Prospecting of metabolic pathways to implement Methacrylic Acid (MAA) production in S. cerevisiae

Microbial Cell Factories 2020

Preparation of class:

1. Go to https://sourceforge.net/projects/reactpred/ and download the Reactpred zip file.
2. Unzip the file into the documents folder of your computer.
3. The program is ready to be used by opening the batch file ReactPred_Ui.bat

Note that you need flash player to work with this program.

Methylacrylic acid (MAA) is an intermediate in the production of methyl methacrylate (MMA), a monomer used in the production of plastics made of poly-MMA or of the PVC-replacer methacrylate-butadiene-styrene. MAA or MMA can be obtained through chemical synthesis, however, microbe-based production of these chemicals has the advantage of bypassing the use and the production of very toxic products for the environment (e.g. MAA synthesis requires acetone cyanide, a fuel-derived compound that is highly toxic), also requiring less energy and less input material and often leading to more pure molecules.

In this class we will focus on *S. cerevisiae* as an interesting host for MAA production, due to its genetic tractability. To this end, we'll search for metabolic routes from a given intermediate in yeast and/or microbial metabolism to MAA.

We will use 4 main methods to search for MAA producing reactions: the databases BRENDA, MINES and Atlas of Biochemistry and the tool Reactpred. MINES and Atlas of Biochemistry are databases that were built with the help of the BNICE tool (discussed in the theoretical class).

Each MAA producing reaction will be evaluated on its feasibility and likelihood to work when implemented in *S. cerevisiae*. Some parameters to consider are:

- Is the substrate of the reaction already present in the yeast's metabolism?
- Is it present in other microbial metabolisms?
- How much do we know about the reaction in question? Do we have a candidate enzyme?
- Is the candidate enzyme from an extremophile? Is it from another yeast or a higher eukaryote?
- How similar is our substrate from the native substrate? (molecular weight and chemical structure)

Fill out the table in the final page during the class, to summarize the pathways found with the different methods. This table will aid the final decision of the best pathways to implement.

Method 1) Search in BRENDA

BRENDA is the main collection of enzyme functional data available to the scientific community, manually curated. We will use this database to search for annotated enzymatic activities that involve MAA – in particular, reactions where MAA is a product. This is the only database that gives us proven reactions, as the other tools will provides us only with reactions **predicted** to produce MAA.

1. Go to https://www.brenda-enzymes.org/index.php
2. Search for MAA as a ligand.
3. Browse through the Ligand page to find reactions where MAA is a product? Note the difference between *in vivo* and *in vitro* reactions.

4. For each of these reactions, click on the reaction page to see name of the enzyme and if there are co-factors or co-substrates.
5. For each of these reactions, research the substrate in question to verify if it is a known metabolite in yeast or other organisms. Use YMDB and KEGG
    a. YMDB → Go to http://www.ymdb.ca/ and search for the compound (upper right). If you get a hit, click on the compound's page and browse through it to verify if this is a metabolite and in which pathway it is inserted.
    b. KEGG → Go to https://www.genome.jp/kegg/

Method 2) Atlas of Biochemistry

This database comprises the *whole theoretical reactome* from the *known metabolome* through expansion of the known biochemistry presented in the KEGG. This was achieved by applying the BNICE framework to the all the compounds present in KEGG, obtaining all possible interconversions between this list of compounds. This means Atlas only works with KEGG compounds. Since MAA is not a KEGG compound (unnatural molecule), it is not present in Atlas. However, you can search for its derivate methyl methacrylate (MMA).

1. Go to http://lcsb-databases.epfl.ch/Home
2. Insert the login info (user: mira_ist ; password: 8za75t)
3. Choose the Atlas option and then BNICE.Atlas reactions. A very large table is presented encompassing all predicted reactions in the Atlas database.
4. Search for reactions that contain MMA by using its KEGG code: C14527. You can check the KEGG code of any compound by looking it up in KEGG and opening the compound page (e.g. the MMA page is https://www.genome.jp/dbget-bin/www_bget?cpd:C14527)
5. How many different molecules can be converted in MMA? Evaluate each precursor by:
    a. Searching them in KEGG and YMDB (step 5 of Method 1)
    b. Doing a new search in Atlas w/ this molecule. Can it be produced from a known metabolite?
6. Now about the reaction. Note that the EC number associated to it only has 3 fields, instead of 4. This is because it is a generalized reaction rule. Use BRENDA to know more about the predicted reaction
    a. Go to the enzyme tree (https://www.brenda-enzymes.org/ecexplorer.php?browser=1) to see the class of enzymatic reaction by looking for the EC number
    b. Search for the EC number with the option "begins with". How many candidate enzymes are annotated in BRENDA?

Method 3) MINE database

This is an extension of known metabolite databases to include molecules that have not been observed, but are likely to occur based on known metabolites and common biochemical reactions. This was achieved using both BNICE and expert-curated reaction rules.

1. Go to http://minedatabase.mcs.anl.gov/#/home
2. Go to structure search, draw the MAA structure and click search.
3. Go to the MAA info page and select the Product tab. This page lists all the reactions in the database where MAA is a product.

4. Browse through the list to check all possible precursors. Since this is a large list, all substrates can't be researched. Choose 1 or 2 reaction that you think are feasible. **TIP:** stay off very large molecules, such as Vernolide and Glaucolide A. Focus on molecules that are already very similar to MAA and require only a small chemical transformation.
5. For the reactions chosen previously, research both the substrate and the reaction involved. Always note if there are co-substrates.
    a. Substrate: search it in KEGG and YMDB, as done previously.
    b. Reaction: look for enzyme class and nº of candidate enzymes, as done previously.

Method 4) Reactpred

This is an open source tool for users to predict biochemical reactions and pathways. It enables the user to give an input molecule and a set of reaction rules and the tool calculates all the possible pathways form where the input molecule can derive, in a retrosynthesis fashion.

1. After installing ReactPred open the file ReactionpredictionUI.jar
2. Go to the "Pathway Predicton System" tab and insert the molecule SMILES.

SMILES (Simplified Molecular Input Line Entry System) is a line notation (a typographical method using printable characters) for entering and representing molecules and reaction.
MAA: CC(=C)C(=O)O

3. On Input Reaction Rules click "Load (default)".
4. Select pathway length 1, prediction tolerance 0.5 and leave the box "rule reversibility" unchecked. Select the option "retro-synthetic" and submit the job.
5. How many pathways were generated?
6. In the log box you should be able to see the name of .txt file where the results were stored. Check the name, go the to the "Pathway analysis System" tab and open the results .txt file. Click Load.
7. Right now you have too many reactions to analyze. How can you decrease this number? Filter and order by Gibbs free energy (we want the reactions with the most negative energy).
8. Now that you have a smaller number of reactions, choose 1 or 2 pathways that you think are good candidates
   Use the same criteria that we used on the previous method: we want precursors that are already very similar to MAA, not very large molecules.
9. For each of the reactions analyze the substrate in question by using the same methods as above. In Reactpred all molecules are written in SMILES. To see their names, go to https://pubchem.ncbi.nlm.nih.gov/search/search.cgi# , select search by identity/similarity and insert the SMILES.

Method	Pathway	Precursor research			Reaction research			Comments
		Present in yeast metabolism?	Present in microbial metabolism?	Which pathway?	EC number	Class of reactions	Nº of candidate enzymes	
