# Main fermentative pathways

We will now combine the overall anabolism with a detailed description of the catabolism. We'll start with anaerobic conditions where substrate level phosphorylation is used to generate **ATP** in the catabolism. The detail modelling of the catabolism will be explained via an example. Let's start with a molar map of anaerobic glucose breakdown:
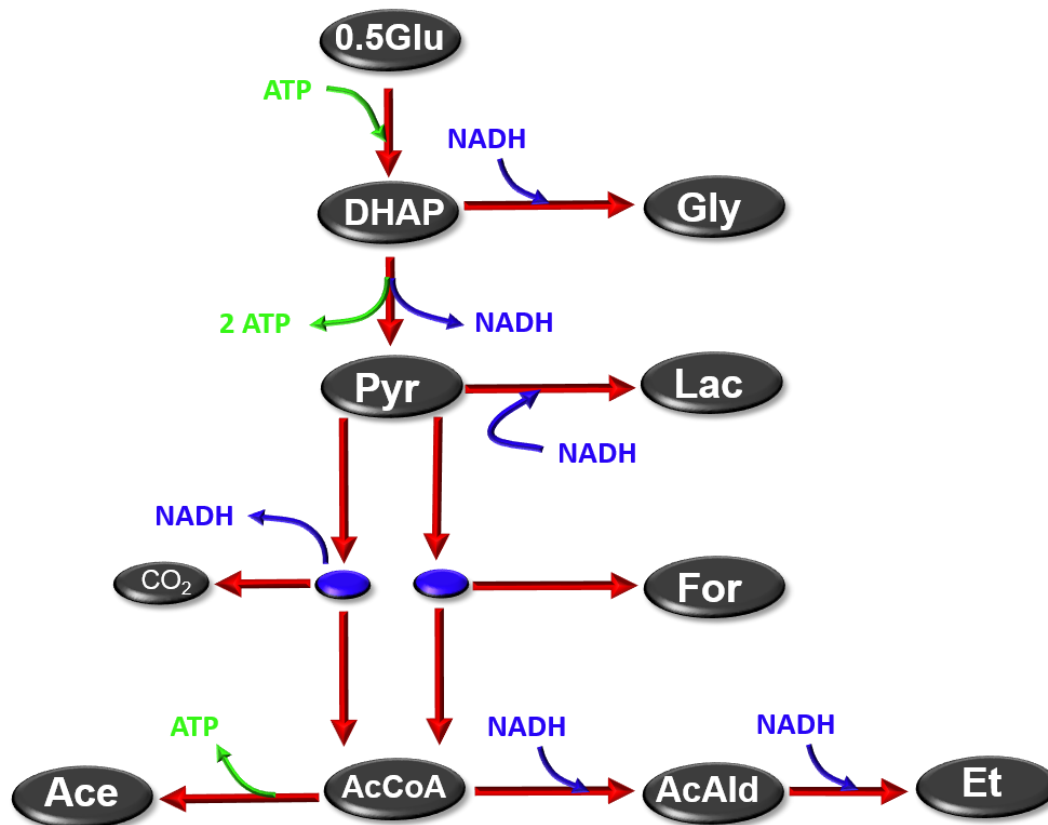
Main catabolic pathways under anaerobic conditions. *The idea of the simplified map is to show the redox (NADH formation or usage) and energy (ATP formation or usage) steps in the catabolic pathway. Accordingly multiple enzymatic steps are lumped in sections of the diagram. All ATP and NADH are given as mol per mol of substrate except for glucose where only a half a mol was considered (as a result of the C3-C3 split in glycolysis). Direction of ATP/NADH arrow indicates consumption/formation. Note that pyruvate (Pyr) oxidation proceeds via two routes to Acetyl-CoA (AcCoA). In the first route on the left side CO2 and NADH are formed as co-products (via the enzyme pyruvate dehydrogenase PDH). In the second route on the right side only formate (For) are formed as co-product (via the enzyme pyruvate formate lyase PFL). Abbreviations used: Glu (glucose), DHAP (dihydroxyacetone phosphate), Gly (glycerol), Pyr (pyruvate), Lac (lactate), For (formate), AcAld (acetaldehyde), AcCoA (acetyl-CoA), Et (ethanol), Ace (acetate)*

Have a close look at the figure and confirm the following observations:

