A method for embedding dynamic pathway models into larger metabolic networks

Wolfram Liebermeister
Institut für Biochemie, Charité – Universitätsmedizin Berlin

Abstract

Metabolic pathways in cells are surrounded by a larger metabolic network, whose dynamics feeds back on the pathways (e.g., through consumption of pathway product) and changes their dynamics. In kinetic models, a pathway is usually bounded by external metabolites whose concentrations are assumed to be fixed and given. To avoid this simplification and to obtain more realistic, dynamic boundary conditions, pathway models can be embedded into larger models decribing the surrounding metabolic network. Such models can be constructed by elasticity sampling, which provides simple rate laws with realistic kinetic constants. The algorithm proposed here starts by mapping the compounds and reactions between the kinetic pathway models and a metabolic network that serves as a scaffold. The next step is to construct a hybrid model with a feasible metabolic state, characterised by flux distribution, metabolite concentrations, and thermodynamic forces. Finally, standardised rate laws are inserted into the surrounding network, resulting in a large kinetic model that realises this metabolic state in a thermodynamically feasible way. The algorithm guarantees a thermodynamically feasible model and allows for a systematic construction of metabolic reference states. It can be used for different purposes: to combine several kinetic models, to turn a metabolic network into a kinetic model with a predefined reference state, to embed kinetic models into such a scaffold model, or to build a kinetic pathway model directly from a list of enzymatic rate laws. Example cases illustrate how the embedding of metabolic pathways can change their dynamics. . Matlab code is provided on github.

Keywords: Metabolic network model, model combination, boundary condition, elasticity sampling

1 Introduction

Kinetic metabolic models could be used to predict the effects of nutrient supply, drugs, genetic variants, or genetic modifications on the metabolic state of cells. However, large kinetic models are difficult to construct because kinetic parameters are often unknown and hard to fit. Some metabolic pathways have been modelled kinetically, but these models cover only small regions of the biochemical networks [1]. To fill gaps between these pathways, or to embed pathways into larger dynamic models, it would be helpful to start with a network of reactions surrounding the pathways, which would then be automatically translated into a dynamic model. Why would this be useful? Can't we just model those pathways that we're actually interested in? Kinetic models need to make assumptions about the dynamics at the pathway boundaries, and they typically assume fixed and given metabolite concentrations or fluxes at the pathway boundaries. In reality, all pathways are interlinked. Since biochemical processes in cells are tightly connected, realistic pathway models would require a dynamic description of their surrounding network [2]. For example, many kinetic models of glycolysis assume fixed ATP levels. This is not a realistic assumption in itself, and it also makes the simulated pathway dynamics less realistic. In reality, a higher glycolytic flux would increase the ATP level; the surrounding network might respond to this and consume more ATP, but this adaptation will take a while: thus, after a sudden increase in glycolytic flux, the ATP level will rise, but may later return to its initial value and may even drop below this value (because of the increased

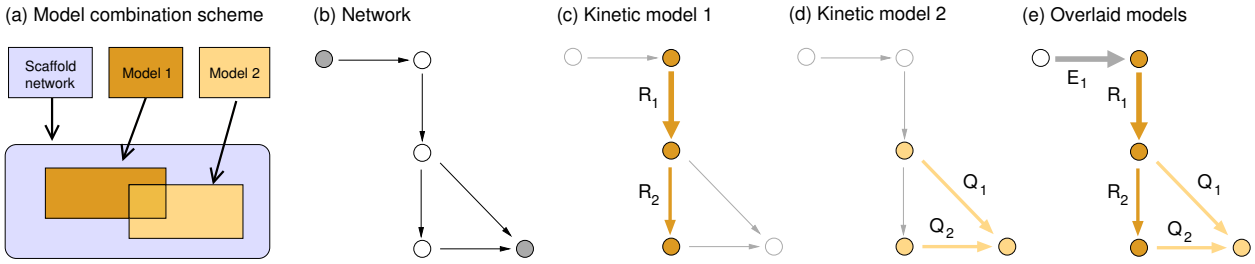

Figure 1: Embedding of kinetic pathway models. In the schematic example, two pathways (blue and yellow) are combined by embedding them into a scaffold network. (a) Example network (metabolites shown as circles, external metabolites shown in grey). (b) The kinetic pathway model R contains the reactions R_1 and R_2; a flux distribution (fluxes: blue arrows) follows from rate laws R_1 and R_2 and from the concentrations in the kinetic model. (c) A second kinetic pathway model Q covers the reactions Q_1 and Q_2. (d) Each of the two kinetic models defines a set of stationary fluxes. To embed the models, a common stationary flux distribution is chosen, matching these fluxes as closely as possible; the flux directions must be preserved. By adjusting the enzyme levels in the pathway models, the reaction rates are made to match the given stationary fluxes; in the rest of the network (grey), the fluxes are realised by simple rate laws with rate constants obtained by elasticity sampling.

consumption) after the glycolytic flux has returned to its original value. In kinetic whole-cell models, all this could be a normal dynamic effect However, an isolated glycolysis model, in which ATP is treated as a boundary metabolite, cannot capture this effect. An analogy to this kind of feedback exists in electrical circuits: when modelling the currents in an electrial circuits, we usually assume a voltage source that provides a fixed voltage (as a "boundary condition" of our circuit model). In reality, larger currents in the circuit may cause the voltage of a battery to drop, which then affects the currents too.

