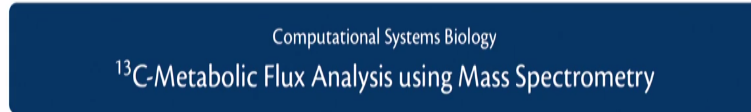


Computational Systems Biology  
Aarthi Ravikrishnan  
Department of Biotechnology  
Indian Institute of Technology – Madras

Lecture - 84  
13C-Metabolic Flux Analysis using Mass Spectrometry

(Refer Slide Time: 00:11)



- ▶ Data Analysis
- ▶ An Example

Aarthi Ravikrishnan

Department of Biotechnology, Bhupat & Jyoti Mehta School of Biosciences  
Initiative for Biological Systems Engineering (IBSE)  
Robert Bosch Centre for Data Science and Artificial Intelligence (RBC DSAI)  
INDIAN INSTITUTE OF TECHNOLOGY MADRAS

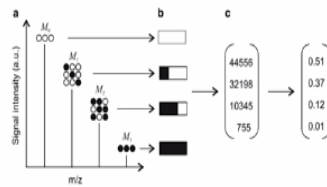


In this lecture, we will look at how to integrate the data obtained from labeled amino acids into metabolic models and see how we can use it to derive information about the fluxes. We will also discuss a small example where <sup>13</sup>C MFA has been carried out to understand the flux distributions in an experimental scenario.

(Refer Slide Time: 00:33)

## Data processing

- From the areas, we get
  1. Relative abundances of each isotopomer
  2. Fractional labelling
- Relative abundances – used to infer the positional labelling
- Fractional labelling – check the quality of data



schmitz et al (2015) Hydrocarbon and Lipid Microbiology Protocols 223-243

$$\text{Fractional labelling} = \frac{\sum_{n=0}^N n \times (\text{Relative abundance of each isotopomer})}{N \times \sum \text{Relative abundance of each isotopomer}}$$

where  $N$  is the total number of carbon atoms in each amino acid

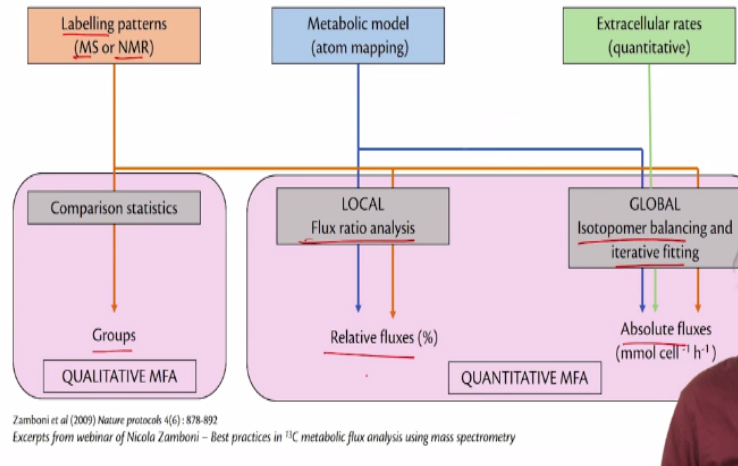
Using these areas, you calculate the relative abundances of each isotopomer and also the fractional labelling. So fractional labelling is nothing but the relative abundance of each isotopomer \* the number of carbon atoms/the total number of carbon atoms \* the sum of all of it. So this is the formula for calculating the fractional labelling and this is basically used to check the quality of the data that you have say for instance.

You feed your experiments with 20% label glucose the fractioning labelling that I should observe should be 20% it should not be more than I mean of course with standard deviation of + or - 0.5 but it should not be more than say it should not be 25 or 30 it means there is something wrong with experiment. Whereas with relative abundances you can find out at what positions the labelling has happened, at say for the example if you have pyruvate has it got labeled at the third position or at the second positions and so on.

So these are the 2 things that you might want to calculate with the areas that you have got from the mass spectrum.

**(Refer Slide Time: 01:50)**

## From data to fluxes



So this is the general overall outline of how do you calculate the fluxes from the data. So you have the labelling patterns which I was talking to you before and I also want to mention that in this lecture we had looked about mass spectrum at least specifically gas chromatography mass spectrometry and there are also other techniques like NMR that can give you ideas of the labelling patterns and with this labelling patterns you can identify them into groups.

And these we call it as qualitative metabolic flux analysis and there is another one called quantitative metabolic flux analysis where you have a metabolic model so by a metabolic model I mean a small model where you know the atom transitions about this we will come to the next slide and you will basically calculate the flux ratio of which pathway has given rise to which labelling and with this you can find out the relative fluxes.

Whereas when you also quantify the extracellular rates such as the substrate uptake rate the product formation rate you can do a global analysis where you do isotopomer balancing and you have a model and you do iterative fitting to find out the flux map for a given particular pathway and these would give you the absolute fluxes. So here you get the relative fluxes and here you get the absolute fluxes.