- The overall route to lactate is redox neutral while a mole of ATP is generated
- The overall route to ethanol is redox neutral if the PDH pathway is taken (two moles of NADH generated and two moles of NADH consumed)
- The overall route to ethanol requires a mole of NADH if the PFL route is taken
- The overall route to ethanol generates a mole of ATP irrespective of the pyruvate oxidation route
- Formate cannot form as standalone product since acetyl-CoA will always form in parallel
- The overall route to acetate generates a mole of NADH and two moles of ATP
- Glycerol formation requires a mole of ATP and a mole of NADH

- Acetaldehyde formation is redox neutral if the PFL route is taken while a mole of ATP is generated in the overall route
- Note the lumping of multiple enzymatic steps for glycolysis

The cmol pathway map

From chapter 3 it should be evident that cmols are preferred above mols when working with biomass. The map above can be converted to a cmol map where all rates are expressed in terms of cmols.
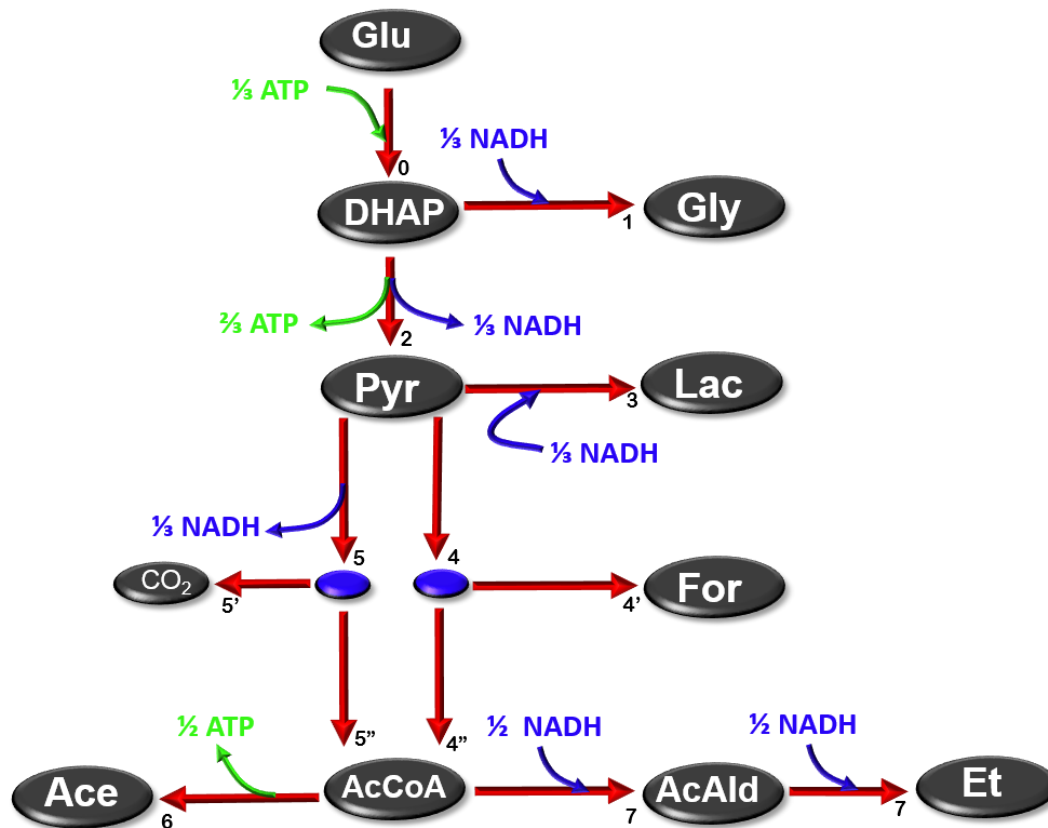
You will see that it looks very similar to the previous graph. The 3 carbon reduction of glucose (0.5xGlucose in the original map) is not required anymore since we are considering only the rates of carbon atoms 'flowing' through the network (with whatever H,O and N is attached to them). But how will we quantify the 'flowing' streams or arrows? What are the units of these arrows referred to as fluxes or rates? The following video addressed this **key concept of the course:**

Video: Lecture 5 Units of metabolic fluxes (https://youtu.be/osKOrHJot-A?
list=PL0EN6xSRgtSStHFlC9TeO8I538CAN2exs)

The H,O and N attachments of the fluxing carbon atoms represents the average from the molecule and not the specific attachments from a specific atom. For example, in lactic acid all carbon atoms will be considered to be attached to two hydrogens and one oxygen, although there is a specific carbon in lactic acid that is not attached to oxygen at all. Try and visualise cmols (carbon atoms with their averaged attachments) 'flowing' through the metabolic network and 'choosing' an outlet stream at the branch points. All carbon atoms originates from glucose and will exit in one of the excretion products.