2 Results

2.1 Model embedding

To obtain more realistic, dynamical boundary conditions for metabolic pathway models, I developed a method for embedding such models into larger metabolic networks. It can be used for different purposes: to merge several kinetic models, to turn a metabolic network into a kinetic model with a predefined reference state [3, 4], to embed kinetic models into such a scaffold model, or to build kinetic pathway models from scratch from a list of enzymatic rate laws. [5] As shown in Figure 1, the pathways are first mapped onto a "scaffold network". The part of the network that surrounds the pathways ("surrounding network") is then automatically populated with kinetic rate laws, based on available data [6, 3, 4, 7]. The pathway models need not refer to individual metabolic pathways, but can also represent single reactions or entire metabolic subnetworks of any size. To obtain the algorithm, I analysed what constraints need be satisfied, what possible conflicts may arise – for example, when combining models with and without cell compartments – and how these conflicts can be solved. There are two main technical issues. The first issue consists in matching the elements, possibly with different names, and detecting inconsistencies between them. The software semanticSBML can be used for this part [8]. Second, there may be more subtle conflicts that would render the output model physically inconsistent. Such conflicts must either be reliably detected and resolved during the model embedding procedure, or the models must be formulated in such a way that these conflicts do not arise. For example, if metabolic models are formulated with standard chemical potentials, and not equilibrium constants, as basic parameters, the Wegscheider conditions between the resulting equilibrium constants will be automatically satisfied, even if these models are merged [9].

The main steps of the algorithm are shown in Figure 2. We first specify a metabolic network and number of

