
Constraint based modelling going multicellular

Patricia do Rosario Martins Conde¹, Thomas Sauter^{1,*} and Thomas Pfau^{1,2}

¹Life Sciences Research Unit, University of Luxembourg, L-1511 Luxembourg, Luxembourg

²Institute of Complex Systems and Mathematical Biology, University of Aberdeen, AB24 3UE Aberdeen, United Kingdom

Correspondence*:

Thomas Sauter

Life Sciences Research Unit, University of Luxembourg, L-1511 Luxembourg, Luxembourg, thomas.sauter@uni.lu

ABSTRACT

Constraint based modelling has seen applications in many microorganisms. For example, there are now established methods to determine potential genetic modifications and external interventions to increase the efficiency of microbial strains in chemical production pipelines. In addition, multiple models of multicellular organisms have been created including plants and humans. While initially the focus here was on modelling individual cell types of the multicellular organism, this focus recently started to switch. Models of microbial communities, as well as multi-tissue models of higher organisms have been constructed. These models thereby can include different parts of a plant, like root, stem or different tissue types in the same organ. Such models can elucidate details of the interplay between symbiotic organisms, as well as the concerted efforts of multiple tissues and can be applied to analyse the effects of drugs or mutations on a more systemic level. In this review we give an overview of the recent development of multi-tissue models using constraint based techniques and the methods employed when investigating these models. We further highlight advances in combining constraint based models with dynamic and regulatory information and give an overview of these types of hybrid or multi-level approaches.

Keywords: Multi-scale modelling, constraint based modelling, multi-tissue modelling, multi-organism modelling, metabolic modelling

1 INTRODUCTION

While genome sequences are now readily available, determining the metabolic properties of organisms is still an open problem. Numerous methods for modelling and analyzing metabolism exist (for a general comparison please refer to (Bordbar et al., 2014)), including a large diversity of multi-scale approaches. E.g. Karr et al. (2012) reconstructed a whole-cell model of *Mycoplasma genitalium*. The overall model combines multiple different modelling techniques for describing different levels of the organism, reaching from regulatory signaling to metabolism and other pathways. However, its comprehensiveness leads to a large number of parameters, making this approach practically quite challenging for more complex organisms. There also exist some applications of kinetic modelling to study different tissues, like liver (Ricken et al., 2015; Zeigerer et al., 2015), heart (Crampin et al., 2004), and brain metabolism (Jolivet et al., 2015),

mainly in the framework of the Virtual Physiological Human Project (Viceconti et al., 2008). A hepatocyte model integrating signaling and regulatory information along with metabolism based on kinetic modelling has recently been published (Ryll et al., 2014). Petri-nets form another approach used for multi-scale modelling (Berestovsky et al., 2013). However, due to the availability of genome annotations and the lack of knowledge on kinetic parameters, stoichiometric network models of metabolism form an widely used and very scalable basis. The concept of constraint based modelling (CBM) allows fast calculations of large networks under the steady state assumption, relying mainly on genetic information easily obtainable by genome sequencing. Recent combinations of these types of models with other modelling techniques show promising results. We will therefore focus in this review on constraint based modelling and its combination with other techniques.

Constraint based modelling (CBM) aims at helping researchers to get a better insight into the complex system of metabolism (Llaneras and Picó, 2008). Additionally, they exploit the property of metabolism that reactions have defined substrates and products. Based on a network of biochemical reactions, a stoichiometric matrix S is created, with columns representing reactions, rows representing metabolites, and entries $S(i, j)$ representing the stoichiometric coefficient of metabolite i in reaction j . For larger networks this requires the application of computational tools usually providing the models in SBML format being an important step towards reproducible science (Pfau et al., 2015). The major constraint imposed by CBM methods is the assumption of an internal quasi steady state of the investigated system. It is assumed that the internal concentrations of metabolites do not change over time (i.e. $S \cdot v = 0$, where v is the flux distribution vector of the system). In addition, three primary constraints are introduced to obtain biologically relevant solutions (Lewis et al., 2012; Orth and Palsson, 2010): mass and charge conservation within reactions, dependency on substrate and enzyme availability and reversibility constraints based on thermodynamics. It is also common to employ an objective function that the system is assumed to be optimized for (Schuetz et al., 2007). The implemented functions range from growth (Feist and Palsson, 2010) over production of ATP to complex combinations of multiple simpler objectives (Vo et al., 2004) and are a particular challenge when multicellular organisms are the target of research. One of the most common approaches to CBM is the use of flux balance analysis (FBA) (Savinell and Palsson, 1992; Varma and Palsson, 1994; Kauffman et al., 2003; Raman and Chandra, 2009; Orth and Palsson, 2010). FBA has been used to investigate the effects of knockouts on metabolism (Segrè et al., 2002; Shlomi et al., 2005) and to design knockout strategies for metabolic engineering (Burgard et al., 2003; Rocha et al., 2010; von Kamp and Klamt, 2014). Basic FBA however, does not account for model dynamics like gene regulation, signaling processes, or metabolic regulation (Orth and Palsson, 2010). Therefore, approaches considering these regulatory elements have been developed (Covert and Palsson, 2002; Mahadevan et al., 2002; Covert et al., 2008). Models for multiple species from all biological kingdoms (prokaryotes, eukaryotes and archaea) have been reconstructed (see <http://systemsbiology.ucsd.edu/InSilicoOrganisms/OtherOrganisms>), and their number is constantly increasing. While the first models used in CBM were mostly aimed at central carbon metabolism, the ever growing availability of genome sequences has let to a rapid increase in genome scale metabolic reconstructions (GSM). The first GSM was *Haemophilus influenzae Rd* (Edwards and Palsson, 1999) in 1999, followed four years later by the first eukaryotic model (*Saccharomyces cerevisiae*) (Förster et al., 2003) and the first mammalian model (*Mus musculus*) in 2005 (Sheikh et al., 2005). The first human reconstruction (Recon1 by Duarte et al. (2007)) was closely followed by the publication of a second human GSM, the Edinburgh Human Metabolic Network, within the same year (Ma et al., 2007). These initial models have seen many improved versions over the past years (e.g. Recon2 (Thiele et al., 2013) and HMR2 (Mardinoglu et al., 2014) for human). While general reconstructions serve as an important knowledge base for our understanding of metabolic capabilities within the reconstructed organism, they

are limited when investigating multicellular organisms exhibiting multiple different tissues. Since different cell types of higher organisms have different functions, and indeed different capabilities, it is therefore necessary to reconstruct tissue specific models. In a recent review, Ryu et al. (2015) give an extensive overview over currently available reconstructions. In addition to manual reconstruction, multiple methods exist for contextualization of metabolic networks (reviewed in Resendis-Antonio et al. (2014), Machado and Herrgård (2014), Robaina Estévez and Nikoloski (2014) and Ryu et al. (2015)). While tissue specificity can help to elucidate important information about a specific tissue, a single tissue model on its own is unable to inform about the complex interactions occurring in a higher organism. This necessitates the combination of models presenting multiple tissues. In this review we will first provide an overview of methods developed for the integration of non metabolic information in constraint based models and give an overview of methods used for model simulation. We will continue by detailing the recent development of constraint based models spanning multiple organisms or tissues along with methods employed when investigating these models. Finally we will present recent advances in combining constraint based models with dynamic and regulatory information and give an overview of these types of hybrid or multi-scale approaches.