You will see that all redox (NADH addition or formation) and energy (ATP addition or formation) interactions are expressed relative to one cmol reacting. Take for example the NADH generated form glycolysis (between DHAP and pyruvate). One mol of NADH is formed per C_3 molecule reacting along the pathway, accordingly $\frac{1}{3}$ mol of NADH is formed per cmol. Also note that the averaged attachment to carbon is changing when redox (NADH) interactions occur. From glucose to pyruvate $\frac{1}{3}$ of NADH is formed, or $2 \times \frac{1}{3}$ hydrogens are released from the reacting cmol, thus pyruvate has a cmol formula of $CH_{\frac{4}{3}}O$ compared to the CH_2O of glucose. This is also reflected in the DOR of glucose (4) and pyruvate ($3\frac{1}{3}$). For ATP

generation or usage the cmol formula won't alter, only the energy state of cmol reacting. Take the example of acetic acid formation where one mole of ATP is formed per mol of acetic acid. Given the two carbons in acetic acid we have $\frac{1}{2}$ mol of ATP per cmol acetic acid that forms.

In this course a distinction will be made between normal nodes and **blue nodes**. Blue nodes represent cmol splits dictated by a single reaction with known stoichiometry. Accordingly the split ratios of the node is fully determined. Normal nodes like the DHAP or pyruvate nodes in the graph above are different to the blue nodes since the split is not determined by stoichiometry. For these nodes we only know that the carbon flux (rate) into the node must equal the carbon flux out of the node.

Sometimes there is more than one option when assigning NADH or ATP. A prime example of this is where pyruvate get oxidised via pyruvate dehydrogenase (left option on figure above). In this enzymatic step CO_2 is released while NADH is formed. All of this occurs in a single enzymatic step and a decision have to be made with regards to what stream (or rate) the NADH will be associated with. From a molar perspective it is known that one mole of NADH is formed in the reaction, accordingly one can say that $\frac{1}{3}$ mol of NADH is formed per pyruvate reacting or a $\frac{1}{2}$ mol of NADH is formed per acetyl-CoA forming. Note that the DOR of acetyl-CoA is 4, one only consider the two carbons that remains from pyruvate (it can be represented by $CHO_{\frac{1}{2}}$).
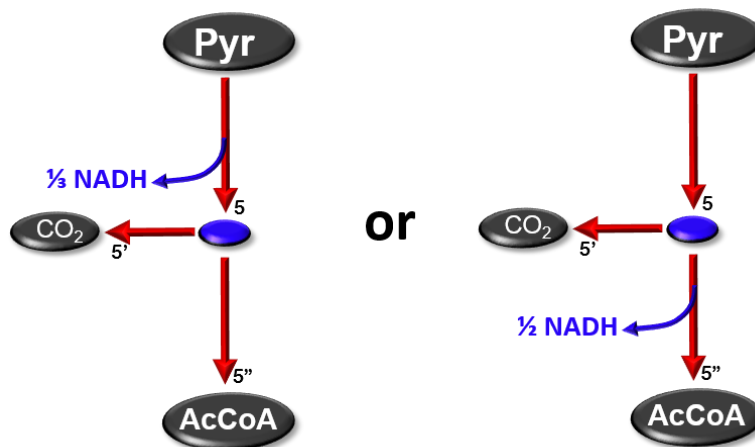
The image above represents a **blue node** with the following rate relationships:

$$r_{5''} = \frac{2}{3}r_5 \quad r_{4''} = \frac{2}{3}r_4 \quad r_{5'} = \frac{1}{3}r_5 \quad r_{4'} = \frac{1}{3}r_4$$