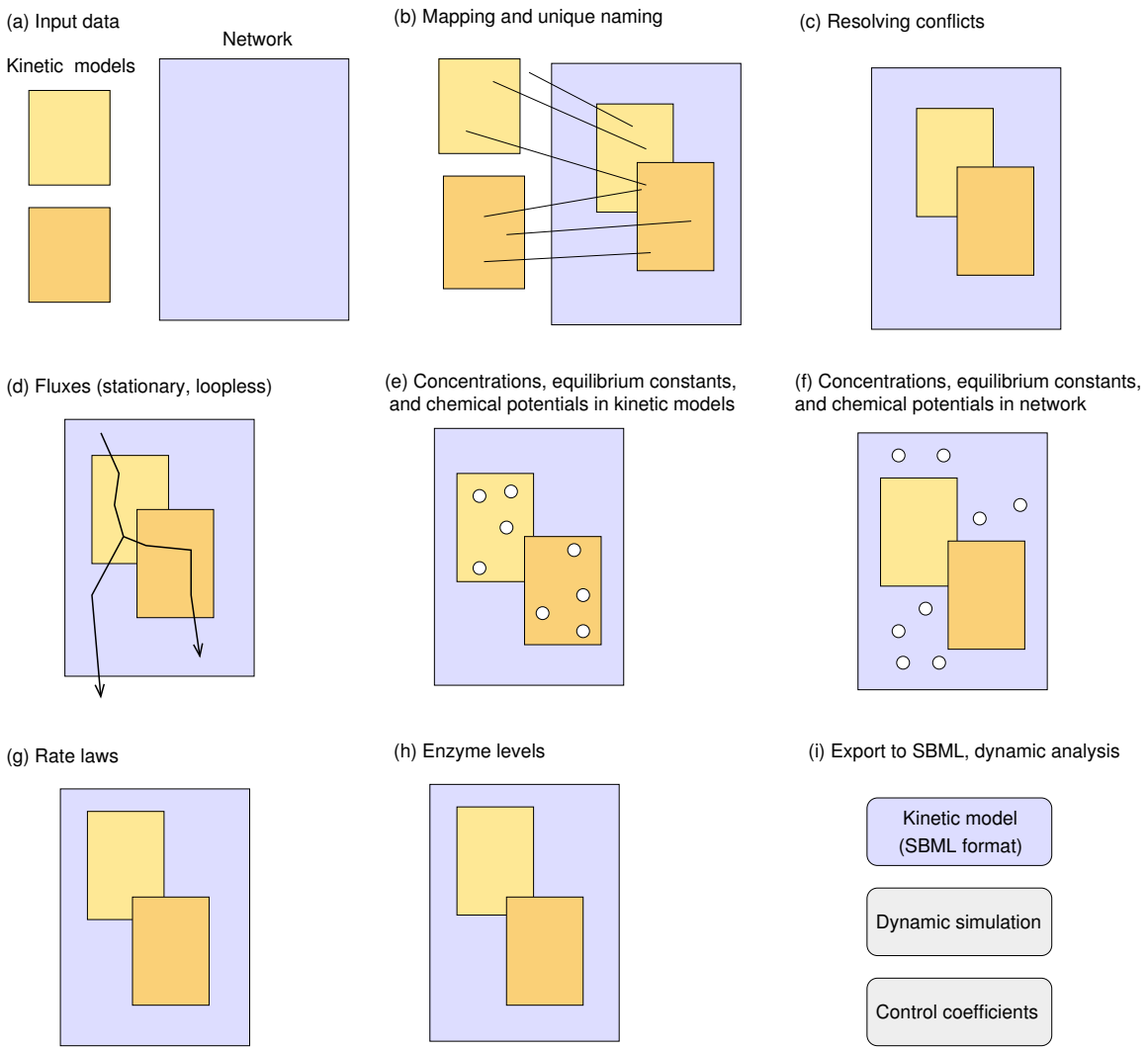

Figure 2: Algorithm for model embedding. The following steps are performed: (a) gather input models and data; (b) map model elements and name them consistently; (c) detect redundant elements (e.g., compounds appearing in several models) and resolve conflicts between them; (d) determine a thermodynamically consistent flux distibution for the entire model; (e) choose consistent metabolite concentrations, equilibrium constants, and chemical potentials within the pathway models; (f) choose consistent metabolite concentrations, equilibrium constants, and chemical potentials in the surrounding network; (g) determine the rate laws and rate constants; (h) determine consistent enzyme levels realising the fluxes; (i) export the combined model und perform dynamical simulations or control analysis.

kinetic pathway models, which may overlap. Each pathway model, when simulated separately, defines a stationary flux distribution. We now determine a flux distribution on the entire network, which must be stationary and must match these fluxes as closely as possible. Importantly, all flux directions from the original pathway models must be preserved; in this way, the local pathway fluxes can later be realized by simply adjusting the enzyme levels. Then, the reactions in the network surrounding the pathway models are realised by to standard rate laws with enzyme parameters chosen to yield the predefined fluxes. The main phases of the algorithm are as follows. (i) A network model and one or several kinetic models are merged into a common network model; all possible conflicts between elements are resolved. (ii) To compute a reference state, a stationary flux distribution is chosen for the entire network; it should match the original flux distributions of the kinetic submodels as closely as possible. (iii) The kinetic pathway models are adjusted to this flux distribution; for the rest of the network, standard rate laws are inserted and adjusted to the flux distribution.

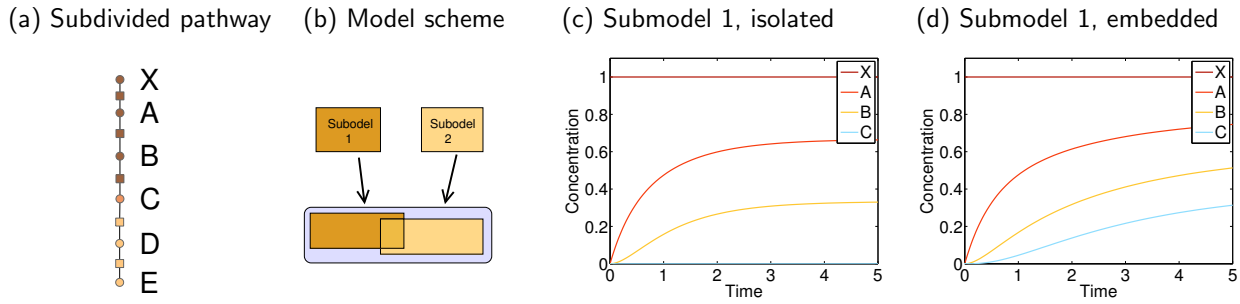

Figure 3: Dynamic effect of a fixed or variable boundary concentration. (a) Simple network model (chain of 5 reactions; metabolites and reactions shown by circles and squares, respectively). For the sake example, the pathway has been split into two kinetic submodels (first three reactions: brown; and last two reactions: yellow), which are then merged again. (b) Model combination scheme. (c) Simulation of the first submodel in isolation, i.e. with a fixed concentration of the boundary metabolite C, which is shared by the two submodels. After a sudden initial increase in A, the concentrations of B and C are increasing, while the concentration of C remains fixed. (d) Simulation of the same submodel, now coupled to the rest of the pathway (submodel 2). Now the concentration of C increases dynamically, which leads to a stronger accumulation of the upstream metabolites in submodel 1.

If several pathway models are combined or embedded, and if these pathway models overlap, there may be conflicts between them. To resolve such conflicts, some extra information may be required: (i) if reactions from different pathway models map to the same stoichiometric reaction, which of the rate laws should be used in the combined model? (ii) For metabolites with different statements about being internal or external, or different concentrations, it needs to be clear which pathway model provides the information to be used. Here, each metabolite or reaction is "owned" by one of the pathway models; for example, there can be a priority order among the pathway models, by which the first model owns all metabolites and reactions it contains, the next model owns all further metabolites and reactions it contains, and so on.