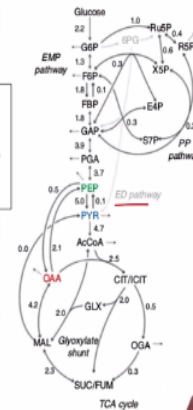
**(Refer Slide Time: 03:24)**

## Global iterative fitting

- Model to comprehensively describe propagation of  $^{13}\text{C}$
- Balances constructed for each isotopomer
- Extracellular rates, flux guesses – simulate this model
- Process is repeated till extracellular rates and  $^{13}\text{C}$  data is matched
- Incomplete/ill-defined networks – wrong fluxes

**Atom transition model for EMP**

$\text{GLC (ABCDEF)} \rightarrow \text{G6P (ABCDEF)}$   
 $\text{G6P (ABCDEF)} \rightarrow \text{F6P (ABCDEF)}$   
 $\text{F6P (ABCDEF)} \rightarrow \text{FBP (ABCDEF)}$   
 $\text{FBP (ABCDEF)} \rightarrow \text{DHAP (CBA) + GAP (DEF)}$   
 $\text{GAP (ABC)} \rightarrow \text{PGA (ABC)}$   
 $\text{PGA (ABC)} \rightarrow \text{PEP (ABC)}$   
 $\text{PEP (ABC)} \rightarrow \text{PYR(ABC)}$



Zamboni et al (2009) Nature protocols 4(6): 878-892

So what do you mean by global iterative fitting. So this is a technique where you know already of a model which has the atom transitions, for instance if you look at glycolysis you have glucose to pyruvate the reactions from glucose to pyruvate and you already know that if there were 6 carbons in glucose there would be 6 carbons in glucose 6 phosphate. So with this you frame rules of this sort I mean this is just an example.

So if you have ABCDEF as different carbons of glucose you have G6P that is ABCDEF and so on and so forth so till pyruvate you have this particular label and this is a simple atom transition model for glycolysis. So like this you have it different pathways that happen inside the cell like say for example ED pathway, pentose phosphate pathway and so on and with the measured rates of substrate uptake product formation you make and also different guess about fluxes.

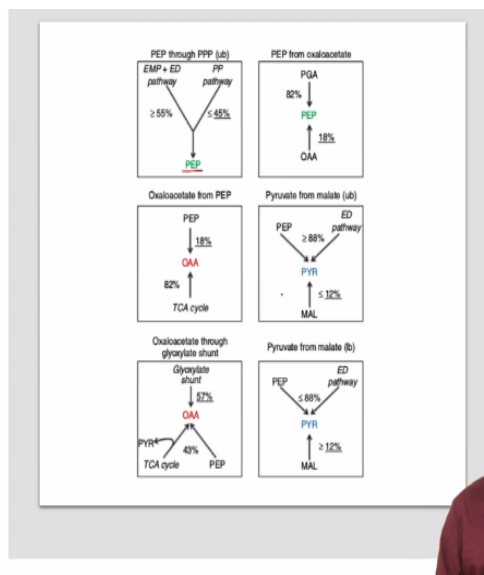
You tend to simulate this model by fitting and different labelling data and you would keep repeating it until your extracellular rates and the C13 data is matched and the only problem with this kind of approach although it gives you a good metabolic flux map is that if you have an incomplete or an ill-defined network say for example you make some mistakes in generating your atom transition model you could obviously lead to wrong fluxes and this could lead to multiple other incorrect and false data.

**(Refer Slide Time: 05:19)**

## Local flux ratio analysis

- Quantifies ratios of fluxes through converging pathways
- Independent interpretation of extracellular rates
- Algebraic equation is written that relates the relative amount of fluxes as a function of  $^{13}\text{C}$  labelling pattern

Zamboni et al (2009) Nature protocols 4(6) : 878-892



The next one is local flux ratio analysis where you try to find out the fluxes through different pathways. You have a ratio of fluxes between different pathways. So here you just have a particular metabolite of interest and with that metabolite you try to frame the rules and using those rules or using those, using the labels that you observe for that particular metabolite you try to see that so much has come from a particular pathway and so much has come from another pathway.

For instance, if you take this particular small example, where you have PEP. So PEP here and based on labelling patterns and all those they have identified that PEP has come from both EMP, ED pathway and PP pathway where the percentage with which it has come through different pathways is different simply because the labels observe the different positions were different. So similarly you could do it for several essential metabolites in the central carbon metabolism and try to understand how the labels have got propagated from the substrate to your product.