Note that that the cmol rate (r_i) does not change if carbon is not split off or joined. The 'attachments' will change if there is redox (NADH) reactions, but the carbon flux will remain.
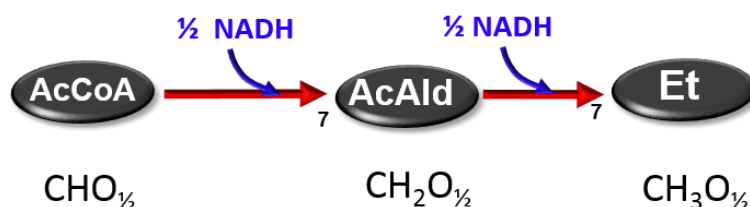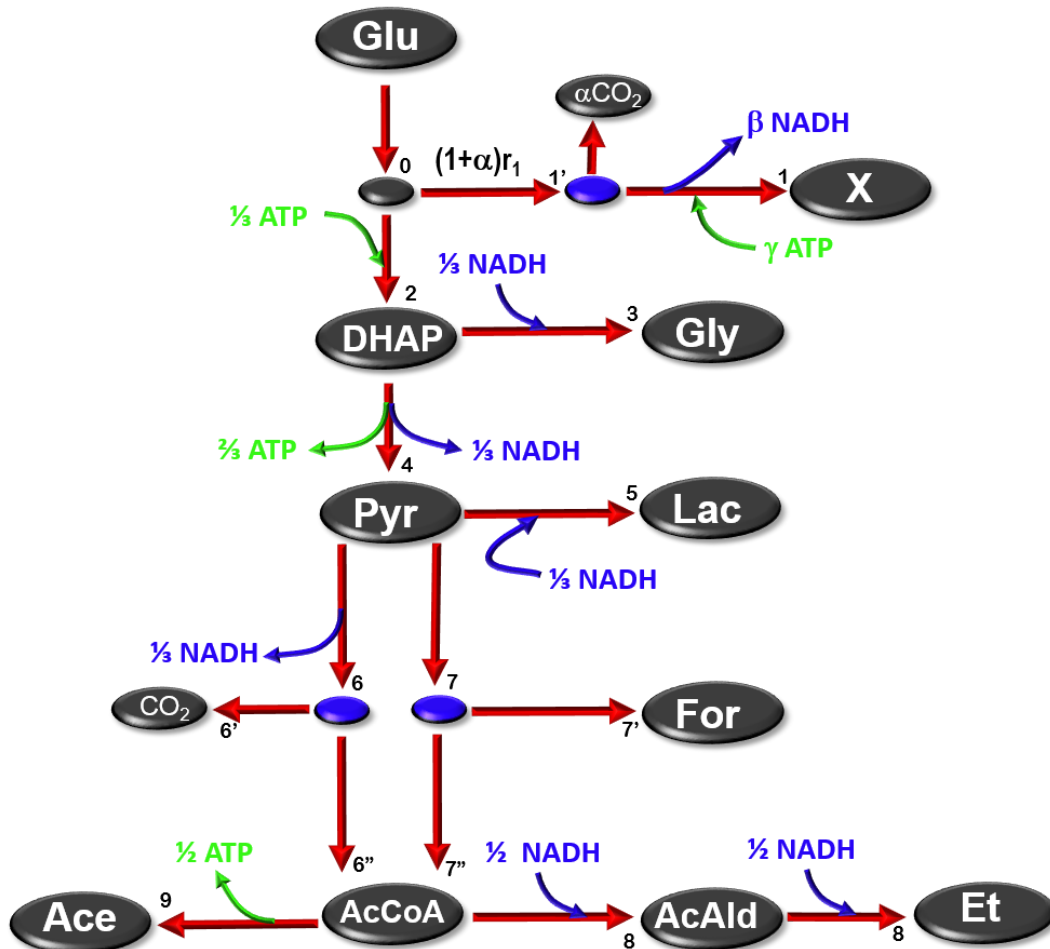

Remember two hydrogens in a NADH, so 1 hydrogen in ½ NADH.

Integrated anabolism and catabolism

We can now combine the catabolism with the anabolism. The overall anabolic reaction was given by:

$$Y_{XS}CH_2O + Y_{XN}NH_3 + \gamma ATP \rightarrow X + \alpha CO_2 + \beta NADH + Y_{XW}H_2O$$
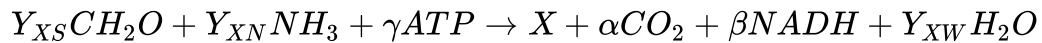
We can add this reaction to the catabolic reactions:

The glucose required for the anabolism is split off right at the start so that r_2 is the glucose rate used for the catabolism. Note that the fluxes was renumbered above. The black node at the anabolism/catabolism split differs from the blue nodes since the split is not governed by the stoichiometry of a single reaction.