The algorithm has been implemented in Matlab and is described in detail in the appendix. As noted above, it can be used to embed one or more given kinetic models into a metabolic network, for model combination (combining several pathway models without embedding them into a network), or for translating a network into a kinetic model (without embedding any pathway models). By analysing the dynamics of pathway models in isolation and embedded in the network, we can see how embedding a pathway in a larger system changes the dynamics of the pathway in simulations (for an example, see Figure 3), how it changes the control properties of the pathway (e.g., described by the metabolic control coefficients), and how embedding two pathways into a common network creates dynamical interactions between the them. Let us now have a look at three example cases.

2.2 Example cases

The following examples show how simple or complex pathway models can be combined and how this changes their dynamical behaviour.

Linear metabolic pathway with fixed or dynamical product level Our first example, shown in Figure 3, demonstrates how the description of a boundary metabolite (assuming either a fixed or dynamical concentration) can alter the internal dynamics of a pathway. In the example, a simple linear pathway has been split in two parts, which are first described by separate kinetic models and merged again. In this case, there is no additional surrounding network. The simulation shows the pathway's response to an initial sudden increase in substrate concentration: the internal metabolite concentrations are rising with a time delay. In the first simulation, the first subpathway is simulated in isolation, assuming a fixed vanishing concentration of the boundary metabolite

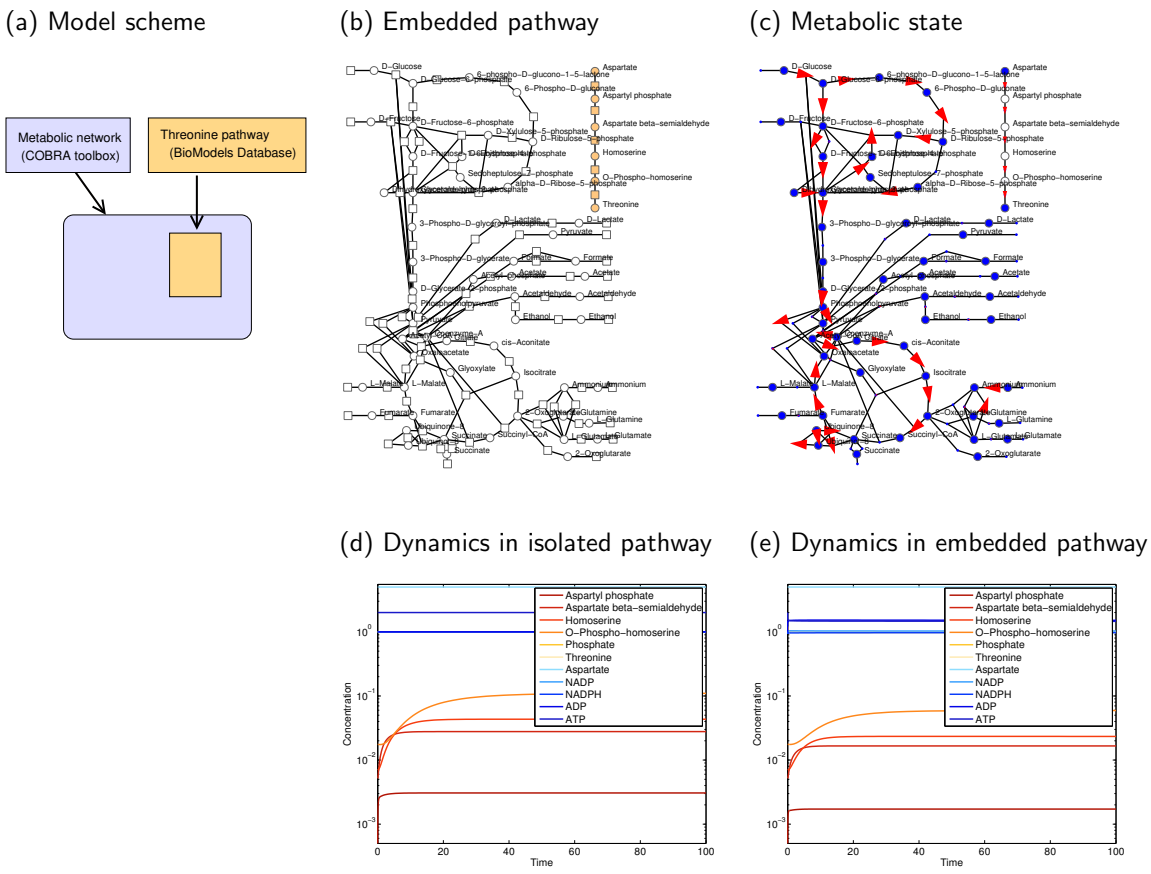

Figure 4: Threonine pathway in *E. coli*, coupled to a model of central metabolism. (a) Model combination scheme. A network model for central metabolism in *E. coli* was obtained from the cobra toolbox [10]. The kinetic threonine pathway model [11] in SBML format was obtained from BioModels Database [12]. (b) Model structure. At first, only the threonine pathway (brown) is described kinetically. (c) Metabolic reference state. Simulation results are shown in (d) for the original (isolated) pathway model and in (e) for the embedded pathway model.

C. In the second simulation, the subpathways are combined; as expected, the concentration of C can rise, and intermediates in the first subpathway accumulate faster.