2 FBA METHOD: DATA INTEGRATION AND EXTENSIONS

Multiple methods can be used to improve constraint based model predictions. Figure 1 provides an overview of these methods, ordered by their time of publication. These methods include ways to integrate regulatory events which can influence the activity of reactions by altering the production of specific proteins necessary to catalyze these reactions. They also provide concepts which allow the use of omics data to determine the availability of enzymatic activities. The first attempts to integrate non metabolic information into a CBM analysis was introduced by Covert et al. (2001), who linked regulatory data to a CBM model of *Escherichia coli* turning reactions on and off (a method termed rFBA). The principle of flux adjustment based on external data was picked up by others and either used to switch reactions on or off (Becker and Palsson, 2008; Vlassis et al., 2014) or to adjust the bounds of fluxes (Colijn et al., 2009; Lee et al., 2012). The former try to derive the activity state of genes and assign reaction availabilities based on these activities using Boolean gene-protein-reaction association rules. These rules are similar to the Boolean rules used by Covert et al. for regulatory constraints, but only use gene activity information, where Covert et al. used additional information, like the availability of a preferred carbon source deactivating a second importer. These methods used for omics integration have received much attention lately and were extensively reviewed in (Machado and Herrgård, 2014) and (Robaina Estévez and Nikoloski, 2014). The reviews focused on flux bound adjustment and network structure contextualization, respectively.

Simultaneous to the development of rFBA, Mahadevan et al. introduced the concept of dynamic FBA (dFBA), which allows the simulation of time courses using FBA (Mahadevan et al., 2002). The idea is to use the resulting outputs (e.g. remaining amount of substrate, generated products) of an earlier time step as inputs for the next time step. Both concepts were later combined to form integrated dynamic FBA (iFBA or idFBA) which uses both time dependent regulatory information and the dFBA approach (Covert et al., 2008; Lee et al., 2008). In addition, Covert et al. used a small dynamic ODE model to dynamically simulate parts of the network based on the results of the earlier time step. This was also the first attempt to combine constraint based and dynamic models in a common framework. In iFBA the ODE model was solved to calculate initial rates, which were subsequently applied to the constraint based model as flux bounds. In addition, gene expression data for each time step was applied to the CBM restricting the active

parts. Finally, the CBM was optimized and the resulting metabolite concentrations used as inputs to the next iteration of the ODE model.

Integration of different data types on the same model, requires the development of new optimization approaches. Each of these approaches is specifically designed for the research problem of interest. In Table 1, an overview on the main FBA-based methods discussed in this review is given. For a broader view on FBA methods, please refer to Lewis et al. (2012). Since 2012, other optimization methods have been developed, and we will briefly present some of them here. Pozo et al. (2012) developed a new optimization algorithm which allows to determine the global optimum of kinetic metabolic models while optimizing for multiple-objectives. Additionally, this algorithm allows the selection of sets of the most efficient optimal alternatives. This optimization approach can be very interesting when being applied to metabolic engineering, i.e. it can be applied when the objective is to optimize the synthesis rate of metabolite X, at minimum cost, and with minimum concentration change of intermediate metabolites. Recently, another approach, developed by Andreozzi et al. (2015), employs CBMs at steady-state and metabolite concentrations to derive feasible kinetic models which are representative of a specific physiological state. As the previous algorithm, this framework is easily applicable to metabolic engineering and can give hints about which enzymes to alter in order to achieve a specific physiological state. Furthermore, it can be used to decrease the uncertainty related with kinetic parameter estimation and to efficiently sample the solution space.

Constraint based metabolic models are commonly aimed to simulate the metabolism of small molecules and tend to include macromolecular biosynthesis and modifications only implicitly by the inclusion of energy or reductant consuming reactions. Recently this neglect has been addressed in studies on *E.coli* and *Thermotoga maritima* which explicitly modelled the macromolecular biosynthesis machinery (Thiele et al., 2012; Lerman et al., 2012). The basis for this approach is a CBM model of the target organism, which was noted as M-matrix and a stoichiometric model of the macromolecular synthesis machinery, noted as E-matrix. The models are combined by forming a large ME matrix and adding coupling constraints that restrict the flux through reactions by requiring the catalyzing enzymes to be available. In addition, biomass is adjusted to reflect the explicit costs of macromolecular synthesis and amino acids, since proteins are allowed to accumulate in the model. In Thiele et al. (2012), the ME-model was simulated and the predictions matched with experimental growth rates and knockout phenotypes. Lerman et al. (2012) used the technique to investigate minimal ribosomal production rates necessary at specific growth rates, and could show changes consistent with experimental data. They also found pathways which become necessary for efficient growth in the ME-model but which are not important in a pure metabolic model. Thiele et al. (2012) mention that performing FBA in the ME-matrix is time consuming, therefore this approach does not easily scale to larger models. In practice, this dimensionality is likely to become too high for many eukaryotic models.

To tackle more complex systems the classic FBA formulation optimizing for a single objective is often insufficient. One approach to handle those multi-objective systems is to use a concept called pareto optimality, where a solution is searched for which any change would lead to the worsening of at least one objective (Oh et al., 2009). The concept has been applied to optimize a human mitochondrial model taking into consideration three objective functions (Vo et al., 2004). There are different ways to combine objectives, e.g. by weighted sums of objective functions, or by successive optimization and fixation of the objective value of one target flux. However, even when multiple objectives are employed, the final solution is not necessarily unique and the number of alternate optima might become so high that is infeasible to calculate them. To analyze those solutions it is often useful to perform Flux Variability Analysis (FVA) (Mahadevan

and Schilling, 2003). In contrast to FBA, FVA determines the allowed ranges of fluxes under the optimal conditions. This can help to indicate flexible reactions, with large ranges of possible flux values, and reactions under tight control, which show only a very small range of possible fluxes.

3 COMBINING MULTIPLE CONSTRAINT BASED MODELS

While originally flux balance analysis was used on small models, partially due to a lack of knowledge about metabolic pathways and also due to a lack of computational resources, ever more complex models have been constructed over the past years. These models can include multiple tissues within an organism, or even multiple distinct organisms which interact metabolically. We will discuss the creation of these models in the following section and provide an overview of the development of these models in Figure 2.

3.1 Constraint Based Models spanning multiple organisms

While models for isolated microbes are now quite advanced, investigations of interactions between organisms are rare. However, there are multiple examples of interactions, either as symbiotic or mutualistic relations (e.g. plant-mycorrhiza interactions) or as parasites (like human-pathogens). To address these types of interactions and investigate their metabolic effects it is necessary to combine multiple organism networks. The first attempt to perform such a combination of metabolic networks of different species was published in 2007 by Stolyar et al. They created a combination of metabolic networks consisting of two symbiotic microbial species and investigated the metabolic exchanges occurring between them. This initial idea has been extended into frameworks for optimization of microbial communities (Zomorodi and Maranas, 2012; Zomorodi et al., 2014). A more detailed review was published by Mahadevan and Henson in 2012.

The concept of combining different organisms in larger models has also been used in simulating host-pathogen interactions. Bordbar et al. (2010) developed a model of the infection of alveolar macrophages by *Mycobacterium tuberculosis* and successfully simulated maximal ATP and nitric oxide (NO) production rates. For the model described by Bordbar et al. (2010), one of the most important constraints was the level of oxygen available to the pathogen, as it is (even in simulations) unable to grow without a minimal amount of oxygen. Furthermore the original biomass function was adjusted to the conditions in the infected macrophage based on expression data. This was achieved in a step-wise manner. First, random sampling was applied to determine the solution space of each individual metabolic component. Second, linear regression was used by iteratively adding and removing metabolic components to and from the biomass function. This process was performed until a new biomass function could better fit the gene expression data. This led to an altered biomass definition with different metabolic components which could reflect the infectious state of the pathogen. The resulting composition is closer to that observed in the infectious state instead of the phase of maximal growth *in vitro*. The combined model with tailored objective functions was interrogated with respect to changes in ATP, NO and NADH production fluxes in the macrophage, and in the biomass function flux for both host and pathogen. In addition, *in silico* gene essentiality studies were performed on the combined model and compared to experimental data, showing a better agreement than the same studies on the disconnected pathogen model. Finally, flux changes in both the macrophage and the pathogen were investigated by mapping gene expression data using the GIMME algorithm (Becker and Palsson, 2008).