The following video might be helpful to cement your understanding:

Video: Lecture 6 The cmol pathway map (https://youtu.be/uqlDHf_UDbw?
list=PL0EN6xSRqtSStHFlC9TeO8I538CAN2exs)

ATP and NADH balances and fundamental physiological rates

The rate of ATP production minus the rate of ATP consumption will be equal to the maintenance requirements of the cell (θ). Using the ATP coefficients of the cmol map (a_i) we can define the following:

$$\sum_{i=0}^{n} a_i r_i = \theta$$

The parameter θ is a physiological property of the microbe that quantifies the mols of ATP spent on maintenance per hour per cmol of biomass:

$$\theta \Rightarrow \frac{mol\ ATP}{cmol\ X \cdot h}$$

Note that the basis of this rate is the same as all the other rates where:

$$r_i \Rightarrow \frac{cmol\ i\ (mol\ i)}{cmol\ X \cdot h}$$

Of particular interest is the formation rate of biomass (μ), another fundamental physiological parameter:

$$r_x = \mu \Rightarrow \frac{cmol\ X}{cmol\ X \cdot h} \Rightarrow \frac{1}{h}$$

The NADH balance states that all produced NADH must be consumed. The NADH coefficients from the cmol map (b_i) can be used so that:

$$\sum_{i=0}^{n} b_i r_i = 0$$

The following video provides more detail on treating ATP and NADH balances:

Video: Lecture 7 ATP balances (https://youtu.be/t4tt28cLm9w?
list=PL0EN6xSRgtSStHFlC9TeO8I538CAN2exs)

Putting it all together

We'll stick to the fermentative pathways but we'll reduce some of the possible fluxes in order to keep the formulation small. Let's consider the following:

Node carbon balances

$$r_0 = (1+\alpha)r_1 + r_2 \qquad (1)$$
$$r_2 = r_3 + r_4 \qquad (2)$$
$$r_4 = r_5 + r_6 \qquad (3)$$
$$r_6 = r_7 + r_8 \qquad (4)$$

Stoicheometric relationship

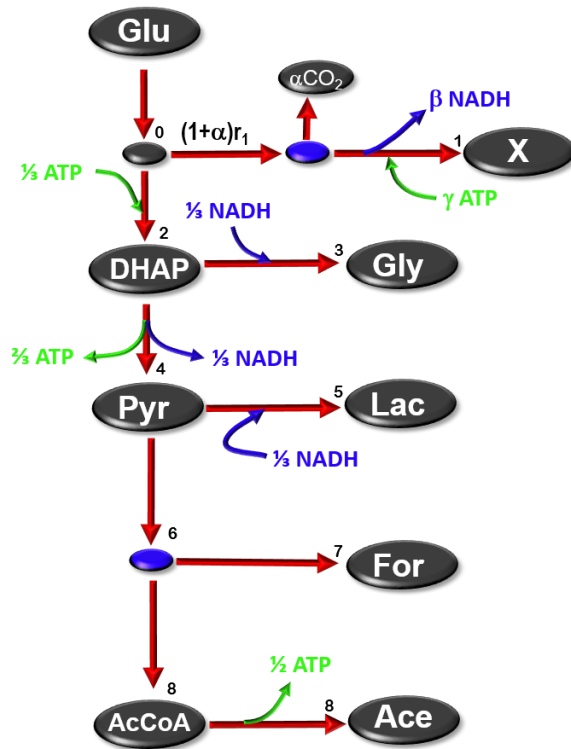$$r_7 = \tfrac{1}{3}\, r_6 \qquad (5)$$

NADH balance

$$\beta r_1 - \tfrac{1}{3}r_3 + \tfrac{1}{3}r_4 - \tfrac{1}{3}r_5 = 0 \qquad (6)$$

ATP balance

$$-\gamma r_1 - \tfrac{1}{3}r_2 + \tfrac{2}{3}r_4 + \tfrac{1}{2}r_8 = 0 \qquad (7)$$

Basis

$$r_1 = \mu \qquad (8)$$

The flux model is given next to the pathway map for ease of reading. The first four equations describe the carbon splits. Equation 5 gives the stoichiometric relationship for formic acid formation. Be careful not to specify an additional relationship (like $r_8 = \frac{2}{3}r_6$). The carbon balance (eq 3) and equation 5 fully specifies the carbon split at the node. An additional equation will be linearly dependant and won't represent an additional fundamental equation. A stoichiometric node that splits into two fluxes can therefore only have 2 equations.