Threonine pathway coupled to network model for ATP and NADH regeneration The threonine synthesis pathway in *E. coli* produces the amino acid threonine from aspartate and consumes ATP and NADH, which need to be regenerated by the surrounding network. In the kinetic model by Chassagnole *et al.* [11], these two cofactors, as well as aspartate and threonine, are treated as external metabolites with fixed concentrations. In reality, the pathway flux would change these concentrations unless there are additional producing and consuming reactions outside the pathway. The dynamics of these reactions would then affect the dynamics within the pathway: whenever the pathway flux increases, the ATP level will decrease; this may increase the ATP production outside the pathway, which makes the ATP level rise again after a while. This ATP dynamics would feed back on the dynamics of the threonine pathway. To model the interplay between pathway flux and ATP regeneration, I embedded the threonine pathway model into a model of central metabolism, which serves as a "surrounding network" (see Figure 4). Unlike a simple ATP-generating reaction (which could have been coupled to the threonine pathway instead), this network model captures time delays, and even the adaption of other pathways or growth to an increased ATP demand in the threonine pathway. In the example, the levels of aspartate and threonine are still treated as controllable parameters. In a simulation, we can perturb them and to study the response of the flux. Of course, it would also be possible to connect them to additional producing and consuming reactions in an

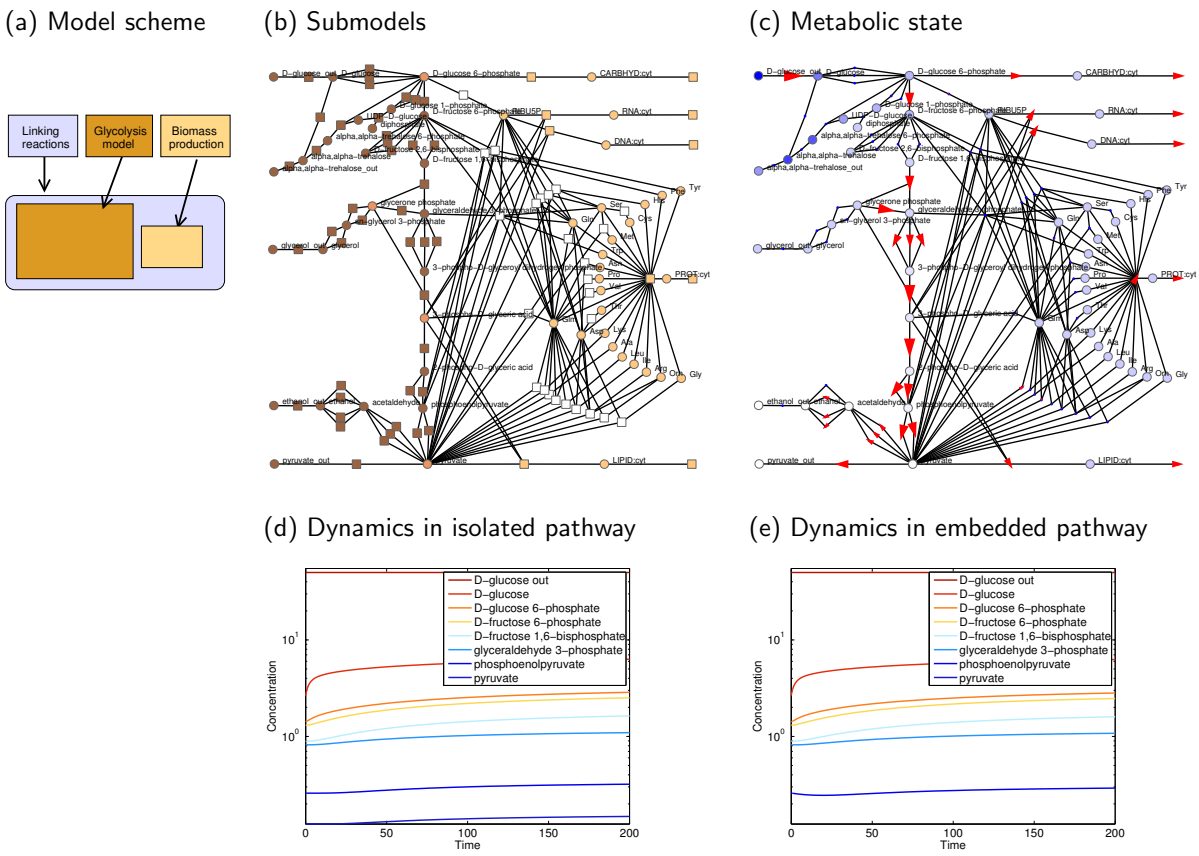

Figure 5: A glycolysis model for the budding yeast *S. cerevisiae*, coupled to biomass producing reactions. (a) Model combination scheme. Glycolysis (brown) and biomass production (yellow) are initially described by separate kinetic models. The models were build automatically from their network structures, using flux analysis and standard rate laws. The models were then coupled through connecting reactions (white squares), which formally play the role of a "surrounding network". (b) Metabolic network structure. (c) Metabolic reference state. Stationary fluxes and concentrations (shades of blue) were derived from the initial kinetic model, other metabolite concentrations were chosen by the algorithm. The panels at the bottom show the dynamics of some glycolytic intermediates after a sudden increase in the glucose level in the isolated (d) or embedded (e) glycolysis pathway.

even larger network model to study their own dynamics.

Yeast glycolysis and biomass production In the previous example, a synthesis pathway was coupled to a metabolic network that provided all necessary cofactors. In the example in Figure 5, we couple two kinetic models, one describing how precursors and cofactors are produced, and another one describing how these precursors and cofactors are used for building macromolecules. The two models are linked by connecting reactions, which are formally treated as a surrounding network. Again, the coupling to other reaction changes the pathway's dynamics after an external increase in substrate level.

3 Discussion

The model embedding algorithm has been implemented in Matlab and is freely available on github[1]. Pathway models and scaffold network can be constructed within matlab or be imported from SBML files. The resulting combined model can be exported in SBML format. However, not all features of SBML are supported.

[1] https://github.com/liebermeister/model-embedding

If a metabolic pathway is embedded into a larger dynamical network, its boundary concentrations can change dynamically, and these changes are affected by the surrounding network dynamics. This alters the dynamics inside the pathway. Accounting for such effects in models is desirable. However, if the surrounding network is large, the simulations will become expensive. If we are only interested in the internal pathway dynamics, details of the surrounding network may not be very important, and even a simplified version of the surrounding network model, which is easier to simulate, may be almost as good. In fact, it is only the dynamical input-output relationship of the surrounding dynamical network that counts. This is why a surrounding model, even if it is not very accurate in its rate laws and kinetic parameters, may still improve the dynamic simulations with a pathway of interest. We can even go one step further, and replace the surrounding dynamical model by a drastically simplified version with a similar input-output relationship. One possibility is the use of reduced black-box models, constructed by linearisation and model reduction by balanced truncation [2].

The treatment of boundary metabolites in our pathway models (as either fixed or dynamic) has an analogy in the comparison between biochemical behaviour *in vitro* and *in vivo*. If a metabolic pathway were constructed in the laboratory by mixing all enzymes and substrates *in vitro*, the resulting *in vitro* system would satisfy simpler boundary constraints (e.g., an exact conservation of all chemical elements) than the corresponding pathway in a living cell. The original *in vivo* pathway in a cell, as an open system, has a different dynamics, and in models it would require a different description of its boundaries. This does not only concern spatial boundaries (fluxes across the cell membrane), but also boundaries between parts of the metabolic network (whether we call those parts pathways or not). This difference between *in vitro* and *in vivo* behaviour can pose problems if one would like to use *in vitro* data for in vivo models. Currently, one can only say that this causes errors and that these errors are hard to quantify. However, a better treatment of boundaries in models could also improve data integration and parameter estimation from *in vivo* data for models of living cells.

Acknowledgements

I thank Bas Teusink, Domenico Bellomo, and Matthias König for insightful discussions. This work was funded by the German Research Foundation (LI 1676/2-1).

References

[1] M. Schulz, F. Krause, N. Le Novère, E. Klipp, and W. Liebermeister. Retrieval, alignment, and clustering of computational models based on semantic annotations. *Molecular Systems Biology*, 7:512, 2011.

[2] W. Liebermeister, U. Baur, and E. Klipp. Biochemical network models simplified by balanced truncation. *FEBS Journal*, 272(16):4034 – 4043, 2005.

[3] R. Steuer, T. Gross, J. Selbig, and B. Blasius. Structural kinetic modeling of metabolic networks. *Proc Natl Acad Sci USA*, 103(32):11868–11873, 2006.

[4] N.J. Stanford, T. Lubitz, K. Smallbone, E. Klipp, P. Mendes, and W. Liebermeister. Systematic construction of kinetic models from genome-scale metabolic networks. *PLoS ONE*, 8(11):e79195, 2013.

[5] B. Teusink, J. Passarge, C.A. Reijenga, E. Esgalhado, C.C. van der Weijden, M. Schepper, M.C. Walsh, B.M. Bakker, K. van Dam, H.V. Westerhoff, and J.L. Snoep. Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. *European Journal of Biochemistry*, 267:5313–5329, 2000.

[6] W. Liebermeister and E. Klipp. Bringing metabolic networks to life: convenience rate law and thermodynamic constraints. *Theor. Biol. Med. Mod.*, 3:41, 2006.

[7] W. Liebermeister. Elasticity sampling links thermodynamics to metabolic control. *Preprint on arxiv.org: arXiv:1309.0267*, page 1309.0267, 2013.

[8] F. Krause, J. Uhlendorf, T. Lubitz, E. Klipp, and W. Liebermeister. Annotation and merging of SBML models with semanticSBML. *Bioinformatics*, 26(3):421–422, 2010.

[9] W. Liebermeister. Validity and combination of biochemical models. In *Proceedings of 3rd International ESCEC Workshop on Experimental Standard Conditions on Enzyme Characterizations*, 2008.