Recently, Heinken et al. (2013) focused on a model of the interactions of the microbium *B. thetaiotaomicron* and the gut cells of a mouse model and could simulate their growth on five different dietary regimes. The coupling between these models was achieved by creating a new compartment (the intestinal lumen)

where the metabolites could be exchanged between the models. This multi-scale model was able to capture the symbiotic growth between mouse and microbe and identify the metabolic crosstalk between the two organisms.

While there have been examples of organism linkage, the main issue present in these attempts is the determination and selection of the links between organisms. Selecting alternative linking compounds could lead to vastly different results and while there are active transporters identified for some compounds, many transport systems allow the exchange of multiple compounds, which makes it difficult to pin down the right selection.

3.2 Reconstruction and analysis of multi-tissue models of higher organisms

Coupling models of different tissues within the same organism, or of specific tissues with pathogens is conceptually similar to coupling multiple organisms. Both situations commonly define interactions by allowing the different models to secrete and consume metabolites provided by the external medium or the other model. The main challenges are again to find which compounds are exchanged and determine the extent to which this exchange occurs. If the aim is to model microbial communities, there is commonly the assumption that the aim of all community organisms is to grow. Thus any interaction between organisms that allows a higher growth rate for both organisms is commonly beneficial for both organisms. In a multicellular organism, with multiple distinct tissues, the aim of each tissue is commonly distinct from growth (with the prominent exception of cancer). Thus it is important to determine the objectives or required activities of each tissue in a multi-tissue model. Those objectives can include e.g. ammonia detoxification in the liver or energy production in the brain. However it is also possible to define certain functionalities that have to be provided by each tissue and assume that the general objective is to perform these tasks most efficiently (e.g. with a minimum amount of wasted energy, or a minimal amount of enzymes required). Thus, there are multiple challenges which have to be addressed when trying to model multicellular organisms using multi-tissue models.

One of the earliest multi-tissue models was a two-tissue model of *Arabidopsis thaliana* by de Oliveira Dal'Molin et al. (2010), describing the interactions between mesophyll and bundle-sheath cells. Other models are often considering the different phases of day and night by creating a dual representation of the model (Cheung et al., 2014). As these models are the result of coupling two different models (one for the day and one for the night), they are conceptually similar to multi-scale models. In these models, the interactions between the day and the night model are implemented as reactions which represent the storage of compounds during either the day or the night. The concept was extended to a true multi-tissue model by Gomes De Oliveira Dal'molin et al. (2015) who created this dual representation for roots, stem and leaves. By investigating the exchanges the authors were able to determine the storage compounds transferred between day and night (Cheung et al., 2014) and elucidate the effect of translocation costs between tissues on the localization of biosynthetic activities (Gomes De Oliveira Dal'molin et al., 2015).

The recent advances in modelling multi-tissue models have also been applied to human tissues, starting with a multi-tissue brain model by Lewis et al. (2010). They created three models for different neuron types in the human brain: glutaminergic, γ -aminobutyrate (GABA)ergic, and cholinergic. The models contained the following submodels: a neuron with a neuronal mitochondria, an astrocyte with an astrocytic mitochondria, an endothelium/blood compartment and a interstitial space. The brain model used was based on Recon1 (Duarte et al., 2007), using only reactions indicated to be localized in the brain by the Human Protein Reference Database (Mishra et al., 2006) or by HINV (Fujii et al., 2004) and additional reaction evidence based on literature research. The reconstructed models could then be used to investigate

downregulated pathways in Alzheimer patients and to obtain a mechanistic overview of the effect of this downregulation. The authors further identified potential routes of acetyl-choline precursor synthesis in the mitochondria. The identification was performed by iteratively removing reactions from Recon1 (Duarte et al., 2007) while retaining acetyl-CoA transport from mitochondria to the cytosol. They could determine three main groups of potential acetyl-CoA biosynthesis using singular value decomposition on a set of 21000 unique minimal reaction sets obtained by this process. The resulting set included the major two pathways in the generated cholinergic model and allowed the qualitative reproduction of multiple regulatory effects on this specific neuron type.

Subsequently, Bordbar et al. (2011) established a multi-tissue model of adipocyte, hepatocyte and myocyte. In the model by Bordbar et al. (2011), each cell model was reconstructed from Recon1 (Duarte et al., 2007) using the SimPheny toolbox to generate initial draft models followed by manual curation. The integration of these cell models was performed in two steps. The first step was renaming all reactions and metabolites according to the compartments they were localized in. In addition, a new blood compartment representing three different fluids, the interstitial fluid, the urine and the blood, was introduced. All cellular exchange reactions were linked to this novel compartment. The second step of integration consisted of adding a bicarbonate buffer reaction in order to achieve proton balancing in the blood compartment. Biomass functions for each cell type were scaled to the average tissue mass and metabolite degradation in the blood compartment was assumed to be insignificant (Bordbar et al., 2011). The analysis of this multi-scale model was performed using FBA and FVA. When there was more than one objective function, the FBA optimization was performed using a pareto optimality approach. The authors were able to analyze specific interactions (like the alanine cycle and Cori cycles between liver and muscle), and could show that those interactions are indeed necessary.

A model similar to the model constructed by Bordbar et al. has been described recently (Kumar et al., 2014). While these two models focus on the same tissues, they differ by their method of reconstruction, integration and analysis. Furthermore, Kumar et al. aimed at obtaining a model for diabetic mice instead of human. However, due to the lack of a genome scale reconstruction of mice at the time, Recon1 was used as a basis for their model. The tissue-specific reconstruction was based on the model building algorithm (MBA) (Jerby et al., 2010). However instead of reconstructing three distinct tissue models, the overall model was reconstructed simultaneously. To achieve this, Recon1 (Duarte et al., 2007), was triplicated to represent all three tissues, and these “tissues” were connected. Subsequently, a high, medium and low confidence set was defined for each part of the triplicated Recon. This extended model was then reduced to its consistent subnetwork (i.e. to a network in which all reactions were able to carry a flux). The resulting model was finally subjected to MBA using the defined confidence sets for the reactions in each “tissue”. While containing the same tissues as the model by Bordbar et al., it consisted of about four times more reactions. The authors then showed that they could transfer the model predictions on mouse phenotypes to diabetic humans. Validation was performed by comparing model predicted phenotypes to known mouse phenotypes based on OMIM (McKusick, 2007). Expression data was incorporated by adapting the flux bounds according to the expression of the gene. The analysis of the physiological and diseased model was performed with FVA using multiple steps:

1. Reactions affected by gene regulation had their flux bounds set to -100 and 100, or 0 and 100, depending on reaction reversibility, while unaffected reaction flux bounds were equal to -1000 and 1000, or 0 and 1000, again depending on reaction reversibility;
2. An FVA of the physiological condition was calculated;

3. The bounds for the disease states were adapted by doubling and halving the bounds for upregulated and downregulated reactions respectively;
4. The diabetic condition was analyzed using FVA;
5. The FVA results were used to calculate the difference between the two conditions.

The resulting exchange fluxes were compared to phenotypes in the OMIM database and showed a better agreement than the analysis on an unmodified Recon1. Finally the model was analyzed for upregulated and downregulated metabolic pathways based on a contextualized subnetwork. This analysis was again performed using FVA and comparing the obtained fluxes with a randomized regulation, resulting in the indication of several tissue specific upregulated and downregulated pathways, which were in accordance with experimental observations.

