Section III

Prospects for Future Development of Biofuels
13

New Tools for Bioprocess Analysis and Optimization of Microbial Fuel Production

Isabelle France George, Philippe Bogaerts, Dimitri Gilis, Marianne Rooman, and Jean-François Flot

Contents

13.1 State-of-the-Art in Biofuel Production by Microbial Cell Factories 427
13.2 Targeted Approach 1: The Rational Design of Optimized Enzymes for Fuel Production 433
13.3 Targeted Approach 2: The Rational Design of Tolerant Strains for Fuel Production 436
13.4 Targeted Approach 3: Analysis and Engineering of Multispecies Microbial Communities Involved in Fuel Production 439
   13.4.1 Study of Community Structure and Dynamics, and Exploitation of Community Properties in the Context of Biofuel Production 440
   13.4.1.1 Why Analyze Community Structure and Dynamics? 440
   13.4.1.2 Tools to Analyze Community Structure and Dynamics and How They Contribute to a Better Understanding of Biofuel-Producing Communities 442
   13.4.2 Management and Genetic Engineering of Microbial Consortia 454
13.5 Targeted Approach 4: Microbial Electrosynthesis of Biofuels 458
13.6 Targeted Approach 5: Dynamic Macroscopic Models of Bioprocesses and Their Application to Microbial Fuel Production 465
   13.6.1 Introduction 465
   13.6.2 Dynamic Macroscopic Models of Bioprocesses and General Kinetic Model Structures 466
   13.6.3 Case Study 1: Bioethanol Production from Potato Peel Wastes 467
   13.6.4 Case Study 2: Microalgae Cultures 470
13.7 Conclusion 473
References 474

13.1 State-of-the-Art in Biofuel Production by Microbial Cell Factories

Biofuels, that is, nonconventional liquid and gaseous fuels derived from renewable sources, such as crop plants, forest products, algae, or waste materials, are widely promoted as a sustainable alternative to fossil fuels and a means to secure our energy supply (Tilman et al. 2009). Other potential benefits of biofuel production are the creation of new, local employment (e.g., in rural areas) and the reduction of emissions of greenhouse gases (Fargione et al. 2008; Duke et al. 2013). A broader definition of
biofuels (that we will use in this chapter) includes fuels produced from other renewable sources, such as carbon dioxide (CO₂), exploited by biodiesel- or hydrogen (H₂)-producing phototrophic organisms.

Presently, thanks to their great metabolic diversity, bacteria, fungi, and microalgae are largely involved in the production of several biofuels, and they are expected to play an even greater role in the future. Most biofuel production processes rely on pure strains or their enzymes.

The most obvious example is first-generation bioethanol produced by fermentation of sugar beet, sugarcane, and corn starch by the yeasts *Saccharomyces cerevisiae* or *Zymomonas mobilis* (Goldemberg 2007; Sánchez and Cardona 2008). Second-generation bioethanol from nonfood feedstocks (wood, sawdust, perennial plants like switchgrass or *Miscanthus*, agricultural residues, or municipal waste [Sánchez and Cardona 2008]) is more challenging to produce because of their lignocellulosic nature (except the last one). Therefore, their conversion requires a combination of physico-chemical and biological treatments to deconstruct lignocellulosic biomass into sugars that can be transformed into end products (Lynd et al. 2002). The standard procedure consists of a physical, chemical, or combined pretreatment to make cellulose and hemicellulose more accessible to hydrolytic enzymes, followed by the biological breakdown of (hemi-)cellulose by multienzyme cocktails, in which cellulase enzymes are the major components (Medie et al. 2012). The fungus *Trichoderma reesei* is the current gold standard for production of commercialized cellulases, but competitive fungal cellulase producers like *Myceliophthora thermophila* have appeared on the market (Gusakov 2013). Finally, monomeric hexose (and pentose) sugars are transformed by the yeast *S. cerevisiae*, the bacterium *Z. mobilis*, or genetically modified *Escherichia coli* into desirable products: mostly ethanol, but also longer-chain alcohols, alkanes, and fatty acid esters (reviewed in Liao et al. 2016). The aforementioned organisms are utilized primarily because (1) their physiology and metabolism have been extensively studied and (2) dedicated genetic engineering toolkits have been developed. As a possible exception, *S. cerevisiae* does have several traits that are advantageous, such as its high fermentative capacity, natural ethanol tolerance, and ability to grow at acidic pH; however, it remains incapable of fermenting C₅ sugars and of surviving at temperatures optimal for exogenous hydrolytic enzymes (Akinosho et al. 2014). Last but not least, lignocellulose degradation and sugar fermentation can be performed by pure bacterial cultures, among which the most exploited strain is thermophilic *Clostridium thermocellum*, which produces an extracellular enzymatic complex—the cellulosome—with remarkable enzymatic hydrolysis efficiency compared with that of free cellulases (Akinosho et al. 2014; Scully and Örlygsson 2014). Unfortunately, it can only tolerate a low ethanol concentration (maximum 5 g/L) and is unable to ferment C₅ sugars.

Production of biofuels needs to be very efficient because of their low unit price. As metabolic pathways in natural microorganisms turn out to be insufficient for conversion of cellullosic biomass to a level that is compatible with industrial applications (i.e., titer > 40 g/L, yield > 90% of the theoretical maximum, and rate > 1 g/L/h [Dien et al. 2003]), they have been optimized or redesigned to improve production efficiency, either by genetic engineering (i.e., the manipulation of an organism’s genome to alter its traits) or by metabolic engineering (i.e., the modification of an organism’s
metabolic network to produce a specific compound) via the optimization of their genes and regulatory processes (Stephanopoulos et al. 1998; Nielsen 2003; Kim et al. 2013). Recent breakthroughs in the high-throughput sequencing of genomes and genome-editing tools have accelerated our capability to modify genetic information (Esvelt and Wang 2013). Recently, the potential of metabolic engineering has been boosted by coupling it with microfluidics, which allows high-throughput selection of the most promising individuals in genetically modified microbial populations (Wang et al. 2014; Huang et al. 2015). So far, microbial engineering for the production of biofuels has largely consisted of overexpressing the desired pathway genes and knocking out competing pathways. Biofuel production pathways have even been introduced in microorganisms that do not natively ferment sugars or hydrolyze cellulose, but whose genetic and metabolic pathways are better known and easier to manipulate, such as *E. coli*, *Lactobacillus* sp., *Pseudomonas putida*, and *Bacillus subtilis* (Liu and Khosla 2010; Zhang et al. 2016). These approaches and more advanced ones based on the artificial regulation of pathways and creation of new pathways have been reviewed elsewhere (Liao et al. 2016).

Microorganisms have also been engineered to perform both cellulose deconstruction and fermentation of the resultant released sugars, a process called consolidated bioprocessing (CBP) (Lynd et al. 2005; Schuster and Chinn 2012; Hasunuma et al. 2013; Mbaneme and Chinn 2015; Salehi Jouzani and Taherzadeh 2015). This way, inhibition of cellulose breakdown by accumulation of its intermediary and final products is avoided. CBP has been implemented either by making efficient cellulose-degrading bacteria like *C. thermocellum* able to produce ethanol (Akinosho et al. 2014; Chung et al. 2014; Huang et al. 2014) or by transferring cellulase-encoding genes into standard strains used for ethanol production, like *S. cerevisiae* (Tsai et al. 2009; Hasunuma and Kondo 2012). Although highly promising, these engineered microbes still harbor insufficient cellulose-to-ethanol conversion yields, and therefore they have not been used for industrial ethanol production yet. *C. thermocellum* is, however, a good CBP candidate, provided that its ethanol yield can be improved (Akinosho et al. 2014).

A lot of effort has also been put into the biological conversion of sugars into various biofuels beyond ethanol. In this respect, bacterial and fungal strains (mostly *E. coli* [Clomburg and Gonzalez 2010], but also *B. subtilis*, *C. thermocellum*, *S. cerevisiae*, *P. putida*, and *Yarrowia lipolytica*) have been engineered to convert sugars into isobutanol (an alternative to ethanol with higher energy content for blending with gasoline or for transformation into kerosene), fatty acids, or the isoprenoid hydrocarbon farnesene (reviewed in Hasunuma and Kondo 2012; Liao et al. 2016).

Four decades of research has strived to produce cheap enzymes and develop robust engineered microorganisms able to operate at high rate and/or at high temperature, to cope with complex hydrolysate streams, to use all sugars produced from (hemi-)cellulose degradation (and not only hexoses), to maintain high viability, and to be resistant to multiple stresses (such as inhibition by high product [ethanol] concentrations, high acid concentrations or high temperature, or contamination by bacteria or phages). However, despite these advances, the process of lignocellulosic biomass conversion to biofuels—primarily bioethanol—remains commercially unviable (Kricka et al. 2015). In the remainder of this chapter, we discuss some of the ways to
circumvent the current obstacles standing in the way of converting efficiently complex biomass into biofuels.

Last but not least, the integration of chemical catalysis into biological catalysis has opened up the possibility to transform products of sugar fermentations other than ethanolic fermentation. For example, acetone and 1-butanol, resulting from solvento-genic acetone-butanol-ethanol (ABE) fermentation (Ezeji et al. 2007), or acetate from acetogenic fermentation, can be converted into long-chain alkanes for use as gasoline, jet fuel, and diesel (Anbarasan et al. 2012; Sreekumar et al. 2015), or into alcohols, acids, esters, or olefins, respectively. Such flexible approaches allowing the synthesis of a variety of bio-based products, in addition to low-value biofuels, are necessary to make biorefineries economically viable (Bozell and Petersen 2010).

Another high-profile example of microbial conversion of a nonfood substrate into biofuel is the production of oil by eukaryotic microalgae or cyanobacteria. Using sunlight, CO₂, and water, they produce biomass, including triglycerides and lipids (up to 70% lipid per dry weight), that can be converted into biodiesel, or sugars that can be fermented into other products. Genetic or metabolic engineering has allowed us to increase the lipid storage capacity in a handful of algal strains, or to expand the range of molecules produced by cyanobacteria (liquid fuels like ethanol, isobutanol, and isopropanol, but also hydrogen, bioplastics, or commodity chemicals). In addition, algal biomass can be converted into biogas through anaerobic fermentation. Algal cultures in ponds or photobioreactors offer great advantages over agricultural oleaginous crops, primarily a higher lipid productivity per ground area, as well as lack of competition for arable land, therefore avoiding the “food or fuel” dilemma (Wijffels and Barbosa 2010). In particular, photobioreactors, although more expensive to set up and operate, produce a much more concentrated algal broth than do open ponds, which decreases the cost of dewatering substantially (Chisti and Yan 2011). Much effort has been devoted to developing pure strains with characteristics of an “ideal” photosynthetic cell factory: high photosynthesis efficiency, ability to grow and produce oil at the same time, excretion of oil outside the cells, and insensitivity to oxygen concentration (Wijffels and Barbosa 2010). A good ability to tolerate heat and salts is desired as well, as many regions of nonarable land where algal growth could extend are arid and have no surface freshwaters, but alkaline or saline water reservoirs beneath them (Hannon et al. 2010). In that respect, cyanobacteria, whose genome is smaller and more easily edited than eukaryotic microalgae, hold great promise for the production of various fuels and chemicals (Angermayr et al. 2015; Liao et al. 2016). However, production of microalgae for biofuels is presently low scale and expensive compared with that of conventional energy crops, and it is therefore not viable at an industrial scale yet (Lardon et al. 2009; Chisti and Yan 2011). Life cycle assessment studies suggest that algal fuels may be better than fossil fuels, but there are no clear answers in this emerging field (Chisti and Yan 2011). Despite large research programs focused on developing microalgal energy production systems since the late 1970s, substantial improvements in productivity are still needed. They are expected to arise in the next 10–15 years from technological improvements (bioreactor design, nutrient and light regimen control, nutrient recycling, biomass harvesting, extraction, and exploitation of all these ingredients into a biorefinery...
infrastructure) and strain improvements (detailed physiological and genomic characterization, dedicated genetic engineering toolkits, and/or exploitation of new strains) (Brennan and Owende 2010; Hannon et al. 2010; Wijffels and Barbosa 2010). With regard to strain improvement, a great step forward is expected with the development of algal species engineered to use atmospheric nitrogen instead of petroleum-derived nitrogen fertilizers (Chisti and Yan 2011). A low-tech, recently revisited alternative is to grow algae in wastewater, which is rich in CO₂ and nutrients, thereby coupling algal biomass production with wastewater treatment (Brennan and Owende 2010).

The two aforementioned examples illustrate the longtime effort devoted to develop liquid biofuels, which offer the advantages of high energy density and compatibility with current vehicle motors. In parallel, gaseous biofuels are of interest as well, either because they already have a market (as in the case of biogas, a mixture of methane [CH₄] and CO₂, with applications similar to natural gas) or because they are expected to find one in the middle-term future (as for molecular hydrogen).

Biogas is produced by complex microbial communities that digest agricultural and municipal organic wastes. In the first stage, organic matter is hydrolyzed and fermented by anaerobic bacteria into acetate, H₂, CO₂, alcohols, and volatile fatty acids, such as propionate or butyrate. The higher volatile fatty acids are further converted into acetate, H₂, and CO₂, by anaerobic acetogenic bacteria. In the second stage, the three latter molecules are transformed into methane and CO₂ by acetotrophic or hydrogenotrophic methane-producing archaea. Both stages need to be balanced in terms of degradation rate: the temperature, pH, ammonia concentration, and retention time must be carefully controlled (Weiland 2010). The by-product of this treatment, the digester residue, can be used as fertilizer. Globally speaking, biogas production has been evaluated as one of the most energy-efficient and environmentally friendly technologies for bioenergy production (Weiland 2010). However, despite a remarkable development of farm-scale or domestic biogas plants and large-scale codigestion plants (where biogas is produced from a mixture of substrates expected to provide a more balanced nutrient pool to microorganisms), more research is still needed to achieve the full potential of biogas as an alternative to fossil fuels (Sárvári Horváth et al. 2016). Indeed, it is mainly used for generating heat and electricity, but its use as a vehicle fuel or as an alternative to natural gas in the grid requires its conversion into a liquid fuel like biodiesel or its upgrade to biomethane, respectively. Conversion of CH₄ to biodiesel is possible—although far from commercially viable—using a group of bacteria called methanotrophs. They use methane as their sole carbon source and convert methane into cellular compounds, among which are biolipids (Strong et al. 2015). Upgrade to near-pure biomethane is hampered by the CH₄ content of biogas, which is highly dependent on the characteristics of feedstocks and the anaerobic microbial community, and can be as low as 50% for carbohydrate substrates. Therefore, the worldwide expansion of biogas production will require technological and biological improvements. The former include new technologies for the pretreatment of recalcitrant substrates, mixing, retention of slow-growing methanogens, process monitoring, and process control. The latter rely on a better understanding of the microbial communities responsible for biogas production. The basic metabolisms in anaerobic digestion processes have been known for decades, but it is only recently
(during the late 1990s) that the great diversity of microbes at play has been progressively unraveled. Such studies have shed light on the highly dynamic nature of communities involved in biogas production, as discussed in Section 13.4.1.2.

