Compartmentalization and Metabolic Channeling for Multienzymatic Biosynthesis: Practical Strategies and Modeling Approaches

ARTICLE *in* ADVANCES IN BIOCHEMICAL ENGINEERING/BIOTECHNOLOGY · AUGUST 2013

Impact Factor: 1.66 · DOI: 10.1007/10_2013_221 · Source: PubMed

4 AUTHORS, INCLUDING:

A-P Zeng
Technische Universität Ham…

3 PUBLICATIONS **27** CITATIONS

SEE PROFILE

Adv Biochem Eng Biotechnol
DOI: 10.1007/10_2013_221

Compartmentalization and Metabolic Channeling for Multienzymatic Biosynthesis: Practical Strategies and Modeling Approaches

U. Jandt, C. You, Y. H.-P. Zhang and A.-P. Zeng

Abstract The construction of efficient enzyme complexes for multienzymatic biosynthesis is of increasing interest in order to achieve maximum yield and to minimize the interference due to shortcomings that are typical for straightforward one-pot multienzyme catalysis. These include product or intermediate feedback inhibition, degeneration, and diffusive losses of reaction intermediates, consumption of co-factors, and others. The main mechanisms in nature to tackle these effects in transient or stable protein associations are the formation of metabolic channeling and microcompartments, processes that are desirable also for multienzymatic biosynthesis in vitro. This chapter provides an overview over two main aspects. First, numerous recent strategies for establishing compartmentalized multienzyme associations and constructed synthetic enzyme complexes are reviewed. Second, the computational methods at hand to investigate and optimize such associations systematically, especially with focus on large multienzyme complexes and metabolic channeling, are discussed. Perspectives on future studies of multienzymatic biosynthesis concerning compartmentalization and metabolic channeling are presented.

Keywords Compartmentalization · Metabolic channeling · Modeling · Multienzymatic synthesis · Synthetic biology

U. Jandt (✉) · A.-P. Zeng
Institute of Bioprocess and Biosystems Engineering, Hamburg University of Technology, Denickestr. 15, 21071 Hamburg, Germany
e-mail: uwe.jandt@tu-harburg.de

A.-P. Zeng
e-mail: aze@tu-harburg.de

C. You · Y. H.-P. Zhang
Virginia Tech, Biological Systems Engineering Department, 210-A Seitz Hall, Blacksburg, VA 24061, USA
e-mail: yoshion@vt.edu

Y. H.-P. Zhang
e-mail: ypzhang@vt.edu

Contents

1 Introduction

Multienzymatic biosynthesis or biotransformation refers to the use of several enzymes and co-enzymes to carry out chemical reactions in a cascade for the production of a desired compound. For in vitro reaction cascades, this mostly leads to typical shortcomings such as product or intermediate feedback inhibition, diffusive losses, consumption of co-factors, and others. On the other hand, such cascade enzyme biosynthesis featuring substrate channeling is ubiquitous inside living cells. Cells face many challenges including rate-limited enzymes, competing metabolic reactions, and labile or toxic intermediates in enzymatic biotransformation (Fig. 1) [174, 175]. For example, some enzymes, such as aldolase (ALD, EC 4.1.2.13) in the glycolysis and gluconeogenesis pathways, suffer from slow turnover, which may result in metabolic flux disequilibrium or rate-limited steps in pathways [35, 36]. Some metabolites can participate in many pathways, leading to their reduced availability for the desired pathway. An example of this is dihydroxy-acetone phosphate (DHAP), an intermediate that is involved in the glycolysis metabolic pathway and the Calvin cycle, and DHAP is also extremely unstable [151].

To deal with these challenges, nature has evolved two main strategies. One is compartmentalization which is the physical separation of biological reactions to bring the cascade enzymes close to each other. Examples of compartmentalization include membrane-bound organelles and bacterial microcompartments [35, 36]. Membrane-bound organelles are common in eukaryotes, such as peroxisomes, which encapsulate reactions that generate or consume the toxic hydrogen peroxide in oxidative reactions [64]. One of the examples of bacterial microcompartments is carboxysome (Fig. 1a), which encapsulates RuBisCO and carbonic anhydrase (CA) to accelerate the reaction rate by providing a high local concentration of carbon dioxide to the enzyme ([21]; Yeates et al. 2008). The membranes of

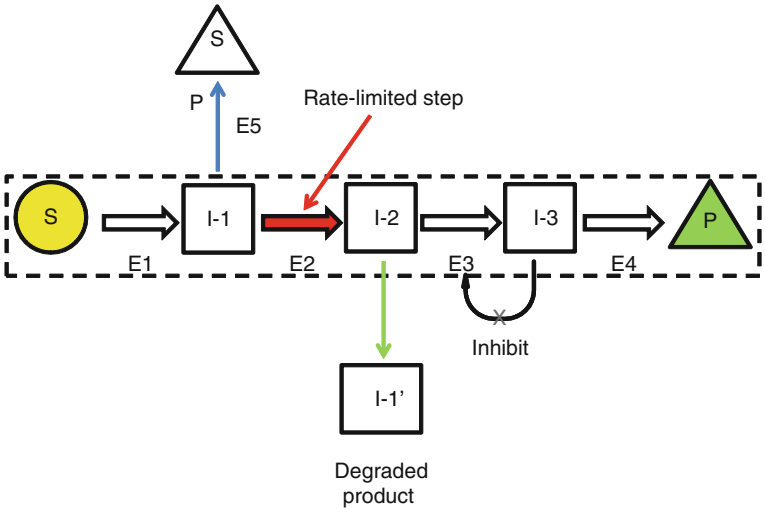

Fig. 1 The scheme of multienzymatic biotransformation and challenges. *S* substrate, I-1, I-2, I-3, I-1′, intermediates, *P* product, *SP* side product, *E* enzyme

organelles and protein shells of bacterial microcompartments can also protect the cell from toxic intermediates [124, 136]. Another way is the formation of metabolic channeling in enzyme complexes, which link cascade enzymes together. Such enzyme complexes found in nature include tryptophan synthase [52], polyketide synthase [159], the family of 2-oxoacid dehydrogenase complexes such as the pyruvate dehydrogenase complex (PDC; [123, 162]) or ketoglutarate dehydrogenase complex (OGDC; [107]), and cellulosomes [13, 14, 48]. A typical example is tryptophan synthase (Fig. 2b), a two-enzyme complex that catalyzes the last two reactions in the biosynthesis of L-tryptophan [52]. The reactive intermediate of insole is channeled from the active site of the α enzyme to the active site of the β enzyme without being released into the surrounding environment.

Enzyme cascades in living cells, cell compartments, and also organisms as seen in nature rely upon a well-orchestrated regulation machinery that features information, energy, and mass transfer at very different time and size scales. These include information dissipation within enzymes, local diffusion of small messenger molecules, and metabolic channeling within temporary or permanently formed enzyme complexes and nanocompartments. The exploitation of many of these evolutionarily developed features for technical application drives the need for a deeper mechanistic understanding of these interactions and thus the development of multiscaled modeling methods [119, 129]. In practice, biologists have devised many nature-inspired synthetic systems to mimic cellular compartmentalization and enzyme complex systems featuring substrate channeling. The key point of these systems is to bring the cascade enzymes spatially close, resulting in an increased reaction rate and reduced diffusion of intermediates to the surrounding

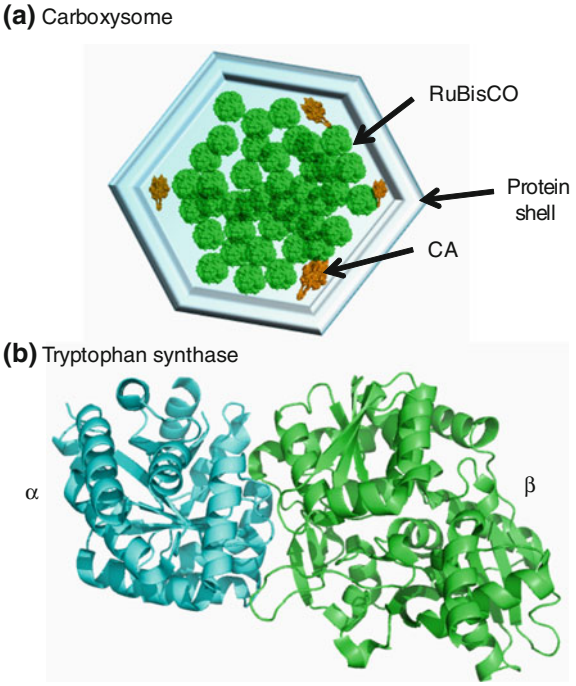

Fig. 2 Nature compartmentalizes enzyme cascade reactions via membrane-bound organelle (Carboxysome) (**a**) and enzyme complexes (Tryptophan synthase) (**b**)

environments. Many of these systems are basis technologies for multienzymatic biotransformation to increase the production of industrially and commercially important chemicals [95, 138, 173].

In this chapter, we describe molecular principles and recently attempted strategies to build multienzymatic compartments and multienzyme complexes, including enzyme encapsulation, fusion proteins, co-immobilization, and protein/nuclear acid scaffolds [4]. For a more technologically oriented review of some of these strategies the reader is referred to the chapter of Ardao and colleagues [4] in this volume. Furthermore, the latest advances of modeling large multienzyme ensembles and complexes, such as directed metabolic channeling, self-assembly, and information transfer, are addressed and discussed. Also we provide perspectives on multienzymatic biosynthesis concerning compartmentalization and synthetic enzyme complexes (metabolons).