The approaches presented show that it has become feasible to simulate multi-tissue systems and that the neglect of metabolic interactions between tissues can hide the true actions of metabolism. However, they also indicate that it is important to define non trivial objectives for each tissue, or to analyze the qualitative changes in flux potentials in order to obtain useful predictions. The former makes it necessary to define the functions and their extend, which is non trivial. The latter approach might be a possibility to hint at potential interactions/exchanges between tissues, as the automatic reconstruction will include only those transporters strictly necessary for the functions of the network. Thus, there is much potential in multi-tissue reconstruction that is likely to become addressed in the near future.

4 MULTI-SCALE MODELLING OF HIGHER ORGANISMS

Up to this point, we have focused mainly on purely stoichiometric models. These models have the advantage of requiring minimal amounts of parameters needing only information about the structure of the underlying network, while still providing the full metabolic capacity. However, they are commonly restricted to simulate non-dynamic processes or compare the situation in two distinct conditions. We will now show recent advances in using constraint based models as parts of larger, multi-scale, frameworks which simulate other cellular processes using dynamic techniques.

The human liver is often a target of metabolic investigations, as it is one of the metabolically most active tissues performing many different tasks. Thus, it comes as no surprise that the metabolic reconstruction of hepatocytes, HepatoNet (Gille et al., 2010), is one of the most frequently employed metabolic networks when attempts at integrating additional processes in human are performed. One of these attempts aimed at integrating regulatory and signaling processes with metabolic networks (Fisher et al., 2013). They used a petri net approach to model the regulatory and signaling network in a dynamic fashion. The petri net was extended by constraint nodes which were linked to the flux bounds in the metabolic model. In addition, objective nodes in the petri net were included, which were used to set the objective in the CBM, and to react to changes in the respective objective values. The CBM was simulated using FBA. The approach employs iterations of the following steps:

1. Calculate the constraint and the objective nodes;
2. Update the metabolic model according to (1);
3. Optimize the metabolic model using FBA;
4. Update the Petri net objective nodes according to the new objective.

Using the generated framework, they investigated bile acid homeostasis, and could generate time courses in agreement with those determined experimentally. In addition, they analyzed genotype-phenotype relationships, from which they could identify several genes which are likely critical in keeping the bile acid homeostasis. The investigation showed a promising way to integrate the effects of regulatory networks on metabolism in a dynamic way. It is however restricted to parts of metabolic networks, where the controlling regulatory network is well understood.

The pharmaceutical industry routinely uses kinetic whole-body models to determine the distribution and effects of drugs on a whole-body scale. These physiologically based pharmacokinetic models (PBPK) aim to describe the absorption, distribution, metabolism and excretion of compounds. They commonly include kinetics for blood flow, and exchange rates for organs or tissues relevant for the biological question. Additionally, they are based on prior knowledge of anatomy, physiology as well as of compounds properties. Furthermore, they can give rise to an understanding of the macroscale effects of a drug treatment. Thus combining these large scale models with the cellular models above provides the potential to understand both whole-body behavior and individual cell responses and can provide information on the potential side effects of a treatment. This concept has been applied to human (Krauss et al., 2012) and arabidopsis (Grafahrend-Belau et al., 2013) and we will discuss both studies in more detail. In the system developed by Krauss et al. (2012) a liver CBM and a whole-body physiologically based pharmacokinetics (PBPK) model were coupled. In order to study, how these compounds affect cellular metabolism, the CBM models were integrated, replacing parts of the tissue specific functions. The model chosen to serve as a CBM was Hepatonet, which was simulated by a static dFBA approach. Since this static method uses linear programming, it is not computationally expensive. The models were coupled using two different approaches. Indirect coupling was used to simulate processes that have no direct effect on the PBPK model e.g. administration of a compound that affects metabolic enzyme activity. For these processes information from the PBPK model was used to constrain the CBM model. For processes that have a direct influence on the dynamics of the PBPK model, direct linkage was used. To achieve this coupling, a feedback update loop was integrated in order to update the fluxes for the next time step in the PBPK model according to those resulting from the simulation of the CBM model. This feedback loop consists of four steps:

1. Clearance and production rates are calculated at the whole-body level using the PBPK model;
2. The upper bounds of coupled reactions in the CBM are adjusted to the calculated values;
3. The CBM is simulated and the coupled rates are determined;
4. The directly coupled PBPK rates are set to the rates obtained in 3 to determine the metabolite levels for the next iteration.

Indirect coupling was used to simulate both the effect of allopurinol treatment on hyperuricemia and a paracetamol overdose causing liver toxification. Direct coupling of the models was employed to simulate impaired ammonia detoxification, where the activity of a specific enzyme in the CBM was impaired. The modelling approach can be applied to study different processes, inferring the effect of cellular alterations on the whole-body level. However, it does not allow a direct assessment of the cellular flux distribution, as the objective used in the CBM model is to maximize the coupled fluxes while minimizing the remaining fluxes, which can lead to odd internal flux distributions.

In an approach similar to the one used by Krauss et al. (2012), Grafahrend-Belau et al. (2013) were simulating a multi-scale model of barley. The multi-tissue plant model was reconstructed from literature-based biochemical reactions and databases. It included leaf, stem, seed and root models which were coupled by adding a phloem compartment that was used for the exchange of carbon and nitrogen sources

between the different plant compartments. The multi-tissue model was coupled to the dynamic whole-plant ProNet-CN model (Mueller et al., 2012). In order to analyze the multi-tissue model, dynamic FBA was used. The optimization of the multi-tissue model was performed in two steps: first the carbon uptake was minimized and the result was used to constrain the second optimization step which consisted of a minimization of all fluxes using the minimal carbon uptake. The coupling approaches were similar to those used by Krauss et al. (2012). The resulting modelling framework was then used to study the seed developmental phase of barley plants.

As could be seen, there are many approaches, to add additional levels of cellular processes to constraint based models, or use constraint based models as parts of complex whole organism simulations. And while recently a whole-cell model of *Mycoplasma genitalium* has been published by Karr et al. (2012), this model requires over 1900 parameters to be fitted or determined. While this has been achieved on a prokaryotic, rather “simple” system, the same would require many more parameters for eukaryotic cells and likely at least an order of magnitude more for multicellular organisms, exhibiting multiple types of tissues. Thus approaches with simplified processes, like CBM models, will serve as a good basis for quite some time to come, but as we showed, it is always possible to combine them with other types of models, to improve our capabilities in simulating the effects of metabolic interventions on larger systems.

5 CONCLUSION

Even with our constantly growing understanding of the kinetic properties of enzymatic processes, the multitude of different conditions and slight variations makes it unlikely that a multi-cellular kinetic model comprising all known metabolic pathways will be established in the near future. Thus we need modelling techniques that can cope with the size of a complex multi-tissue network. We have presented several examples of purely constraint based modelling approaches that aim at providing these techniques. These methods however, lack the capability to include effects like distribution limits, or the dynamic response to the changes in metabolite levels generated by the CBM models. In the recent years, methods to combine CBM models with other types of models have been proposed to address these issues. These techniques could allow us to achieve a better understanding of the interactions between dynamic processes on the whole-body scale, and cellular metabolic processes. They provide ways to integrate additional levels of control influencing the behavior of the metabolic network, and could provide means to decipher the intricate interactions of the different regulatory processes in a higher organism. Thus, further development and improvement of these techniques is likely to be an important step in our voyage to a complete understanding of metabolism.