Another promising fuel is molecular hydrogen, a nearly ideal energy carrier because of its high energy content, high efficiency of conversion to usable power, and nonpolluting oxidation product (water). Moreover, it has wider industrial applications than CH₄ (such as the synthesis of ammonia and hydrogenation of fossil fuels). Biohydrogen has the potential to replace current energy-intensive hydrogen production technologies based on fossil fuels (Elsharnouby et al. 2013). Four biological pathways can lead to hydrogen production: (1) biophotolysis by green microalgae (where solar energy is directly converted to H₂ via photosynthetic reactions), (2) photofermentation by anoxic phototrophs (where organic substrates like fatty acids or alcohols are converted to H₂ and CO₂ thanks to the energy of sunlight), (3) dark fermentation by heterotrophs (where organic substrates are converted to simpler substrates and CO₂ in anaerobiosis), and (4) microbial electrolysis cells (a very recent technology discussed in Section 13.5). The use of biophotolysis is challenged by the low transfer efficiency of light, complexity in reactor design, and low hydrogen production rates. Harnessing photofermentation to produce H₂ from organic acids is hindered by its need for light, by its low solar conversion efficiency, by the sensitivity of the nitrogenase enzyme (responsible for H₂ production) to ammonia and oxygen, and by its high ATP demand. As far as we know, photofermentation has been mostly tested in lab reactors inoculated with pure strains; therefore, its efficiency when upscaled to industrial bioreactors remains uncertain. In contrast, hydrogen production using dark fermentation is flexible in regard to substrate range, cheaper, and easier to implement, but several parameters need to be considered to ensure a high hydrogen yield and production rate: (1) the inoculum composition (as end products will depend on the bacterial metabolisms involved in the process); (2) the substrate complexity and biodegradability; and (3) operational bioreactor parameters, such as pH (pH < 4.5 negatively affects hydrogen production), nutrient concentration, partial H₂ pressure, and temperature (typically, [hyper-]thermophilic cultures exhibit superior performance to mesophilic ones because the hydrolysis rate is favored by high temperatures). Despite its potential, the development of dark fermentation for hydrogen production at the industrial scale is presently limited by its low hydrogen yield. One of the reasons is that this process relies on undefined inocula (e.g., from anaerobic digesters, compost piles, or soils) with interconnected degradation pathways, some of them producing little or no hydrogen. Diverting the process exclusively toward high-yield hydrogen-producing pathways remains challenging (Ghimire et al. 2015). A solution could be the use of monocultures or cocultures of pure strains, which are less robust than complex communities, but offer the possibility to detect more easily metabolic shifts to find conditions that are optimal for biohydrogen production (Elsharnouby et al. 2013) (see Section 13.4.2). Another hurdle in dark fermentation is the production of residues besides H₂, primarily volatile fatty acids, that need to be utilized to achieve complete conversion of the organic biomass. For this purpose, dark fermentation can be coupled with anaerobic digestion or photofermentation to convert the by-products to CH₄ or H₂, respectively. In particular,
photofermentation harbors high substrate-to-\(\text{H}_2\) conversion efficiency; therefore, the combined hydrogen yield is usually greatly improved in comparison with the sole dark fermentation (Ghimire et al. 2015).

In conclusion, despite a tremendous effort and public money investment over the last decades to improve biological processes of liquid and gaseous biofuels, most of them are still not commercially viable. There is fierce debate regarding their ability to compete economically with fossil fuels and, if so, when (e.g., Biello 2011; Ramos et al. 2016); their environmental impact (Zah et al. 2007); and how to measure the latter properly (McKone et al. 2011). In an era of exploding biofuel-oriented research and development (R&D) teams, it is important to remember the basic criteria that should drive research in this area: cost in the short term and energy efficiency in the long term (Zhang 2011). Nevertheless, there is a general consensus that we need to develop more efficient or new conversion technologies for biofuel production. As mentioned briefly in this chapter, a vast literature deals with engineering pure strains for fuel production (see Chapter 10 for a more detailed discussion). In the present chapter, we focus on four specific approaches that could contribute to further improving microbial biofuel production and that reflect the authors’ scientific interests: (1) the rational design of robust enzymes and tolerant strains, (2) the use of natural or engineered microbial consortia for fuel production, (3) the production of biofuels by microbial electrosynthesis (MES), and (4) the development of more efficient dynamic macroscopic models of conversion bioprocesses.

13.2 Targeted Approach 1: The Rational Design of Optimized Enzymes for Fuel Production

The term biofuel encompasses a wide variety of products: bioethanol, biodiesel, biogas, biohydrogen, biomethanol, biomethyl-ether, biobutanol, and so on. Among them, bioethanol—and biodiesel to a lesser extent—represents the largest industrial production. Their bioproduction relies on microbial or enzymatic reactors, and uses several types of sources.

The enzymes involved in biofuels are various and numerous, depending on the type of biofuel required and the starting material. For instance, biodiesel can be obtained by transesterification of vegetal, animal, or waste oils (triglycerides and fatty acids), using lipases as catalyst. In the case of first-generation bioethanol production from arable crops, enzymes such as \(\alpha\)-amylases or glucoamylases are used (Harris et al. 2014), whereas the second-generation production exploits enzymes such as endoglucanases, \(\beta\)-glucosidases, or xylanases (Harris et al. 2014; Uday et al. 2016) to degrade lignocellulosic sources into fermentable sugars. Note that laccases or lytic polysaccharide monoxygenases can help the work of the latter enzymes (Harris et al. 2014; Walton and Davies 2016). This is not an exhaustive list of enzymes used in this field, but it illustrates the large number of pathways to obtain biofuels with biomolecules as catalysts.

In general, the physicochemical conditions of these industrial processes are remote from the physiological conditions: higher temperatures, acidic or basic conditions,
The enzymes must therefore be carefully selected or modified to be efficient in these conditions. In particular, the optimization of the thermal or thermodynamic stability, the solvent stability, and the substrate specificity is often required. Concerning the latter, the adaptation of the enzyme to a substrate slightly different from the original one, or to a larger range of substrates, can be aimed. The optimization of these enzymes can be achieved by directed evolution (Jemli et al. 2016; Kaushik et al. 2016), by rational design, or by combination of rational and random mutation approaches. In directed evolution (Dalby 2011), the natural processes that operate during evolution (i.e., random mutations followed by selection) are mimicked in the laboratory. Several molecular biology techniques are used for that purpose: mutations are randomly introduced in the gene coding for the enzyme using, for instance, error-prone polymerase chain reactions (PCRs); the most optimal enzyme variants according to the desired property are screened and selected; and this is iteratively repeated until the desired degree of optimization is reached. This method relies on random mutations and is quite powerful, but it is also quite expensive and requires high-throughput screening and selection. In the case of rational design (Kaushik et al. 2016), mutations likely to endow the enzyme with interesting properties are selected on the basis of a sequence–structure–function relationship analysis using several bioinformatics tools. These computational tools range from structural analysis and sequence alignments to more dedicated software that predict the effect of mutations on the thermodynamic stability (Guerois et al. 2002; Zhou and Zhou 2002; Capriotti et al. 2005; Cheng et al. 2006; Yin et al. 2007; Masso and Vaisman 2008; Dehouck et al. 2009, 2011; Worth et al. 2011; Pires et al. 2014; Laimer et al. 2015), on the thermal stability (Masso and Vaisman 2008, 2014; Pucci and Rooman 2014; Pucci et al. 2016), or on the solubility (Tian et al. 2010; Agostini et al. 2012) (Figure 13.1).

As a first example of rational design of enzymes with improved thermostability for bioethanol production, let us consider the works of Bayram Akcapinar et al. (2015) and Larsen et al. (2015). Lignocellulosic biomass is a possible source of fermentable sugars, but it is difficult to degrade. Working at higher temperatures is generally needed to decrease the viscosity of the medium, to increase the solubility of the substrate, and to reduce the risk of microbial contamination. Enzymes with improved thermostability are therefore desirable. In a study aimed at thermostabilizing endo-glucanase I from *T. reesei*, which degrades cellulose (Bayram Akcapinar et al. 2015), a first step was to select candidate mutations using three bioinformatics tools that compute the effect of single-point mutations on the thermodynamic stability of the protein: I-mutant (Capriotti et al. 2005), PoPMuSiC (Dehouck et al. 2009, 2011) (Figure 13.1), and SDM (Worth et al. 2011). A subset of possible mutations was then subjected to molecular dynamic simulations at 310 K and 550 K to confirm, using this different approach, the effect of the variants on the thermostability of the protein. Finally, three mutants were characterized experimentally and showed a better thermostability than the wild type, but unfortunately a lower specific activity (Figure 13.1). Another study (Larsen et al. 2015) focused on endo-1,4-β-galactanase from *Talaromyces stipitatus*. Their approach combined sequence alignments with thermostable enzymes, the use of PoPMuSiC (Dehouck et al. 2009, 2011), and an analysis of crystallographic
B-factors to identify possible variants. The sequence alignments aimed to identify residues that were conserved in the thermostable enzymes and different in the meso-stable ones. Then PoPMuSiC was used to compute the stability changes triggered by the introduction of these mutations in the endo-1,4-β-galactanase, and the B-factor analysis revealed the most flexible residues. The authors finally obtained mutations that increase the thermostability of the enzyme.

The stability of an enzyme in acidic conditions is another property that could be tuned. Beliën et al. (2009) worked on the endo-β-1,4-xylanase from B. subtilis, an enzyme that hydrolyzes xylans. This hydrolysis does not produce fermentable sugars but helps other enzymes access and degrade other parts of the lignocellulosic material, thereby generating sugars that can be turned into ethanol. They combined molecular modeling techniques with software that predicts the mutations that will change the pKa value of a given residue (Tynan-Connolly and Nielsen 2006). The different variants that have been designed showed experimentally an increase in their functional stability at lower pH.

Rational design is also helpful to improve the catalytic activity and the enantioselectivity, or to change the substrate specificity of an enzyme. For instance, secondary
alcohols are poor substrates of the lipase from *Burkholderia cepacia*, but Ema et al. (2012) engineered a double mutation that overcomes that. For that purpose, they modeled the transition state of the enzyme–substrate reaction and performed docking simulations of the substrate on this transition state model. Other authors (Santarossa et al. 2005) modified the substrate specificity of a *Pseudomonas fragi* lipase. In this study, they modeled the structure of the lipase and combined visual inspection and structural superimpositions to design the variant. This mutated lipase showed an increased relative activity on C8 substrates and a higher thermostability.

Note finally that enzyme production is expensive, and having a process that permits the reuse of the enzymes can make it more economically viable. Working with immobilized enzymes on a solid support is an alternative to reactors with free enzymes that will facilitate the separation from the products and thus their recovery. We will not describe in this chapter the wide variety of immobilization supports and techniques, which are detailed in different reviews (Franssen et al. 2013; Sheldon and van Pelt 2013; Eş et al. 2015). We will just mention that the immobilization process can improve some of the physicochemical properties of the enzyme, such as its thermostability or its stability in nonaqueous solvents. Together with the bioinformatics technique described above, immobilization has been reported to yield very efficient biocatalysts for use in biofuel production (Zhang et al. 2012; Franssen et al. 2013; Singh et al. 2013; Poppe et al. 2015).

13.3 Targeted Approach 2: The Rational Design of Tolerant Strains for Fuel Production

Among the various features that are expected to be improvable by genetic engineering, the increase of biofuel tolerance in engineered microbes appears as an important challenge for the success of biofuel production at the industrial scale (Nicolaou et al. 2010; Dunlop 2011; Fu et al. 2016). Indeed, the accumulation of biofuel molecules in the cell envelope of producing microorganisms is in general toxic to them. These molecules are highly hydrophobic; hence, they modify the membrane properties and influence the physiological processes that are essential to the cell. This leads to inhibition of cell growth and even to cell death. Efflux pumps increase the tolerance, but burden the cells when overexpressed. Therefore, a trade-off must be found between the toxicity of the biofuel and the cost of expressing the corresponding efflux pumps.

Efflux pumps are membrane proteins that export toxic compounds such as bile salts, solvents, and antimicrobial drugs, and thus play an important role in cell survival (Paulsen et al. 1996; Putman et al. 2000; Nikaido and Takatsuka 2009) (Figure 13.2a). Some are quite specific to certain molecules, whereas others are able to transport a large range of them. The expression of efflux pumps has been shown to be a promising engineering strategy for the production of long-chain alcohols (Dunlop et al. 2011); in contrast, there is evidence that efflux pumps are not effective for exporting short-chain alcohols (Ankarloo et al. 2010; Minty et al. 2011). Clearly, the tolerance mechanisms through efflux pumps are specific to particular classes of biofuels
Another aspect that complicates the increase of tolerance is that the overexpression of efflux pumps—of membrane proteins in general—changes the membrane content and can therefore be toxic for the cell, or at least inhibit cell growth (Wagner et al. 2007). Moreover, all types of biofuels are not equally toxic (Sikkema et al. 1994; Isken and de Bont 1998; Ramos et al. 2002; Nicolaou et al. 2010); in particular, longer-chain alcohols are usually more toxic than short-chain alcohols. Finally, toxicity also depends on the microorganism. For example, *P. putida* can tolerate higher butanol concentrations (Rühl et al. 2009).

In general, microorganisms that exhibit high tolerance levels use multiple mechanisms in synergy: several types of efflux pumps (Segura et al. 2003), as well as other mechanisms, such as altered membrane properties and energy metabolism (Weber and de Bont 1996; Ramos et al. 2002; Nicolaou et al. 2010), and changes in cell size and shape to decrease the cell surface-to-volume ratio (Neumann et al. 2005). Finally, note that while tolerance improvement is a necessary condition to improve production yields, it is far from being a sufficient condition (Dunlop 2011).
A first approach to improve biofuel tolerance consists in using genomic DNA from microorganisms that are not (or less) affected by hydrocarbon-rich environments, such as *Marinobacter aquaeolei* (Singer et al. 2011), and transplant them in, for example, *E. coli* (Tomko and Dunlop 2015).

Mathematical models that take into account the detrimental effects of both biofuel toxicity and pump overexpression have been designed with the objective of improving the understanding of microbial biofuel production and robustness against intrinsic noise (i.e., random variability of the number of molecules) and parameter variations (Dunlop et al. 2010). These include feedback loops similar to those that are commonly used by organisms as versatile control mechanisms, for example, in response to changing conditions and to maintain homeostasis (Alon 2006). An example of a negative feedback loop is given by cells producing a biofuel that inhibits their growth; positive feedback would instead increase their growth. In synthetic biology, such loops, whether positive or negative, are used as key building blocks for synthetic rewiring of gene circuits (Nandagopal and Elowitz 2011).

For example, Dunlop’s model (Dunlop et al. 2010) consists of a set of differential equations. The first one models the growth of the cellular population, with a degradation term that depends on the quantity and type of biofuel. The second equation quantifies biofuel production, which is proportional to the number of cells in the population. The introduction of efflux pumps in the cellular membrane adds an additional degradation term to the first equation, accounting for the toxicity of the pumps, as well as a degradation term in the second equation, which represents the export of biofuel to the extracellular medium. A third equation models the expression and degradation of the pump itself. This simple model includes a negative feedback mechanism due to the toxicity of the biofuel for the cell, which limits biofuel production. It describes the case where the pump’s expression is controlled by a constitutive promoter and is thus constant (Figure 13.2b). The number of cells and the intracellular biofuel level reach a constant value in the stationary state, whereas the extracellular biofuel levels increase.