The NADH and ATP balances are a direct application of the formulas given above. Equation 8 is the choice of basis, where r_X (r_1) is used instead of the usual $-r_S$. Also note that the actual rate of biomass formation was specified instead of the arbitrary $-r_S = 1$ assignment used before. For the 9 unknown rates we have 8 equations, so one specification is required.

The formulation is very similar to the set of linear equations generated in chapter 3. We can also employ the matrix formulation to solve this set of linear equations:

$$
\begin{bmatrix}
 & r_0 & r_1 & r_2 & r_3 & r_4 & r_5 & r_6 & r_7 & r_8 \\
eq_1 & -1 & (1+\alpha) & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
eq_2 & 0 & 0 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\
eq_3 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & 0 & 0 \\
eq_4 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 \\
eq_5 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{3} & 1 & 0 \\
eq_6 & 0 & \beta & 0 & -\frac{1}{3} & \frac{1}{3} & -\frac{1}{3} & 0 & 0 & 0 \\
eq_7 & 0 & -\gamma & -\frac{1}{3} & 0 & \frac{2}{3} & 0 & 0 & 0 & \frac{1}{2} \\
eq_8 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
eq_9 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0
\end{bmatrix}
\times
\begin{bmatrix}
r_0 \\ r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \\ r_6 \\ r_7 \\ r_8
\end{bmatrix}
=
\begin{bmatrix}
0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \theta \\ \mu \\ 0
\end{bmatrix}
$$

Note that lactic acid (v_5) was specified to be 0.4 $\frac{cmol\ Lac}{cmol\ Gluc}$. This becomes equation number 9.

We can now solve the system of linear equations in Python. We'll use the average value of $\alpha = 0.1 \frac{cmol\ CO_2}{cmol\ X}$. Given this specification we can solve for $\beta = 0.1$ if we take the generic biomass formula of $CH_{1.8}O_{0.5}N_{0.2}$ (remember Chapter 4.1 (https://nbviewer.jupyter.org/github/willienicol/Biochemical-engineering-notes/blob/master/4%20Stoichiometry%20that%20includes%20internal%20reactions/Modelling%20the%20a

We further define a γ value of 1.8 $\frac{mol\ ATP}{cmol\ X}$ for this anaerobic system.

In [2]:

```
import numpy

beta, alpha, gamma = 0.1, 0.1,1.8
S = numpy.matrix([[-1,(1+alpha),1,0,0,0,0,0,0],
[0,0,1,-1,-1,0,0,0,0],
[0,0,0,0,1,-1,-1,0,0],
[0,0,0,0,0,0,1,-1,-1],
[0,0,0,0,0,0,-1/3,1,0],
[0,beta,0,-1/3,1/3,-1/3,0,0,0],
[0,-gamma,-1/3,0,2/3,0,0,0,0.5],
[0,1,0,0,0,0,0,0,0],
[0,0,0,0,0,1,0,0,0]])

from numpy.linalg import matrix_rank  #use rank function to determine if all equations
 are linearly independant
a=matrix_rank(S)
a
```

Out[2]:

9

A rank less than 9 would have implied that the 9 equations were not independent. For now we are good to go.

All that remains is to use the physiological parameters of the organism and one additional specification. Assume that $\theta = 0.02$ and $\mu = 0.3$ while $r_{LA} = 0$

In [16]:

```
theta, mu, rlac = 0.02,0.3,0.4
C = numpy.matrix([[0,0,0,0,0,0,theta,mu,rlac]]).T
Y = numpy.linalg.solve(S,C)
Y
```

Out[16]:

```
matrix([[ 3.56      ],
        [ 0.3       ],
        [ 3.23      ],
        [ 1.46      ],
        [ 1.77      ],
        [ 0.4       ],
        [ 1.37      ],
        [ 0.45666667],
        [ 0.91333333]])
```