2 Practical Strategies

Inspired by natural bacterial microcompartments [84], microcompartments have been heterologously expressed and could be used to encapsulate foreign pathways, including enzymes and substrates (Fig. 2a). Carboxysomes from cyanobacteria were heterologously expressed in *Escherichia coli,* and the encapsulated RuBisCO

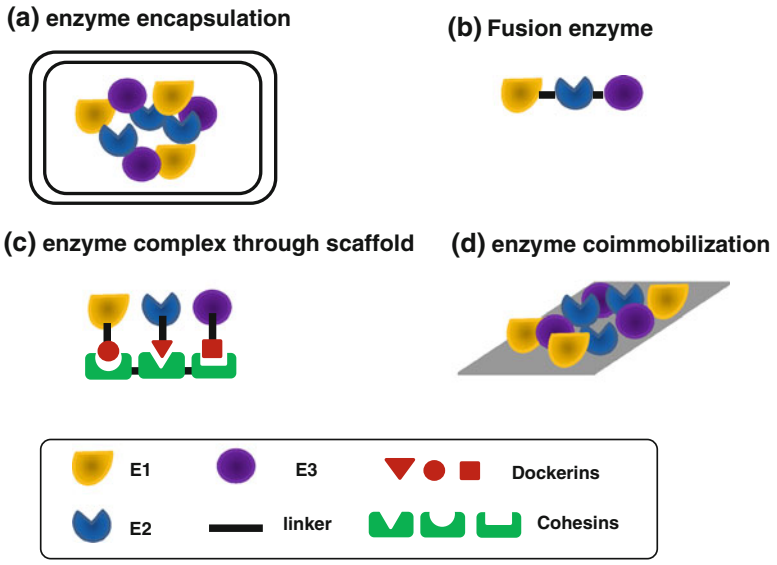

Fig. 3 Strategy for compartmentalization of cascade enzymes: enzyme encapsulation by membrane or protein shell (**a**), a fusion protein of multiple functions (**b**), an enzyme complex through the interaction with the scaffold (**c**), co-immobilization of several enzymes (**d**)

in purified synthetic carboxysomes was still capable of fixing carbon dioxide in vitro [21]. The ethanolamine utilization (Eut) and 1,2-propanediol utilization (Pdu) microcompartment from *Salmonella* have also been expressed in *E. coli* [41, 122]. These heterologously expressed microcompartments have the potential to increase the metabolic flux in enzymatic biosynthesis to the desired product, making this method a useful tool for biotransformation. Another compartmentalization strategy utilizes naturally existing membrane-bound organelles. The advantages of using membrane-bound organelles is that the localization mechanisms are well studied [147] and they allow us to engineer some model microorganisms such as *E. coli* and yeast. For example, a methyl halide transferase was introduced into yeast vacuoles, resulting in the increase of methyl iodide production [16]. The introduction of terpenoid production pathway components to the mitochondria significantly increased product yield [57] (Fig. 3).

Another encapsulation device was recently introduced using a lumazine synthase capsid [168]. Native lumazine synthase, an enzyme catalyzing riboflavin synthesis, does not encapsulate any other enzymes. However, directed evolution was used to engineer electrostatic interactions between this capsid and encapsulated proteins, rendering this capsid the ability to capsulate a toxic enzyme, HIV protease. The ability to insulate toxic compounds from the cell by using a designable compartmentalization system could be very important for living-cell biotransformation.

There are many challenges, such as targeting of more foreign enzymes to microcompartments, and controlling the architecture of artificial microcompartments and stoichiometries of encapsulated enzymes before heterologously expressed microcompartments can be used for multienzymatic biosynthesis. Another problem is the limited knowledge of the mechanism of substrate transport and secretion of final products across the microcompartment shell [90]. A better understanding of these factors and studies in natural compartmentalization systems may help develop better synthetic microcompartments [35, 36].

2.1 Fusion Proteins

The fusion of multiple proteins together is the creation of chimeric proteins, in which two or more cascade enzymes are combined by a linker to form a multifunctional single polypeptide (Fig. 2b) [43, 174]. For example, Bulow et al. constructed a bifunctional enzyme containing cascade reactions mediated by *E. coli* β-galactosidase (LacZ) and galactokinase (GalK) [27]. The resulting fusion protein displayed the enzymatic activity of both gene products. Later, the same group produced another fusion protein containing LacZ and galactose dehydrogenase from *Pseudomonas fluorescens* for the sequential hydrolysis of lactose followed by the oxidation of the galactose, forming the corresponding lactone. This synthetic fusion enzyme displayed kinetic advantages (1.5 to 2.4-fold) over free enzyme mixtures. Such enhancement increased when lactose concentration decreased [97]. This suggests that the proximity conferred by the artificial fusion can enable some amount of synergistic action of the cascade reaction. Such effect has also been revealed in other fusion enzymatic systems [104, 120, 132, 142, 145].

The construction of multifunctional fusion proteins has been proposed to have a potential application in industrial production because it may be beneficial in directing metabolic flux to a preferred pathway [28]. Two genes encoding glycerol 3-phosphate dehydrogenase and glycerol 3-phosphatase were reported to be spontaneously fused into an open reading frame encoding a bifunctional enzyme with a 4-AA linker by removing a small C-terminal fragment of glycerol-3-phosphate dehydrogenase [109]. The high efficiency of this fusion protein enables partial glycerol 3-phosphate channeling between two active sites, resulting in higher glycerol yields. By utilizing a chemostat, the evolved *E. coli* strains that have a spontaneous mutation of this fusion enzyme can produce glycerol from glucose in high yield, concentration, and productivity. The effects of different linker lengths (e.g., 2, 14, 24, 46, and 104 amino acids) of a fusion of hydrogenase and ferredoxin for hydrogen production were tested by Silver and her coworkers [2]. They found that the optimal linker length was 14 amino acids in vivo with an enhanced factor of more than four and no substrate channeling was observed by in vitro tests.

Although building fusion protein is a very simple way to build an enzyme complex [77], the misfolding of large multidomain proteins often happens in hosts [34, 117], resulting in large uncertainty regarding this strategy.

2.2 Building Enzyme Complexes Using Protein Scaffolds

Inspired by extracellular hydrolase enzyme complexes, cellulosomes [15, 49, 146, 170, 171], a synthetic scaffold containing different cohesins from different microorganisms was first used for the assembly of cascade metabolic enzymes (Fig. 2c) [169]. Triosephosphate isomerase (TIM), ALD, and fructose 1,6-bisphosphatase (FBP) were engineered to have a dockerin at their C-terminals. These three dockerin-containing enzymes can be self-assembled into a static trifunctional enzyme complex through interaction with a miniscaffold protein consisting of three different matching cohesins. The formation of this three-enzyme complex was validated by the result of native PAGE. The particle size of this synthetic enzyme complex was increased approximately fivefold relative to the scaffold only (the particle sizes were estimated by their Zeta potentials with the Malvern Zetasizer Nano ZS system). Such an enzyme complex features substrate channeling because of the spatial proximity of the cascade enzymes [174]. The synthetic metabolon showed more than one order of magnitude enhancements on reaction rates compared to the noncomplexed TIM, ALD, and FBP mixture. This result suggests that using cohesins and dockerins from cellulosomes would be a powerful tool for building an enzyme complex for multienzymatic biotransformation.

Other protein scaffolds from metazoan signaling proteins have been tested for synthetic metabolons for enhancing product titers [51]. The mevalonate pathway suffers from a toxic intermediate and flux imbalance due to unmatched enzyme activities in the pathway. The best-designed synthetic scaffoldin increased product yields 77-fold [51]. The enzymes responsible for the synthesis of the glucaric acid pathway were scaffolded in the same way, resulting in fivefold enhancement in the product titer [114]. Another pathway is the biological hydrogen pathway containing [Fe–Fe]-hydrogenase, ferredoxin, and pyruvate–ferredoxin oxidoreductase (PFOR), which suffers from competing reactions. The synthetic enzymatic pathway containing a protein scaffold, hydrogenase, and PFOR has an approximately fivefold reaction rate enhancement [2].

The degree of enhancement in product yield or reaction rate is believed to be related to enzyme/enzyme ratios, their orientation, and scaffold stoichiometry. However, there is no clear rule to estimate it before testing. Increasing the number of the interaction domain in the scaffold for rate-limited enzymes can dramatically increase the product titers [51]. Via it, optimized stoichiometries and geometries of synthetic protein scaffolds may make multienzyme biotransformation more effective.

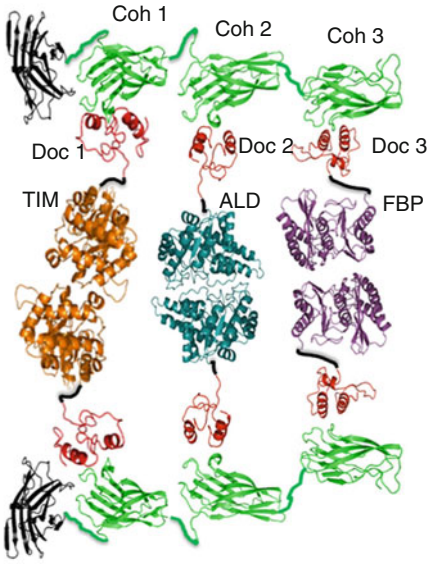

Fig. 4 Enzyme complexes are hypothesized to oligomerize into large complexes due to the oligomerization of enzymes

Understanding the structure of enzyme complexes may help elucidate the mechanism pertaining to substrate channeling among enzyme cascades. It has been hypothesized that they may form higher-order complexes due to enzyme oligomerization [94]. For example, in the synthetic TIM, ALD, and FBP enzyme complex, it is known that only dimeric TIM is active [101]; all monomeric, dimeric, and tetrameric ALDs are active [144]; and both dimeric and tetrameric FBPs are active [65]. So this three-enzyme complex may form a dimer with TIM, ALD, and FBP facing one another, as hypothesized previously (Fig. 4) [28].