In the second model (Figure 13.2c), the constitutive promoter of the pump is replaced by a promoter that depends on the intracellular biofuel level: the higher the level, the higher the number of expressed pumps, with saturation happening above some threshold value of intracellular biofuel concentration. A slightly more complex model (Figure 13.2d) introduces an additional gene, the expression of which depends on the biofuel concentration, and acts as a repressor on the efflux pump expression. The fourth model (Figure 13.2e) is a mixture of the latter two models: a feedforward loop is introduced so that both the biofuel and the transcription repressor protein control pump expression.

The four controllers were simulated for different parameter values and were shown to be capable of producing similar levels of biofuel, although the first model (with constant pump expression) is less effective. Moreover, this simplest model was shown to be more sensitive to the parameters, and to produce biofuel only for a limited range of parameter values. Moreover, the last, most complex model, with a feedforward loop, appeared more robust to noise. These models were improved using a slightly more complex feedback mechanism from the biofuel concentration by introducing
a biosensor, which delays the production of pumps until necessary (Harrison and Dunlop 2012). More complete models could include diffusion, several types of efflux pumps, or other tolerance mechanisms.

Among the other possible tolerance mechanisms that are controlled by genes and could be added to the model, changes in the cellular wall or membrane are interesting candidates (Sikkema et al. 1995). For example, the overexpression of targeted genes such as *mur* that are involved in cell wall biosynthesis increases ethanol tolerance by up to more than 100% in *E. coli* (Goodarzi et al. 2010; Nicolaou et al. 2012). Heat shock proteins are also key targets, as they are involved in the response to many types of environmental stress, among which is stress caused by high biofuel concentrations. Indeed, proteins of this family appear to be upregulated in response to biofuel stress (Rutherford et al. 2010), and their overexpression has been shown to increase biofuel tolerance in *E. coli* (Zingaro and Terry Papoutsakis 2013) and *Clostridium acetobutylicum* (Mann et al. 2012).

### 13.4 Targeted Approach 3: Analysis and Engineering of Multispecies Microbial Communities Involved in Fuel Production

For decades, synthetic biologists have selected and tweaked molecular pathways to make cells produce useful products like biofuels in sufficient amounts. Genetic and metabolic engineering tools have been extensively applied to a handful of workhorse model strains (such as *E. coli* or *S. cerevisiae*), but they are not yet easily transferrable to other potentially interesting microorganisms (Liao et al. 2016). In addition, engineered strains are often fragile, might suffer from self-intoxication by the molecules they produce (an issue discussed in Section 13.3), and are vulnerable to environmental changes. Last but not least, the use of genetically modified microorganisms (GMMs) is strictly regulated in most countries. In Europe, they are solely allowed in confined environments (like fermenters), and prior risk assessment for humans and the open environment is required (European Council Directive 2009/41/EC). The U.S. legislation (the Toxic Substances Control Act of the Environmental Protection Agency) is less strict, and GMMs are allowed for environmental processes (like cleanup of contaminated soils).

For all these reasons, exploiting the properties of multispecies communities might be an interesting alternative. In nature, microbes do not live as pure strains, but in complex, dynamic communities in which they interact with other members via the exchange of genes (Frost et al. 2005; Thomas and Nielsen 2005; Choi and Kim 2007; Boto 2010) and communication signals and metabolites (Seth and Taga 2014). Living in the vicinity of other microorganisms can have detrimental effects (increased competition for substrates, and antagonistic interactions through production of antimicrobials) (Hibbing et al. 2010), but a growing number of examples of species synthesizing molecules that are beneficial to other community members (e.g., amino acids, cofactors, and scavengers of reactive oxygen species) suggest that positive interactions are ubiquitous (Seth and Taga 2014). Very frequently, the benefit for the donor (if any) has not been clearly identified. Identification of the exact nature of interactions can be tricky, because there
is a whole range of interactions between competition (where both species are negatively
affected) and cooperation (where both species are positively affected) (Faust and Raes
2012), and because interactions can involve more than two species. In addition, detect-
ing metabolite cross-feedings is difficult due to their intrinsically dynamic nature and
the complexity of communities (Ponomarova and Patil 2015). To unravel the balance
between competition and cooperation, simple models based on metabolic pathways
have been proposed for pairs of species (Freilich et al. 2011). In complex communities,
the best documented example of positive interactions within complex communities is
the sequential transformation of organic matter by metabolic cross-feeding in anaerobic
digestion. It involves a cascade of metabolic reactions (hydrolysis, acidogenic and ace-
togenic fermentations, and methane production) performed synergistically by different
functional guilds of microbes. Their metabolic cooperation is thermodynamically con-
strained, and one of the steps, acetogenesis (performed by so-called synthrophs), would
be impossible without methanogens consuming its products ($H_2$, formate [Stams 1994;
Schink 1997], or—as recently discovered—electrons [Kouzuma et al. 2015]). Examples
of cross-feeding have been reported in other microbial systems as well (Morris et al.
2013), such as the human gut (Louis et al. 2014; Rakoff-Nahoum et al. 2016) or consortia-
degrading xenobiotics (Dejonghe et al. 2003; Lykidis et al. 2011; Men et al. 2011).

In this chapter, we discuss how a better understanding of the functioning of the
multispecies communities involved in biofuel production could contribute to better
process performance, and how genetic and metabolic engineering toolkits initially
developed for pure strains are progressively being applied to low-complexity consortia
to enhance process yield and stability.

13.4.1 Study of Community Structure and Dynamics, and Exploitation
of Community Properties in the Context of Biofuel Production

13.4.1.1 Why Analyze Community Structure and Dynamics?
There is a vast literature on the beneficial properties of mixed communities compared
with pure strains, which range from their greater functionality (i.e., their potential to
be more performative or to realize more diverse functions) (Bell et al. 2005) to their
greater biomass productivity and their superior robustness (i.e., their ability to cope
with perturbations) (Stenuit and Agathos 2015).

Robustness and its variants, resistance (the degree to which a community remains
unchanged after a perturbation) or resilience (the degree to which a community changes after a perturbation but eventually returns to its predisturbance state), are
key properties of community dynamics because they guarantee the maintenance of
the community’s structural organization and/or functional performance in the face of (un-)predictable disturbances (Shade et al. 2012; Stenuit and Agathos 2015). These
terms apply to the phylogenetic composition of a community and/or to its functional
traits (encoded in a set of metabolic genes). Interestingly, both can be uncoupled, and
there are many examples of fluctuations or shifts in taxonomic composition while
metabolic traits are preserved (Fernández et al. 1999; Bull 2010; Kraft et al. 2014). Such
shifts are not necessarily the result of disturbances, as they are frequently observed
in undisturbed conditions. They reflect a certain degree of functional redundancy among community members for a broad function, for example, denitrification (Kraft et al. 2014), heterotrophic activity (Ofiţeru et al. 2010) in wastewater treatment plants, or fermentation in biogas plants (Werner et al. 2011). As a result, such communities are more prone to stochasticity with regard to their taxonomic composition.

On the other hand, in systems with cross-feedings between community members, the performance of the community is tightly linked to its structure, that is, the number of taxa (richness) and their relative abundance (evenness). In such cases, community dynamics is either quite predictable (as in the case of anaerobic digestion [Vanwonterghem et al. 2014]) or chaotic (as in the case of nitrification, which relies on a fragile interaction between two partners sensitive to environmental conditions [Graham et al. 2007]). In a remarkable large-scale comparative study of anaerobic reactors, Werner et al. (2011) demonstrated that bioreactor performance (methane production) was more related to phylogenetic community structure than to operating conditions. They showed that syntrophic populations behaved—as expected—as functional specialists, and that they were stable, resilient after changes in operating conditions (an important characteristic for functional maintenance), and very site (i.e., reactor) specific. In contrast, the communities of fermentative Clostridia and Bacteroidetes were highly dynamic, relying on functional redundancy to maintain the overall community function. Better reactor performances (methane production and organic solid removal) were observed when total communities were more even and more dynamic over time, respectively. The latter suggests that community dynamics helped maintain efficient performance.

Finally, initial community assembly is often a critical step in community dynamics. In closed systems like reactors, the same initial assemblage of microbes can lead to different community structures, a phenomenon called “ecological drift” (e.g., wastewater microbes colonizing electrodes in microbial fuel cells [MFCs] [Zhou et al. 2013]). Drift corresponds to stochastic changes in the relative abundance of different taxa over time and is favored in communities with low diversity and when selection is weak. In open systems, community assembly is also largely dependent on immigration or dispersal processes, that is, movements across space (Ofiţeru et al. 2010). Beyond niche differentiation (a deterministic process), drift (a neutral process), and dispersal, diversification (i.e., the evolution of community members’ genome via mobile genetic elements or mutations) also influences community structure (Nemergut et al. 2013). Unfortunately, diversification remains largely understudied. As pointed out before, it is often more appropriate to study community assembly at the functional than taxonomical level, as the dynamics of both can be nonsynchronized (Burke et al. 2011). Because of these four factors in action (niche differentiation, drift, dispersal, and diversification), assessing “benchmark” values of community stability (i.e., in undisturbed conditions) is difficult—but necessary—before analyzing community robustness. Generally speaking, a reasonable range of variation around a reference value is tolerated. This reference value is easier to define for microbial functions than for phylotypes, as the identification of a “core” microbiome responsible for the function of interest can be difficult in communities with high phylotype turnover and functional redundancy.
When facing a stress, the response of a community is a function of (1) the community properties (species’ plasticity, species’ physiological state, community taxonomic richness and evenness, interspecies connectivity, and functional redundancy) and (2) the nature of disturbances applied, such as frequency, intensity, and duration (“pulse” vs. “press” disturbance) (Bender et al. 1984; Shade et al. 2012; Stenuit and Agathos 2015). For example, a greater richness (Cook et al. 2006) or evenness (Wittebolle et al. 2009) was shown to favor functional robustness after a perturbation. Beyond robustness, the response of a community can be evolution to one or multiple alternative stable states (with a different community structure and/or function) or to irremediable dysfunctionality and collapse. The existence of multiple stable states and possibly fluctuations in between has been reported in microbial communities (Shade et al. 2012). For example, in bioreactors, a new community configuration was shown to positively (Luo et al. 2015) or negatively (Bürgmann et al. 2011) impact the targeted function (biogas production and nitrogen removal, respectively). The magnitude of robustness, a key parameter for microbial communities in engineered ecosystems, can be assessed using various metrics (Stenuit and Agathos 2015).

The recent development of high-throughput molecular and computational tools to analyze the metagenome (collective set of genes), transcriptome (collective set of expressed genes), proteome (collective set of expressed proteins), and metabolome (collective set of secreted molecules) of microbial communities has undoubtedly boosted our understanding of complex microbial ecosystems. These techniques and their recent development are discussed in Section 13.4.1.2. We also provide examples of useful input from such techniques for biofuel production, from the discovery of new enzymes to the manipulation of environmental parameters for increased microbial fuel production based on advanced knowledge of the communities in action. Finally, we also discuss how community-integrated “omics” and high resolution time and space analyses of microbial ecosystems, which hold promise to unravel community functioning and dynamics, must be integrated into a more holistic approach that also includes mathematical predictive modeling. Such a unifying perspective, obviously useful to understand complex, natural ecosystems, could be translated into practical applications in renewable energy supply, but also in environmental safety and health, management of environmental biogeochemical cycles, and provision of new materials, a concept globally named “microbial resource management” (Verstraete et al. 2007).

13.4.1.2 Tools to Analyze Community Structure and Dynamics and How They Contribute to a Better Understanding of Biofuel-Producing Communities

One of the greatest challenges in environmental microbiology today is undoubtedly unraveling and modeling microbial interactions in complex and often dynamic communities. This task is hampered by our inability to cultivate 90%–99% of environmental microbes in laboratory conditions (Rappe and Giovannoni 2003) and to recreate complex environmental conditions in the lab. Although novel cultivation techniques led to remarkable successes in the isolation of recalcitrant microbes (Pham and Kim 2012; Stewart 2012), it is above all the spectacular development over the last 15 years of high-throughput sequencing, “meta-omics” techniques (Figure 13.3), and
**Figure 13.3** Different techniques used to characterize the phylogenetic and functional diversity of a microbial community. Metagenomics provides information about the taxa present in the samples and their functional potential (alternatively, only 16S rRNA marker genes are sequenced when the goal is to obtain deep phylogenetic information—not shown). The functional activity of the community is measured by metatranscriptomics, metaproteomics, and/or metabolomics. Metatranscriptomics measures gene expression, which can be compared across samples, and helps identify key metabolic pathways. Metabolic pathways may also be inferred from metaproteomics, which reflects functional protein expression, and metabolomics, which focuses on metabolites. In single-cell sequencing, individual cells are selected by optical tweezers or fluorescence-activated cell sorting (FACS) before DNA sequencing, allowing genome analysis of single community members. Finally, stable isotope probing (SIP) is based on the incorporation of a $^{13}$C-radiolabeled substrate into the active fraction of the community, followed by recovery of labeled DNA (or less often RNA) by isopycnic centrifugation and sequencing. When integrated, omic techniques can be mapped against each other (dotted arrows, each arrow pointing to the reference dataset), thereby providing more powerful analyses. (Adapted from Vanwonterghem, I. et al., *Curr. Opin. Biotechnol.*, 27, 55–64, 2014.)
integrated bioinformatic pipelines that has nurtured recent breakthroughs in quantifying microbial community composition, function, and spatiotemporal dynamics (Zengler and Palsson 2012; Segata et al. 2013).

The central focus of all studies on microbial relationships is to understand the link between community diversity, community function, and community robustness. For this purpose, emerging technologies like metagenomics, metatranscriptomics, and metaproteomics (Figure 13.3) are precious allies to analyze the composition of complex natural communities in pre- and postdisturbance situations (the latter being mostly changes in physicochemical parameters: temperature, oxygen concentration, and nutrient concentration). Over two decades (since the early 1990s), these technologies were preceded by the application of low-throughput, culture-independent approaches to study the diversity of phylogenetic marker genes (e.g., 16S rRNA genes for bacteria and archaea and 18S rRNA genes for eukaryotes) (Schmidt et al. 1991; Hugenholtz et al. 1998) or, to a lesser extent, specific functional genes. They included clone libraries, quantitative PCR (qPCR), and fingerprinting techniques such as DGGE, FAME, ARISA, or (T-)RFLP (reviewed in Nocker et al. 2007). Their wide application to microbial ecosystems has shed light on the tremendous and ubiquitous diversity of microbes (Hugenholtz et al. 1998; Huisman and Weissing 1999; Cases and de Lorenzo 2002) and provided a partial glimpse of microbial community dynamics. In the mid-2000s, the popularity and effectiveness of 16S-based surveys grew dramatically with the increased throughput of next-generation sequencing (NGS) methods, like 454 and then Illumina sequencing, and the concomitant development of bioinformatic pipelines to analyze the 16S data (Schloss 2009; Caporaso et al. 2010, 2012; Logares et al. 2012; Bálint et al. 2016). Today, “routine” analyses consist of sequencing simultaneously tens of thousands of 16S rRNA marker genes per sample using NGS. A remarkable discovery resulting from deep sequencing of 16S rRNA genes was the existence, in most environments, of a “rare” biosphere made of numerous taxa in low relative abundance (Lynch and Neufeld 2015). In the specific case of anaerobic digesters, the application of NGS has allowed a better characterization of microbial communities in action than with older methods (reviewed in Sárvári Horváth et al. 2016). It has also provided a better understanding of how reactor performance (substrate removal and methane production) is correlated with community structure (Werner et al. 2011). Anaerobic digesters were shown to be typically dominated by previously uncharacterized microorganisms (Sundberg et al. 2013), which may influence the degradation process. In addition, phylotypes could be classified into core taxa common to most of the digesters, shared taxa among a few digesters, and specific taxa to each digester (Rivière et al. 2009; Stolze et al. 2016).