DISCLOSURE/CONFLICT-OF-INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

TP and TS designed the study. PRMC, TP and TS wrote, edited, and refined the manuscript.

Funding: PRMC, TS and TP are supported by the University of Luxembourg.

REFERENCES

- Andreozzi, S., Miskovic, L., and Hatzimanikatis, V. (2015). ischrunk-in silico approach to characterization and reduction of uncertainty in the kinetic models of genome-scale metabolic networks. *Metab. Eng.* doi:10.1016/j.ymben.2015.10.002
- Beard, D. A., Babson, E., Curtis, E., and Qian, H. (2004). Thermodynamic constraints for biochemical networks. *J. Theor. Biol.* 228, 327–333. doi:10.1016/j.jtbi.2004.01.008
- Becker, S. A. and Palsson, B. Ø. (2008). Context-specific metabolic networks are consistent with experiments. *PLoS Comput. Biol.* 4, e1000082. doi:10.1371/journal.pcbi.1000082
- Berestovsky, N., Zhou, W., Nagrath, D., and Nakhleh, L. (2013). Modeling integrated cellular machinery using hybrid petri-boolean networks. *PLoS Comput. Biol.* 9, e1003306. doi:10.1371/journal.pcbi.1003306
- Bordbar, A., Feist, A. M., Usaite-Black, R., Woodcock, J., Palsson, B. Ø., and Famili, I. (2011). A multi-tissue type genome-scale metabolic network for analysis of whole-body systems physiology. *BMC Systems Biology* 5, 180. doi:10.1186/1752-0509-5-180
- Bordbar, A., Lewis, N. E., Schellenberger, J., Palsson, B. Ø., and Jamshidi, N. (2010). Insight into human alveolar macrophage and *M. tuberculosis* interactions via metabolic reconstructions. *Mol. Syst. Biol.* 6, 422. doi:10.1038/msb.2010.68
- Bordbar, A., Monk, J. M., King, Z. A., and Palsson, B. Ø. (2014). Constraint-based models predict metabolic and associated cellular functions. *Nat. Rev. Genet.* 15, 107–120. doi:10.1038/nrg3643
- Burgard, A. P., Pharkya, P., and Maranas, C. D. (2003). OptKnock: A Bilevel Programming Framework for Identifying Gene Knockout Strategies for Microbial Strain Optimization. *Biotechnol. Bioeng.* 84, 647–657. doi:10.1002/bit.10803
- Cheung, C. Y. M., Poolman, M. G., Fell, D. A., Ratcliffe, R. G., and Sweetlove, L. J. (2014). A Diel Flux Balance Model Captures Interactions between Light and Dark Metabolism during Day-Night Cycles in C3 and Crassulacean Acid Metabolism Leaves. *Plant Physiol.* 165, 917–929. doi:10.1104/pp.113.234468
- Colijn, C., Brandes, A., Zucker, J., Lun, D. S., Weiner, B., Farhat, M. R., et al. (2009). Interpreting Expression Data with Metabolic Flux Models: Predicting *Mycobacterium tuberculosis* Mycolic Acid Production. *PLoS Comput. Biol.* 5, e1000489. doi:10.1371/journal.pcbi.1000489
- Covert, M. W. and Palsson, B. Ø. (2002). Transcriptional regulation in constraints-based metabolic models of *Escherichia coli*. *J. Biol. Chem.* 277, 28058–28064. doi:10.1074/jbc.M201691200
- Covert, M. W., Schilling, C. H., and Palsson, B. Ø. (2001). Regulation of gene expression in flux balance models of metabolism. *J. Theor. Biol.* 213, 73–88. doi:10.1006/jtbi.2001.2405
- Covert, M. W., Xiao, N., Chen, T. J., and Karr, J. R. (2008). Integrating metabolic, transcriptional regulatory and signal transduction models in *Escherichia coli*. *Bioinformatics* 24, 2044–2050. doi:10.1093/bioinformatics/btn352
- Crampin, E. J., Smith, N. P., and Hunter, P. J. (2004). Multi-scale modelling and the IUPS physiome project. *J. Mol. Histol.* 35, 707–714. doi:10.1007/s10735-004-2676-6
- de Oliveira Dal’Molin, C. G., Quek, L.-E., Palfreyman, R. W., Brumbley, S. M., and Nielsen, L. K. (2010). AraGEM, a genome-scale reconstruction of the primary metabolic network in Arabidopsis. *Plant Physiol.* 152, 579–589. doi:10.1104/pp.109.148817
- Duarte, N. C., Becker, S. A., Jamshidi, N., Thiele, I., Mo, M. L., Vo, T. D., et al. (2007). Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proc. Natl. Acad. Sci. U. S. A.* 104, 1777–1782. doi:10.1073/pnas.0610772104

- Edwards, J. S. and Palsson, B. Ø. (1999). Systems properties of the *Haemophilus influenzae Rd* metabolic genotype. *J. Biol. Chem.* 274, 17410–17416. doi:10.1074/jbc.274.25.17410
- Feist, A. M. and Palsson, B. Ø. (2010). The biomass objective function. *Curr. Opin. Microbiol.* 13, 344–349. doi:10.1016/j.mib.2010.03.003
- Fisher, C. P., Plant, N. J., Moore, J. B., and Kierzek, A. M. (2013). QSSPN: dynamic simulation of molecular interaction networks describing gene regulation, signalling and whole-cell metabolism in human cells. *Bioinformatics* 29, 3181–3190. doi:10.1093/bioinformatics/btt552
- Förster, J., Famili, I., Palsson, B. Ø., and Nielsen, J. (2003). Large-scale evaluation of in silico gene deletions in *Saccharomyces cerevisiae*. *OMICS* 7, 193–202. doi:doi:10.1089/153623103322246584
- Fujii, Y., Imanishi, T., and Gojobori, T. (2004). H-Invitational Database: integrated database of human genes. *Tanpakushitsu Kakusan Koso.* 49, 1937–1943
- Gille, C., Bölling, C., Hoppe, A., Bulik, S., Hoffmann, S., Hübner, K., et al. (2010). HepatoNet1: a comprehensive metabolic reconstruction of the human hepatocyte for the analysis of liver physiology. *Mol. Syst. Biol.* 6, 411. doi:10.1038/msb.2010.62
- Gomes De Oliveira Dal'molin, C., Quek, L.-E., Saa, P. A., and Nielsen, L. K. (2015). A multi-tissue genome-scale metabolic modelling framework for the analysis of whole plant systems. *Front. Plant. Sci.* 6. doi:10.3389/fpls.2015.00004
- Grafahrend-Belau, E., Junker, A., Eschenröder, A., Müller, J., Schreiber, F., and Junker, B. H. (2013). Multiscale metabolic modeling: dynamic flux balance analysis on a whole-plant scale. *Plant Physiol.* 163, 637–647. doi:10.1104/pp.113.224006
- Heinken, A., Sahoo, S., Fleming, R. M. T., and Thiele, I. (2013). Systems-level characterization of a host-microbe metabolic symbiosis in the mammalian gut. *Gut Microbes* 4, 28–40. doi:10.4161/gmic.22370
- Jerby, L., Shlomi, T., and Ruppin, E. (2010). Computational reconstruction of tissue-specific metabolic models: application to human liver metabolism. *Mol. Syst. Biol.* 6, –. doi:10.1038/msb.2010.56
- Jolivet, R., Coggan, J. S., Allaman, I., and Magistretti, P. J. (2015). Multi-timescale modeling of activity-dependent metabolic coupling in the neuron-glia-vasculature ensemble. *PLoS Comput. Biol.* 11, e1004036. doi:10.1371/journal.pcbi.1004036
- Karr, J. R., Sanghvi, J. C., Macklin, D. N., Gutschow, M. V., Jacobs, J. M., Bolival, B., et al. (2012). A whole-cell computational model predicts phenotype from genotype. *Cell* 150, 389–401. doi:10.1016/j.cell.2012.05.044
- Kauffman, K. J., Prakash, P., and Edwards, J. S. (2003). Advances in flux balance analysis. *Curr. Opin. Biotechnol.* 14, 491–496. doi:10.1016/j.copbio.2003.08.001
- Krauss, M., Schaller, S., Borchers, S., Findeisen, R., Lippert, J., and Kuepfer, L. (2012). Integrating cellular metabolism into a multiscale whole-body model. *PLoS Comput. Biol.* 8, e1002750. doi:10.1371/journal.pcbi.1002750
- Kumar, A., Harrelson, T., Lewis, N. E., Gallagher, E. J., LeRoith, D., Shiloach, J., et al. (2014). Multi-tissue computational modeling analyzes pathophysiology of type 2 diabetes in MKR mice. *PLoS One* 9, e102319. doi:10.1371/journal.pone.0102319
- Lee, D., Smallbone, K., Dunn, W. B., Murabito, E., Winder, C. L., Kell, D. B., et al. (2012). Improving metabolic flux predictions using absolute gene expression data. *BMC Systems Biology* 6, 73. doi:10.1186/1752-0509-6-73
- Lee, J. M., Gianchandani, E. P., Eddy, J. A., and Papin, J. A. (2008). Dynamic Analysis of Integrated Signaling, Metabolic, and Regulatory Networks. *PLoS Comput. Biol.* 4, e1000086. doi:10.1371/journal.pcbi.1000086

- Lerman, J. A., Hyduke, D. R., Latif, H., Portnoy, V. A., Lewis, N. E., Orth, J. D., et al. (2012). In silico method for modelling metabolism and gene product expression at genome scale. *Nat. Commun.* 3, 929. doi:10.1038/ncomms1928
- Lewis, N. E., Nagarajan, H., and Palsson, B. Ø. (2012). Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods. *Nat. Rev. Microbiol.* 10, 291–305. doi:10.1038/nrmicro2737
- Lewis, N. E., Schramm, G., Bordbar, A., Schellenberger, J., Andersen, M. P., Cheng, J. K., et al. (2010). Large-scale in silico modeling of metabolic interactions between cell types in the human brain. *Nat. Biotechnol.* 28, 1279–1285. doi:10.1038/nbt.1711
- Llaneras, F. and Picó, J. (2008). Stoichiometric modelling of cell metabolism. *J. Biosci. Bioeng.* 105, 1–11. doi:10.1263/jbb.105.1
- Ma, H., Sorokin, A., Mazein, A., Selkov, A., Selkov, E., Demin, O., et al. (2007). The Edinburgh human metabolic network reconstruction and its functional analysis. *Mol. Syst. Biol.* 3, 135. doi:10.1038/msb4100177
- Machado, D. and Herrgård, M. (2014). Systematic evaluation of methods for integration of transcriptomic data into constraint-based models of metabolism. *PLoS Comput. Biol.* 10, e1003580. doi:10.1371/journal.pcbi.1003580
- Mahadevan, R., Edwards, J. S., and Doyle, F. J., 3rd (2002). Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophys. J.* 83, 1331–1340. doi:10.1016/S0006-3495(02)73903-9
- Mahadevan, R. and Henson, M. A. (2012). Genome-based Modeling and Design of Metabolic Interactions in Microbial Communities. *Comput. Struct. Biotechnol. J.* 3, e201210008. doi:10.5936/csbi.201210008
- Mahadevan, R. and Schilling, C. (2003). The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab. Eng.* 5, 264–276. doi:10.1016/j.ymben.2003.09.002
- Mardinoglu, A., Agren, R., Kampf, C., Asplund, A., Uhlen, M., and Nielsen, J. (2014). Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease. *Nat. Commun.* 5, 3083. doi:10.1038/ncomms4083
- McKusick, V. A. (2007). Mendelian Inheritance in Man and its online version, OMIM. *Am. J. Hum. Genet.* 80, 588–604. doi:10.1086/514346
- Mishra, G. R., Suresh, M., Kumaran, K., Kannabiran, N., Suresh, S., Bala, P., et al. (2006). Human protein reference database–2006 update. *Nucleic Acids Res.* 34, D411–D414. doi:10.1093/nar/gkj141
- Mueller, J., Eschenroeder, A., Christen, O., Junker, B. H., and Schreiber, F. (2012). ProNet-CN model: A dynamic and multi-scale process network combining photosynthesis, primary carbon metabolism and effects of leaf nitrogen status. In *Plant Growth Modeling, Simulation, Visualization and Applications (PMA), 2012 IEEE Fourth International Symposium on.* 289–296. doi:10.1109/PMA.2012.6524848
- Oh, Y.-G., Lee, D.-Y., Lee, S. Y., and Park, S. (2009). Multiobjective flux balancing using the NISE method for metabolic network analysis. *Biotechnol. Prog.* 25, 999–1008. doi:10.1002/btpr.193
- Orth, J. D. and Palsson, B. Ø. (2010). Systematizing the generation of missing metabolic knowledge. *Biotechnol. Bioeng.* 107, 403–412. doi:10.1002/bit.22844
- Pfau, T., Pacheco, M. P., and Sauter, T. (2015). Towards improved genome-scale metabolic network reconstructions: unification, transcript specificity and beyond. *Brief. Bioinform.* doi:10.1093/bib/bbv100
- Pozo, C., Guillén-Gosálbez, G., Sorribas, A., and Jiménez, L. (2012). Identifying the preferred subset of enzymatic profiles in nonlinear kinetic metabolic models via multiobjective global optimization and pareto filters. *PLoS One* 7, e43487. doi:10.1371/journal.pone.0043487
- Raman, K. and Chandra, N. (2009). Flux balance analysis of biological systems: applications and challenges. *Brief. Bioinform.* 10, 435–449. doi:10.1093/bib/bbp011

- Resendis-Antonio, O., González-Torres, C., Jaime-Muñoz, G., Hernandez-Patiño, C. E., and Salgado-Muñoz, C. F. (2014). Modeling metabolism: A window toward a comprehensive interpretation of networks in cancer. *Semin. Cancer Biol.* doi:10.1016/j.semcancer.2014.04.003
- Ricken, T., Werner, D., Holzhütter, H., König, M., Dahmen, U., and Dirsch, O. (2015). Modeling function-perfusion behavior in liver lobules including tissue, blood, glucose, lactate and glycogen by use of a coupled two-scale PDE-ODE approach. *Biomech. Model Mechanobiol.* 14, 515–536. doi:10.1007/s10237-014-0619-z
- Robaina Estévez, S. and Nikoloski, Z. (2014). Generalized framework for context-specific metabolic model extraction methods. *Front. Plant. Sci.* 5, 491. doi:10.3389/fpls.2014.00491
- Rocha, I., Maia, P., Evangelista, P., Vilaca, P., Soares, S., Pinto, J., et al. (2010). OptFlux: an open-source software platform for in silico metabolic engineering. *BMC Systems Biology* 4, 45. doi:10.1186/1752-0509-4-45
- Ryll, A., Bucher, J., Bonin, A., Bongard, S., Goncalves, E., Saez-Rodriguez, J., et al. (2014). A model integration approach linking signalling and gene-regulatory logic with kinetic metabolic models. *Biosystems* 124, 26 – 38. doi:http://dx.doi.org/10.1016/j.biosystems.2014.07.002
- Ryu, J. Y., Kim, H. U., and Lee, S. Y. (2015). Reconstruction of genome-scale human metabolic models using omics data. *Integr. Biol.* doi:10.1039/c5ib00002e
- Savinell, J. M. and Palsson, B. Ø. (1992). Optimal selection of metabolic fluxes for in vivo measurement. I. Development of mathematical methods. *J Theor Biol* 155, 201–214
- Schuetz, R., Kuepfer, L., and Sauer, U. (2007). Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. *Mol. Syst. Biol.* 3, 119. doi:10.1038/msb4100162
- Segrè, D., Vitkup, D., and Church, G. M. (2002). Analysis of optimality in natural and perturbed metabolic networks. *Proc. Natl. Acad. Sci. U. S. A.* 99, 15112–15117. doi:10.1073/pnas.232349399
- Sheikh, K., Förster, J., and Nielsen, L. K. (2005). Modeling hybridoma cell metabolism using a generic genome-scale metabolic model of *Mus musculus*. *Biotechnol. Prog.* 21, 112–121. doi:10.1021/bp0498138
- Shlomi, T., Berkman, O., and Ruppin, E. (2005). Regulatory on/off minimization of metabolic flux changes after genetic perturbations. *Proc. Natl. Acad. Sci. U. S. A.* 102, 7695–7700. doi:10.1073/pnas.0406346102
- Stolyar, S., Van Dien, S., Hillesland, K. L., Pinel, N., Lie, T. J., Leigh, J. A., et al. (2007). Metabolic modeling of a mutualistic microbial community. *Mol. Syst. Biol.* 3, 92. doi:10.1038/msb4100131
- Thiele, I., Fleming, R. M. T., Que, R., Bordbar, A., Diep, D., and Palsson, B. Ø. (2012). Multiscale modeling of metabolism and macromolecular synthesis in *E. coli* and its application to the evolution of codon usage. *PLoS One* 7, e45635. doi:10.1371/journal.pone.0045635
- Thiele, I., Swainston, N., Fleming, R. M. T., Hoppe, A., Sahoo, S., Aurich, M. K., et al. (2013). A community-driven global reconstruction of human metabolism. *Nat. Biotechnol.* 31, 419–425. doi:10.1038/nbt.2488
- Varma, A. and Palsson, B. Ø. (1994). Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Appl. Environ. Microbiol.* 60, 3724–3731
- Viceconti, M., Clapworthy, G., and Jan, S. V. S. (2008). The virtual physiological human — a european initiative for *in silico* human modelling. *J. Physiol. Sci.* 58, 441–446. doi:10.2170/physiolsci.RP009908
- Vlassis, N., Pires Pacheco, M., and Sauter, T. (2014). Fast reconstruction of compact context-specific metabolic network models. *PLoS Comput. Biol.* 10, e1003424. doi:10.1371/journal.pcbi.1003424

- Vo, T. D., Greenberg, H. J., and Palsson, B. Ø. (2004). Reconstruction and functional characterization of the human mitochondrial metabolic network based on proteomic and biochemical data. *J. Biol. Chem.* 279, 39532–39540. doi:10.1074/jbc.M403782200
- von Kamp, A. and Klamt, S. (2014). Enumeration of smallest intervention strategies in genome-scale metabolic networks. *PLoS Comput. Biol.* 10, e1003378. doi:10.1371/journal.pcbi.1003378
- Zeigerer, A., Bogorad, R., Sharma, K., Gilleron, J., Seifert, S., Sales, S., et al. (2015). Regulation of liver metabolism by the endosomal {GTPase} rab5. *Cell Rep.* 11, 884 – 892. doi:http://dx.doi.org/10.1016/j.celrep.2015.04.018
- Zomorodi, A. R., Islam, M. M., and Maranas, C. D. (2014). d-OptCom: Dynamic Multi-level and Multi-objective Metabolic Modeling of Microbial Communities. *ACS Synth. Biol.* 3, 247–257. doi:10.1021/sb4001307
- Zomorodi, A. R. and Maranas, C. D. (2012). OptCom: a multi-level optimization framework for the metabolic modeling and analysis of microbial communities. *PLoS Comput. Biol.* 8, e1002363. doi:10.1371/journal.pcbi.1002363

FIGURES

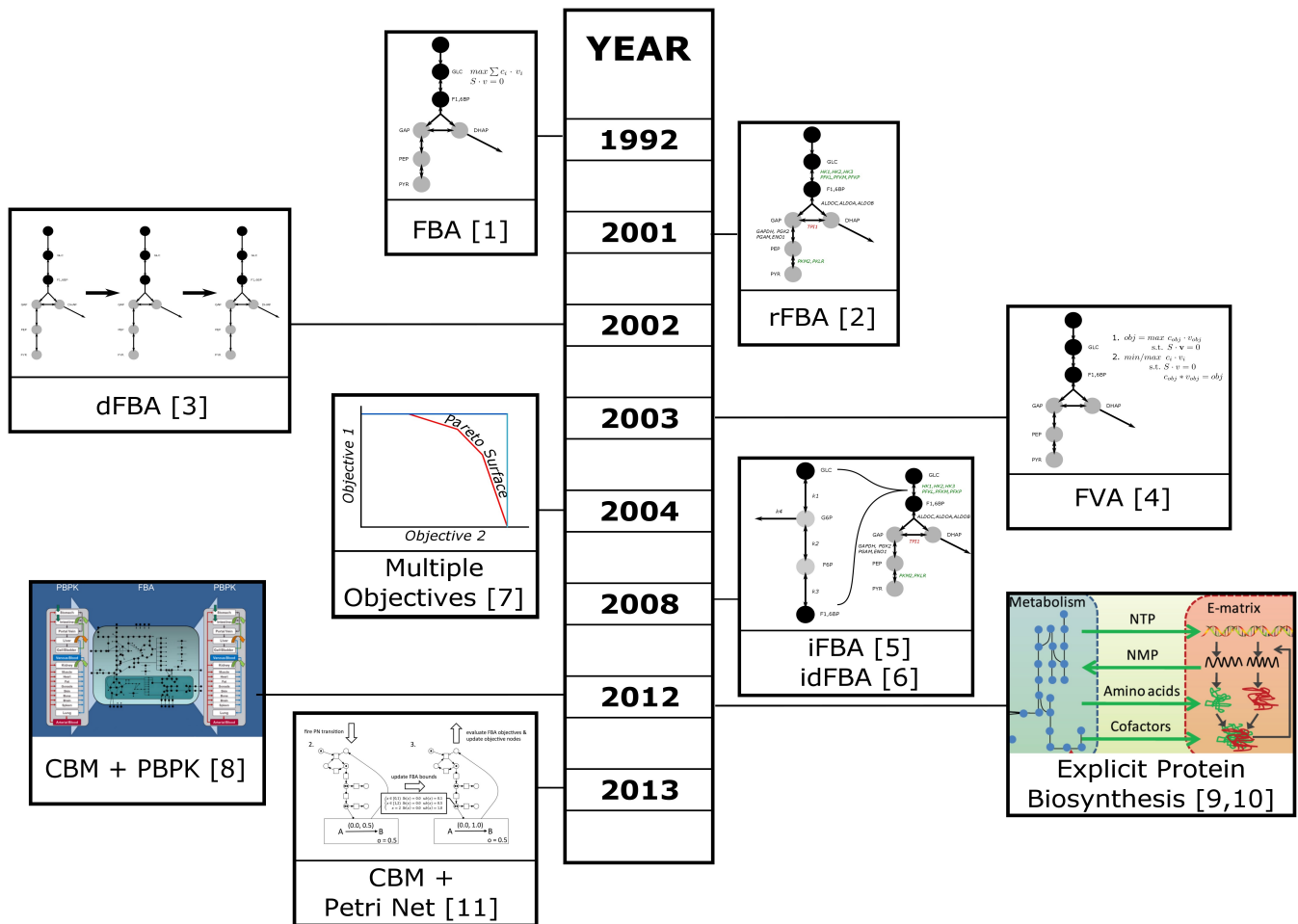


Figure 1. Timeline of the development of techniques for the integration of data and the simulation and analysis of complex systems. Please refer to the main text for details. ([1] Savinell and Palsson (1992); [2] Covert et al. (2001); [3] Mahadevan et al. (2002); [4] Mahadevan and Schilling (2003); [5] Covert et al. (2008); [6] Lee et al. (2008); [7] Vo et al. (2004); [8] Krauss et al. (2012); [9] Thiele et al. (2012); [10] Lerman et al. (2012); [11] Fisher et al. (2013)), Images for [8],[9] and [10] are derived from images taken from the respective publications which are provided under a Creative Commons attribution license (<https://creativecommons.org/licenses/by/2.0/>)

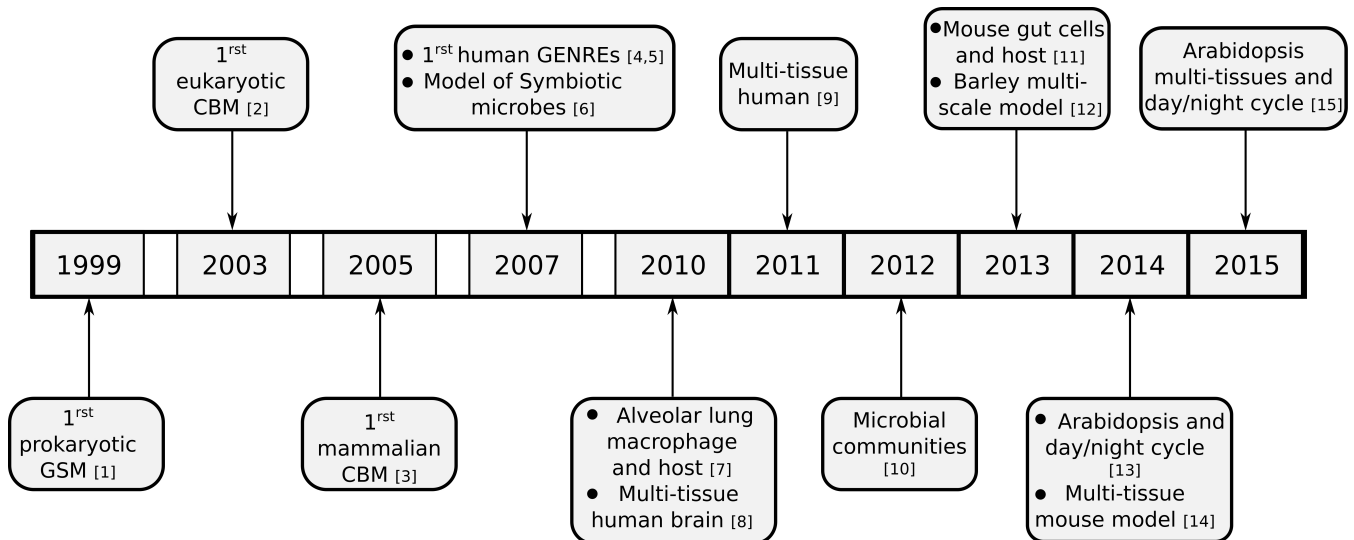


Figure 2. Timeline of development of reconstruction of metabolic models and realisation of different models spanning multiple tissues or organisms. Except for the development of the initial genome scale reconstructions from various kingdoms of life, only multi-tissue or multi-compartment models are listed. ([1] Edwards and Palsson (1999); [2] Förster et al. (2003); [3] Sheikh et al. (2005); [4] Duarte et al. (2007); [5] Ma et al. (2007); [6] Stolyar et al. (2007); [7] Bordbar et al. (2010); [8] Lewis et al. (2010); [9] Bordbar et al. (2011); [10] Zomorodi and Maranas (2012); [11] Heinken et al. (2013); [12] Grafahrend-Belau et al. (2013); [13] Cheung et al. (2014); [14] Kumar et al. (2014); [15] Gomes De Oliveira Dal'molin et al. (2015))

Table 1. Overview of different FBA-based methods.

Approach	Concept	Description	Mathematical concepts
FBA [1]	Optimize an objective function	Keep the internal system in steady state, and optimize a given objective using Linear Programming (LP)	$Min/Max \ w = \sum c_i v_i$ s.t. $S \cdot v = 0$ $lb_i \leq v_i \leq ub_i$
rFBA [2]	Optimization of a metabolic and regulatory network	Transcription regulation integrated as time dependent constraints (0,1) using Boolean rules to compare multiple time points.	FBA s.t. $v_i = 0$ if rule for reaction i inactive
dFBA-DOA [3]	Optimize multiple time-steps simultaneously	Simultaneous optimization over the entire time period	Non Linear Problem, combining end point objective and intermediate objectives
dFBA-SOA [3]	Optimize multiple time-steps iteratively	Divide the time period in intervals Use initial conditions or result from the previous time step to solve the dynamic equations Fix the metabolic model fluxes to the fluxes obtained when solving the ODEs Optimize the model at the next time step	FBA s.t. $lb_{i,t} = f_{i,lb}(\vec{v}_{t-1})$ $ub_{i,t} = f_{i,ub}(\vec{v}_{t-1})$ f includes changes to the initial conditions (i.e. available concentrations).
Thermodynamic FBA [4]	Integration of thermodynamical constraints into a metabolic network	Changing of reaction reversibility dependent on free energy changes	Adjust ub_i and lb_i according to thermodynamic considerations
iFBA [5]	Optimization of an integrated metabolic, regulatory and kinetic network	Combination of dFBA-SOA, and rFBA	
idFBA [6]	Optimization of an integrated metabolic, regulatory and signaling network.	dFBA-SOA with additional phenotypic data After each time step, update phenotype variables Constrain next time step using updated phenotype data	dFBA with additional time dependent constraints $v_{t,i} = 0 \forall i$ with $I_{t,i} = 0$ I is the incidence matrix with one column per time point
ME-Matrix [7,8]	Optimization of a coupled metabolic and cellular machinery model	Addition of explicit biosynthesis of cellular machinery necessary for metabolic reactions Coupling constraints that couple protein synthesis to flux rate	Coupling: $v_i - c_{max} \cdot v_s \leq 0$ $v_i - c_{min} \cdot v_s \geq -s, \ s \geq 0$ with $c_{min,max}$ the coupling coefficient. and the latter equation enforcing more protein synthesis for higher flux
QOSPN (Quasi-steady state Petri nets) [9]	Optimization of an integrated metabolic, regulatory and signaling network using Petri-nets	Calculate constraint and objective nodes for each time step Update the metabolic model according to the previous step Optimize the metabolic model Update the Petri net objective nodes according to the new objective	

Integration of different approaches with CBMs lead to the development of new optimization frameworks. In this table, some FBA-based approaches are described.

Abbreviations: c, optimization weight; Max, maximization; Min, minimization; lb, lower bound; s, accumulation variable; ub, upper bound; w, objective function

References: [1] Savinell and Palsson (1992); [2] Covert et al. (2001); [3] Mahadevan et al. (2002); [4] Beard et al. (2004); [5] Covert et al. (2008); [6] Lee et al. (2008); [7] Thiele et al. (2012); [8] Lerman et al. (2012); [9] Fisher et al. (2013)