL, Ras C, van Winden WA, van Gulik WM, Heijnen JJ: **Quantitative analysis of the microbial metabolome by isotope dilution mass spectrometry using uniformly C-13-labeled cell extracts as internal standards**. *Anal Biochem* 2005, **336**:164-171.

29. Nasution U, van Gulik WM, Proell A, van Winden WA, Heijnen JJ: **Generating short-term kinetic responses of primary metabolism of *Penicillium chrysogenum* through glucose**

perturbation in the bioscope mini reactor. *Metab Eng* 2006, **8**:395-405.

30. Wu L, van Dam J, Schipper D, Kresnowati M, Proell AM, Ras C, van Winden WA, van Gulik WM, Heijnen JJ: **Short-term metabolome dynamics and carbon, electron, and ATP balances in chemostat-grown *Saccharomyces cerevisiae* CEN.PK 113-7D following a glucose pulse**. *Appl Environ Microbiol* 2006, **72**:3566-3577.

31. Jaqaman K, Danuser G: **Linking data to models: data regression**. *Nat Rev Mol Cell Biol* 2006, **7**:813-819.

32. Moles CG, Mendes P, Banga JR: **Parameter estimation in biochemical pathways: a comparison of global optimization methods**. *Genome Res* 2003, **13**:2467-2474.

33. Rodriguez-Fernandez M, Mendes P, Banga JR: **A hybrid approach for efficient and robust parameter estimation in biochemical pathways**. *Biosystems* 2006, **83**:248-265.

34. Hoops S, Sahle S, Gauges R, Lee C, Pahle J, Simus N, Singhal M, Xu L, Mendes P, Kummer U: **COPASI- A COmplex PAthway SImulator**. *Bioinformatics* 2006, **22**:3067-3074.

35. Ji X, Xu Y: **libSRES: a C library for stochastic ranking evolution strategy for parameter estimation**. *Bioinformatics* 2006, **22**:124-126.

36. Zi Z, Klipp E: **SMBL-PET: a systems biology markup language-based parameter estimation tool**. *Bioinformatics* 2006, **22**:2704-2705.

37. Haunschild MD, Freisleben B, Takors R, Wiechert W: **Investigating the dynamic behavior of biochemical networks using model families**. *Bioinformatics* 2005, **21**:1617-1625.

38. Kristensen NR, Madsen H, Jørgensen SB: **A method of systematic improvement of stochastic grey-box models**. *Comput Chem Eng* 2004, **28**:1431-1449.

39. Torres NV, Voit EO, Glez-Alcon C, Rodriguez F: **Method and application to the maximization of the rate of ethanol, glycerol, and carbohydrate production on *Saccharomyces cerevisiae***. *Biotechnol Bioeng* 1997, **55**:758-772.

40. Visser D, Schmid JW, Mauch K, Reuss M, Heijnen JJ: **Optimal re-design of primary metabolism in *Escherichia coli* using linlog kinetics**. *Metab Eng* 2004, **6**:378-390.

41. Vital-Lopez FG, Armaou A, Nikolaev EV, Maranas CD: **A computational procedure for optimal engineering intervention using kinetic models of metabolism**. *Biotechnol Prog* 2006, **22**:1507-1517.

42. Canada KA, Iwashita S, Shim H, Wood TK: **Directed evolution of toluene *ortho*-monooxygenase for enhanced 1-naphthol synthesis and chlorinated ethene degradation**. *J Bacteriol* 2002, **184**:344-349.

43. Meyer A, Schmid A, Held M, Westphal AH, Röthlisberger M, Kohler H-PE, van Berkel WJH, Witholt B: **Changing the substrate reactivity of 2-hydroxybiphenyl 3-monooxygenase from *Pseudomonas azelaica* HBP1 by directed evolution**. *J Biol Chem* 2002, **277**:5575-5582.

44. Yun H, Hwang BY, Lee JH, Kim BG: **Use of enrichment culture for directed evolution of the *Vibrio fluvialis* JS17 omega-transaminase, which is resistant to product inhibition by aliphatic ketones**. *Appl Environ Microbiol* 2005, **71**:4220-4224.

45. Bornscheuer UT: **Trends and challenges in enzyme technology**. *Adv Biochem Eng Biotechnol* 2005, **100**:181-203.

46. Di Ventura B, Lemerle C, Michalodimitrakis K, Serrano L: **From *in vivo* to *in silico* biology and back**. *Nature* 2006, **443**:527-533.

47. Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L: **The FoldX web server: an online force field**. *Nucleic Acids Res* 2005, **33**:W382-W388.

48. Selzer T, Albeck S, Schreiber G: **Rational design of faster associating and tighter binding protein complexes**. *Nat Struct Biol* 2000, **7**:537-541.

49. Fischer M, Thai QK, Grieb M, Pleiss J: **DWARF — a data warehouse system for analyzing protein families**. *BMC Bioinformatics* 2006, **7**:495.

50. Qian Z, Fields CJ, Yu Y, Lutz S: **Recent progress in engineering α/β hydrolase-fold family members**. *Biotechnol J* 2006, **1**:1-9.

51. Knoll M, Müller M, Pleiss J, Pohl M: **Factors mediating activity, selectivity, and substrate specificity for the thiamin diphosphate-dependent enzymes benzaldehyde lyase and benzoylformate decarboxylase**. *ChemBioChem* 2006, **7**:1928-1934.

52. Looger LL, Dwyer MA, Smith JJ, Hellinga HW: **Computational design of receptor and sensor proteins with novel functions**. *Nature* 2003, **423**:185-190.

53. Carlqvist P, Svedendahl M, Branneby C, Hult K, Brinck T, Berglund P: **Exploring the active-site of a rationally redesigned lipase for catalysis of Michael-type additions**. *ChemBioChem* 2005, **6**:331-336.

54. Kuhlman B, Dantas G, Ireton GC, Varani G, Stoddard BL, Baker D: **Design of a novel globular protein fold with atomic-level accuracy**. *Science* 2003, **302**:1364-1368.

55. •• Dwyer MA, Looger LL, Hellinga HW: **Computational design of a biologically active enzyme**. *Science* 2004, **304**:1967-1971.
This paper describes a sophisticated computational and experimental approach to convert a ribose-binding protein into a triose phosphate isomerase by exchanging up to 22 amino acids. The resulting protein had excellent kinetic properties.