Likewise, high-throughput sequencing of targeted functions has revealed a much greater diversity of genes potentially important in biogeochemical cycles and depollution than previously realized (Iwai et al. 2010; Bowen et al. 2013). However, a better understanding of the metabolic capabilities of these uncultured microorganisms, the level of functional redundancy within communities, and the fundamental mechanisms of interspecies interactions was still needed. In the past 10 years, continuous advances in sequencing have made it possible to (almost) achieve this goal using omic technologies to analyze entire collections of genomes
and transcriptomes from a sample, and thereby infer community composition and function. Coupled to technological advances in protein and metabolite analysis, imaging, and labeling techniques, such as stable isotope probing (Chen and Murrell 2010), these provide altogether a solid technological basis for a better understanding of complex microorganisms. Recent examples include the soil microbiome, the wastewater microbiome (Muller et al. 2014), the microbiome of anaerobic digesters (Vanwonterghem et al. 2014), and the highly publicized human gut microbiome (Franzosa et al. 2015), which might be manipulated to treat gastrointestinal disorders (reviewed in Grogan et al. 2015).

13.4.1.2.1 Metagenomics. Metagenomics is the analysis of the metagenome, that is, the collection of genomes, of a given microbial community at a given point in time. DNA is extracted from the environment, and either directly sequenced using NGS or cloned into a vector transferred to a host to build large metagenomic libraries. Clone libraries are then screened for target genes (genetic screening) or for metabolic activities (functional screening) resulting from heterologous expression of cloned DNA by the host. Library-based metagenomics is tedious (especially the cloning step of large DNA fragments) and time-consuming, and suffers from barriers to heterologous gene expression (that can be partly overcome by the use of phylogenetically diverse hosts) and slow advances in screening technology. Despite such limitations, it has allowed the discovery of novel gene products that lacked sequence similarity to genes of known function (e.g., Craig et al. 2009; George et al. 2010; Jacquiod et al. 2014). However, this approach now lags behind direct sequencing of environmental DNA, also referred to as “shotgun metagenomics,” which has been boosted by the advent of high-throughput sequencing (NGS) and is now widely adopted in environmental microbiology laboratories worldwide. Metagenomics goes beyond 16S rRNA gene surveys, as it provides concrete linkages between microbial community composition (“who is there”) and potential function (“who is able to do what”). In addition, because it is not based on PCR amplification of specific genes, common PCR biases, like those observed in 16S rRNA gene-based surveys, are avoided and no prior knowledge of the community is required. This latter point explains why metagenomics has enabled unprecedented discovery of new taxa and genes. The ultimate goal of metagenomics is to reconstruct large DNA fragments and complete genomes (Sharon and Banfield 2013). Historically, this was not possible because of insufficient sequencing depth (except for low-diversity samples) and the challenge of assembling sequence data originating from complex genomic mixtures; hence, gene-centric approaches based on unassembled data were favored. Thanks to continuing advances in sequencing depth at decreasing cost, read binning (classification), and read assembly (reviewed in Thomas et al. 2012), genome reconstruction is now within reach, even in natural and engineered environments harboring high taxonomical diversity (Albertsen et al. 2013). Another technique, meta3C (see Section 13.4.1.2.2), is emerging as a powerful alternative to sequence microbial genomes from complex environments. Last but not least, the issue of assembling (near-)complete genomes from complex communities can now be tackled using a third technique called single-cell genomics. This technique has been reviewed elsewhere (Blainey 2013; Lasken and McLean 2014).
Briefly, the workflow consists of (1) cell isolation by flow cytometry and fluorescence-activated cell sorting or by use of optical tweezers, (2) whole-genome amplification of single cells by multiple displacement amplification (random DNA amplification using phi29 polymerase and random hexamers) (Binga et al. 2008), and (3) NGS of amplicons and the read assembly (Rinke et al. 2014). Single-cell genomics is still facing technical challenges, such as physical separation of cells from the environmental matrix, uneven amplification of genomic fragments, and subsequent assembly of sequencing data with variable read depth. This being said, it has already proved its potential to access new genomes of uncultivated strains across the tree of life (Rinke et al. 2013) and to reveal intraspecies genetic variations (Lasken and McLean 2014) or unknown interactions between organisms (Yoon et al. 2011). In addition, single-cell genomics and metagenomics can be used on the same samples, thereby allowing cross-validation of the results.

By analyzing all the genomes in a community, it is possible to build up a catalogue of reference genomes for other meta-omic approaches (e.g., metatranscriptomics and metaproteomics), and find out whether the different microbial guilds harbor specific metabolism or if there are vital interactions between them or some functional redundancy.

For example, metagenomics has been extensively applied to characterized natural communities degrading cellulosic plant material (Hess et al. 2011; Lelie et al. 2012). The main drivers of such studies have been (1) the fundamental understanding of species- and community-level mechanisms responsible for efficient cellulose degradation, in order to help advance biofuel production, and (2) the discovery of as yet uncharacterized lignocellulosic biomass-degrading enzymes. Indeed, the paucity of known enzymes that efficiently deconstruct plant polysaccharides represents a major bottleneck for industrial-scale conversion of lignocellulosic biomass into biofuels. Analyzing the enzyme-encoding genes in the metagenome of an anaerobic poplar wood-degrading community allowed us to identify numerous (hemi-)cellulases, debranching enzymes, and enzymes homologous to fungal lignolytic enzymes that could be assigned to Bacteroidetes and Clostridiales (Lelie et al. 2012). Unexpectedly, enzymes assigned to Magnetospirillum were also identified, which were hypothesized to play a role in the degradation of lignin-derived aromatic compounds, thereby preventing these toxic compounds from accumulating in the environment. More recently, an analysis of the metagenome of the sugarcane bagasse microbiome brought to light the existence of a large repertoire of lignocellulolytic enzymes and newly defined auxiliary activity proteins (Mhuantong et al. 2015). In addition, comparison with other lignocellulolytic metagenomes revealed the existence of a core of conserved genes despite very different species composition. Besides looking at communities that are found on the surface of lignocellulosic material, several metagenomic studies were focused on host-associated microbes living in the gut of cellulose-degrading organisms. For example, the metagenomic analysis of the cow rumen microbiome provided a substantially expanded catalogue of genes coding for carbohydrate-active enzymes, and allowed us to generate draft genomes of 15 uncultured strains involved in cellulose deconstruction (Hess et al. 2011). Similar studies were performed on the panda...
microbiome (Zhu et al. 2011), leaf-cutter ant microbiome (Suen et al. 2010), and
termite microbiome (Scharf 2015).

13.4.1.2.2 Meta3C, a New Technique to Sequence Microbial Genomes in Complex Communities. Obtaining genomic information on the microorganisms involved in the production of microbial fuels is nowadays the fastest way to characterize their metabolic activities, thereby paving the way for improvements in yield, titer, and productivity (Orphan 2009; Warner et al. 2009; Gowen and Fong 2011). However, most naturally occurring members of microbial communities are very difficult, if not impossible, to cultivate in isolation. Although sequencing the genomes of bacterial species from pure cultures is a relatively simple task that can be performed efficiently and at a relatively low cost using the latest sequencing techniques (Koren and Phillippy 2015), sequencing the genomes of mixed bacterial species is much more complicated: one ends up with hundreds of thousands of genome fragments that are mixed with one another in the assembly software output, with little to no indication of which microorganism they come from.

For this reason, people have long contented themselves with “gene-centric” metagenomic analyses in which the main aim is to produce a catalogue of the genes present in a given environment (Tringe et al. 2005). Complete or near-complete genome assemblies of noncultivable members of microbial communities have been sometimes obtained using iterative assembly procedures (Pelletier et al. 2008); however, this only works for members that are largely overrepresented in the community. Recently, the availability of new methods to “bin” metagenomes into sets of groups of contiguous sequences (“contigs” [Staden 1980]) hypothetically coming from different species has triggered a shift from gene-centric to “genome-centric” approaches, where the aim is now to assemble and characterize the genome of each member of the community in order to understand its metabolic activities (Waldor et al. 2015). The most popular binning approaches are based on the guanine + cytosine (GC) content and/or on coverage (i.e., how often a given stretch of DNA is represented among the sequence reads), assuming that the abundance of each species in the mix should endow it with a characteristic coverage signature (Albertsen et al. 2013; Alneberg et al. 2014). However, microbial genomes have varying and overlapping GC contents, and the coverage signal can be obscured by the presence of repeated genome regions or by natural and/or artefactual variations in coverage along genomes, notably as a function of the distance to its origins of replication (Semova et al. 2012; Hawkins et al. 2013).

To overcome these problems, a promising new approach has been to turn to “contact genomics” (Flot et al. 2015) (Figure 13.4). This term encompasses several approaches that rely on the use of a cross-linking agent (generally formaldehyde) to “record” the physical contacts between chromosomes in cells. The first studies using such a technique, under the name “chromosome conformation capture” (3C), aimed to investigate the tridimensional structure of yeast chromosomes (Dekker et al. 2002): first, a genomic library was generated from a cross-linked chromatin fragment, and then the frequency of cross-linking events between chosen genome regions was investigated using semiquantitative PCRs. Nowadays, an overall picture of the
physical interactions in a genome can be gleaned in a single step by applying NGS to the genomic library mentioned above. These genome-wide approaches are called “Hi-C” (Lieberman-Aiden et al. 2009) or “3C-seq” (chromosome conformation capture sequencing [Marbouty et al. 2014]), depending on whether an additional enrichment step is performed to ensure that most read pairs contain information on the three-dimensional (3D) interactions of the chromosomes.

Contact genomic libraries contain information regarding the physical linkage of DNA regions on chromosomes across very long distances (typically at least 150–250 kb); as a result, they are very useful to scaffold contigs and improve assemblies, with the potential to yield chromosome-scale scaffolds if the data are in sufficient amount and quality (Flot et al. 2015). For example, Marie-Nelly et al. (2014) used 3C-seq to finish the genome of *T. reesei*, a fungus used extensively in the biomass-to-biofuel industry (Gusakov 2011); starting from a published draft assembly fragmented into 77 scaffolds (Martinez et al. 2008), they obtained seven large scaffolds containing 99.8% of
the initial assembly, in perfect agreement with karyotypic data, suggesting that this species possesses seven chromosomes. Several programs have been devised to scaffold contigs using contact genomic information: dnaTri (Kaplan and Dekker 2013), Lachesis (Burton et al. 2013), GRAAL (Marie-Nelly et al. 2014), and HiRISE (Putnam et al. 2016). Among these, the latter two are the most appealing, as they are based on explicit probabilistic models. At present, there is no program available to assemble directly the reads obtained by sequencing contact genomic libraries; instead, a preliminary assembly must first be obtained, and then 3C data are used to improve it. This two-step process is a major limitation, as generating a decent preliminary assembly can consequently be a daunting task when dealing with highly repetitive and/or heterozygous genomes. These problems could potentially be solved by taking the contact genomic information into account already during contig formation (Flot et al. 2015).

Contact genomics also empowers metagenomic analyses: as cell membranes prevent the chromosomes of different organisms from interacting spatially (even if they live in close symbiotic relationships), quantifying 3D interactions makes it possible to divide metagenomic datasets into bins reflecting faithfully the different components of the microbial community under investigation (Beitel et al. 2014; Marbouty et al. 2014; Marbouty and Koszul 2015). When dealing with microbial communities (either natural or evolved in the laboratory), a possible workflow is therefore to first generate a metagenomic assembly using a suitable program such as IDBA-UD (Peng et al. 2012), and then bin the resulting assembly by performing a network analysis of 3C data (Marbouty et al. 2014), and finally reassemble each genome separately using GRAAL (Marie-Nelly et al. 2014). In the case of diploid microorganisms, this may be followed by a final step during which 3C data are used to phase the scaffolds, that is, to determine which genetic variants are parts of the same haplotype (Selvaraj et al. 2013; Flot et al. 2015). Although such meta3C approaches have not been applied yet to microbial communities relevant to biofuel production, this represents a promising research direction.

13.4.1.2.3 Metatranscriptomics. Metatranscriptomics is the analysis of all mRNAs (expressed genes) in a microbial community at a given point in time. It involves mRNA extraction, followed by reverse transcription and high-throughput sequencing of cDNAs. This technique focuses on active members in a community, and therefore the level of complexity of analyses is reduced compared with metagenomics. To infer differences in levels of gene expression, metatranscriptomic reads are mapped against reference genomes or a metagenome of the same environment. This method has advantageously replaced microarrays to measure in situ gene expression, as it is faster and more sensitive, and allows the identification of expressed genes without prior knowledge of their sequence. Technical challenges include the recovery of sufficient amounts of high-quality mRNA (a task made difficult by the overwhelming contribution of rRNA to the total RNA pool), the inherent liability of mRNA, and the biases linked to cDNA synthesis and amplification. The first application of metatranscriptomics to a production-scale biogas plant dates back from 2012 (Zakrzewski et al. 2012). In that study, the comparison of metatranscriptome data and 16S rRNA...
gene data indicated a high transcriptional activity of archaeal species. More recently, several studies of the metagenome and metatranscriptome of biogas reactors have shown that they contain a tremendous number of as yet uncharacterized taxa, and that the highly abundant taxa are not necessarily the most active ones (Campanaro et al. 2016; Maus et al. 2016; Stolze et al. 2016). In some digesters, high functional redundancy was detected—especially in the early steps of anaerobic digestion (Cai et al. 2016)—raising the question of whether this phylogenetic diversity is needed to ensure efficient anaerobic degradation of organic matter. In all, there are still few metatranscriptomic studies of communities involved in biofuel production, but this technique has great potential to identify the most active pathways in different operating conditions (and therefore drive the community toward pathways of interest), and to measure the immediate regulatory response of the entire community to a perturbation. On the other hand, metatranscriptomics is the method of choice for the genetic analysis of eukaryotic communities, as illustrated by recent metatranscriptomic studies on organic matter degradation by fungi in forest soils (Damon et al. 2012; Kuske et al. 2015).

13.4.1.2.4 Metaproteomics. Metaproteomics is the large-scale characterization of the entire protein complement of environmental microbiomes at a given point in time (Wilmes et al. 2015). Compared with metatranscriptomics, it takes into account posttranslational events. Proteins are extracted from the environmental sample, fractionated, separated using liquid chromatography (LC), and detected with tandem mass spectrometry (MS/MS) (VerBerkmoes et al. 2009). This protocol has not changed much for the last decade, after spectacular progress in analytical capacities. In about a year, proteomics moved from separation of proteins by two-dimensional (2D) polyacrylamide gel electrophoresis, in-gel enzymatic digestion, identification of a few proteins by LC-MS/MS, and de novo peptide sequencing to the identification of thousands of proteins by the above-mentioned pipeline (Wilmes et al. 2015). As other omic techniques, metaproteomics is facing technical challenges, including (1) cell lysis and quantitative protein extraction from complex matrices, (2) standardized protein fractionation methods exploiting both molecular weight and charge, (3) the need for mass spectrometers with faster scan speeds and higher mass accuracies, and (4) the need for good reference databases (sample-derived metagenomic and metatranscriptomic datasets rather than public protein databases) (Wilmes et al. 2015). Solutions to these challenges already exist, but they are not routinely applied to metaproteomic studies yet (Wilmes et al. 2015). One exception might be coupling metaproteomics to metagenomics to analyze the same sample, an approach called “community proteogenomics” (VerBerkmoes et al. 2009). This procedure is now widely applied, since its usefulness to improve protein identification was proved over and over. It is based on the construction of a protein sequence database from the sequenced metagenome. In silico trypsin digest is performed on the predicted proteins, resulting in a peptide database. MS/MS spectra are then matched to the peptides in the database, and after filtering, a list of identified peptides is obtained. Peptides that are present in only one protein in the whole database can be unambiguously tracked back to their corresponding protein, and thus permit reliable protein identification, and hopefully the
microbial species it is derived from. In addition, in case the protein mixture was fractionated (into extracellular, soluble, and membrane fractions) after the initial extraction, information about protein localization is available. There are other difficulties linked to the application of metaproteomics to complex communities, like detection of strain variability (amino acid substitutions) and follow-up of protein turnover in highly dynamic microbial assemblages (VerBerkmoes et al. 2009).

Metaproteomics has been successfully applied to many natural environments with a moderate level of complexity (the last problematic environment might be soils); in engineered ecosystems, activated sludge has been extensively studied with this technique from the perspective of polymeric substance production and biological phosphorus removal (Wilmes et al. 2015). The metaproteome expressed by thermophilic communities producing methane from cellulose (paper) has been investigated too (Lü et al. 2014); syntrophic acetate oxidation and hydrogenotrophic methanogenesis were revealed to be the dominant pathways for methane production. One unexpected result was, however, the high proteolytic activity in that community. The potential of metaproteomics to understand microbial ecology of biogas plants has been recently reviewed in Heyer et al. (2015).

13.4.1.2.5 Metabolomics. Metabolomics is the qualitative and quantitative analysis of low-molecular-weight molecules produced by a microbial community at a given point in time. It relies on mass spectrometry methods. Noninduced metabolite production was historically studied by adding the cell-free culture filtrate of a donor strain to the recipient culture, assaying filtrate activity, and identifying the molecules in the spent medium. Recently, imaging mass spectrometry (Watrous et al. 2012) has been applied to explore the metabolite production of strains stimulated by the presence of other organisms in petri dishes. A recent analysis of the secreted metabolome of *Streptomyces coelicolor* growing in the vicinity of five other actinomycetes revealed that for each interaction, a specific chemical response was produced (Traxler et al. 2013). Other recent instrumentation includes microfluidics coupled to MS. The range of small molecules produced by microbes is vast, from amino acids to nucleosides, polyketides, alkaloids, signaling molecules, enzymes, or cofactors (Phelan et al. 2012). They are suspected to be produced by small subsets of the community, either passively by cell lysis or leaking or actively by efflux, and benefit a large number of neighbors. How these neighbors contribute to community stability or diversity is usually unknown. Microbial interactions not only influence the survival of community members, but also control their behavior and differentiation. Identifying metabolic exchanges is feasible for simple cocultures (Nakanishi et al. 2011), but it is an overwhelming task in natural communities with hundreds of species living in fluctuating conditions (Ponomarova and Patil 2015): first, because metabolites cannot be directly attributed to a particular species; second, because a large proportion of microbes have never been characterized in regard to their metabolic needs and biosynthetic capabilities; and third, because only a few percent of spectra can currently be annotated. Fortunately, new computational tools have been recently developed to annotate spectra from unknown molecules based on automated comparison with chemical structures found in public chemistry databases (Silva et al. 2015).
13.4.1.2.6 Perspectives in the Analysis of Complex Communities. When combined, metagenomics, metatranscriptomics, metaproteomics, and/or metabolomics provide a formidable suite of tools (meta-omics) to study the direct link between microbial community composition or genetic potential and final phenotype. Many authors predict that such integrated platforms will become the future standard for large-scale characterization of microbial consortia, including those performing important biological processes (Raes and Bork 2008; Narayanasamy et al. 2015). For example, meta-omics is now revealing the complexity of prokaryotic degradative activity in lignocellulose-rich environments (Cragg et al. 2015; Scharf 2015). Its application to reactor communities has just begun, from anaerobic digestors (Maus et al. 2016) to wastewater activated sludge (Roume et al. 2015). The goal of these two studies was to harness such communities in the future for optimal CH4 production and lipid biofuel production (Muller et al. 2014), respectively. Transitioning from research that produces lists of genes, proteins, and genomes to research that unlocks their functional significance is not an easy task, though, and more emphasis on function is badly needed (Cragg et al. 2015; Scharf 2015).

Another limitation is that at present, such techniques are mostly used to provide a “snapshot” of a community, which does not say much about its adaptive response to disturbances, potential for self-stabilization, or phylogenetic and functional dynamics (all key parameters in biotechnological applications). Therefore, in a foreseeable future, they will hopefully be integrated in space- and time-series analyses in order to unravel structure–function relationships within microbial communities in stable and/or perturbed conditions. For example, meta-omic analysis of microbial communities over time in a larger number of biogas plants operating in different conditions would make a significant contribution to a better understanding of the relationship between microbial community structure, operating conditions (operational parameters and nature of substrates), process efficiency (e.g., CH4 yield), and process stability. In addition, metatranscriptomics and metaproteomics could be used to indirectly determine which environmental stressors or growth-limiting factors affect each species’ growth (Konopka and Wilkins 2012). Indeed, determining such factors in most environments is analytically difficult, as nutrient concentrations are extremely low and biomass increase is usually modest. Practically speaking, it would consist of sampling across environmental gradients and querying the physiological state of the different microbes, that is, searching for transcripts or proteins correlated with gradient intensity.

Despite the impressive accumulation of experimental data over the last decade using meta-omic approaches, several authors warn that an important gap remains: the identification of broadly applicable principles that can then be used to develop conceptual (and eventually predictive) models of microbial community dynamics (Konopka et al. 2014; Prosser 2015). Therefore, one could question the usefulness of describing in detail community structure (“who is there” and “who is doing what”), as broad principles to describe interactions between microbes (“who is doing what with whom”) and community-level properties derived from such interactions are still largely missing. We think that this assertion must, however, be nuanced.
First, taking into account microbial community structure without knowledge of microbial interactions can significantly increase the accuracy of community function prediction. In a recent paper, Graham et al. (2016) performed a comprehensive analysis of 82 literature datasets on carbon and nitrogen cycling across different ecosystems and showed that incorporating microbial diversity metrics in statistical analyses could improve model accuracy of processes like denitrification, nitrogen mineralization, or carbon respiration rates, compared with analyses where only environmental parameters or gross physiological parameters (like biomass) were used to predict the function.

Second, this knowledge gap is presently being filled by the development of tailored mathematical tools to model community dynamics based on high-resolution meta-omics datasets coupled to extensive ecological time series. The development of such tools is still in its infancy but is going fast, because it is driven by (1) a fundamental need to understand the common mechanistic bases that explain how the collection of species properties drive high-order community properties and generate the endogenous community dynamics (Prosser et al. 2007; Little et al. 2008; Konopka et al. 2014; Widder et al. 2016), and (2) an applied need to predict the functioning of specific microbial systems (i.e., communities performing useful tasks for human-kind: waste treatment, nutrient cycling, bioremediation, etc.), and how to design more robust communities. Mathematical tools to exploit large microbial datasets and identify patterns of (non-)association of certain ecotypes and forces that maintain community function have been reviewed elsewhere (Stenuit and Agathos 2015). In summary, they are based on (1) the comparison of independent datasets from different environments (via multivariate statistical tools, calculation of interaction coefficients based on correlations in phylotype abundance, buildup of cooccurrence networks [Faust et al. 2015] to identify keystone species, etc.) and (2) the analysis of time-series datasets from the same environment for dynamic modeling, detection of cause-and-effect relationships, and predictive purposes. Last but not least, such tools are progressively being integrated with engineering and biochemistry models to predict fluxes of molecules (like flux balance analysis, a computational method for analyzing the steady-state flow of metabolites through the metabolic network of an organism or a community and evaluating the production of targeted metabolites—see Section 13.6).

Third, a more recent approach is being implemented to unravel species interactions. It consists of manipulating the structure of the community to understand the function of each member, its connections with other members, and its relative importance in the response to disturbances. To answer such questions, the most common experimental setup is to work with simplified, artificial communities in controllable reproducible systems, like batches or chemostats (Bull 2010; Goers et al. 2014), via the initial addition or substraction of species in the mix. Such model systems are oversimplified, but they (1) are highly flexible and reproducible; (2) offer the possibility to apply single or multiple perturbations to the targeted community, from changes in physicochemical parameters to invasion by foreign, potentially competing microbes or by predators (protozoa, phages, etc.); and (3) allow us to easily study community heterogeneity, redundancy, modularity, internal control (via feedback loops), and
species- and multispecies-level responses to perturbations (Stenuit and Agathos 2015), all features that could hopefully be extrapolated to more complex microbiomes. More recently, it has been suggested that alteration of the microbiomes in their natural context would be possible in the future by in situ microbiome engineering (Sheth et al. 2016). This approach is based on a suite of genetic tools to manipulate community composition and functions: probiotics (addition of foreign microbes), antibiotics (unspecific killing of microbes), plasmids or insertion sequences (for the stable propagation of exogenous DNA in the community [Springael et al. 2002]), or phages (for the elimination of specific strains). In particular, phage host range could be modularly engineered by swapping phage tail components, and used to deliver a lethal CRISP-Cas RNA-guided nuclease system in the target strain. However, such engineered vectors have not been successfully implemented to natural microbial communities yet and will not be in the near future.

13.4.2 Management and Genetic Engineering of Microbial Consortia

Years ago, scientists advocated the benefits of using complex multispecies microbial communities for bioenergy production (Rittmann et al. 2008). The idea was to benefit from their low cost and high resilience toward fluctuations in environmental conditions or intrusions by microbial invaders (Table 13.1). In regard to this last point, working with mixed communities represents a considerable advantage: it eliminates the need to work under stringent sterile conditions. In addition, the authors suggested ways to manage mixed communities for maximal production of desired metabolites: (1) CH₄ production by funneling the electron flow through H₂ and acetate in the anaerobic digestion process, (2) H₂ production by biomass fermentation in the absence of methanogens coupled to rapid harvest of H₂, and (3) algal lipid production by coupling algae with heterotrophic bacteria that recycle the organic soluble compounds released by algae into inorganic nutrients and CO₂ made available again for the algae. Other benefits from algae–bacteria interactions in the context of biofuel production include a better protection of algae against harmful bacteria and fungi (therefore avoiding algal population collapse), and initiation of algal flocculation by bacteria (therefore improving biomass harvesting) (Ramanan et al. 2016).

Less than a decade later, it must be recognized that the scientific community did not pursue much in that direction. One exception is the study of Kohn and Kim (2015), where the product spectrum in a mixed culture fermentation process was determined based on thermodynamic considerations. The authors succeeded in favoring the growth of microorganisms from the cow’s rumen microbiome that were able to produce high concentrations of fuel alcohols and alkanes, or H₂ and CO₂, by manipulating the ratio of H₂ to CO₂ and the total gas pressure. But apart from biomass conversion into CH₄ production, natural complex communities have been largely underutilized in biofuel production applications, because they are often too tricky to manage when the production of one specific molecule is targeted. A more successful alternative is the development of artificial consortia of a few wild-type or (most often) engineered strains. Indeed, the lower level of complexity of artificial
### TABLE 13.1 Comparative SWOT Analysis of Monocultures versus Mixed Consortia for Biofuel Production

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Weaknesses</th>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Monocultures</strong></td>
<td><strong>Mixed Consortia</strong></td>
<td><strong>Monocultures</strong></td>
<td><strong>Mixed Consortia</strong></td>
</tr>
<tr>
<td>- The most widely used strains are (almost) perfect cell factories: metabolic processes are understood, modeled, and controlled</td>
<td>- Sensitivity to stresses (change in substrate type, temperature, viral invasion etc.); therefore, maintaining monocultures at large scale can be impractical and cost-prohibitive</td>
<td>- Usually greater functionality (faster growth, greater performance, or greater flexibility in terms of substrates used) than pure strains</td>
<td>- Current techniques like metabolomics are not powerful enough to study mixed communities in their whole complexity</td>
</tr>
<tr>
<td>- Available genetic tools developed to engineer specific strains are not easily transferrable to other strains</td>
<td>- Increased robustness, resilience, and self-stability when facing perturbations</td>
<td>- Enables cultivation and exploitation of the metabolism of microbes that are unable to grow as pure strains</td>
<td>- No general methods are available to engineer consortia of different organisms</td>
</tr>
<tr>
<td>- Can accommodate complex waste streams as substrates</td>
<td>- Mixed communities may harbor community-level functional traits that are not found in individual members</td>
<td>- Mixed consortia may harbor community-level functional traits that are not found in individual members</td>
<td>- Final products are rarely pure</td>
</tr>
</tbody>
</table>

**Opportunities**
- Omics revolution could rapidly lead to more diverse genetic toolkits
- Possibility to tweak environmental parameters to direct community activity toward a desired function

**Threats**
- None
- Omics techniques do not provide a basic mechanistic understanding of community functioning and its emergent properties
- Techniques to manipulate consortia are still in their infancy; their future is uncertain

**Note:** SWOT, strengths, weaknesses, opportunities, and threats.
Microbial Fuels

consortia is easier to handle, and they can be analyzed with current technologies like metabolomics (Ponomarova and Patil 2015) with enough resolution and coverage. In contrast, tools developed for genome-scale editing (Esvelt and Wang 2013) are not transposable yet to metagenome engineering of complex communities in order to add, remove, or modify their functional profile and alter their community-level properties (Sheth et al. 2016).

Compared with monocultures, metabolic exchange in multispecies consortia can be viewed as a strategy for group success: they favor the emergence of community properties, such as enhanced process performance and faster growth (Morris et al. 2013). In biotechnological applications, compartmentalization of parts of metabolic pathways in different community members (“division of labor”) can enhance the production of a valuable metabolite, as recently demonstrated for the synthesis of a precursor of the antitumor drug paclitaxel (Zhou et al. 2015). Division of labor was also reported in the conversion of pollutants with inhibitory intermediates (Dejonghe et al. 2003). However, predicting which of the two strategies (engineering several strains to catalyze specific steps of a pathway vs. engineering one strain to catalyze the complete pathway) maximizes the performance of a process is difficult (Lindemann et al. 2016). In that respect, understanding the biochemical causes of metabolic specialization (i.e., the fact that species often specialize at metabolizing only subsets of the available substrates rather than all of them) could provide a foundation for rational design of consortia in synthetic ecology (Johnson et al. 2012).

Emergence and maintenance of metabolic exchanges depend on particular circumstances, such as the spatial structure of the microbial community, nutrient availability, diffusion constraints, and cost-effectiveness of the biosynthetic processes in question (Morris et al. 2013). In particular, the existence of a spatial structure was shown to stabilize consortia. Using microfluidics, Kim et al. (2008) showed indeed that imposing a microscopic spatial structure was both necessary and sufficient for the coexistence of three soil bacterial species involved in reciprocal syntrophy and forming a consortium that was otherwise unstable. The authors foresaw that controlling spatial structure, rather than matching metabolic and growth rates, would expand the range of systems amendable to synthetic biology approaches. In addition, aggregation in flocs or biofilms maximizes the efficiency of metabolite transfer and stimulates otherwise thermodynamically unfavorable metabolic processes (Agapakis et al. 2012). For example, syntrophic hydrogen-producing fermenters and methanogens produce large aggregative flocs in anoxic lake sediments or sewage sludge. On the other hand, cell populations that communicate through diffusible molecules may be separated to a larger degree without disrupting communication; in this case, coculture systems can be treated as comprising “monoculture modules” that can be connected with each other (Goers et al. 2014). Nutrient-poor environments favor interspecies relationships because reciprocal exchange is more likely to be necessary. Remarkably, it can arise between unrelated strains providing that complementation of each other’s biosynthetic capabilities becomes vital for the survival of the consortium (Hom and Murray 2014).

There are many examples of consortia tailored for specific functions. In practice, cocultures rarely exceed two or three species, because larger numbers of interacting
populations, although potentially more valuable for industrial applications, often show unstable and unpredictable behavior (Goers et al. 2014).

In the context of cellulosic biofuels, several studies have explored the issue of using mixed cultures for ethanol conversion compared with engineered monocultures for CBP. The majority of published papers reported that fermentation consortia produced a higher yield of ethanol than did a single microorganism (Alper and Stephanopoulos 2009). For example, dilute acid softwood hydrolysate (which contains glucose and xylose as the dominant sugars) was fermented to ethanol with exceptional yield by a coculture of \textit{S. cerevisiae}, \textit{Pachysolen tannophilis}, and a genetically modified \textit{E. coli} strain that carried both \textit{pdc} and \textit{adhB} genes derived from \textit{Z. mobilis} (Qian et al. 2006). At the enzymatic level, an enzyme cocktail expressed by the fungus \textit{Hypocrea jecorina} and a cellulosome preparation secreted from \textit{C. thermocellum} showed great activity on pretreated biomass and purified cellulose, respectively (Resch et al. 2013). When combined, these systems displayed dramatic synergistic activity of cellulose deconstruction. Xia et al. (2012) set up a consortium of three \textit{E. coli} strains able to remove acetate (a growth inhibitor contained in lignocellulosic hydrolysates) and to degrade simultaneously glucose, xylose, and arabinose. Kato et al. (2004) reported improved cellulose degradation by a coculture of an anaerobic cellulolytic \textit{Clostridium} and aerobic noncellulolytic bacteria. The latter probably consumed all available oxygen and degraded inhibitory substances toward \textit{Clostridium}. In another study, the same authors observed that a cellulose-degrading community of five strains showed functional and structural stability (Kato et al. 2005). Interactions between community members were unraveled by removing each community member separately and recording community dynamics and degradation efficiency. Finally, Minty et al. (2013) designed and tested a synthetic two-species consortium for the production of isobutanol from cellululosic biomass. In this coculture, the fungus \textit{T. reesei} secreted cellulase enzymes to hydrolyze lignocellulosic biomass into soluble saccharides, which were metabolized by a genetically engineered bacterium \textit{E. coli} into desired products. Isobutanol titer and yield reached 1.88 g/L and 62% of theoretical maximum, respectively, which are low in absolute value but greater than those reported to date for CBP production of advanced biofuels.

Regarding algal biofuels, several recent papers reported that algal growth and lipid accumulation were enhanced when algae were cocultured with specific strains of bacteria compared with axenic cultures (Bashan et al. 2002; Higgins and VanderGheynst 2014). Cocultivation of microalgae with filamentous fungi was shown to enhance total biomass production and lipid yield as well, but also biomass flocculation (Wrede et al. 2014; Muradov et al. 2015). The authors suggested that oil composition could be optimized via such interactions, as shifts in the fatty acid profile were observed. Finally, a modeling study of cocultivation of an algal strain with an oleaginous yeast predicted that the algae–yeast coculture could produce biodiesel at competitive prices (Gomez et al. 2016). The model was built on genome-based metabolic networking. In this system, the algae benefited from a local source of CO$_2$ (the respiration of lignocellulosic sugars by yeast metabolism). At the same time, the yeast benefited from the O$_2$ produced by algal photosynthesis, and therefore increased lipid production.
Regarding the production of H₂ by dark fermentation, cocultures were reported to achieve systematically better results than monocultures (reviewed in Bader et al. 2010; Elsharnouby et al. 2013). In the first type of coculture, a strict anaerobe is mixed with a facultative one, which consumes traces of oxygen and therefore allows stable H₂ production without any added reducing agents. In the second type of coculture, a cellulolytic strain (e.g., *C. thermocellum* or *C. acetobutylicum*) is grown with a noncellulolytic high-hydrogen-producing strain, allowing efficient hydrogen production from cellulotic waste without expensive hydrolysis treatment. In the third type of coculture, a hydrogen-producing aciduric strain is grown with a high-hydrogen-producing strain, which allows us to extend the range of pH values at which biohydrogen can be produced, and therefore eliminates the need for buffering the growth medium.

The question that remains, however, is how to design productive microbial consortia from building blocks (microbial populations) in a rational way. Mechanistic understanding of interactive behaviors in cocultures remains very limited, and to date, the design of consortia more productive than monocultures was either due to luck or based on intuitive ecological principles like division of labor (Lindemann et al. 2016). In the future, advancements in the field will increasingly rely on mathematical modeling, either population-based modeling for the prediction of the interspecies dynamics without detailed information of intracellular metabolism, or metabolic network modeling based on energy and material fluxes within and between cells (Lindemann et al. 2016).

13.5 Targeted Approach 4: Microbial Electrosynthesis of Biofuels

MES is the production of chemical commodities by microbes using electrons derived from the cathode of a bioelectrochemical system. This technology is increasingly focusing attention on a context of transition from an oil economy to an electric economy. Electrobiocommodities include acetate, but also biofuels like H₂, CH₄, ethanol, butanol, and 1,3-propanediol.

In their natural habitats, microorganisms use various forms of electron acceptors (O₂, nitrate, sulfate, ferric iron oxide, and heavy metals), as well as of electron donors (organic compounds, ferrous iron oxide, H₂S, and H₂), to sustain their energy requirements. In particular, some of them have the capacity to use natural insoluble electron donors or acceptors. The best-characterized ones are *Geobacter sulfurreducens* and *Shewanella oneidensis*, reducing ferric iron oxides, and *Acidithiobacillus ferroxidans*, oxidizing ferrous iron minerals. In addition, there is growing evidence that a variety of anaerobic microbes can oxidize metallic iron (Fe(0)), a form of iron uncommon in nature that has become widespread due to human activity during the last millennium. The best documented Fe(0)-oxidizing microbes are sulfate-reducing bacteria corroding metallic pipes (Dinh et al. 2004; Venzlaff et al. 2013), but some methanogenic strains have been shown recently to use metallic iron as the sole electron donor as well (Dinh et al. 2004).

Microorganisms exploiting insoluble compounds for their oxidoreduction reactions have been harnessed in microbial electrochemical technologies, because
they can give and receive electrons to and from electrodes. Initially, in the early 2000s, such technologies were developed and operated to produce electrical current (reviewed in Logan 2009). The basic configuration of these so-called MFCs consists of two chambers containing an electrode, filled in with an electrolyte (an aqueous solution), and separated by an ion exchange membrane. Both electrodes are connected to an external electric circuit. Current is generated at the anode from the oxidation of organic compounds from wastewater or, less frequently, the conversion of toxic $\text{H}_2\text{S}$ into elemental sulfur or sulfate. At the cathode, oxygen is reduced chemically or biologically. It is also possible to expose the cathode directly to air and therefore eliminate the need of a cathode chamber. MFCs basically work as a voltage source with internal resistance; the difference in potential between the oxidation reaction at the cathode and the reduction reaction at the anode leads to a spontaneous electrical current from the anode to the cathode. The actual potential is lower than this difference because of various losses related to the electrochemical properties of the electrodes and the biofilm, the accumulation of reaction products, and the depletion of reactants near the two electrodes, and ohmic losses in the electrical circuit itself (reviewed in Rinaldi et al. 2008). Over a decade, the performance level of MFCs has been raised by several orders of magnitude, reaching 2400 mW/m$^2$ (Logan 2009) and 200–250 W/m$^3$ (Rinaldi et al. 2008). This was made possible thanks to substantial research effort to optimize aspects related to materials, cell architecture, and buffer solution (Rinaldi et al. 2008), but also thanks to a better understanding of the biological aspects of MFCs. In this context, electron transfer mechanisms between bacteria and the electrodes have been extensively studied. Historically, they were mainly examined in two strains: *G. sulfurreducens* and *S. oenidensis* (reviewed in Lovley 2012). First, direct contact between electrogenic microbes and the electrode is possible via c-type cytochromes or “nanowires.” Nanowires are either type IV electroconductive pili (Reguera et al. 2005) made of the aromatic amino acids phenylalanine and tyrosine, or outer membrane and periplasmic extensions containing cytochromes used as “hopping” sites by electrons. Microbes able to make direct contact with an electrode usually develop thick biofilms at its surface. Second, transfer can be mediated by electron shuttles (also called mediators) secreted by the microbes (phenazine derivatives or flavins). In mixed communities, shuttles produced by one species can help other species in their interactions with the electrode (Rabaey et al. 2004). Alternatively, artificial shuttles (like neutral red, methylene blue, methyl viologen, antheraquinone-2,6-disulfonate [AQDS], and resazurin) can be added in the medium. A new generation of putative electron transfer molecules, called conjugated oligoelectrolytes, has been recently introduced into this research field (Philips et al. 2015). These molecules integrate into the microbial membrane, making it leakier, which favors the release of electroactive components. In a great number of studies, artificial shuttles have been successfully supplied to nonelectroactive microbes in order to allow them to exchange electrons with the electrode without the need for attachment to it. However, long-term use of artificial shuttles is questionable; they represent an extra cost, are nonselective (they can be used by other microorganisms), may be unstable (e.g., methyl viologen) or toxic, and can be lost in the medium on their way to the electrode surface (Philips et al. 2015). In some
cases, MFCs have been extended to two biological compartments; the electrons microbiologically generated at the anode can be used by electrotrophic microbes at the cathode (for the elimination of \( \text{NO}_3^- \) and \( \text{NO}_2^- \) from wastewater, for the precipitation of soluble U(VI) into insoluble U(IV), etc.) (reviewed in Logan 2009).

Today, MFCs remain noncompetitive compared with chemical fuel cells in terms of efficiency and power output. However, they represent a more sustainable way to produce energy, which explains the tremendous effort that has been devoted so far to developing upscaled MFC systems to combine electricity production and wastewater treatment. Interestingly, another potentially more successful application of MFC is emerging: their integration into robotic or electronic systems as sole power source (i.e., in remote sensors, autonomous robots, or electronic gadgets). For this purpose, different strategies are being tested to miniaturize MFCs (while keeping enough power output) and improve their autonomy (Ieropoulos et al. 2012, 2016; Liu et al. 2016; Tommasi et al. 2016). For example, Ren et al. (2015) recently developed a miniaturized MFC equipped with an anode made of carbon nanotubes that reached a power density of 3320 W/m\(^3\).

MES cells are a variant of microbial electrochemical cells, where focus is on the reductive microbial reactions taking place at the cathode to produce specific chemical compounds (Rabaey and Rozendal 2010; Lovley and Nevin 2013; Tremblay and Zhang 2015). Various electron acceptors can be reduced at the cathode, ranging from \( \text{O}_2 \) to protons and \( \text{CO}_2 \). In particular, MES cells based on electroautotrophic microbes developing at the cathode have received increasing attention over the last 5 years due to their ability to transform \( \text{CO}_2 \), a greenhouse gas, into various organic compounds: \( \text{CH}_4 \), acetate, or other multicarbon compounds. For \( \text{CH}_4 \) production, another advantage of MES cells is that the gas they produce is enriched in methane (about 85%) compared with biogas resulting from anaerobic digestion, because of the consumption of carbon dioxide in MES cells. The utility of autotrophic MES cells for biological production of \( \text{H}_2 \) has been reported as well. Their initial functioning is a bit different, though, as the first prototypes (in the early 2000s) to generate \( \text{H}_2 \) relied on coating the cathode with hydrogenase enzymes. Such enzymatic cells typically possess power densities that are generally orders of magnitude greater than MES cells using whole microbes, but they are not self-regenerating and therefore losing activity. Therefore, as far as we know, they have been replaced by MES cells based on living microbes in all applications.

In MES cells, electrotrophic microbes are inoculated either as mixed communities (usually undefined cultures from wastewater, sludge, or sediment) or pure cultures. Mixed-community reactors were reported to have a remarkable production rate, but most of the time they produce a mix of compounds (e.g., alcohols–\( \text{H}_2 \)–organic acids or \( \text{CH}_4 \)–acetate–\( \text{H}_2 \)), illustrating the difficulty to generate a single specific product (Marshall et al. 2012; Tremblay and Zhang 2015). Therefore, most studies have focused on pure strains.

Compared with electron transfer to the anode in MFCs, electron transfer from the cathode to microbes is very poorly characterized. It is currently the object of much study, as understanding the underlying mechanisms of electron transfer could lead to major breakthroughs in the development of more efficient MES cells. Similarly
Evidence for both electron shuttle-mediated and direct electron transfer mechanisms has been found (Figure 13.5), although discrimination between both is often challenging.

The first MES cells for methane, acetate, or H₂ production relied on the addition of artificial shuttles (reviewed in Tremblay and Zhang 2015). Later, studies showed that the same processes could take place without artificial shuttles (Rozendal et al. 2008; Cheng et al. 2009; Villano et al. 2010). In the study of Rozendal et al. (2008), H₂ was produced after reverting the polarity of an anode, on which a mixed biofilm of acetate and hydrogen-oxidizing microbes had developed, to make it a biocathode. Their strategy was based on the well-known reversibility of hydrogenases. This experimental setup raised questions on the autotrophy of microbes producing H₂ (as they could have been exploiting organic compounds derived from the biofilm), but a recent study showed that it is possible to grow a self-regenerating electroactive biofilm at the cathode that uses CO₂ as the sole carbon source to produce H₂ as the sole compound (Jourdin et al. 2015). Indirect transfer via other exogenous shuttles (formate, Fe(II), NH₄⁺, and H₂) or via shuttles secreted by bacteria (phenazine, riboflavin, B12 vitamin, or DNA) has been reported as well (Philips et al. 2015). Among these, molecular hydrogen is a unique shuttle in MES cells. First, it can be used as an electron donor by all mixed communities and pure strains producing CH₄ or the multicarbon compounds that have been studied so far (Philips et al. 2015). Second,
it is abiotically generated in the cathode vicinity from the protons migrating from the anodic chamber, as long as cathode potential is sufficiently negative (i.e., lower than the formal potential of the $2\text{H}^+/\text{H}_2$ couple). This potential determines, in theory, which mechanism is used by bacteria and archaea to acquire the necessary electrons for their metabolism, that is, $\text{H}_2$-mediated transfer versus direct electron transfer (Tremblay et al. 2016). Direct electron transfer via c-type cytochromes has been demonstrated for metal-oxidizing bacteria, but the latter do not produce biofuels (unless they can be engineered to do so). It is suspected for electromethanogenic archaea and electroautotrophic bacteria (including electroacetogenic ones) can get electrons directly from a cathode, but the underlying mechanism is unknown. In practice, discrimination between $\text{H}_2$-mediated and direct electron transfer is challenging. For example, depending on the potential set at the cathode, $\text{CH}_4$ can be produced by direct transfer of electrons to the methanogenic archaea (Cheng et al. 2009; Villano et al. 2010) and/or by use of the $\text{H}_2$ abiotically produced at the cathode (Villano et al. 2010). The hypothesis of direct electron transfer is supported by observations in nature that some methanogens can corrode Fe(0) much faster than related hydrogenotrophic methanogens (Dinh et al. 2004; Uchiyama et al. 2010), suggesting the existence of a direct mechanism to oxidize metallic iron. In addition, a methanogenic strain ($\text{Methanococcus maripaludis}$) mutated to lack hydrogenases could still produce $\text{CH}_4$ (although at a low rate) (Lohner et al. 2014). However, some researchers (Deutzmann et al. 2015) suspect that Fe(0)-oxidizing methanogens actually exploit $\text{H}_2$ and formate in MES cells by releasing extracellular redox enzymes, such as hydrogenases and formate dehydrogenases, which associate with the electrode surface and catalyze the generation of $\text{H}_2$ and formate. Therefore, it is likely that in MES cells, most of the electrons used for $\text{CH}_4$ production are indirectly derived from $\text{H}_2$ produced at the cathode, either abiotically or biotically via extracellular enzymes or other microbes (in mixed biofilms). For example, in their study, Villano et al. (2010) used an uncharacterized methanogenic culture in which interspecies $\text{H}_2$ transfer between electroactive $\text{H}_2$-producing microorganisms and $\text{H}_2$-utilizing methanogenic archaea could have sustained $\text{CH}_4$ production. Future research is needed to investigate the involvement of mechanisms of direct electron transfer, production of redox enzymes, and interspecies interactions in biomethane production.

The same holds true for the production of acetate and other multicarbon compounds in MES cells. The latter include ethanol, butanol, or 2-3 butanediol, which can be produced by natural or genetically engineered Clostridia under appropriate conditions (Schiel-Bengelsdorf and Dürre 2012; Claassens et al. 2016). Several pure acetogenic cultures ($\text{Sporomusa}$ species, $\text{Clostridium}$ species, or $\text{Morella thermoacetica}$), as well as mixed acetogenic communities (dominated by $\text{Acetobacterium}$ species), were shown to convert CO$_2$ into acetate with electrons drawn from an electrode, but the mechanism of electron transfer was unclear (Nevin et al. 2011; Marshall et al. 2012). In pure cultures, direct electron uptake was hypothesized based on the absence of detectable concentrations of molecular $\text{H}_2$ and the observation of attachment of bacteria on the cathode surface (Nevin et al. 2011). However, all the aforementioned bacteria are capable of using $\text{H}_2$ as an electron donor for the reduction of CO$_2$, and therefore they could use an indirect electron uptake mechanism mediated...
by H₂ abiotically generated by the cathode (and not detected). In addition, Clostridia, Acetobacteria, and M. thermoacetica are gram-positive bacteria, and it is unclear how electrons could be transferred through their thick cell wall; in addition, some have no cytochromes encoded in their genome. However, the idea of direct electron transfer has recently returned to the fore by the recent description of an acetogenic strain capable of using metallic iron as a sole electron donor (Kato et al. 2015). This strain was highly related to Sporomusa sphaeroides and was capable of significantly enhancing Fe(0) corrosion in comparison with abiotic corrosion, while hydrogenotrophic acetogens could not. This suggested a direct, as yet unraveled electron uptake mechanism. As far as acetogenic communities enriched on cathodes are concerned, they most likely rely on (1) hydrogen abiotically generated at the cathode when low electrode potentials are used, and (2) hydrogen produced by other, nonacetogenic bacteria drawing electrons from the cathode and then used by hydrogenotrophic acetogens (Patil et al. 2015). Interspecies H₂ transfer within a community characterized by a high rate or efficiency of acetate synthesis was recently reported as well (Jourdin et al. 2016). The authors suggested that such a high rate of efficiency was because consumption of all H₂ and its transformation into acetate took place within the biofilm, that is, between closely located microbes, and that H₂-producing bacteria improved the electrode surface by producing copper nanoparticles.

Besides their capacity to use CO₂ directly as feedstock, a major strength of MES cells is their versatility with regard to the source of electrons. Indeed, the necessary electricity can be generated from multiple renewable sources, like solar energy, wind, or wastewater. Critics of renewable energy often cite the fact that such electricity sources are intermittent. In this context, MES becomes a complementary technology that allows direct storage of electricity into value-added chemical commodities. Another strength is their efficiency. Tremblay and Zhang (2015) compared the efficiency of plant photosynthesis versus MES coupled with solar cells. In nature, plants convert sunlight to biomass with an efficiency of 1%–2% (3% for microalgae). The conversion efficiency of electricity to chemical commodities by MES processes is about 80%–90%. As photovoltaic solar panels are at least six times more efficient at capturing the sun energy than plants, they concluded that powering MES with electricity from solar cells was a more potent strategy for storing the sun energy into multicarbon compounds than photosynthesis. A similar conclusion was drawn by Claassens et al. (2016), who compared the solar-to-product energy conversion efficiency of plant biomass fermentation, for example, into ethanol (in total <1%) and of solar panels coupled to electrocatalysis (to produce H₂) and to chemoautotrophic MES cells fed with H₂ and CO₂ (in total ≈7.6%).

Despite all its potential, MES is not an economically viable technology yet, because of the high electron demand for the reduction of CO₂, and therefore low production rates. Recent advances could, however, help their transition into commercial applications (Rosenbaum and Franks 2014). First, to minimize internal power losses, cathodes could be engineered to improve their specific surface area, topography, electric conductivity, and biocompatibility. To date, stainless steel is the most efficient cathode material, but cathodes are generally made of cheaper carbonaceous material (plain graphite, graphite granules, or carbon felt, cloth or granules, and vitreous
Microbial Fuels

There are different possibilities to increase electron transfer at moderate cost (reviewed in Rinaldi et al. 2008; Tremblay and Zhang 2015). One of them is to increase the electrode surface available for bacterial colonization using nanostructured materials like carbon nanotubes or graphene. Another is coating the electrode with positively charged compounds (e.g., chitosan) to favor attachment of bacteria (which are negatively charged). Optimization of other parts of the electrochemical hardware is also ongoing, such as the removal of the expensive ion exchange membrane or the improvement of the electrical circuit. Such technical improvements have contributed to the recent 400- to 500-fold increase in organic compound production rates (in mM/day/m²) and electron transfer rates (in mA/m²) (Tremblay and Zhang 2015). Nevertheless, this trend will only continue if technical improvements are coupled with careful selection of strains and a better understanding of all the interactions between microorganisms and electrodes. In that respect, the wide variety of autotrophs in nature provide a largely untapped pool of potential hosts that could be converted into autotrophic cell factories, as illustrated by the recent contribution of metallic iron-oxidizing strains to MES. Their performance could be improved by genetic engineering, providing that the required genetic tools are available (Rosenbaum and Henrich 2014; Claassens et al. 2016). Alternatively, the ability to interact with electrodes could be introduced into strains that are both metabolically versatile and easy to engineer, as attempted in *E. coli* (Jensen et al. 2010). In addition, as mentioned earlier, a lot of research is currently dedicated to characterizing electron transfer between microbes and the cathode, as this represents a cornerstone for more efficient MES production of chemical commodities. New transfer mechanisms could be discovered, but there is plenty of space to improve known ones as well. For example, Tan et al. (2016) recently succeeded in modifying a strain of *G. sulfurreducens* so that it could produce nanowires containing the amino acid tryptophane, which can promote fast electron transport, resulting in a 2000-fold increase in nanowire conductivity.

Another way of diversifying MES applications and obtaining sufficient productivity rates is electrofermentation (reviewed in Rabaey and Rozendal 2010; Schievano et al. 2016), a hybrid metabolism in which electrons from the cathode are utilized as coreducing equivalents in addition to carbon source-derived reducing equivalents (e.g., NADH) during heterotrophic fermentation. This way, fermentation is redirected toward more electron-dense desired products (e.g., propionate, glutamic acid, butyric acid, ethanol, or butanol). In early electrofermentation based on pure strains, the addition of shuttles was necessary due to the lack of an electroactive heterotroph. For example, Steinbusch et al. (2010) tested different shuttles to stimulate acetate reduction into ethanol by a mixed inoculum growing at the cathode. Mediators are not necessary anymore: Choi et al. (2014) showed that the heterotroph *Clostridium pasteurianum* could simultaneously use a soluble organic substrate (glucose or glycerol) and a cathode as electron donors. Usually, *C. pasteurianum* produces acetate and butyrate (and little butanol) from glucose, and butanol and 1,3-propanediol from glycerol. When grown in the presence of glucose or glycerol and a small electrode supply by the cathode, its fermentation profile was steered to produce a greater amount of electron-dense products (butanol and 1,3-propanediol, respectively). The mechanism
of direct electron transfer in this gram-positive bacterium was not unraveled. The authors speculated that such results were opening up the possibility of using electricity to efficiently produce electron-dense chemicals and fuels.

13.6 Targeted Approach 5: Dynamic Macroscopic Models of Bioprocesses and Their Application to Microbial Fuel Production

13.6.1 Introduction

Many dynamic models have been used for modeling key component concentration time profiles in bioprocesses. Most of these bioprocess models can be found in the framework of biopharmaceutical and agro-food applications where the products of interest are obtained through cultures of microorganisms (bacteria and yeasts) or mammalian cells in bioreactors. Dynamical models at the macroscopic scale allow reproducing the concentration time profiles of the cells, main substrates, metabolites, and products of interest (Bastin and Dochain 1990). Models that consider the cells as a global entity (without paying attention to intracellular metabolism) and their population as homogeneous are generally classified as unstructured and unsegregated. This kind of macroscopic model typically consists of a set of differential equations that represent the mass balances of the above-mentioned main components. Such a macroscopic description allows the building of a model-based control and monitoring tools. For instance, bioreactor feeding time profiles can be determined for optimizing the bioprocess (e.g., maximum yield of substrate to cells or maximum cell productivity) (Banga et al. 2005; Amribt et al. 2014; Richelle and Bogaerts 2014). Closed-loop controllers can be tuned for maintaining some variables at their optimal set point despite the presence of process disturbances and/or modeling errors (e.g., maintaining microorganism cultures at the limit of overflow metabolism) (Van Impe and Bastin 1998; Smets et al. 2004; Dewasme et al. 2010). Software sensors (also called state observers) can be used for online estimation of key variable concentrations that are not directly measured through sensor probes (e.g., a cell density software sensor based on output gas analysis and base feeding flow rate) (Bogaerts and Vande Wouwer 2003; Ali et al. 2015). All these engineering tools have the common feature of being model-based solutions, and they all require a mathematical description of the process that is sufficiently detailed for accurately reproducing the basic biological phenomena linking the essential process variables, but sufficiently simple so that they can be efficiently used for reaching the engineering goal.

With the increasing number of microbial fuel production processes, the use of these dynamic macroscopic models (and the associated model-based engineering tools) has naturally been extended to this new application field. Although these processes have, of course, their own specificities, the general framework described above remains applicable with some adaptations. Section 13.6.2 provides to the reader some basic concepts about dynamic macroscopic models of bioprocesses and shows how general kinetic laws (which are the key issues in this kind of model) can provide very flexible and generic model structures. Sections 13.6.3 and 13.6.4 present two case
studies (bioethanol production from potato peel wastes and microalgae cultures) that illustrate how the general framework of Section 13.6.2 can be adapted to specific applications.

13.6.2 Dynamic Macroscopic Models of Bioprocesses and General Kinetic Model Structures

The first step in macroscopic modeling of bioprocesses is to summarize the main biological phenomena (e.g., cell growth, maintenance, and death) that link the cells and the main extracellular components (substrates, metabolites, and products of interest) (Bastin and Dochain 1990). This leads to the definition of a macroscopic reaction scheme, which contains \( M \) reactions involving \( N \) products:

\[
\sum_{i \in R_k} K_{ik} \xi_i \xrightarrow{\varphi_k} \sum_{j \in P_k} K_{jk} \xi_j \quad k \in [1,M], \quad i,j \in [1,N] \tag{13.1}
\]

where \( \varphi_k \) is the rate of the \( k \)-th reaction, \( K_{ik} \) and \( K_{jk} \) are pseudo-stoichiometric (or yield) coefficients, \( \xi_i \) and \( \xi_j \) are the extracellular components (substrates, metabolites, products, or cells), \( R_k \) is the set of component indices that are reactants (or catalysts) in the \( k \)-th reaction, and \( P_k \) is the set of component indices that are products (or [auto-]catalysts) in the \( k \)-th reaction.

Note that there exist methods for determining the minimal number of reactions \( M \) and even for determining the global reaction scheme on the basis of experimental measurements of the different component concentrations (Bernard and Bastin 2005a, 2005b; Hulhoven et al. 2005). The general dynamic macroscopic model (Bastin and Dochain 1990) is made of the ordinary differential equations (ODEs) describing the mass balances of each of the \( N \) components involved in the reaction scheme (Equation 13.1):

\[
\frac{d\xi(t)}{dt} = K \varphi(\xi(t)) - D(t)\xi(t) + F(t) - Q(t) \tag{13.2}
\]

where \( \xi(t) \in \mathbb{R}^N \) is the vector of the \( N \) component concentrations, \( \varphi(t) \in \mathbb{R}^M \) is the vector of the \( M \) reaction rates, \( K \in \mathbb{R}^{N \times M} \) is the (pseudo-)stoichiometric or yield matrix \( (K_{ij} \) is the yield coefficient of the \( i \)-th component in the \( j \)-th reaction), \( D(t) \in \mathbb{R} \) is the dilution rate (feeding flow rate divided by the volume), \( F(t) \in \mathbb{R}^N \) is the vector of external inflow rates, and \( Q(t) \in \mathbb{R}^N \) is the vector of gaseous outflow rates.

Given some assumptions, the yield matrix \( K \) (or at least part of it) can be identified on the basis of experimental measurements of \( \xi(t) \) through simple linear regressions and without any knowledge about the reaction rates \( \varphi(t) \) (Chen and Bastin 1996; Bogaerts et al. 2003). The key issue in Equation 13.2 is the mathematical structure of the reaction rates \( \varphi(t) \), which are typically nonlinear functions (of the different component concentrations \( \xi(t) \)) representing different biological phenomena. There exist
plenty of kinetic models representing particular phenomena (activation, saturation, and inhibition) for particular components (substrates, metabolites, etc.) with particular mathematical functions. In order to circumvent the different choices that have to be made for a given application, it is possible to use general kinetic model structures that allow representing in a systematic way the different biological phenomena due to any component in any reaction (Savageau 1969a, 1969b; Haag et al. 2005; Grosfils et al. 2007; Richelle and Bogaerts 2015). A first example is the extended Monod law framework, where each reaction rate is modeled as

$$
\phi_k(\xi) = \mu_{\text{max},k} \prod_{l \in A_k} \frac{\xi_l}{\xi_l + K_{S,kl}} \prod_{m \in I_k} \frac{1}{\xi_m + K_{I,km}} X
$$

where $\mu_{\text{max},k}$ is the maximum specific rate of the $k$-th reaction, $\xi_l$ is the concentration of the $l$-th component in vector $\xi$, $X$ is the cell density, $A_k$ is the set of indices of the components that activate (as reactants or catalysts) the $k$-th reaction, $I_k$ is the set of indices of the components that inhibit (as metabolites or reactants) the $k$-th reaction, $K_{S,kl}$ is the saturation constant of the $l$-th component in the $k$-th reaction, and $K_{I,km}$ is the inhibition constant of the $m$-th component in the $k$-th reaction.

A second example of a general kinetic model is given by (Grosfils et al. 2007; Richelle and Bogaerts 2015)

$$
\phi_k(\xi) = \mu_{\text{max},k} \prod_{l \in A_k} \xi_l^{\gamma_{lk}} \prod_{m \in I_k} \xi_m^{\beta_{mk}} X
$$

where $\gamma_{lk}$ is the activation coefficient of the $l$-th component (reactant or catalyst) in the $k$-th reaction, and $\beta_{mk}$ is the inhibition coefficient of the $m$-th component (metabolite or reactant) in the $k$-th reaction.

Note that this latter general kinetic model structure has the advantage that it can be rigorously linearized, up to a logarithmic transformation, hence allowing the identification of the kinetic parameters $\mu_{\text{max},k}$, $\gamma_{lk}$, and $\beta_{mk}$ based on simple linear regressions. It is also worth noting that it is possible to go back to the extended Monod formalism (Equation 13.3) based on a model identified with Equation 13.4, as shown by Richelle and Bogaerts (2015).

The concept of the macroscopic reaction scheme (Equation 13.1), the mathematical model consisting of the corresponding mass balances (Equation 13.2), and general kinetic model structures like Equations 13.3 and 13.4, can be used for modeling bioprocesses in many different fields of applications. The next two sections focus on two particular case studies in the framework of microbial fuels.

**13.6.3 Case Study 1: Bioethanol Production from Potato Peel Wastes**

In this case study, potato peel residues are used as a source of fermentable sugars that are subsequently used as carbon source in *S. cerevisiae* batch cultures for
producing bioethanol (Richelle et al. 2015). The additional use of ammonium as an inorganic nitrogen source provides a possibility to boost the fermentation metabolism. However, too high ammonium concentrations may lead to an inhibition (or at least saturation) effect. A macroscopic model of the process has been proposed and validated and allowed, determining the optimal initial ammonium concentration to be provided to a batch reactor for maximizing the bioethanol production, and which will be fully depleted at the end of the batch run.

The initial macroscopic reaction scheme, corresponding to the general form given in Equation 13.1, that was proposed by Richelle et al. (2015), contained three reactions involving four components (fermentable sugars, ammonium, yeast cells, and ethanol), and the corresponding dynamic model (Equation 13.2) (using extended Monod kinetics [Equation 13.3]) presented 14 parameters to be identified. Based on experimental data coming from six batch cultures (differing by their initial ammonium concentrations) and a multistep parameter identification procedure, it was shown that the model could be strongly reduced (based on the parameter uncertainty analysis), leading to the final very simple macroscopic model:

\[
\begin{align*}
S & \xrightarrow{\varphi_1} \alpha_1X + \beta_1E \\
N & \xrightarrow{\varphi_2} \beta_2E
\end{align*}
\]  
(13.5)

where \(S, N, X,\) and \(E\) represent, respectively, the fermentable sugars, ammonium, yeast cells, and ethanol; \(\alpha_1, \beta_1,\) and \(\beta_2\) are yield coefficients; and \(\varphi_1\) and \(\varphi_2\) are the reaction rates.

Note that the first reaction represents the classical fermentation of the sugars, leading to yeast growth and ethanol production (not coupled with a respiration reaction after model reduction), and that the second reaction represents the boost provided by nitrogen consumption on the ethanol production (not coupled with biomass growth after model reduction).

The dynamic model corresponding to the mass balances of the components involved in (Equation 13.5) is given by

\[
\begin{align*}
\frac{dS}{dt} &= -\varphi_1 \\
\frac{dN}{dt} &= -\varphi_2 \\
\frac{dX}{dt} &= \alpha_1\varphi_1 \\
\frac{dE}{dt} &= \beta_1\varphi_1 + \beta_2\varphi_2
\end{align*}
\]  
(13.6)
The reaction rates (still after model reduction) use finally simple Monod laws:

\[
\varphi_1 = \mu_{\text{max},S} \frac{S}{S + K_S} X \\
\varphi_2 = \mu_{\text{max},N} \frac{N}{N + K_N} X
\]

(13.7)

The values of the six parameters, \(\alpha_1, \beta_1, \beta_2, \mu_{\text{max},S}, \mu_{\text{max},N}\), and \(K_N\), were identified based on different subsets of the six batch cultures. Note that an arbitrarily low value has been fixed for \(K_N\), as it was associated to high variation coefficients but could not be canceled, so that the model stops the nitrogen consumption when nitrogen is fully depleted. The identified values can be found in the paper by Richelle et al. (2015).

Figure 13.6 presents cross-validation results, that is, comparisons between concentrations estimated with Equations 13.6 and 13.7, compared with experimental data that were not used for the parameter identification. For instance, the first two rows corresponding to Cross-Validation 1 present a comparison of model estimates with experimental data of, respectively, Experiments 1 and 4, the model parameters having been identified on the basis of the measurements from Experiments 2, 3, 5, and 6. The model leads to satisfactory predictions of the process behavior.

Figure 13.6  Comparison between model simulation (blue curves) and measurements of all subsets of two different batch experiments (Exp.) used for cross-validation (read circles with 95% confidence intervals). (Reproduced from Richelle, A. et al., Bioprocess Biosyst. Eng., 38(9), 1819–33, 2015.)
Using the model identified with the whole set of experiments, the final ethanol concentration was computed for different initial ammonium concentrations, all the other experimental variables being fixed (batch duration of 17 hours and initial concentrations of fermentable sugars, yeast cells, and ethanol corresponding to the mean values of the six experiments). The results are plotted in Figure 13.7 and show that there is an optimal initial ammonium concentration (170 mg/L) that maximizes the final bioethanol production and above which no fermentation boost can be obtained anymore under these experimental conditions.

This case study illustrates how a very simple dynamic macroscopic model (two reactions, four components, and six parameters to be identified) can exhibit very satisfactory prediction abilities, and how it can be used to determine optimal experimental conditions.

13.6.4 Case Study 2: Microalgae Cultures

This second case study concerns microalgae cultures that are used, besides other applications, to produce large quantities of lipids (even above 50% dry weight, at least at the lab scale), which can then be transformed into biodiesel through a transesterification reaction. Dynamic macroscopic models of microalgae cultures are generally based on the well-known Droop model (Droop 1968, 1983). The key concept of this model is to assume that the uptake of the limiting substrate (e.g., a nitrogen source) leads to the accumulation of an internal quota, and the microalgae growth depends on (and starts only above a minimal value of) that internal quota. Note that by
considering this intracellular quota, the model should not be classified as “unstructured” anymore and should be considered a “structured and unsegregated” model. Still making the parallel with the general framework described in Section 13.2, the reaction scheme (Equation 13.1) becomes

\[
S \xrightarrow{\phi_1} Q \\
Q \xrightarrow{\phi_2} Q + X
\]  

(13.8)

where \( S, Q, \) and \( X \) represent, respectively, the limiting substrate, the internal quota, and the microalgae, and \( \phi_1 \) and \( \phi_2 \) are the reaction rates.

The first reaction represents the substrate uptake leading to internal quota accumulation. The second one represents the microalgae growth catalyzed by the internal quota.

The dynamic model corresponding to the mass balances of the components involved in Equation 13.8 is given by

\[
\frac{dS}{dt} = -\phi_1 - DS + DS_{in}
\]

\[
\frac{d(QX)}{dt} = \phi_1 - D(QX)
\]

\[
\frac{dX}{dt} = \phi_2 - DX
\]

(13.9)

where \( S \) and \( X \) are the volumetric concentrations of the limiting substrate (mass of substrate per unit of volume) and of the biomass (mass of microalgae cells per unit of volume), \( Q \) is the intracellular quota concentration (mass of intracellular substrate per unit of microalgae cell mass), \( D \) is the dilution rate, and \( S_{in} \) is the substrate concentration in the feeding.

The reaction rates are given by

\[
\phi_1 = \rho(S)X = \rho_{max} \frac{S}{S + K_s} X
\]

\[
\phi_2 = \mu(Q)X = \mu_{max} \left(1 - \frac{Q_{min}}{Q}\right) X
\]

(13.10)

where \( \rho_{max} \) and \( \mu_{max} \) are the maximum specific rates of substrate uptake and biomass growth, \( K_s \) is a saturation constant, and \( Q_{min} \) stands for the minimum quota for microalgae growth. Bernard and Gouzé (1995) showed that the internal quota is bounded as

\[
Q_{min} \leq Q \leq Q_{max}
\]

(13.11)
where the maximum internal quota is given by

\[ Q_{\text{max}} = Q_{\text{min}} + \frac{\rho_{\text{max}}}{\mu_{\text{max}}} \quad (13.12) \]

Injecting Equation 13.10 into Equation 13.9 and developing the time derivative of the product \( Q \cdot X \) leads to the usual Droop model:

\[
\begin{align*}
\frac{dS}{dt} &= -\rho(S)X - DS + DS_{\text{in}} \\
\frac{dQ}{dt} &= \rho(S) - \mu(Q)Q \\
\frac{dX}{dt} &= \mu(Q)X - DX 
\end{align*}
\quad (13.13)
\]

Benavides et al. (2015) proposed a dedicated methodology for the parameter identification of the model (Equations 13.10 and 13.13) (with four kinetic parameters, \( \rho_{\text{max}}, K_S, \mu_{\text{max}}, \) and \( Q_{\text{min}} \), and three initial conditions, \( S(0), Q(0), \) and \( X(0) \)), based on two batch experiments. A cross-validation test (validation with two other batch experiments, not used for parameter estimation) is reproduced in Figure 13.8.

**Figure 13.8** (a and b) Are two independent batch cultures used to test the predictive capability of the model by cross-validation. Comparison between model simulation (blue curves) and measurements of two different batch experiments used for cross-validation (red curves for online measurements and red dots with 95% confidence intervals for off-line measurements). (Reproduced from Benavides, M. et al., *Bioprocess Biosyst. Eng.*, 38(9), 1783–93, 2015.)
The goodness of fit is quite acceptable, apart from the internal quota $Q$, which decays experimentally and not in the model simulation. The authors propose a solution to that problem by considering a consumption of the internal quota with the biomass growth (leading to an additional negative term in the mass balance of $Q$).

The Droop model can be extended in many ways for taking into account light limitation effects in photobioreactors, temperature limitation effects, and so forth (Bernard et al. 2016). It can also be extended for reproducing the production of lipids, like triacylglycerols, which can be transformed into biodiesel (Bernard et al. 2016).

The Droop model has been used for building software sensors, for example, interval observers for estimating substrate and internal quota concentrations based on microalgae concentration measurements (Goffaux et al. 2009). An extended version was also used to test in simulation adaptive control laws aimed at maximizing the microalgae production (Mairet et al. 2015).

13.7 Conclusion

In this chapter, we have provided an overview of different ways to improve biofuel production that seemed relevant to us. Genetic engineering and bioinformatic modeling tools are presently driving substantial improvements in the design of efficient enzymatic catalysts and whole-cell factories. In parallel, the application of molecular biology tools (omic techniques) to unravel the diversity and dynamics of complex communities historically used for biofuel production (i.e., fermentative and methanogenic microbes used in anaerobic digestion) has led to a much better understanding of their functioning that can help optimize their performance. These tools have also shed light on an untapped diversity of enzymes and microorganisms in natural environments that could be exploited for biofuel production. In the forthcoming years, we expect biofuels to be increasingly produced by low-complexity consortia rather than pure strains, as the latter have repeatedly proved to be more robust and more versatile. Their internal functioning is, however, neither fully characterized nor mastered. In this context, omic techniques could provide a solid analytical platform to understand the interactions between partners and guide their metabolism toward targeted products. Finally, the versatility of microbial metabolism is currently exploited in bioelectrochemical applications for the production of a variety of chemical commodities, including biofuels. This technology is young and is still being investigated at the laboratory scale, but it could be scaled up and integrated into biorefineries, providing that the current spectacular improvements in production rate and electron transfer rates observed over the last couple of years persist. Such improvements imply a better characterization of electron transfer between microbes and electrodes, improved electrode materials, and better management of microbial metabolisms involved in such processes. Time will tell, among these different ways of exploiting microbial metabolism, which ones will lead to commercial production of microbial biofuels. At the macroscopic level of the biofuel production process, several techniques allow the building of dynamic models that are able to predict the concentration time profiles of the main culture species (substrates, cells, metabolites, and products of interest). They can be successfully used for...
model-based process optimization, control, and soft sensing. Future research will help to build efficient bridges between this macroscopic level of description and the available knowledge at the metabolic level (e.g., metabolic networks). Although many research efforts have already been devoted to some microorganisms and mammalian cell lines, many open problems remain, for instance, in the field of microalgae.

References


Choi, Okkyoung, Taeyeon Kim, Han Min Woo, and Youngsoon Um. 2014. Electricity-Driven Metabolic Shift through Direct Electron Uptake by Electroactive Heterotroph *Clostridium pasteurianum*. *Scientific Reports* 4: 6961.