[10] J. Schellenberger, R. Que, R.M.T. Fleming, I. Thiele, J.D. Orth, A.M. Feist, D.C. Zielinski, A. Bordbar, N.E. Lewis, S. Rahmanian, J. Kang, D.R. Hyduke, and B.Ø. Palsson. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA toolbox v2.0. *Nature Protocols*, 6:1290–1307, 2011.

[11] C. Chassagnole, B. Raïs, E. Quentin, D. A. Fell, and J. Mazat. An integrated study of threonine-pathway enzyme kinetics in Escherichia coli. *Biochem J*, 356:415–423, 2001.

[12] N. Le Novère, B. Bornstein, A. Broicher, M. Courtot, M. Donizelli, H. Dharuri, L. Li, H. Sauro, M Schilstra, B. Shapiro, J.L. Snoep, and M. Hucka. Biomodels database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Research*, 34:Database Issue:D689–91, 2006.

A Algorithm for model embedding

A.1 Input data and preconditions

To run the algorithm for model embedding, the following input data need to be provided.

1. **Kinetic pathway models.** One or several kinetic pathway models, each defined by (i) unique identifiers for all metabolites and reactions, (ii) a stoichiometric matrix, (iii) rate laws with all necessary kinetic constants, and (iv) a list of external metabolites.

2. **Metabolic states of pathway models.** Each submodel shows a (stationary or non-stationary) metabolic state, defined by the desired metabolite concentrations. Rate laws and metabolite concentrations together determine the fluxes. Later, during model embedding, it may be impossible to preserve these predefined fluxes precisely; if the fluxes need to be modified, reaction-specific penalty weights can be used to define how strongly each reaction flux may deviate from its original value. The penalty weights need to be defined.

3. **Stoichiometric metabolic model as a scaffold network.** A stoichiometric metabolic model, defined by (i) a list of metabolites and reactions with unique identifiers, (ii) a stoichiometric matrix, and (iii) a list of external metabolites.

4. **Changes in which metabolites are internal or external** If some of the external metabolites (e.g., boundary metabolites of the pathway models) should be set to be internal while embedding the models (or internal metabolites should be set to be external), this needs to be specified.

5. **Criteria for choosing fluxes and concentrations in the network.** To adjust fluxes and concentrations between pathway models and scaffold network, some defined criterion, based on additional constraints or objective functions, must be used. These may be flux objectives concerning the production of energy, valuable compounds, or biomass, or a fit to measured flux values. Concentrations or equilibrium constants may also be constrained or optimised. These objectives and constraints must be predefined and all necessary data (e.g., flux bounds) must be provided.

6. **Element mapping.** A mapping between elements (metabolites and reactions) in the pathway models and the corresponding elements in the scaffold network. If elements from the pathway models have no corresponding elements in the scaffold network, the missing network elements will be created automatically.

7. **Redundant elements.** If a metabolite appears in several pathway models, these model elements are called redundant, and their properties, assigned by different pathway models, may be in conflict (e.g., unique IDs or concentrations; metabolites being treated as external or internal). Each redundant metabolite is "owned" by one of the pathway models, which then determines the properties of this metabolite. Redundant reactions are defined similarly, and their properties (e.g., unique IDs and rate laws) are determined by the pathway model that owns the reaction. To determine which pathway model owns which metabolites and reactions, a priority order among the pathway models must be defined. The first model owns all metabolites and reactions it contains, the next model owns all further metabolites and reactions it contains, and so on. By default, the priority order is given by the order of pathway models in the list.

8. **Layout information.** If desired, layout information (positions and properties of metabolite and reaction elemenets in the network graphics) should be given. The list of elements should cover the entire scaffold network, as well as all elements of the pathway models to be added to the scaffold network.

A.2 Algorithm

The algorithm can be used to embed kinetic pathway models into a surrounding network, to couple different pathway models, or to turn metabolic networks into dynamical models. Note that the term "pathway model" is used here for simplicity only: what I call "pathway model" can be any kinetic metabolic model describing a single reaction, a pathway, or an entire metabolic subnetwork.

Phase 1: Map the elements between pathway models and scaffold network In this phase of the algorithm, model elements are mapped between models and conflicts between are detected and resolved.

- **1.1 Check element mapping** Map the metabolites and reactions between pathway models and scaffold network, based on unique names or identifiers. Check whether all metabolites and reactions from the pathway models exist in the scaffold network. If some are missing, add them to the scaffold network.
- **1.2 Unique naming.** If necessary, rename the metabolites and reactions in the scaffold network to match their names and identifiers in the pathway models.
- **1.3 Find redundant elements.** Check whether reactions, parameters, or metabolites in different pathway models are redundant (the check is based on unique names or identifiers).
- **1.4 Find and resolve conflicting model statements.** If redundant metabolite elements exist, check for conflicts between the pathway models (e.g., between metabolite concentrations or between a metabolite being tagged as external or internal); in case of a conflict, decide which submodel "owns" the metabolite, or let the user decide, and adjust the pathway models accordingly. If redundant metabolite elements with different concentrations are found, issue a warning and tell the user that the concentrations will be altered. If redundant parameter elements with different values are found, issue a warning and tell the user that parameter values will be altered. If redundant reaction elements with different rate laws or local parameters are found, issue a warning and tell the user that rate laws or local parameter values will be altered. Tell the user that these changes can change the reaction rates in pathway models.
- **1.5 Set metabolites internal or external as specified by the user** Based on information given by the user, set some internal metabolites to be external or *vice versa*.