Note that:

$$-r_s = r_0 = 3.56 \frac{cmol\ S}{cmol\ X \cdot h}$$

$$r_X = r_1 = 0.3 \frac{1}{h}$$

$$r_{Gly} = r_3 = 1.46 \frac{cmol\ X}{cmol\ X \cdot h}$$

$$r_{AA} = r_8 = 0.913 \frac{cmol\ X}{cmol\ X \cdot h}$$

We can express the rates as yields with respect to the incoming glucose flux (Y_{Si}) by merely dividing all the actual rates by the glucose uptake rate:

In [30]:
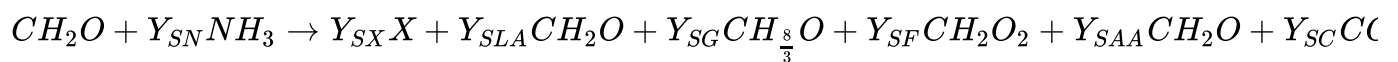
```
Ysi=Y/Y[0]
Ysi
```

Out[30]:

```
matrix([[ 1.        ],
        [ 0.08426966],
        [ 0.90730337],
        [ 0.41011236],
        [ 0.49719101],
        [ 0.11235955],
        [ 0.38483146],
        [ 0.12827715],
        [ 0.25655431]])
```

We can check if these yields are consistent with the mass balances performed in the previous chapter. The overall (only uptake and excretion) equation will be as follow:

$$CH_2O + Y_{SN}NH_3 \rightarrow Y_{SX}X + Y_{SLA}CH_2O + Y_{SG}CH_{\frac{8}{3}}O + Y_{SF}CH_2O_2 + Y_{SAA}CH_2O + Y_{SC}CO$$

Given the 9 components, 4 elemental balances and a glucose basis, 4 specifications will be required (as compared to 1 in the previous example). We'll define biomass, lactic acid, glycerol and formic acid from our previous answer and check whether acetic acid is calculated correctly as **this will confirm mass balance consistency!** Let's set up the matrix:
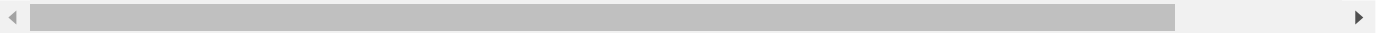
In [33]:

```python
# Order of components: Gluc NH3 X Gly Lac For Ace CO2 H20

S2 = numpy.matrix([[-1,0,1,1,1,1,1,1,0],
[-2,-3,1.8,8/3,2,2,2,0,2],
[-1,0,0.5,1,1,2,1,2,1],
[0,-1,0.2,0,0,0,0,0,0],
[1,0,0,0,0,0,0,0,0],    #glucose basis
[0,0,1,0,0,0,0,0,0],    #biomass
[0,0,0,1,0,0,0,0,0],    #glycerol
[0,0,0,0,1,0,0,0,0],    #lactic
[0,0,0,0,0,1,0,0,0]])#formic

# Specifications

Ysx=Ysi[1]; # from flux model yields
Ysg=Ysi[3]; # from flux model yields
Ysl=Ysi[5]; # from flux model yields
Ysf=Ysi[7]; # from flux model yields

C2 = numpy.matrix([[0,0,0,0,1,Ysx,Ysg,Ysl,Ysf]]).T
Y2 = numpy.linalg.solve(S2,C2)
Y2
```

Out[33]:

```
matrix([[ 1.        ],
        [ 0.01685393],
        [ 0.08426966],
        [ 0.41011236],
        [ 0.11235955],
        [ 0.12827715],
        [ 0.25655431],
        [ 0.00842697],
        [-0.09456929]])
```

Let's get the acetic acid yield from both models:

In [34]:

```python
Ysaa1=Ysi[8,0].tolist()  #from flux model -extracting matric element to list
Ysaa2=Y2[6,0].tolist()   #from mass balance
[Ysaa1, Ysaa2]
```

Out[34]:

```
[0.2565543071161049, 0.25655430711610494]
```

The check is done!

What have you learned? With the flux model 3 less specs was required to obtain the same answer. Do you now understand why we did the flux model?

Back to Contents page (https://nbviewer.jupyter.org/github/willienicol/Biochemical-engineering-notes/blob/master/List%20of%20contents.ipynb)

In [ ]: