How to use the K-FIT code

The code requires three input files in XLS(X) format: the metabolic model, reaction mechanisms, and the experimental data. Please use the example files in the folder as a template for creating your own input files.

The metabolic model file (model.xlsx) contains two tabs: “Reactions” and “Metabolites”.

In the “Reactions” tab:

RxnID is the abbreviated name of the reaction. E.g. PGI

Rxn name is the full name of the reaction. E.g. Phosphoglucose isomerase for PGI

Rxn Formula contains the stoichiometric reaction formula for the overall reaction expressed using abbreviated metabolite names. Stoichiometric coefficient is specified in parenthesis before the metabolite E.g. (1) G6P.

Reversible reactions are indicated using a double-headed arrow (⇔) and irreversible reactions are indicated using a single headed arrow (→). A default lower bound of -10^8 is set for reversible reactions. Irreversible reactions have a lower bound of 0.

The default upper bound for all reactions is set to 10^8. Alternatively, if the direction of flux through a reversible reaction is known to be in the reverse direction only, the upper bound for such reactions can be set to 0.

Macro/bm is not used for FBA. It is used to exclude reactions with non-integer stoichiometric coefficients (such as biomass reaction, protein synthesis) when checking for correctness of the atom mapping model.

Subsystem describes the pathway(s) to which the reaction belongs. It is included for convenience of data analysis once fluxes have been elucidated.

The metabolites tab contains details on the list of metabolites in the model.

“metid” contains the list of abbreviated metabolite names. The corresponding metabolite names are contained under “metnames”.

Only “metid” needs to be specified for constructing the metabolic model and performing analyses using stoichiometric methods (any type of FBA except thermodynamic-FBA, FBA, OptKnock, etc…). The supplied code checks for consistence between the list of metabolites specified in the “Metabolites” tab and the metabolite abbreviations used in the “Reactions” tab and returns an error if there are missing/extra metabolites. The error log also reports the list of missing/extra metabolites.

“metnames” is an optional input in this excel file. While it is not used in the code, this information allows the user to compare models from different sources as different sources use different metabolite IDs (e.g. BiGG, modelSEED, KEGG, etc).

“metformula” is also an optional input. It allows the user to check for elemental balancing in their constructed models but is not used in this code. Additional troubleshooting features will be introduced in subsequent versions of this code.

The second input file required for kinetic parameterization is the specification of the mechanism of enzyme catalysis along with the regulatory properties.

“mechanism” specifies whether the enzyme catalysis is associated with ordered binding of metabolites to the enzyme or whether it follows a ping-pong mechanism (such as transaldolase, transketolase, and transaminase reactions).

“SBO” is the binding order for reactants. If a reaction has multiple reactants (such as PFK: F6P + ATP → FBP + ADP), the reactants must be listed and separated by a semicolon (F6P;ATP). If the stoichiometric coefficient of a metabolite is greater than 1 (say 2), then it must be listed multiple times (twice in case the stoichiometric coefficient is 2).

“PRO” is the product release order and is filled out similar to “SBO”.

“CI”, “UCI”, and “NCI” correspond to competitive, uncompetitive, and non-competitive inhibitors, respectively.

“act” lists the activators for a particular enzyme.

“exch” indicates whether a particular reaction is an exchange reaction or not.

“sub” indicates whether a particular reaction represents the uptake of a metabolite (glucose, oxygen, ammonia, etc…).

“exch” reactions that are not uptake reactions are modeled as irreversible to ensure that metabolites that are not present in the growth media are not considered as uptaken nutrients by the model.

Finally, the experimental data used for estimating the elementary kinetic parameters is provided in the third input file. This input file contains the fluxes in the WT strain and all mutant strains. The flux through the reactions under “ID” are reported under “flx” with a standard deviation in “SD”. “rxns” refers to the reaction ID in the model file (model.xlsx) to which the measurement under “ID” corresponds to.

The “Mutant” tab contains the reaction(s) that are knocked out in the mutant strains. “WT” is the Wild-Type strain with no mutations. Fluxes in the “WT” are excluded from fitting. Instead, they are used to anchor kinetic parameters in the K-SOLVE step. Therefore, the “SD” for the “WT” strain does not affect the parameterization procedure.

Running K-FIT:

the first step is the compile the information from the three input files into a MATLAB model.

This model is created with a set of default options. The most important of these options is the initialization from a randomized starting point (model.options.reinit). By default, this is set to true, meaning that the parameterization with begin from a random initial guess.

The next step is to estimate the kinetic parameters by calling “kineticestimate.m”. kineticestimate takes the “model” created in “modcompile.m” as a required input and a second optional input, and returns an output file containing the minimum SSR, predicted fluxes, predicted metbaolite concentrations, fit statistics, and values of the variables at the optimal solution.

When model.options.reinit is set to true (initialize from random initial guess), the second input is not required. However, when starting from a previous solution, model.options.reinit must be set to false and the previous result must be provided as a second input. This second input contains the previous solution which serves as the initial guess.