**(Refer Slide Time: 06:33)**

## Software to process the data

- IMS2Flux<sup>1</sup> – PERL interface that processes mass spectrometric measurements and generates data for 13C-MFA
- INCA<sup>2</sup> – MATLAB code – full isotopomer distributions of measured metabolites
- FiatFlux<sup>3</sup> – MATLAB code – supports flux analysis
- SUMOFlux<sup>4</sup> – Predicts flux ratios using <sup>13</sup>C–data by combining surrogate modelling and machine learning

1. Poskar et al (2012) *BMC Bioinformatics* 13:295

2. Young (2014) *Bioinformatics* 30(9):1333-1335

3. Zamboni et al (2005) *BMC Bioinformatics* 6:209

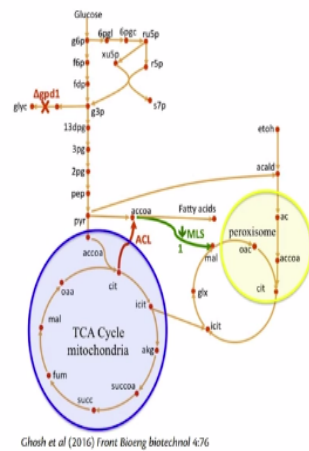
4. Kogadeeva & Zamboni (2016) *PLoS Comp Bio* <https://doi.org/10.1371/journal.pcbi.1005109>

So these are few software. So all these global fitting and the local flux ratio analysis these are already done using different software and these are few softwares which can process your data that is directly coming from GC-MS. I can give you data for performing C13 flux analysis. So the first one is IMS2Flux which is a PERL interface and this processes your areas from a mass spectrometry measurement and it also generates data for C13-MFA and another one is INCA, where it is basically a MATLAB code where you have the full isotopmer distributions.

And it gives you the full isotopmer distributions of measured metabolites and this is also used for generating data for C13-MFA and there is also FiatFlux and SUMOFlux. So FiatFlux is one of the old methods, but still it supports it is a very user friendly interface and supports flux analysis and SUMOFluxes are very recently developed method which combines surrogate and machine learning techniques to find out different flux ratios. So these are few of the softwares which I used.

**(Refer Slide Time: 07:52)**

## <sup>13</sup>C MFA for overproduction of FA in yeast



- Non-native ATP citrate lyase (ACL) expressed
- Use of MFA suggested Acetyl-CoA was lost through malate
- Downregulating this enzyme increased fatty acid (FA) production 26%
- Flux analysis also suggested knocking out glycerol production (GPD1)
- FA production improved by 33%

And this is one of the applications that I want to mention where <sup>13</sup>C MFA has been of great potential. So here they have taken yeast the baker's yeast and they have found out the fluxes to different pathways and they wanted to over produce fatty acids and for this they have expressed a non-native ATP citrate lyase and there is a metabolic flux analysis and when they did a metabolic flux analysis they found out Acetyl-CoA that is the fatty acid that they wanted to overproduce was lost through malate.

So they wanted to down regulate this enzyme and see how the system works and after down regulating they found that the fatty acid production increased by 26% and also they again they would have flux metabolic flux analysis and this suggested that they knocked on the glycerol production and once when they did this they found out the fatty acid production was improved by 33%. So this is one of the applications of MFA where you try to see how the flux ratio fluxes varies through different pathways and over produce the product of interest.

**(Refer Slide Time: 09:09)**

## Summary

- $^{13}\text{C}$ -MFA describes the origin and fate of metabolites
- Choice of tracers depends on the type of application
  - Positionally labelled substrates – Resolve alternate pathways
  - Uniformly labelled substrates – Understand different types of reactions
- Prior knowledge of the substrate uptake rates/product formation rates is essential
- Steady state condition is essential to perform  $^{13}\text{C}$ -MFA
- Involves analysis of isotopomer distributions from the mass spectra

So to summarize  $^{13}\text{C}$  MFA describes the origin and fate of different metabolites say for example you have a particular carbon source and you want to trace how it gets into different metabolites you might want to use a labeled substrate and try to find out different pathways through which this carbon source gets incorporated and the choice of the tracer depends on the type of application so as I was already saying positionally labeled substrates are used to resolve alternative pathways.

And uniformly labeled substrate are used to understand different types of reactions that happen in a particular cell and for carrying out this analysis prior knowledge of the substrate uptake rates and the product formation rates are very essential simply because you do not want to take a sample at a time point when the cell has not reached a steady state which I did not mention in my slide, but it is very important that you carry out this analysis in steady state.

Because the labels have to be incorporated inside the cell for you to be able to identify these labelling patterns and these could vary with different phases of the growth. For instances in the initial lag phase you do not find any big label incorporation because the cells have still trying to adapt to the environment and grow whereas in a late exponential phase there is a sort of steady state that is observed in a batch culture and this is essential to perform  $^{13}\text{C}$  MFA because you know after a particular time point.



You do not observe that the cells multiply because of which you do not observe any increase in the labelling patterns and the labels are at steady state and this analysis to you know basically do a C13 MFA you try to identify different isotopomer distributions which are saying from the mass spectra and back trace it to different flux ratios and also generate the flux map and using this flux map you can try to see if you want to over express a particular metabolite or down regulate a particular gene or so on. So with this I end my lecture.

**(Refer Slide Time: 11:37)**



Recap

Topics covered

- ▶ Data Analysis
- ▶ An Example

In the next video ...

- ▶ Recap
- ▶ Simple Exercises

In this video, we covered the following topic that is the data analysis where we looked at how to integrate the data that we got from GC-MS analytical technique into a metabolic model and look at different pathways through which the labelling patterns came from. We also discussed a small example where <sup>13</sup>C MFA has been applied. In the next video we will look at simple exercises which deal with data from GC-MS analytical technique and also see how to do deal with this data to carry out further analysis.