Phase 2: Choose a consistent metabolic state (flux distribution, concentrations, and equilibrium constants In this phase of the algorithm, consistent fluxes and metabolite concentrations are determined for the entire network. The fluxes and concentrations must be thermodynamically feasible. If possible, the values from the original pathway models should be preserved; where this is not possible, adjustments can be made, but the changes should be as small as possible.

- **2.1 Stationary fluxes.** Compute a stationary, thermodynamically feasible flux distribution for the entire network model, while preserving the fluxes from the pathway models, e.g., by flux balance analysis with flux minimisation. If preserving the fluxes is impossible, determine a stationary flux distribution in the network that preserves the fluxes in the pathway models as closely as possible (e.g., minimising the Euclidean distance or a distance with reaction-specific penalty weights). In any case, the flux *directions* from the pathway models must be preserved. If this is impossible, stop and issue an error message.
- **2.2 Concentrations in pathway models.** Update the pathway models such that these fluxes are realised: choose metabolite concentrations in the pathway models. The concentrations from the pathway models should be preserved as clsely as possible. The fluxes are recalculated in the pathway models; if any of the fluxes changes its direction, the algorithm should stop with an error message.

- **2.3 Equilibrium constants in pathway models.** To compute the equilibrium constants within the pathway models, determine a set of equilibrium concentrations within each pathway model. Equilibrium concentrations can be computed numerically by setting all metabolites internal and integrating the model until a steady state has been reached.
- **2.4 Chemical potentials in pathway models** Compute all chemical potentials within the pathway models (using the previously determined concentrations and equilibrium concentrations).
- **2.5 Concentrations and equilibrium constants in surrounding network.** Determine metabolite concentrations and equilibrium constants in the surrounding network, while satisfying all predefined objectives and constraints (upper and lower bounds, target values to be approximated, etc). The equilibrium constants or formation energies may either be predefined or be determined together with the metabolite levels. They must yield feasible thermodynamic forces, i.e. their signs must agree with the predefined flux directions. If this is possible, thermodynamic parameter balancing can be used to obtain realistic values (metabolite levels and equilbrium constants). If it is not possible, stop and tell the user to change the assumptions about the flux distribution; or to continue and accept that the model will not be thermodynamically balanced.

Phase 2 (variant in which network fluxes are imposed on pathway models) There is a variant of the algorithm in which we do not start fluxes in the pathway models and impose them onto the scaffold network, but in which we do exactly the opposite: the pathway models are updated such that this flux distribution $\mathbf{v}^*$ is realised within the pathway models, and concentrations may have to be adjusted within the submodels. The fluxes need not be stationary within the pathway models. Therefore, adjusted concentrations are determined within the pathway models by minimising

$$\min \stackrel{!}{=} \Delta \mathbf{c} \qquad \text{s.t.} \quad \mathbf{v}^* = \mathbf{v}(\mathbf{c}, \mathbf{k}) \tag{1}$$

for each pathway model. The pathway models are updated according to their priority order. In each model, the concentrations of metabolites that belong to a higher-priority model are kept fixed. Then we start the workflow again with the updated pathway models.

Phase 3: Choose or adjust the rate laws In this phase of the algorithm, rate laws are chosen or adjusted such that all fluxes are kinetically realised with the chosen metabolite levels; then the final dynamical model is built.

- **3.1 Insert rate laws into the surrounding network.** Run elasticity sampling [7] for the surrounding network to obtain rate laws and kinetic constants that realise the metabolic flux directions defined in phase 2.
- **3.2 Adjust enzyme levels in the embedded pathway models.** If the flux distributions in the pathway models have been adjusted in phase 2, the rate laws in the pathway models must be adjusted, too, to realise these new fluxes. Since the flux directions have been preserved, this can be simply done by adjusting the $v^{\max}$ values (or the enzyme levels). To do so, the original $v^{\max}$ values or enzyme levels are scaled by a factor $v^{\text{flux}}/v^{\text{kin}}$, i.e., the ratio between the desired flux and the reaction rate computed from the previously adjusted metabolite concentrations.
- **3.3 Build the model.** Construct a kinetic model for the entire network and export it as a matlab data structure or as an SBML model. Conflicts between the kinetic pathway models are resolved based on their priority order or based on explicit information given by the user.

Phase 4 (optional): Replace the surrounding network by reduced models. In this last, additional phase of the algorithm, the surrounding dynamical network may be replaced by a simplified black-box model as mentioned in the discussion section and described in [2].

A.3 Matlab implementation

Matlab code and example models are freely available on github[2]. More documentation and results for the example models can be found on www.metabolic-economics.de/model-embedding.

[2] https://github.com/liebermeister/model-embedding