2.3 Building Enzyme Complexes Using Nucleic Acid Scaffolds

Nucleic acids, both DNA and RNA, can be used as scaffolds to recruit enzymes to form multienzyme complexes. DNA or RNA can be designed to fold into various structures in vitro, forming simple structures such as sheets to more complicated structures such as tubes and capsules [3, 165]. Therefore it is more convenient to synthesize DNA or RNA scaffolds in vitro than protein scaffolds. However, the cost of DNA and RNA is more expensive than that of protein scaffolds, which may impair their application on a large scale.

Several multienzyme systems have been successfully assembled using a DNA scaffold. NADH-flavin mononucleotide (FMN) oxidoreductase and luciferase were assembled onto DNA scaffolds by using streptavidin–biotin linkages, leading to enhanced enzymatic activity relative to the free enzyme mixture [118]. Also,

plasmids can be used as a DNA scaffold, which recruits enzymes via zinc finger domains that bound to specific motifs on the scaffold [44]. Several enzymes were tested using this system, including pathways for the production of resveratrol, 1,2-propanediol, and mevalonate [44]. Yields were found to increase as a function of scaffold architecture, similar to protein scaffolds.

Recently, RNA has drawn huge attention as a scaffold to recruit enzymes for enzyme complexes [1, 46, 71]. The RNA molecules can comprise multiple different aptamer motifs, which can recruit proteins, as well as have complementation regions, which form scaffolds. Such scaffolds are relatively easy to build and characterize, and their expression level can be controlled. So RNA scaffolds are suitable for the expression of specific stoichiometries of enzymes in one RAN chain. For the proof-of-concept experiment, [Fe–Fe]-hydrogenase and ferredoxin, enzymes involved in the biological production of hydrogen, were recruited to RNA scaffolds for hydrogen production [50]. This cascade suffers from side reactions and requires the enzymes to be in close proximity to increase the metabolite flux [2]. Hydrogen production was also found to be related to scaffold architecture.

Regardless of protein or nucleic acid scaffolds, the generality of such scaffolds needs more investigation than only a few proof-of-concept experiments. The mechanism of scaffold action must be further elucidated [35, 36, 94].

2.4 Co-Immobilization of Enzymes

Another approach for facilitating multienzymatic biotransformation is immobilizing cascade enzymes by tethering enzymes on one solid support (Fig. 3d) [20, 55, 160, 172] or direct cross-linking (Mateo et al. 2006; Moehlenbrock et al. 2010). The solid support can be synthetic organic polymers (e.g., Amberlite XAD-7 and Eupergit®C) (Katchalski-Katzir and Kraemer 2000; Kirk and Christensen 2002), inorganic polymers (e.g., silica, zeolites, and mesoporous silicas; Díaz and Balkus 1996; Petri et al. 2005; Zhou and Hartmann 2012), and natural polymers (e.g., cellulose, starch, agarose, and chitosan; Krajewska 2004). Regenerated amorphous cellulose (RAC) made from Avicel (Zhang et al. 2006) is an excellent solid support with a large external surface for enzyme immobilization (Hong et al. 2008; Liao et al. 2012). The entire binding surface of RAC is externally accessible to the large-size enzymes, so the enzymes immobilized on RAC rarely lose their apparent activity (Hong et al. 2008; [169]). An enzyme complex containing TIM, ALD, and FBP immobilized on RAC through a CBM-containing scaffold showed higher activity towards the simple enzyme complex linked by the scaffolding [172]. Multiple enzymes can be precipitated by adding salts, organic solvents, or nonionic polymers followed by cross-linking, resulting in cross-linked enzyme aggregates (CLEAs; Sheldon and van Pelt 2013). A combined CLEA containing an S-selective hydroxynitrile lyase from *Manihot esculenta* and a nonselective nitrilase from *Pseudomonas fluorescens*, catalyzed one-pot conversion of

benzaldehyde to S-mandelic acid (Mateo et al. 2006). This combined CLEA showed higher activity than the mixture of the two separate CLEAs.

Enzyme components during co-immobilization can be randomly distributed [20, 55, 160], positionally assembled [87, 161, 163], and even the active site of an enzyme face that of another enzyme [103]. The enhanced reaction rates among co-immobilization of cascade enzymes have been observed for several systems ([103]; Mateo et al. 2006; [115, 152]), but direct cross-linking could lead to loss of enzyme activity [174].

3 Modeling Approaches

The vastly differing time and size scales of regulation and interaction processes that are necessary for specific and efficient biocatalysis within (and between) living cells prohibited extensive computational simulations in the past (Fig. 5). Generally, strongly reductive approaches were (and are still) widely used to describe only partial processes. With ever-increasing computational power, simulations on a systems level of larger time- and size-scaled regions even up to cells become more realistic in the foreseeable future [158]. Many efforts have been made in systems biology in recent years in order to understand living cells better and to build whole cell models [63, 83].

Despite the fact that computational power will still increase in the next years, pure computation power alone will not be sufficient. Instead, versatile theoretical and mathematical frameworks to describe multilevel interactions are needed and are often yet to be developed [119]. In-depth discussions of various modeling approaches (phenomenological vs. mechanistic, continuous vs. stochastic) describing enzyme interactions [167], followed by detailed considerations of the combination of different scales and methods [129], for example, to capture the hierarchical organization of cells [81] contribute to that development. It has become clear in the last decades that the dynamical self-organization of cells and cell metabolism is fundamentally driven by functional enzymatic associations, catalytic reactions produced during metabolite channeling, microcompartmentalization, and the development of dissipative networks [105, 106]. In this section, the focus is therefore laid on modeling aspects to describe dynamics in multienzyme complexes on a molecular level (Sect. 3.1), in microcompartments (Sect. 3.2), and the determination of fluctuating enzyme–enzyme collaborations from a systems level perspective (Sect. 3.3).

Various spatiotemporally resolved modeling approaches on a larger scale, covering different aspects in mammalian cell compartments and whole cells, have been described elsewhere [47, 79, 150].

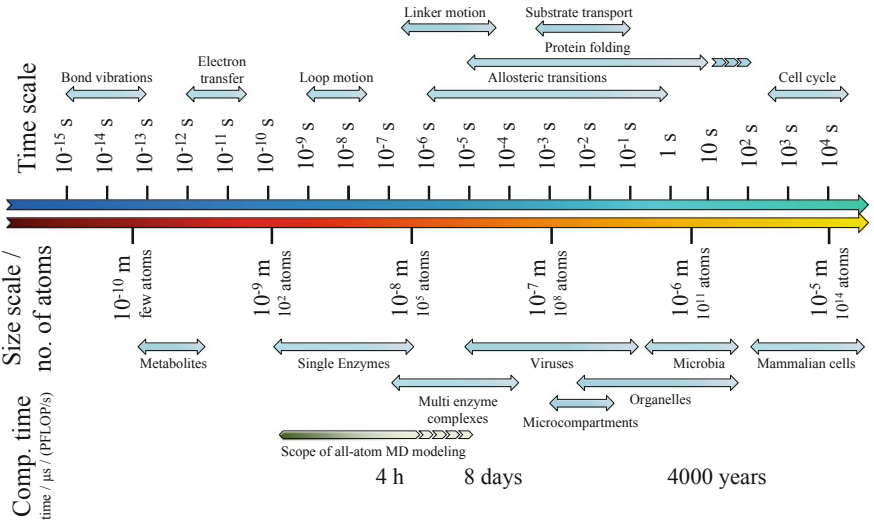

Fig. 5 Typical time and size scales of biologically relevant events and corresponding estimated computation time for coverage of 1 µs using extensive simulation on a molecular level per available petaFLOPS/s (1 petaFLOPS/s = 10^{15} floating point operations per second) computation power (Sect. 3.1.4). Computation performance with respect to system size and parallel computation. Units has been estimated to scale by O(n log n)

3.1 Molecular Level

Molecular modeling methods represent a powerful tool to elucidate the regulation and interplay of different molecular regions within enzymes and partly also enzyme complexes. The classical approaches perform all-atom simulations of biomolecules in explicit solvents, usually applying several restraints, simplifications, and cut-offs to accelerate the simulation [108]. The numerically simulated timestep represents the fastest relevant motion (bond vibrations) and usually ranges between 1–4 fs [93]. The most popular methods and program packages are CHARMM [23], AMBER [33], GROMACS [73], and NAMD [128]. Structure data at atom resolution (5 Å) employed in these simulations can be determined using, for example, X-ray crystallography, although it is usually not directly available for (large) enzyme complexes. The availability of atomic resolution structure data has been drastically increasing in recent years. For example, the stable pyruvate dehydrogenase multienzyme complex (PDC) consists of multiple copies of three main enzymes (not considering regulatory kinases and phosphatases attached to the complex) with a total weight of 9 MDa. For the main constituting enzymes, more than 80 different datasets are currently available.

3.1.1 Structure Determination

An important point to be considered in the context of molecular-level modeling is the structure determination of enzyme complexes under physiologic conditions, especially after application of variations such as point mutations. Classical atom-resolution structure determination methods, such as X-ray spectrography [17, 40], rely on the ability to establish a regular crystal structure of a single enzyme. Since many available structure data originate from variable species and also variable tissues [123], the data are usually homogenized using homology modeling methods [86].

The structure extraction of crystallized enzymes is not possible in every case because it can hamper the ternary or quarternary conformation of the enzyme and it is not possible for larger multienzyme complexes. Methods such as small-angle X-ray scattering (SAXS; [18, 69]) and small-angle neutron scattering (SANS; [162]) provide coarse structure data also for large complexes, but with no information about atomic distribution. The gap can be closed with fitting algorithms that systematically optimize orientations and positions of subcomponent atomic structure data to match the low-resolution measured data of the whole complex, obtained, for example, from electron microscopy [56] or SAXS [126]. The dynamics of proteins and protein folding can be observed with NMR spectroscopy, especially in-cell NMR of labeled biomolecules [19].

3.1.2 Molecular Dynamics Simulations

It can be estimated from Fig. 4 that all-atom simulations of higher-order structure changes of enzymes, such as involved in allosteric regulation, protein folding, and also substrate transport via channeling are yet beyond reach. Instead, simplified models, such as steered molecular dynamics simulations, umbrella sampling methods, and implicit solvent models among others, are often employed to reach longer timescales close to the millisecond range [70, 91]. For example, the model-based examination of fast-folding proteins (100 μs–3 ms) showed a set of common principles ruling the folding process of 12 structurally different proteins. Such long simulations, however, increasingly reveal shortcomings of the simplifications applied to the underlying force fields [91], as has been suspected for the apparent over-stabilization of helices during long timescale rearrangements compared to experimental data [62].

For comparably small enzyme systems, molecular modeling methods are already widely used for rational protein design [37, 40], which is also addressed in the chapter by Bornscheuer et al. in this volume. Efforts and successes in design and characterization of synthetic compartmentalization have been reviewed in Chen and Silver [35, 36]. Molecular modeling methods as well as other bioinformatics methods were, for example, applied for medical applications such as antiviral research [88] and the identification of novel anticancer targets [102]. Recent successful attempts include, for example, the introduction of novel

catalytic reactions including the Diels–Alder reaction [148], design of retro-ALDs [80], or the reduction of allosteric product inhibition of aspartokinase for lysine production [38, 39]. The complete dynamic protein-inhibitor binding process for trypsinbenzamidine has been elucidated based on numerous molecular dynamics simulations, revealing three different and partially metastable intermediate binding states between the free and bound states [26]. Generally, the transfer of MD simulation predictions to experimental procedures is challenging, as has been discussed extensively for the example of protein folding [139].

Large enzyme complexes have been less subjected to systematic reconnaissance of structure–function relationships, partly due to their size and corresponding computation power demands and partly due to their complex function involving metabolic channeling, information transfer between subcomponents, self-assembly, and others. Novel inhibitory compounds against the E1 subcomponent of the PDC have been identified with molecular docking methods [130]. It was shown that the dynamic of loops in the active center of the E1 component of PDC is likely to control the decarboxylation step rates of this component, which was revealed by ^{19}F NMR analysis [9, 82]. The influence of structure on the self-assembly of protein complexes has been addressed for pore-forming complexes including the TatA complex, identifying an electrostatic charge zipper mechanism that is also proposed for other membrane-bound proteins [60, 164]. Mostly, symmetry plays a pivotal role in the assembly of large multienzyme structures that are constituted of multiple copies of enzymes (review in [92]). For example, it has been shown earlier that Euclidean geometric considerations predicted the formation of the dodecahedric shape of the PDC E2p component almost exactly [78]. Specific efforts to evaluate the mechanisms of substrate channeling in multienzyme complexes are discussed separately in the following section.

3.1.3 Metabolic Channeling

Not many, although increasing, efforts have been made to elucidate systematically mechanisms of metabolic channeling within enzymes and enzyme complexes by means of structural modeling methods. A straightforward approach to facilitate metabolic channeling is to bring the active sites of corresponding enzymes close to each other, an effect referred to as “channeling by proximity” [12]. In this context, the optimal orientation leading to partial channeling between subsequent enzymes of the glycolytic pathway has been quantitatively simulated [72]. One of the most extensively studied natural channeling mechanisms is in tryptophan synthase [74]. Dunn et al. resolved the chemical background controlling the allosteric interactions responsible for activity switching of the site depending on the stage of the subunit catalytic cycle and also the conformation changes preventing the escape of substrate [52, 53]. These efficiently coordinated cycles obviously rely on directed information transfer for synchronization between the subunits, which may be more systematically elucidated with recently developed directed energy dissipation models [100], especially because long-range interactions seem to play an

important role for allosteric networks in tryptophan synthase [5]. Shorter-range synchronization processes between different active sites, such as proton channels (or "wires") in thiamine enzymes, have also been described [59]. The mechanisms of aldehyde channeling involving two gating residues (His-20 and Tyr-290) in BphI-BphJ, an aldolase–dehydrogenase complex, have been examined [31], as well as the translocation of glutamate or aspartate in a glutamate transporter along an energetically preferred permeation pathway of $\approx$23 Å and two phases [70]. Further molecular dynamics simulation efforts in transporters and membrane channels have been reviewed in Khalili-Araghi et al. [85] and Bahar et al. [8].

3.1.4 Aspects of Computation Power

With growing computational power and availability of multiscale structure data, the structure–function relationships of single parts or aspects of multienzyme complexes can be increasingly elucidated, such as structure stability and self-assembly properties of PDC [162]. However, comprehensive structure-based modeling of enzyme complexes as large as PDC are still beyond the scope of available computation power, instead extensive simplifications (e.g., normal mode analysis, NMA, or a quantized elastic deformational model, QEDM; [8, 24, 68, 110] can be applied to evaluate global conformation variations.

With presently available average-scale computation clusters (2–4 teraFLOPS/s, 1 teraFLOPS/s = 10^{12} floating point operations per second), the computation time for a molecular dynamics calculation of a 30,000-atom system covering 100 ns is roughly 1 day [62, 91]. In theory, an all-atom simulation of a large multienzyme complex (PDC: 9 MDa \approx 10^6 atoms) covering 1 μs assuming an approximate O (n log n) scaling of computation efficiency would take more than 450 days, which is impractical. On the other hand, the available computation power is still increasing more or less exponentially. The classic notion of "Moore's law," the postulated doubling of computation capacity every 18–24 months, seems to have held true, at least until the end of this decade [25]. For comparison: the fastest computer today (Spring 2013) reaches a computation speed of $\approx$17,500 teraFLOPS/s ("Titan," Oak Ridge National Laboratory, USA), whereas 1 exaFLOPS/s (1 million teraFLOPS/s) performance computers are already in the pipeline for 2018–2020 approximately [25, 131].

However, the faster nominal computation speeds usually do not translate linearly to decreased simulation times: up-scaling of molecular dynamics calculations to larger, that is, more parallelized, computer systems are far from trivial because of the increased communication overhead between the parallel computation threads. This is mainly caused by long-range and slowly decaying electrostatic interactions that need to be considered in the simulations and prohibit simple subdivision of the calculated systems. More efficient approaches involve fast Ewald mesh/multigrid methods [143] or the reaction field (RF) method that allows improved scaling behavior [O (n) instead of O (n log n)] up to more than 30,000 cores, yielding 30 ns/day for 3–5 million atom systems [140].

Another problem that emerges during massively parallel simulations of very large molecular systems is the limited available memory (few gigabytes) per single computation core, principally increasing the communication overhead drastically. Sophisticated communication management and load balancing methods, however, make the simulation of very large systems (up to 100 million atoms) tractable and efficiently scalable [108].

Extensive comparisons of widely used simulation packages (among others: AMBER, CHARMM, GROMACS) have been performed by Loeffler and Winn [98], with total system sizes (including solvent and lipids) ranging from 2×10^5 to 3×10^6 atoms.

In a few singular cases, biological simulations at scales comparable to the largest multienzyme complexes have already been performed, for example, for the comparably small and simply structured Satellite tobacco mosaic virus (STMV; 50 ns with \approx 10^6 atoms at a computation speed of 0.7–1.1 ns/day) [61].

3.2 Modeling of Microcompartments

In natural systems, another widespread mechanism to minimize diffusive losses, metabolic competition, and the like in densely crowded environments such as the cytosol is the formation of subcompartments of usually only few dozen nanometers in size. With respect to the organization of membranes, this refers to functional rafts of 50 nm diameter with specialized functions [149] whose formations are highly dynamic [96]. Also within compartments, microcompartmentalization of substrates and metabolites occurs and is increasingly considered relevant [134]. The motion of large solutes is hampered by molecular crowding and sieving, whereas the mobility of small solutes is strongly restricted by binding [121]. In some cases, subcompartments and a microheterogeneity can be visualized [11] very well, for example, for local Ca^{2+} domains [29, 133]. Other examples are the regulated ADP [112] and ATP [113] diffusion within isolated brain mitochondria and cardiomyocytes.

A correct mechanistic model understanding of enzymatic reaction and metabolite transfer within or between such microcompartments consequently enforces descriptions beyond bulk ordinary differential equation systems that imply perfectly mixed compartments without diffusive barriers at "mesoscopic" scale (i.e., of some μm size), and a very high (principally infinite) number of target molecules. This has led to fundamentally different views on how to approach subcompartmentalization effects from the modeling side [10, 134]. Stochastic methods are becoming more prevalent. For example, the diffusion restrictions of macromolecules and small signaling molecules in cells have been studied with stochastic simulation [89]. Also, enzyme reaction kinetics can be described more accurately with Monte Carlo methods [66, 67] if few substrate molecules are present. For this purpose, stochastic simulation tools for biological systems have been developed, such as Stochsim [58] or MCell [32].

3.3 Interactome Network

A vast variety of multienzyme combinations has developed in nature. However, many of them form temporary metabolic channels (metabolons) between mutually communicating enzymes or enzyme cascades [99, 153, 154], which are hard to detect and isolate. It seems therefore worthwhile to scan available omics datasets systematically—especially covering the interactome [76, 137]—on a systems level for candidates forming such temporary channels. Recently, large-scale protein–protein interaction (PPI) databases have become available, indicating potential interactions between protein pairs, in some cases weighted by reliability of that interaction [42, 157]. The amount of potential interactions often exceeds 10^5; for the human interactome, it has been estimated to be around 650,000 [155]. Usually, curation based on the literature is performed to reduce the number of false positives; however, the benefit of such curations has been questioned [45].

Generally, the identification of protein complex candidates is assumed to correlate with connectivity in the interactome and can be performed with high-throughput clustering and network analysis methods by means of neighborhood analysis (e.g., ClusterONE for weighted networks; [116], MCL or molecular complex detection, MCODE, based on unweighted networks; [6], and others). Such methods may further be used for global identification of regulatory enzymes at metabolic branching points, involving metabolic channeling [75]. Also, metabolic flux rates and the degree of complexification can be estimated [54]. Furthermore, based on interactome analysis, it has been argued recently that numerous protein–protein channeling interactions do not necessarily rely on physically close location of the enzymes, instead nonenzymatic mediator proteins might be responsible for associating subsequent enzymes in a pathway [125].

The principal problem of such methods is the yield of many false positives (among a smaller amount of false negatives). A comprehensive comparison of frequently used methods of interactome network analysis with respect to robustness and sensitivities is given in Brohee and Van Helden [22].

4 Conclusions and Perspectives

In this contribution, practical and modeling aspects of understanding and exploiting the principles of compartmentalization and metabolic channeling have been discussed. It shows that many methods and approaches exist to facilitate artificial multienzymatic biosynthesis, however, the details are often not systematically understood and metabolic channeling is often simply triggered by proximity of two active sites. Because naturally evolved systems, such as tryptophane synthase or the family of stable and big 2-oxoacid dehydrogenase complexes, exhibit much more sophisticated channeling and regulation behavior, it is worthwhile to dig deeper to reveal their secrets.

It is expected that in the coming years, breakthroughs will presumably be made in the following areas: (1) development of novel and powerful algorithms for systematic identification of multienzyme complexes by integrating different networks such as metabolic and interactome networks; (2) multiscale modeling approaches from the atomic to compartment level, increasingly covering biologically relevant timescales; (3) computational design of new functional biomolecules with desired metabolic channeling; (4) development of novel nanomaterials that are capable of encapsulating multienzymes through co-immobilization; (5) discovery of new scaffolds or development of scaffold-free technologies for the construction of multienzyme complexes.

References

1. Afonin KA, Bindewald E, Yaghoubian AJ, Voss N, Jacovetty E, Shapiro BA, Jaeger L (2010) In vitro assembly of cubic RNA-based scaffolds designed in silico. Nat Nano 5(9):676–682
2. Agapakis CM, Ducat DC, Boyle PM, Wintermute EH, Way JC, Silver PA (2010) Insulation of a synthetic hydrogen metabolism circuit in bacteria. J Biol Eng 4:3
3. Aldaye FA, Palmer AL, Sleiman HF (2008) Assembling materials with DNA as the guide. Science 321(5897):1795–1799
4. Ardao I, Hwang E-T, Zeng A-P (2013) Multi-enzymatic bioreaction systems at micro/nano scale. Adv Biochem Eng/Biotechnol (Accepted)
5. Axe JM, Boehr DD (2013) Long-range interactions in the alpha subunit of tryptophan synthase help to coordinate ligand binding, catalysis, and substrate channeling. J Mol Biol 425(9):1527–1545
6. Bader G, Hogue C (2003) An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics 4(1):2
7. Badger MR, Price GD (2003) CO_2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. J Exp Bot 54(383):609–622
8. Bahar I, Lezon TR, Bakan A, Shrivastava IH (2010) Normal mode analysis of biomolecular structures: functional mechanisms of membrane proteins. Chem Rev 110(3):1463–1497
9. Balakrishnan A, Nemeria NS, Chakraborty S, Kakalis L, Jordan F (2012) Determination of pre-steady-state rate constants on the escherichia coli pyruvate dehydrogenase complex reveals that loop movement controls the rate-limiting step. J Am Chem Soc 134(45): 18644–18655
10. Barros L, Martínez C (2007) An enquiry into metabolite domains. Biophys J 92(11):3878
11. Barros LF (2010) Towards single-cell real-time imaging of energy metabolism in the brain. Front Neuroenergetics 2:4
12. Bauler P, Huber G, Leyh T, McCammon J (2010) Channeling by proximity: the catalytic advantages of active site colocalization using Brownian dynamics. J Phys Chem 1(9):1332–1335
13. Bayer EA, Belaich JP, Shoham Y, Lamed R (2004) The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. Annu Rev Microbiol 58:521–554
14. Bayer EA, Morag E, Lamed R (1994) The cellulosome—a treasure-trove for biotechnology. Trends Biotechnol 12(9):379–386
15. Bayer EA, Shimon LJ, Shoham Y, Lamed R (1998) Cellulosomes-structure and ultrastructure. J Struct Biol 124(2–3):221–234

16. Bayer TS, Widmaier DM, Temme K, Mirsky EA, Santi DV, Voigt CA (2009) Synthesis of methyl halides from biomass using engineered microbes. J Am Chem Soc 131(18): 6508–6515
17. Berger I, Blanco A, Boelens R, Cavarelli J, Coll M, Folkers G, Nie Y, Pogenberg V, Schultz P, Wilmanns M et al (2011) Structural insights into transcription complexes. J Struct Biol 175(2):135–146
18. Bernadó P, Svergun D (2011) Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. Mol BioSyst 8(1):151–167
19. Bertini I, Felli IC, Gonnelli L, Kumar MVV, Pierattelli R (2011) 13 C direct-detection biomolecular NMR spectroscopy in living cells. Angew Chem Int Ed Engl 50(10): 2339–2341
20. Betancor L, Berne C, Luckarift HR, Spain JC (2006) Coimmobilization of a redox enzyme and a cofactor regeneration system. Chem Comm 0(34):3640–3642
21. Bonacci W, Teng PK, Afonso B, Niederholtmeyer H, Grob P, Silver PA, Savage DF (2012) Modularity of a carbon-fixing protein organelle. Proc Natl Acad Sci U S A 109(2):478–483
22. Brohee S, Van Helden J (2006) Evaluation of clustering algorithms for protein–protein interaction networks. BMC Bioinformatics 7:488
23. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 4(2):187–217
24. Brooks G (2009) Cell-cell and intracellular lactate shuttles. J Physiol 587(23):5591–5600
25. Brown M (2010) DARPA plans exaflop supercomputer for 2018. Wired
26. Buch I, Giorgino T, De Fabritiis G (2011) Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. Proc Natl Acad Sci U S A 108(25):10184–10189
27. Bulow L, Ljungcrantz P, Mosbach K (1985) Preparation of a soluble bifunctional enzyme by gene fusion. Nat Biotech 3(9):821–823
28. Bülow L, Mosbach K (1991) Multienzyme systems obtained by gene fusion. Trends Biotechnol 9(1):226–231
29. Cancela JM, Van Coppenolle F, Galione A, Tepikin AV, Petersen OH (2002) Transformation of local Ca2+ spikes to global Ca2+ transients: the combinatorial roles of multiple Ca2+ releasing messengers. EMBO J 21(5):909–919
30. Cannon GC, Shively JM (1983) Characterization of a homogenous preparation of carboxysomes from Thiobacillus neapolitanus. Arch Microbiol 134(1):52–59
31. Carere J, Baker P, Seah SY (2011) Investigating the molecular determinants for substrate channeling in BphI-BphJ, an aldolase-dehydrogenase complex from the polychlorinated biphenyls degradation pathway. Biochemistry 50(39):8407–8416
32. Casanova H, Berman F, Bartol T, Gokcay E, Sejnowski T, Birnbaum A, Dongarra J, Miller M, Ellisman M, Faerman M, Obertelli G, Wolski R, Pomerantz S, Stiles J (2004) The virtual instrument: support for grid-enabled MCell simulations. Int J High Perf Comput Appl 18:3–17
33. Case DA, Cheatham TE III, Darden TA, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling CL, Wang B, Woods RJ (2005) The amber biomolecular simulation programs. J Comput Chem 26(16):1668–1688
34. Chang H-C, Kaiser CM, Hartl FU, Barral JM (2005) De novo folding of GFP fusion proteins: high efficiency in eukaryotes but not in bacteria. J Mol Biol 353(2):397–409
35. Chen AH, Silver PA (2012) Designing biological compartmentalization. Trends Cell Biol 22(12):662–670
36. Chen AH, Silver PA (2012) Designing biological compartmentalization. Trends Cell Biol 22(12):662–670
37. Chen Z, Jandt U, Rappert S, Zeng A-P (2013) Proteindesign fuer die Entwicklung von industriellen Mikroorganismen. BIOspektrum 19(1):99–101

38. Chen Z, Meyer W, Rappert S, Sun J, Zeng A (2011) Coevolutionary analysis enabled rational deregulation of allosteric enzyme inhibition in corynebacterium glutamicum for lysine production. Appl Environ Microbiol 77(13):4352
39. Chen Z, Rappert S, Sun J, Zeng A (2011) Integrating molecular dynamics and co-evolutionary analysis for reliable target prediction and deregulation of the allosteric inhibition of aspartokinase for amino acid production. J Biotechnol 154(4):248–254
40. Chen Z, Wilmanns M, Zeng A (2010) Structural synthetic biotechnology: from molecular structure to predictable design for industrial strain development. Trends Biotechnol 28(10):534–542
41. Choudhary S, Quin MB, Sanders MA, Johnson ET, Schmidt-Dannert C (2012) Engineered protein nano-compartments for targeted enzyme localization. PLoS ONE 7(3):e33342
42. Collins SR, Kemmeren P, Zhao X-C, Greenblatt JF, Spencer F, Holstege FC, Weissman JS, Krogan NJ (2007) Toward a comprehensive atlas of the physical interactome of Saccharomyces cerevisiae. Mol Cell Proteomics 6(3):439–450
43. Conrado RJ, Varner JD, DeLisa MP (2008) Engineering the spatial organization of metabolic enzymes: mimicking nature's synergy. Curr Opin Biotechnol 19(5):492–499
44. Conrado RJ, Wu GC, Boock JT, Xu H, Chen SY, Lebar T, Turnšek J, Tomšič N, Avbelj M, Gaber R, Koprivnjak T, Mori J, Glavnik V, Vovk I, Benčina M, Hodnik V, Anderluh G, Dueber JE, Jerala R, DeLisa MP (2012) DNA-guided assembly of biosynthetic pathways promotes improved catalytic efficiency. Nucleic Acids Res 40(4):1879–1889
45. Cusick ME, Yu H, Smolyar A, Venkatesan K, Carvunis A-R, Simonis N, Rual J-F, Borick H, Braun P, Dreze M et al (2008) Literature-curated protein interaction datasets. Nature 6(1):39–46
46. Delebecque CJ, Lindner AB, Silver PA, Aldaye FA (2011) Organization of intracellular reactions with rationally designed RNA assemblies. Science 333(6041):470–474
47. Dinh A-T, Pangarkar C, Theofanous T, Mitragotri S (2007) Understanding intracellular transport processes pertinent to synthetic gene delivery via stochastic simulations and sensitivity analyses. Biophys J 92(3):831–846
48. Doi RH, Kosugi A (2004) Cellulosomes: plant-cell-wall-degrading enzyme complexes. Nat Rev Microbiol 2(7):541–551
49. Doi RH, Kosugi A, Murashima K, Tamaru Y, Han SO (2003) Cellulosomes from mesophilic bacteria. J Bacteriol 185(20):5907–5914
50. Ducat DC, Sachdeva G, Silver PA (2011) Rewiring hydrogenase-dependent redox circuits in cyanobacteria. Proc Natl Acad Sci U S A 108(10):3941–3946
51. Dueber JE, Wu GC, Malmirchegini GR, Moon TS, Petzold CJ, Ullal AV, Prather KLJ, Keasling JD (2009) Synthetic protein scaffolds provide modular control over metabolic flux. Nat Biotech 27(8):753–759
52. Dunn MF (2012) Allosteric regulation of substrate channeling and catalysis in the tryptophan synthase bienzyme complex. Arch Biochem Biophys 519(2):154–166
53. Dunn M, Niks D, Ngo H, Barends T, Schlichting I (2008) Tryptophan synthase: the workings of a channeling nanomachine. Trends Biochem Sci 33(6):254–264
54. Durek P, Walther D (2008) The integrated analysis of metabolic and protein interaction networks reveals novel molecular organizing principles. BMC Syst Biol 2:100
55. El-Zahab B, Jia H, Wang P (2004) Enabling multienzyme biocatalysis using nanoporous materials. Biotechnol Bioeng 87(2):178–183
56. Fabiola F, Chapman MS et al (2005) Fitting of high-resolution structures into electron microscopy reconstruction images. Structure 13(3):389–400
57. Farhi M, Marhevka E, Masci T, Marcos E, Eyal Y, Ovadis M, Abeliovich H, Vainstein A (2011) Harnessing yeast subcellular compartments for the production of plant terpenoids. Metab Eng 13(5):474–481
58. Firth C, Bray D (2004) Stochastic simulation of cell signaling pathways. In Bower JM, Bolouri H (eds) Computational modeling of genetic and biochemical networks, pp 263–286. MIT Press, Cambridge

59. Frank RA, Titman CM, Pratap JV, Luisi BF, Perham RN (2004) A molecular switch and proton wire synchronize the active sites in thiamine enzymes. Science 306(5697):872–876
60. Frank R, Pratap J, Pei X, Perham R, Luisi B (2005) The molecular origins of specificity in the assembly of a multienzyme complex. Structure 13(8):1119–1130
61. Freddolino P, Arkhipov A, Larson S, McPherson A, Schulten K (2006) Molecular dynamics simulations of the complete satellite tobacco mosaic virus. Structure 14(3):437–449
62. Freddolino P, Liu F, Gruebele M, Schulten K (2008) Tenmicrosecond molecular dynamics simulation of a fast-folding WW domain. Biophys J 94(10):L75
63. Freddolino PL, Tavazoie S (2012) The Dawn of virtual cell biology. Cell 150(2):248–250
64. Gehrmann W, Elsner M (2011) A specific fluorescence probe for hydrogen peroxide detection in peroxisomes. Free Radic Res 45(5):501–506
65. Gerbling K-P, Steup M, Latzko E (1985) Fructose-1,6-bisphosphatase from Synechococcus leopoliensis. Eur J Biochem 147(1):207–215
66. Gibson MA, Bruck J (2000) Efficient exact stochastic simulation of chemical systems with many species and many channels. J Phys Chem 104(9):1876–1889
67. Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. J Phys Chem 81(25):2340–2361
68. Gorba C, Miyashita O, Tama F (2008) Normal-mode flexible fitting of high-resolution structure of biological molecules toward one-dimensional low-resolution data. Biophys J 94(5):1589–1599
69. Gorynia S, Bandeiras T, Pinho F, McVey C, Vonrhein C, Round A, Svergun D, Donner P, Matias P, Carrondo M et al (2011) Structural and functional insights into a dodecameric molecular machine - the RuvBL1/RuvBL2 complex. J Struct Biol 176(3):279
70. Gu Y, Shrivastava IH, Amara SG, Bahar I (2009) Molecular simulations elucidate the substrate translocation pathway in a glutamate transporter. Proc Natl Acad Sci U S A 106(8):2589–2594
71. Guo P (2010) The emerging field of RNA nanotechnology. Nat Nano 5(12):833–842
72. Hakobyan, D. and Nazaryan, K. 2010. Molecular dynamics study of interaction and substrate channeling between neuron-specific enolase and b-type phosphoglycerate mutase. Proteins Struct Funct Bioinformatics 78(7):1691–1704
73. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput 4:435–447
74. Huang X, Holden H, Raushel F (2001) Channeling of substrates and intermediates in enzyme-catalyzed reactions. Annu Rev Biochem 70(1):149–180
75. Huthmacher C, Gille C, Holzḧ tter H (2008) A computational u analysis of protein interactions in metabolic networks reveals novel enzyme pairs potentially involved in metabolic channeling. J Theor Biol 252(3):456
76. Ito T, Chiba T, Ozawa R, Yoshida M, Hattori M, Sakaki Y (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. Natl Acad Sci 98(8):4569–4574
77. Iturrate L, Sanchez-Moreno I, Doyaguez EG, Garcia-Junceda E (2009) Substrate channelling in an engineered bifunctional aldolase/kinase enzyme confers catalytic advantage for C–C bond formation. Chem Commun 2009(13):1721–1723
78. Izard T, Aevarsson A, Allen MD, Westphal AH, Perham RN, de Kok A, Hol WGJ (1999) Principles of quasi-equivalence and Euclidean geometry govern the assembly of cubic and dodecahedral cores of pyruvate dehydrogenase complexes. Proc Natl Acad Sci U S A 96(4):1240–1245
79. Jandt U, Zeng A-P (2012) Modeling of intracellular transport and compartmentation. In: Hu W-S, Zeng A-P (eds) Advances in biochemical engineering and biotechnology: genomics and systems biology of mammalian cell culture, pp 221–249
80. Jiang L, Althoff EA, Clemente FR, Doyle L, Röthlisberger D, Zanghellini A, Gallaher JL, Betker JL, Tanaka F, Barbas CF et al (2008) De novo computational design of retro-aldol enzymes. Science 319(5868):1387–1391

81. Joshi H, Singharoy A, Sereda Y, Cheluvaraja S, Ortoleva P (2011) Multiscale simulation of microbe structure and dynamics. Prog Biophys Mol Biol 107(1):200–217
82. Kale S, Ulas G, Song J, Brudvig GW, Furey W, Jordan F (2008) Efficient coupling of catalysis and dynamics in the e1 component of escherichia coli pyruvate dehydrogenase multienzyme complex. Proc Natl Acad Sci U S A 105(4):1158–1163
83. Karr JR, Sanghvi JC, Macklin DN, Gutschow MV, Jacobs JM, Bolival B, Assad-Garcia N, Glass JI, Covert MW (2012) A whole-cell computational model predicts phenotype from genotype. Cell 150(2):389–401
84. Kerfeld CA, Heinhorst S, Cannon GC (2010) Bacterial microcompartments. Annu Rev Microbiol 64:391–408
85. Khalili-Araghi F, Gumbart J, Wen P-C, Sotomayor M, Tajkhorshid E, Schulten K (2009) Molecular dynamics simulations of membrane channels and transporters. Curr Opin Struct Biol 19(2):128–137
86. Kiefer F, Arnold K, Künzli M, Bordoli L, Schwede T (2009) The u SWISS-MODEL repository and associated resources. Nucleic Acids Res 37(suppl 1):D387
87. Kim MI, Kim J, Lee J, Jia H, Na HB, Youn JK, Kwak JH, Dohnalkova A, Grate JW, Wang P, Hyeon T, Park HG, Chang HN (2007) Crosslinked enzyme aggregates in hierarchically-ordered mesoporous silica: a simple and effective method for enzyme stabilization. Biotechnol Bioeng 96(2):210–218
88. Kirchmair J, Distinto S, Markt P, Schuster D, Wolber G (2011) Development of anti-viral agents using molecular modeling and virtual screening techniques. Infectious Disorders-Drug Targets 11(1):64
89. Klann MT, Lapin A, Reuss M (2009) Stochastic simulation of signal transduction: impact of the cellular architecture on diffusion. Biophys J 96(12):5122–5129
90. Klein MG, Zwart P, Bagby SC, Cai F, Chisholm SW, Heinhorst S, Cannon GC, Kerfeld CA (2009) Identification and structural analysis of a novel carboxysome shell protein with implications for metabolite transport. J Mol Biol 392(2):319–333
91. Klepeis J, Lindorff-Larsen K, Dror R, Shaw D (2009) Longtimescale molecular dynamics simulations of protein structure and function. Curr Opin Struct Biol 19(2):120–127
92. Lai Y-T, King NP, Yeates TO (2012) Principles for designing ordered protein assemblies. Trends Cell Biol 22(12):653–661
93. Larsson P, Hess B, Lindahl E (2011) Algorithm improvements for molecular dynamics simulations. Wiley Interdiscip Rev Comput Mol Sci 1(1):93–108
94. Lee H, DeLoache WC, Dueber JE (2012) Spatial organization of enzymes for metabolic engineering. Metab Eng 14(3):242–251
95. Lee SK, Chou H, Ham TS, Lee TS, Keasling JD (2008) Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels. Curr Opin Biotechnol 19(6):556–563
96. Lingwood D, Ilya Levental H-JK, Simons K (2009) Lipid rafts as functional heterogeneity in cell membranes. Biochem Soc Trans 37(Pt 5):955–960
97. Ljungcrantz P, Carlsson H, Mansson MO, Buckel P, Mosbach K, Bulow L (1989) Construction of an artificial bifunctional enzyme, beta-galactosidase/galactose dehydrogenase, exhibiting efficient galactose channeling. Biochemistry 28(22):8786–8792
98. Loeffler HH, Winn M (2012) Large biomolecular simulation on HPC platforms III. AMBER, CHARMM, GROMACS, LAMMPS and NAMD. Technical report, STFC Daresbury Laboratory, Warrington WA4 4AD, UK
99. Lyubarev A, Kurganov B (1989) Supramolecular organization of tricarboxylic acid cycle enzymes. Biosystems 22(2):91–102
100. Ma C, Xiu Z, Zeng A (2011) A new concept to reveal protein dynamics based on energy dissipation. PLoS ONE 6(10):e26453
101. Maes D, Zeelen JP, Thanki N, Beaucamp N, Alvarez M, Thi MH, Backmann J, Martial JA, Wyns L, Jaenicke R, Wierenga RK (1999) The crystal structure of triosephosphate isomerase (TIM) from Thermotoga maritima: a comparative thermostability structural analysis of ten different TIM structures. Proteins 37(3):441–453

102. Mak L, Liggi S, Tan L, Kusonmano K, Rollinger JM, Koutsoukas A, Glen RC, Kirchmair J (2013) Anti-cancer drug development: computational strategies to identify and target proteins involved in cancer metabolism. Curr Pharm Design 19(4):532–577
103. Mansson MO, Siegbahn N, Mosbach K (1983) Site-to-site directed immobilization of enzymes with bis-NAD analogues. Proc Natl Acad Sci U S A 80(6):1487–1491
104. Mao Q, Schunk T, Gerber B, Erni B (1995) A String of Enzymes, Purification and Characterization of a Fusion Protein Comprising the Four Subunits of the Glucose Phosphotransferase System of Escherichia coli. J Biol Chem 270(31):18295–18300
105. Martínez de la Fuente I (2010) Quantitative analysis of cellular metabolic dissipative, self-organized structures. Int J Mol Sci 11(9):3540–3599
106. Martínez de la Fuente I, Cortes JM, Perez-Pinilla MB, Ruiz-Rodriguez V, Veguillas J (2011) The metabolic core and catalytic switches are fundamental elements in the self-regulation of the systemic metabolic structure of cells. PLoS ONE 6(11):e27224
107. McLain AL, Szweda PA, Szweda LI (2011) Ketoglutarate dehydrogenase: a mitochondrial redox sensor. Free Radic Res 45(1):29–36
108. Mei C, Sun Y, Zheng G, Bohm E, Kale L, Phillips J, Harrison C (2011) Enabling and scaling biomolecular simulations of 100 million atoms on petascale machines with a multicore-optimized message-driven runtime. In: High performance computing, networking, storage and analysis (SC), 2011 International conference for, pp 1–11. IEEE
109. Meynial Salles I, Forchhammer N, Croux C, Girbal L, Soucaille P (2007) Evolution of a Saccharomyces cerevisiae metabolic pathway in Escherichia coli. Metab Eng 9(2):152–159
110. Ming D, Kong Y, Lambert MA, Huang Z, Ma J (2002) How to describe protein motion without amino acid sequence and atomic coordinates. Proc Natl Acad Sci U S A 99(13):8620–8625
111. Miyazaki H, Kusano Y, Shinjou N, Shoji F, Yokokawa M, Watanabe T (2012) Overview of the K computer system. FUJITSU Sci Tech J 48(3):255–265
112. Monge C, Beraud N, Kuznetsov AV, Rostovtseva T, Sackett D, Schlattner U, Vendelin M, Saks VA (2008) Regulation of respiration in brain mitochondria and synaptosomes: restrictions of ADP diffusion in situ, roles of tubulin, and mitochondrial creatine kinase. Mol Cell Biochem 318(1):147–165
113. Monge C, Grichine A, Rostovtseva T, Sackett D, Kuznetsov A, Saks V (2009) Compartmentation of ATP in cardiomyocytes and mitochondria kinetic studies and direct measurements. Biophys J 96:241
114. Moon TS, Dueber JE, Shiue E, Prather KLJ (2010) Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered E. coli. Metab Eng 12(3):298–305
115. Mosbach K, Mattiasson B (1970) Matrix-bound enzymes. II. Studies on a matrix-bound two-enzyme-system. Acta Chem Scand 24(6):2093–2100
116. Nepusz T, Yu H, Paccanaro A (2012) Detecting overlapping protein complexes in protein–protein interaction networks. Nat Methods 9(5):471–472
117. Netzer WJ, Hartl FU (1997) Recombination of protein domains facilitated by co-translational folding in eukaryotes. Nature 388(6640):343–349
118. Niemeyer CM, Koehler J, Wuerdemann C (2002) DNA-directed assembly of bienzymic complexes from in vivo biotinylated NAD(P)H:FMN oxidoreductase and luciferase. ChemBioChem 3(2–3):242–245
119. Noble D (2010) Biophysics and systems biology. Philos Transact A Math Phys Eng Sci 368(1914):1125–1139
120. Orita I, Sakamoto N, Kato N, Yurimoto H, Sakai Y (2007) Bifunctional enzyme fusion of 3-hexulose-6-phosphate synthase and 6-phospho-3-hexuloisomerase. Appl Microbiol Biotechnol 76(2):439–445
121. Ovádi J, Saks V (2004) On the origin of intracellular compartmentation and organized metabolic systems. Mol Cell Biochem 256–257(1–2):5–12
122. Parsons JB, Frank S, Bhella D, Liang M, Prentice MB, Mulvihill DP, Warren MJ (2010) Synthesis of empty bacterial microcompartments, directed organelle protein incorporation, and evidence of filament-associated organelle movement. Mol Cell 38(2):305–315

123. Patel M, Korotchkina L (2006) Regulation of the pyruvate dehydrogenase complex. Biochem Soc Trans 34(2):217–222
124. Penrod JT, Roth JR (2006) Conserving a volatile metabolite: a role for carboxysome-like organelles in Salmonella enterica. J Bacteriol 188(8):2865–2874
125. Pérez-Bercoff A, McLysaght A, Conant GC (2011) Pate e terns of indirect protein interactions suggest a spatial organization to metabolism. Mol BioSyst 7(11):3056–3064
126. Petoukhov MV, Svergun DI (2005) Global rigid body modeling of macromolecular complexes against small-angle scattering data. Biophys J 89(2):1237–1250
127. Pettersson H, Pettersson G (2001) Kinetics of the coupled reaction catalysed by a fusion protein of (beta)-galactosidase and galactose dehydrogenase. Biochim Biophys Acta: Protein Struct Mol Enzymol 1549(2):155–160
128. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K (2005) Scalable molecular dynamics with NAMD. J Comp Chem 26(16):1781–1802
129. Qu Z, Garfinkel A, Weiss J, Nivala M (2011) Multi-scale modeling in biology: how to bridge the gaps between scales? Prog Biophys Mol Biol 107(1):21–31
130. Ren Y, He J, Feng L, Liao X, Jin J, Li Y, Cao Y, Wan J, He H (2011) Structure-based rational design of novel hit compounds for pyruvate dehydrogenase multienzyme complex E1 components from Escherichia coli. Bioorg Med Chem 19(24):7501–7506
131. Resch MM (2013) Beyond exaflop computing: reaching the frontiers of high performance computing. In: Sust Sim Perf 2012, pp 3–12. Springer
132. Riedel K, Bronnenmeier K (1998) Intramolecular synergism in an engineered exo-endo-1,4-β-glucanase fusion protein. Mol Microbiol 28(4):767–775
133. Rizzuto R, Pozzan T (2006) Microdomains of intracellular Ca2+: molecular determinants and functional consequences. Physiol Rev 86(1):369–408
134. Saks V, Beraud N, Wallimann T (2008) Metabolic compartmentation- a system level property of muscle cells: real problems of diffusion in living cells. Int J Mol Sci 9(5):751
135. Saks V, Monge C, Guzun R (2009) Philosophical basis and some historical aspects of systems biology: from Hegel to noble-applications for bioenergetic research. Int J Mol Sci 10(3):1161–1192
136. Sampson EM, Bobik TA (2008) Microcompartments for B12-dependent 1,2-propanediol degradation provide protection from DNA and cellular damage by a reactive metabolic intermediate. J Bacteriol 190(8):2966–2971
137. Sanchez C, Lachaize C, Janody F, Bellon B, Röder L, Euzenat J, Rechenmann F, Jacq B (1999) Grasping at molecular interactions and genetic networks in Drosophila melanogaster using FlyNets, an internet database. Nucleic Acids Res 27(1):89–94
138. Savage DF, Way J, Silver PA (2008) Defossiling fuel: how synthetic biology can transform biofuel production. ACS Chem Biol 3(1):13–16
139. Schaeffer RD, Fersht A, Daggett V (2008) Combining experiment and simulation in protein folding: closing the gap for small model systems. Curr Opin Struct Biol 18(1):4–9
140. Schulz R, Lindner B, Petridis L, Smith J (2009) Scaling of multimillion-atom biological molecular dynamics simulation on a petascale supercomputer. J Chem Theor Comp 5(10):2798–2808
141. Selivanov V, Krause S, Roca J, Cascante M (2007) Modeling of spatial metabolite distributions in the cardiac sarcomere. Biophys J 92(10):3492–3500
142. Seo HS, Koo YJ, Lim JY, Song JT, Kim CH, Kim JK, Lee JS, Choi YD (2000) Characterization of a bifunctional enzyme fusion of trehalose-6-phosphate synthetase and trehalose-6-phosphate phosphatase of Escherichia coli. Appl Environ Microbiol 66(6):2484–2490
143. Shan Y, Klepeis J, Eastwood M, Dror R, Shaw D (2005) Gaussian split Ewald: a fast Ewald mesh method for molecular simulation. J Chem Phys 122:054101
144. Sherawat M, Tolan DR, Allen KN (2008) Structure of a rabbit muscle fructose-1,6-bisphosphate aldolase A dimer variant. Acta Cryst D 64(5):543–550

145. Shibuya I, Tamura G, Shima H, Ishikawa T, Hara S (1992) Construction of an α-amylase/glucoamylase fusion gene and its expression in Saccharomyces cerevisiae. Biosci Biotechnol Biochem 56(6):884–889
146. Shoham Y, Lamed R, Bayer EA (1999) The cellulosome concept as an efficient microbial strategy for the degradation of insoluble polysaccharides. Trends Microbiol 7(7):275–281
147. Siddiqui MS, Thodey K, Trenchard I, Smolke CD (2012) Advancing secondary metabolite biosynthesis in yeast with synthetic biology tools. FEMS Yeast Res 12(2):144–170
148. Siegel JB, Zanghellini A, Lovick HM, Kiss G, Lambert AR, Clair JLS, Gallaher JL, Hilvert D, Gelb MH, Stoddard BL et al (2010) Computational design of an enzyme catalyst for a stereoselective bimolecular diels-alder reaction. Science 329(5989):309–313
149. Simons K, Ikonen E (1997) Functional rafts in cell membranes. Nature 387:569–572
150. Slepchenko BM, Schaff JC, Macara I, Loew LM (2003) Quantitative cell biology with the virtual cell. Trends Cell Biol 13(11):570–576
151. Ślepokura K, Lis T (2010) Dihydroxyacetone phosphate, DHAP, in the crystalline state: monomeric and dimeric forms. Carbohyd Res 345(4):512–529
152. Srere PA, Mattiasson B, Mosbach K (1973) An immobilized three-enzyme system: a model for microenvironmental compartmentation in mitochondria. Proc Natl Acad Sci U S A 70(9):2534–2538
153. Srere P (1987) Complexes of sequential metabolic enzymes. Annu Rev Biochem 56(1):89–124
154. Srere P, Ovadi J (1990) Enzyme-enzyme interactions and their metabolic role. FEBS Lett 268(2):360–364
155. Stumpf MP, Thorne T, De Silva E, Stewart R, An HJ, Lappe M, Wiuf C (2008) Estimating the size of the human interactome. PNAS 105(19):6959–6964
156. Tanaka S, Kerfeld CA, Sawaya MR, Cai F, Heinhorst S, Cannon GC, Yeates TO (2008) Atomic-level models of the bacterial carboxysome shell. Science 319(5866):1083–1086
157. Tarassov K, Messier V, Landry CR, Radinovic S, Molina MMS, Shames I, Malitskaya Y, Vogel J, Bussey H, Michnick SW (2008) An in vivo map of the yeast protein interactome. Science 320(5882):1465–1470
158. Tomita M (2001) Whole-cell simulation: a grand challenge of the 21st century. Trends Biotechnol 19(6):205–210
159. Tran L, Broadhurst RW, Tosin M, Cavalli A, Weissman KJ (2010) Insights into protein–protein and enzyme-substrate interactions in modular polyketide synthases. Chem Biol 17(7):705–716
160. van de Velde F, Lourenço ND, Bakker M, van Rantwijk F, Sheldon RA (2000) Improved operational stability of peroxidases by coimmobilization with glucose oxidase. Biotechnol Bioeng 69(3):286–291
161. van Dongen SFM, Nallani M, Cornelissen JJLM, Nolte RJM, van Hest JCM (2009) A three-enzyme cascade reaction through positional assembly of enzymes in a polymersome nanoreactor. Chem Eur J 15(5):1107–1114
162. Vijayakrishnan S, Kelly S, Gilbert R, Callow P, Bhella D, Forsyth T, Lindsay J, Byron O (2010) Solution structure and characterisation of the human pyruvate dehydrogenase complex core assembly. J Mol Biol 399(1):71–93
163. Vriezema DM, Garcia PML, Sancho Oltra N, Hatzakis NS, Kuiper SM, Nolte RJM, Rowan AE, van Hest JCM (2007) Positional assembly of enzymes in polymersome nanoreactors for cascade reactions. Angew Chem Int Ed 46(39):7378–7382
164. Walther TH, Gottselig C, Grage SL, Wolf M, Vargiu AV, Klein MJ, Vollmer S, Prock S, Hartmann M, Afonin S et al (2013) Folding and self-assembly of the tata translocation pore based on a charge zipper mechanism. Cell 152(1):316–326
165. Wei B, Dai M, Yin P (2012) Complex shapes self-assembled from single-stranded DNA tiles. Nature 485(7400):623–626
166. Welch G, Easterby J (1994) Metabolic channeling versus free diffusion: transition-time analysis. Trends Biochem Sci 19(5):193

167. Westerhoff HV, Kolodkin A, Conradie R, Wilkinson SJ, Bruggeman FJ, Krab K, van Schuppen JH, Hardin H, Bakker BM, Moné MJ et al (2009) e Systems biology towards life in silico: mathematics of the control of living cells. J Math Biol 58(1):7–34
168. Wörsdörfer B, Woycechowsky KJ, Hilvert D (2011) Directed evolution of a protein container. Science 331(6017):589–592
169. You C, Myung S, Zhang YHP (2012) Facilitated substrate channeling in a self-assembled trifunctional enzyme complex. Angew Chem Int Edit 51(35):8787–8790
170. You C, Zhang XZ, Sathitsuksanoh N, Lynd LR, Zhang Y-HP (2012) Enhanced microbial utilization of recalcitrant cellulose by an ex vivo cellulosome-microbe complex. Appl Environ Microbiol 78(5):1437–1444
171. You C, Zhang X-Z, Zhang Y-HP (2012) Mini-scaffoldin enhanced mini-cellulosome hydrolysis performance on low-accessibility cellulose (Avicel) more than on high-accessibility amorphous cellulose. Biochem Eng J 63:57–65
172. You C, Zhang YHP (2012) Self-assembly of synthetic metabolons through synthetic protein scaffolds: one-step purification, co-immobilization, and substrate channeling. ACS Synth Biol 2:102–110
173. Zhang Y-HP (2010) Production of biocommodities and bioelectricity by cell-free synthetic enzymatic pathway biotransformations: challenges and opportunities. Biotechnol Bioeng 105(4):663–677
174. Zhang YHP (2011) Substrate channeling and enzyme complexes for biotechnological applications. Biotechnol Adv 29(6):715–725
175. Zhang YHP, Myung S, You C, Zhu Z, Rollin JA (2011) Toward low-cost biomanufacturing through in vitro synthetic biology: bottom-up design. J Mater Chem 21(47):18877–18886