

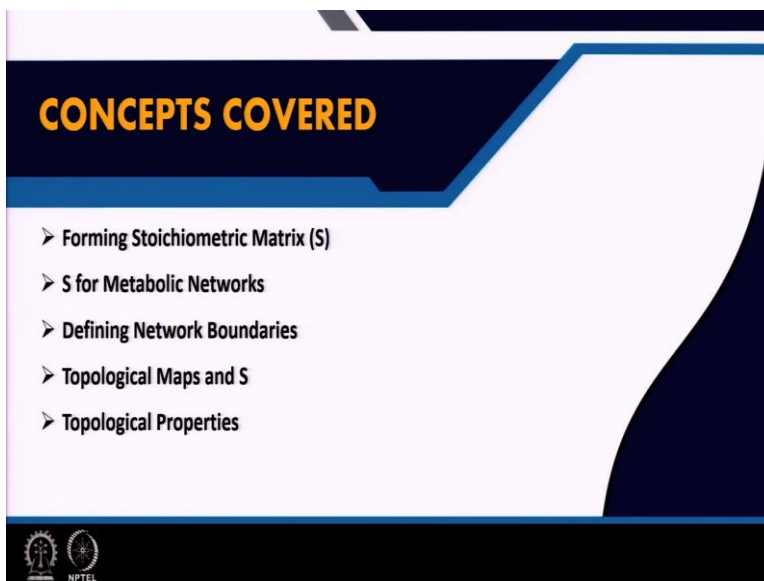
**Metabolic Engineering**  
**Prof. Amit Ghosh**  
**School of Engineering Science and Engineering**  
**Indian Institute of Technology – Kharagpur**

**Lecture – 16**

**The Stoichiometric Matrix: Representing Reconstructed Network Mathematically**

Welcome to metabolic engineering course, today we will talk about this stoichiometric matrix which is very important part in metabolic network which represent the reconstructed and networks mathematically. So, mathematically if you want to build a network that can be represented mathematically and it will have all the information all the reaction which you can actually operate for different purpose.

**(Refer Slide Time: 01:00)**



So, this lecture today it will have following subheading, we will learn about, how to form a stoichiometric matrix and what is S for a metabolic network? Defining network boundaries and topological maps and stoichiometric matrix and different topological properties for S.

**(Refer Slide Time: 01:25)**

**Mathematical Representation of Reconstructed Networks**

Key questions:

- How do we go about taking information from chemical reactions in our network reconstruction and putting it into some kind of formal mathematical framework?
- What are the characteristics (properties) of this mathematical framework?
- What does this mathematical formulation tell us about the state of the biological/chemical network? Insights?
- Does this modeling represent reality?

The slide features a background with various icons related to science and technology, including a gear, a tree, a network diagram, and a chemical flask. A small video inset in the bottom right corner shows a man speaking. The NPTEL logo is visible in the bottom left corner.

So, mathematically representation of reconstructed network we have following questions and how do we go about talking information from chemical reaction in our network reconstruction and putting in into some kind of mathematical framework. So, the entire metabolic network or the network reconstruction we put in into some kind of mathematical framework, what are the characteristics of this mathematical framework that we would like to know when you actually bring all the reaction in a network framework.

And what does these mathematical formulation tell us about the state of the biological or chemical network, what are the insights you get from this network? Mathematically you can represent any network in the form of a stoichiometric matrix and then you can calculate different properties of these network, network properties can be evaluated from the matrix itself because matrix is easy to operate.

Because matrix algebra is very well known, many problems can be addressed when you actually put in a matrix form matrix algebra, matrix optimization and other things we can play around and find many insight you can bring out from the network that is why the network is always represented in the form of a matrix. And today we are going to learn how you can build this matrix that is stoichiometric matrix  $S$ . So, the different properties of  $S$  we will see how? How we can bring out new feature from the matrix itself.

**(Refer Slide Time: 03:10)**

**Metabolic Modeling**

Substrate

Reactions

- $v_1$  Substrate  $\rightarrow$  A
- $v_2$  A  $\rightarrow$  B
- $v_3$  A  $\rightarrow$  C
- $v_4$  B  $\rightarrow$  product1
- $v_5$  C  $\rightarrow$  product2

$$\frac{d[A]}{dt} = v_1 - v_2 - v_3$$

$$\frac{d[B]}{dt} = v_2 - v_4$$

$$\frac{d[C]}{dt} = v_3 - v_5$$

S =	$\begin{matrix} v_1 & v_2 & v_3 & v_4 & v_5 \\ \begin{matrix} A \\ B \\ C \end{matrix} & \begin{pmatrix} 1 & -1 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{pmatrix} \end{matrix}$
-----	--

3x5

So, in a previous class also I gave this example where you have a small network, you can see on the left hand side the substrate suppose this is a small network where some carbon source is entering the network and then I have 2 product; product 1 and 2. And this is a very simple network where you have only 3 metabolites A, B, C the internal metabolites inside the suppose you consider this as a cell and it has extra solar fluxes which is entering.

That is V 1 and some product flux like V 4 and V 5 which is going outside the cell which is forming product 1 and product 2. So, finally we have 5 reactions V 1, V 2 , V 3, V 4 and v 5 and we have internal metabolite A, B, C. So, now we want to write a differential equation that is the time derivative of this concentration  $dA / dt$ ,  $dB / dt$  and  $dC / dt$  how do you write that? I think right now you will be able to write easily.

This is if you write in terms of fluxes V 1 , V 2 , V 3 , V 4 , V 5 then your  $dA / dt$  will become  $dA / dt$  it will be  $v_1 - v_2 - v_3$ . So, the component which is coming in and the component which is going out, so whichever is going out is minus and whichever is coming in it will be plus, so plus  $v_1 - v_2 - v_3$  for example for  $dA / dt$ . So, for  $dA / dt$  you have  $v_1 - v_2 - v_3$  for  $dB / dt$  we have  $v_2 - v_4$  and for  $dC / dt$  we have  $v_3 - v_5$ .

So, these are the time derivative of the concentration that is  $dA / dt$ ,  $dB / dt$ ,  $dC / dt$  and that is a change in metabolite A as a function of time. So, this can be represented this equation can be

represented in a stoichiometric matrix form. So, this stoichiometric matrix for these 3 equations will be written as  $S$ . So,  $S$  will have this is there are 5 reactions, so that is why you have 5 columns this is called column 1 for  $v_1$ ,  $v_2$ ,  $v_3$ ,  $v_4$  and  $v_5$ .

So, we have 5 columns; so column 1 is for  $v_1$ , column 2 for  $v_2$ , column 3 for  $v_3$ , column 4 for  $v_4$  and column 5 for  $v_5$  and the metabolites are in the row. So, metabolites we have 3 metabolites A, B, C that is why I have only 3 rows, so now we can see the metabolite A is actually involved in how many reactions, so metabolite A is involved in that is  $v_1$ ,  $v_2$ ,  $v_3$ . So, that is why I have nonzero term for  $v_1$ ,  $v_2$ ,  $v_3$  whereas  $v_4$ ,  $v_5$  are 0.

Because it does not participate the metabolite A does not take part in the reaction  $v_4$  and  $v_5$ . So that is why this component is 0 and which is  $v_1$  is coming in that is why this +1 and  $v_2$ ,  $v_3$  are going out. That is why this -1, -1. Similarly for metabolite B we have only  $v_2$  and  $v_4$  are actually active for metabolite B. So, we can see that  $v_2$  is +1 and  $v_4$  is -1. And similarly for metabolite C we have  $v_3$  +1 and  $v_4$ ,  $v_5$  -1.

So, in this way you can write the network in mathematically in the form of a stoichiometric matrix  $S$ . And how much ever big the network is you can compress those information into a matrix form suppose in a metabolic network the dimension of the matrix maybe 1000 by 2000 like that. So, here we have only 3 by 5 matrix. So, we have 3 rows and 5 columns but actually in real case in a metabolic network we may have 1000 metabolite and 2000 reactions it become 1000 by 2000 matrix.

And more important thing these matrix is a sparse matrix, so most of the elements will be 0 and few elements will be nonzero. These kinds of matrix what you get from the metabolic network is basically a sparse matrix where most of the elements are 0. Now I will describe how we can actually form these stoichiometric matrix are the chemical equation and this stoichiometric coefficient how you determine this stoichiometric coefficient.

**(Refer Slide Time: 08:13)**

## Stoichiometric Coefficients

Integral numbers  
 Universal biochemical constants  
 Constants: time-invariant

chemical reaction:  $aA + cC \xrightarrow{V_i} eE + hH$

Representation as a column in a matrix:

A	-a
B	0
C	-c
D	0
E	+e
F	0
G	0
H	+h

compounds

So, these stoichiometric coefficients are actually determined by the chemical reaction coefficient. So, if the coefficient for metabolite A is 'a' and the coefficient for metabolite C is 'c' that is the number it is an integer number as 'a' correspond to the stoichiometric coefficient of metabolite A like that we have 'e', 'h' and that is stoichiometric coefficient of metabolite E and H respectively. So that we can for reaction  $V_i$  and this is the reaction which I want to represent in stoichiometric matrix.

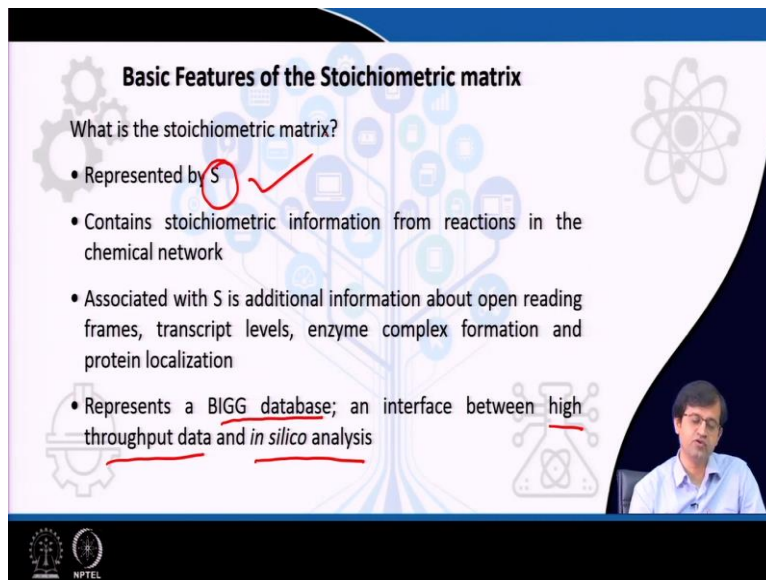
And you can see that since it is in the stoichiometric coefficient for metabolite A is 'a' that is why minus 'a' which is a reactant that is why its minus and then so on for reactant C we will have the coefficient minus 'c' and the product we have plus 'e' and plus 'h'. This way you can represent the reaction in a matrix where the rows are basically the metabolite. So A, B, C, D, E, H are the metabolite these are each row for each metabolite you see how many reaction it participate and also how many reaction you represent it will come under the columns and this way this stoichiometric coefficients are added in the matrix.

**(Refer Slide Time: 09:35)**

### Basic Features of the Stoichiometric matrix

What is the stoichiometric matrix?

- Represented by S ✓
- Contains stoichiometric information from reactions in the chemical network
- Associated with S is additional information about open reading frames, transcript levels, enzyme complex formation and protein localization
- Represents a BIGG database; an interface between high throughput data and in silico analysis



Now the basic feature of stoichiometric matrix is what is this stoichiometric matrix mean? It is generally represented by S. So, you generally represented by stoichiometric by S. And the stoichiometric information they contain stoichiometric information from reaction in the chemical reactor network and the network have this stoichiometric information associated with S is additional information about open reading frames, transcript level, enzyme complex formation and protein localization.

So, S can also have the transcript level and the enzyme complex formation and protein localization information as well. Because S is actually determined based on this information and represent and you can represent them in a BIGG database. As a BIGG database is another database which is actually available online that you can check all the reaction and how this stoichiometric data is stored in the database also you can look at.

An interface between high throughput data and in silico analysis. So, you can integrate different high throughput data and correlate with the matrix. So, this matrix is looks very simple but It is very powerful in the sense that it can integrate the high throughput data in the matrix. And then performing silico analysis. Today is we can see the biologists are actually generating lot of data high throughput data.

And sometime we do not know what to do with the data. So, these high throughput data are generated and it is in huge amount and sometimes we do not know how to use the data. So, using the, you can use the data in these to integrate the data in this stoichiometric matrix and do in silico analysis which can predict new behavior or you can understand the biology better by these data and where this stoichiometric matrix plays an important role.

The center half of these in a network analysis is basically the stoichiometric matrix. So, stoichiometric matrix is given great attention sometime building these stoichiometric matrix can be one PhD, so it is for a given organism suppose you want to construct these stoichiometric matrix. So, this stoichiometric matrix are readily available for many microorganism for human cell lines are also available.

We have Recon 1 Recon 2 Recon 3 for 3<sup>rd</sup> generation network is also available where you the entire metabolic network of the human cell is also available where the S is available and the construction of S is very laborious in case of a multicellular organism where you have many compartments for microbial cell it is quite easy. Since there are no compartments but as you go to higher organism you will see that we have many compartments inside the cell that has to be taken care properly when you build the network.

**(Refer Slide Time: 12:33)**

**Basic Features of the Stoichiometric matrix**

- Formed from the stoichiometric coefficients of the reactions that comprise a reaction network
- Entries are integers
  - Columns of S correspond to reactions
  - Rows of S correspond to a compound. By looking across the rows, one observes all the reactions a given compound participates in, and how the reactions are interconnected
- S transforms the flux vector (reaction rates) into a new vector that contains the time derivatives of the concentrations; therefore S is a linear transformation of the flux vector

$\frac{dC}{dt} = S \cdot v$

NPTEL

Other basic features of the stoichiometric matrix that is the stoichiometric coefficient that comprise the reaction network, this stoichiometric coefficient the matrix have that comprise the reaction network entries are integers you can see the column of S is represent to reaction and the rows of S corresponds to the compound. So, we have the as I told earlier also the columns always correspond to reaction and the rows always correspond to compound by looking across the rows one observe all the reaction a given compound participant.

So, if you just consider the row you will see that how many reaction a given compound participates and how the reaction are interconnected. So, just by looking into the stoichiometric matrix you know how the networks connection how they, what the basic properties of the network you can understand. For example given a compound you know how many reactions it participates that is how many reaction it is involved.

The S transforms the flux vector into a new vector that contains the time derivative of the concentration and therefore S is a linear transformation of the flux vector. So, in the previous the small network was also I have shown that that the concentration vector that is  $dC / dt = S \cdot v$ . So, this time derivative of the concentration is actually proportional or transformed the time derivative or transformed to contain the time derivative the S actually transform the flux vector.

So, this is the flux vector v, so the v flux vector is actually transform, the S matrix actually transforms the flux vector into time derivative of the concentration and this is a linear transformation. So, S is a linear transformation of the flux vector that is what the S transform the flux vector into time derivative of the concentration and it is a linear transformation.

**(Refer Slide Time: 14:55)**



### All Elements Must be Balanced During a Chemical Reaction


All elements have to balance during a chemical conversion (i.e. the number of C, H, O, etc. has to be equal on both sides of the reaction equation). The elemental balance of a stoichiometric reaction vector can be checked using the elemental matrix, E.

Example: Biochemistry textbook definition of glucokinase:

$$\text{GLC} + \text{ATP} \xrightarrow{v_1} \text{G6P} + \text{ADP}$$

	GLC	ATP	G6P	ADP
C	6	10	6	10
H	12	13	11	13
O	6	13	9	10
P	0	3	1	2
N	0	5	0	5

$$\mathbf{S} = \begin{matrix} v_1 \\ \text{GLC} \\ \text{ATP} \\ \text{G6P} \\ \text{ADP} \end{matrix} \begin{matrix} -1 \\ -1 \\ 1 \\ 1 \end{matrix}$$



So, now another important thing there the reaction you import from various database like KEGG, MetaCyc, Brenda those reaction may not be mass balanced or charge balanced all the elements have to be balanced during chemical conversion and the number of C, H, O has to be equal on both sides of the reaction you have seen that how in the school level also you have balanced the reaction like the number of atoms on the left hand side should be equal to the number of atoms in the right hand side.

But for a metabolic network we have about 2000, 3000 reactions manually correcting the elements is actually laborious and it may have a lot of error also. So, here in this class we will know how mathematically using matrix algebra you would be able to correct the reaction mass balance, so the mass balance, so this is the scheme which is used for balancing the reaction where you can balance the carbon atoms, hydrogen atom, oxygen atom.

And you can balance in both sides both left hand right hand sides you can balance for a given reaction. For example here you can see the left hand side has glucose and ATP. And the product we have glucose 6 phosphate and ADP. And the reaction is  $v_1$ . So, for this simple 1 reaction you can make a stoichiometric matrix, since we have only one reaction that is why we have only one column and we have 4 metabolites that is why you have 4 rows. So, it is a 4 by 1 metrics.

And the reactants are -1 and the product is +1. So, minus for glucose, minus for ATP and product glucose 6 phosphate is +1 and ADP is +1 very simple. And once you construct the stoichiometric matrix then the next step is to construct the elemental matrix. So, the elemental matrix which I have defined here the elements balance of the stoichiometric reaction vector can be checked using elemental matrix.

So, let us construct the elemental matrix what is elemental matrix? Elemental matrix is again the rows are actually the atom. So, I represent the rows are carbon, hydrogen, oxygen, phosphorus and nitrogen. And the columns are the metabolite. So, for a glucose molecule how many carbons are there? 6, how many hydrogen's are there? 12. And how many oxygen's are there? 6 remaining atoms like phosphorus and nitrogen or 0.

So, in this way and give the number of in each of the rows and then for ATP and glucose 6 phosphate and ADP similarly you write the number of elements present in the molecule that is the number of carbon, number of hydrogen, number of oxygen, number of phosphorus, number of nitrogen. So, this we write down in a column. And your construct will be elemental matrix. And this elemental matrix and stoichiometric matrix when you make and then you take the dot product.

**(Refer Slide Time: 18:07)**

**Elemental Balancing During Chemical Reactions**

The stoichiometric reaction vector must be orthogonal to all the rows in E.

Example continued:

$$ES = \begin{pmatrix} 6 & 10 & 6 & 10 \\ 12 & 13 & 11 & 13 \\ 6 & 13 & 9 & 10 \\ 0 & 3 & 1 & 2 \\ 0 & 5 & 0 & 5 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \begin{matrix} \text{C} \\ \text{H} \\ \text{O} \\ \text{P} \\ \text{N} \end{matrix}$$

The H<sup>+</sup> row is not balanced → a proton is missing on the right hand side:

$$GLC + ATP \longrightarrow G6P + ADP + H^+$$

$\checkmark E \cdot S = 0$

514 411

So, the dot product  $E \cdot S = 0$  then you say that the reaction is mass balanced. So, if  $E \cdot S$  is not 0 then there is a problem. So, here you can see for the same reaction you can see glucose plus ADP plus glucose 6 phosphates and ADP we saw after taking it when you multiply and the matrix E and S you should know how to multiply this matrix multiplication you should know to perform this calculation. So, matrix are multiplied A and B you can multiply the matrix this is the matrix E has the dimension of 5/4.

And this one this stoichiometric matrix has a dimension of 4/1. So that is why you can multiply these 2 matrix and you get another 4/1 matrix 4/1 while the 4/1 matrix is the all the elements are not 0. So, one of the elements you can see that this is not 0 and since it is not 0 that is why the reaction is not balanced the element mass balance is not there. So, to make this reaction mass balance since this is a minus 1, it is the dot product is minus 1 that is there is extra hydrogen in the left hand side.

In the left hand side we have more hydrogen so that is why I have added one proton on the right hand side to balance the reaction. So, if I add one more hydrogen at the product side then this reaction is balanced. So, by taking a dot product  $E \cdot S$  you can actually able to find the balance reaction that is the mass balance you can do so that the reactions are elementary balance and this you can do it by running a small code or you can run a code to actually able to find how much elements are on the left hand side and how many of them are on the right hand side.

And if there is any imbalance then you can add accordingly to balance the reaction. So, elemental balance is required when you form the metabolic network.

**(Refer Slide Time: 20:35)**

### Charge Balance During Chemical Reactions

Similar to the elemental balancing, electric charge must also be conserved:

**ES = 0**  
E: contains electrical charges

Example:  
Superoxide dismutase reaction

$$2 \text{O}_2^- + 2 \text{H}^+ \rightarrow \text{H}_2\text{O}_2 + \text{O}_2$$

$\text{O}_2$	$\text{H}$	$\text{H}_2\text{O}_2$	$\text{O}_2$
$E = e^-$	$-1$	$1$	$0$
$0$	$1$	$0$	$0$

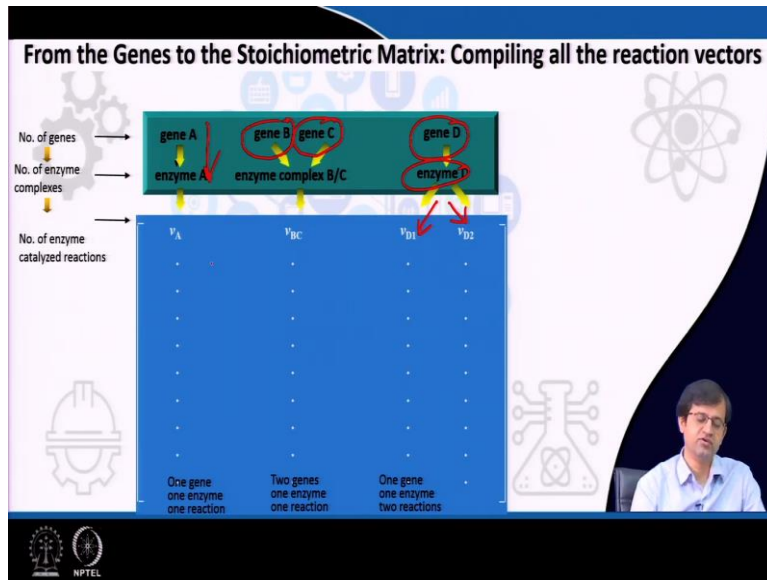
$S = \begin{bmatrix} -2 \\ -2 \\ 1 \\ 1 \end{bmatrix}$

$$\begin{bmatrix} -1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} -2 \\ -2 \\ 1 \\ 1 \end{bmatrix} = 0$$

Similarly charge balance also required charge balance also you can construct these stoichiometric matrix for this reaction this peroxide dismutase is a reaction while you have the oxygen combining with proton giving rise hydrogen peroxide plus oxygen. So, this reaction you can construct the stoichiometric matrix which is very easy where it is just -2 , -2, +1, +1 and elemental matrix for the charge also you can make by the while you can see the oxygen has -1 charge and then the higher proton has +1 charge that is why it has 1 and hydrogen peroxide and oxygen has 0, 0.

So, you should take a dot product E.S then you see that the dot product is 0 that means it is charged balance. So, it do not have to do anything, so in this way you can balance the charge and you can balance the mass make sure all the reactions are actually mass balance charge balance and then you can apply laws of physics another other mass conservation formula or many other techniques can be applied provided you have elemental and charge balance. This way you can make the network accurate and which is also you can publish.

**(Refer Slide Time: 22:00)**



This stoichiometric matrix for metabolic reaction network. That is from the genes to the now you have to add the gene. So, for reaction the reactions stoichiometric you have added but on top of that we have the gene information. So, in the metabolic network the very important part that they include the gene information, so every reaction is mapped back to a gene. So, we have this stoichiometric matrix which has many reaction  $v_A$ ,  $v_{BC}$ ,  $v_{D1}$ ,  $v_{D2}$ .

And each of the columns that is each column is a reaction and each reaction is mapped back we have another mapping table. So, the matrix may not have the mapping but we have a parallel file which actually take care of the mapping and it says that the reaction  $v_A$  is actually correspond to enzyme A and then it is corresponding to a gene. So, this way you can have the mapping for each reaction.

So, this is a one to one, one gene one reaction one enzyme mapping which is a very simple, gene A give rise to enzyme A and it is enzyme A catalyzing reaction  $v_A$ . So, this one gene one enzyme one reaction is very simple but in biology this is sometimes become very complicated in the sense like you can have 2 genes one enzyme one reaction. So, this is a case where there are 2 gene; gene B and gene C which from A complex protein complex.

And we have basically gene B and gene C give rise to protein which are the subunits of a protein and they catalyze reaction BC. So, in this way you can see that 2 genes that are giving rise to one

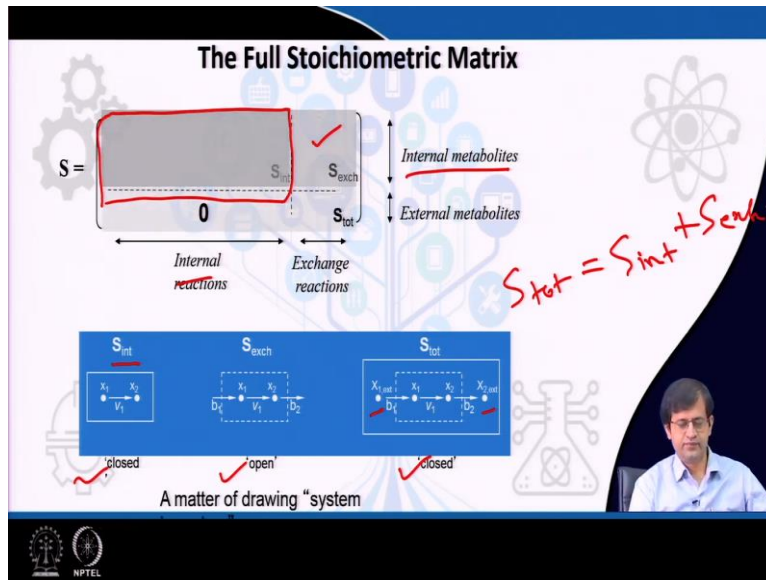
enzyme one protein and then one reaction and there are cases like we have one gene one enzyme but it is catalyzing 2 reactions. So, this is also a possibility where you have one gene one enzyme and 2 reactions.

So, this then the mapping you have to do where based on the annotation on the enzyme gene annotation and try to figure out how the and this is also known as GPR relationship, last class also I told that gene protein reaction association GPR relations are also required for each and every reaction you have to map to gene. And that is from gene to stoichiometric matrix and compiling all the reaction vector. So, you have to compile all the reaction vectors so that each reaction is connected to some gene so this way the network is complex.

So, then the next step is defining the network boundaries this is another important thing you should be able to decide that the network boundaries in the sense that each reaction has a flux and flux has a value and absolute value it ranges from minus infinity to plus infinity. If you do not know about the flux value then you put a boundary which goes from minus infinity to plus infinity. Now in metabolic network it is assume that when the flux value is negative then it is in going in the reverse direction and if it is positive it is going in the forward direction.

So, this is a notation used in metabolic network, so if a reaction is forward going and reverse going you still represent the same reaction. So, we do not represent 2 reaction for forward and reverse. So, this is the notation using the metabolic network where this stoichiometric matrix is defined the single reaction can be forward going and reverse going depending on the flux value. And that is if the flux value is negative then it is going in the reverse direction if it is a positive value it is going in the forward direction.

**(Refer Slide Time: 25:42)**



So, this is the convey notation used in the metabolic network where the reactivity depend on the value of the flux. The flux is stoichiometric matrix you can see it can be a closed system or open system or a closed system. So, the internal metabolites suppose you construct a stoichiometric matrix. So, there is stoichiometric matrix in this quadrant you see that it has only internal metabolites, so this internal metabolite and internal reaction so this box correspond to your all internal reaction.

So, the  $S_{int}$  so all inter solar reaction which are inter cellular metabolite, inter cellular reactions that correspond to the  $S_{int}$  and then you have the  $S_{exch}$ .  $S_{exch}$  is the stoichiometric matrix which involve the exchange reaction and also the internal metabolite. So, the exchange reactions are actually the reaction which involved in importing or exporting some metabolite for example glucose is important inside the cell and ethanol is exported out or carbon dioxide is exported out.

So, this all exchange equation is on clapped in this region where you can say that it is exchange stoichiometric for matrix for the exchange reaction. And when you consider a closed system that is  $S_{int}$ , so  $S_{int}$  is becoming closed network where nothing is going out and nothing is coming in. So that is the internal reaction, so all the reactions are internal and the metabolites are also internal and  $S_{exch}$  is last one to this region of the matrix where you can see that the metabolites are actually imported and exported.

And it is an open system because we do not know the concentration of metabolite which is coming out from outside. So, they will also do not know the concentration of the metabolite which is going out that is how much is formed outside. So, this is an open system where open systems are actually sometimes defined in this way where you do not know the concentration of the flux  $v_1$  and  $v_2$ .

Similarly for you can make this system as a closed system by giving an approximate concentration like  $X_1$  and  $X_2$ .  $X_1$  external and  $X_2$  external and then you can become a closed system and here are how you can apply laws of physics laws of chemistry that is why most of the metabolite network we make a closed system so that nothing is going in nothing is coming out suppose glucose is entering than we assume there is a defined concentration outside.

And if ethanol is what is outside you define a concentration outside. So, this way you can define  $X_1$  external and  $X_2$  external and make it closed system, the closed system networks are also you define the stoichiometric matrix for the  $S$  total which involve  $S$  exchange and  $S$  internal. So  $S$  total is basically  $S$  internal and  $S$  exchange.

**(Refer Slide Time: 28:58)**

**Partitioning of the Flux Vector into Internal and External (Exchange) Fluxes**

External fluxes are those fluxes that flow across the cellular boundary.

These are denoted by  $b_j$ . These fluxes are often accessible to measurement or can be estimated based on experimental data.

The sign convention adopted for these fluxes is that they are positive if mass is flowing out of the cell.

Internal fluxes are those that take place within the cell (within our system boundary).

These fluxes are hard to measure, but often we will know their maximum value.

The slide features a background with scientific icons like a brain, a gear, and a molecular structure. A small video inset in the bottom right shows a man speaking. The NPTEL logo is visible in the bottom left corner.

So, the partitioning of the flux vector into internal and external flux are the external fluxes are those fluxes that flows across the cell boundary as I told the external fluxes are those fluxes that



flows across the cell boundary. So, it is coming from outside to inside or inside to outside those are external fluxes, those are denoted by  $b$ , the external fluxes are denoted by  $b$  and these fluxes are often accessible to measurement.

So, you can measure these fluxes and can estimate based on the experimental data. So, the external fluxes you can measure just by measuring the concentration at 2 different time points in HPLC. So, the internal fluxes you cannot measure by external fluxes you can measure externally using a HPLC, the sign convention adopted for this flux is that they are positive the masses are flowing out of the cell.

So, this is a very important notation that is used by the metabolic network whenever the flux is going outside the cell suppose the ethanol is going outside the cell then the value of the flux is actually always positive. And if it is negative that means it is coming inside, so glucose is glucose flux which is exchange flux is always negative that is because it is coming from outside. So, this sign convention is used in the metabolic network whenever the flux is going out and that is positive and whenever the fluxes is coming inside the cell then it is negative.

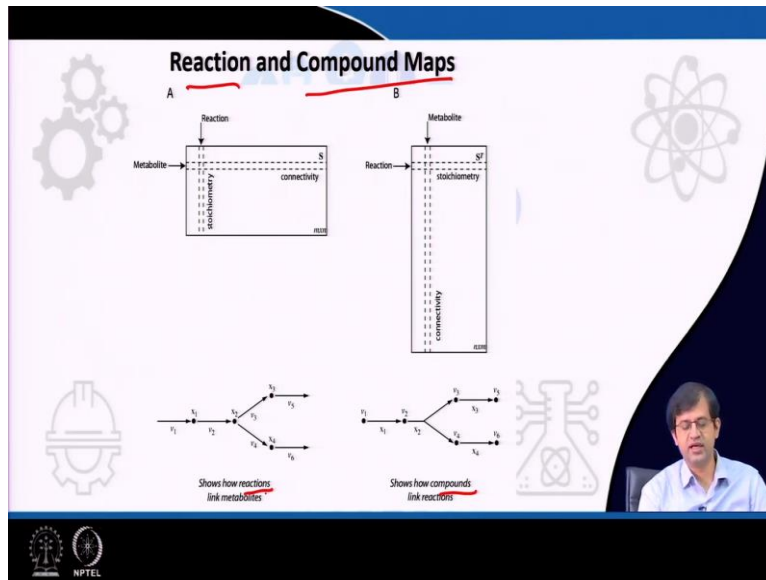
And internal fluxes are those that take place within the cell and the internal fluxes can be positive or negative depending whether it is forward or reverse going. So, these are the sign convention using metabolic network whenever exchange flux is positive we say that the flux is going outside the cell and whenever the exchange flux is negative that means that flux is entering the cell but if the flux of the internal reactions are positive and negative then we say that they that reaction is going if it is negative then it is reverse going and if it is positive is a forward going

So, these are the sign convention using the metabolic network. These fluxes are hard to measure, the internal fluxes are hard to measure but often you know their maximum value from kinetics. The maximum value of the flux you can get from kinetics but most of the time you cannot measure these fluxes the internal fluxes. That is why you need to do know the metabolic network and then you can do modeling and get the flux value.

And because the internal fluxes are actually if you are not able to measure the internal fluxes then the whole systems biology valid is play a crucial role to understand how the fluxes are flowing.

The network maps and this stoichiometric matrix, the 4th part of the stoichiometric matrix and is how you draw the network maps.

**(Refer Slide Time: 31:50)**



The network maps are of 2 types that is the reaction maps and compound maps. So, what are reaction map and what are compound maps? The reaction maps are basically the network where the nodes are actually metabolite and links are the reaction. So, we are here you can see the reactions  $v_1, v_2, v_3, v_4, v_5, v_6$  are the links and the nodes are the metabolite  $x_1, x_2, x_3, x_4, x_6$ . So, this way you can define the network. Similarly you can have a compound map, in the compound map it is just the reverse where you can see that the nodes are the reaction and the links are the metabolite.

So, here you can see  $v_1, v_2, v_3, v_4, v_5, v_6$  are actually the reaction whereas in the reaction map the reaction map we have the nodes are the metabolite. And how do you convert from a reaction map to metabolite map we just take it  $S$  transpose to now you take  $S$  transpose then using the  $S$  transpose if you draw the network map then it will become a compound map. So, there are 2 kinds of network reaction map and the compound map.



the compound map. So, you do little bit practice you close this part and try to draw the map on your own on a piece of paper and just do it as exercise today. And similarly once you have the map for the reaction map then you can draw the compound map but do not see the reaction map you draw on your own do some practice for your exam.

**(Refer Slide Time: 35:40)**

**Connectivity properties of the stoichiometric matrix**

Reactions (j) →

For every metabolite

Metabolites (i) ↓

$$\begin{bmatrix} S_{11} & \dots & S_{1n} \\ \vdots & \ddots & \vdots \\ S_{m1} & \dots & S_{mn} \end{bmatrix}$$

$$S_{ij}$$

$$J_i = \sum_j |S_{ij}|$$

$$I_j = \sum_i |S_{ij}|$$

$$\bar{S}_{ij} = 0 \text{ if } S_{ij} = 0$$

$$\bar{S}_{ij} = 1 \text{ if } S_{ij} \neq 0$$

$$J_i = \text{the number of reactions in which a metabolite participates (metabolite connectivity)}$$

$$I_j = \text{the number of metabolites that participate in a reaction}$$

Reaction network diagram showing metabolites: ANS1, POR, ACIBS, ACLS, ALATA, L, DHDPS, PPDK.

Some topological properties of the stoichiometric matrix. These are the topology the connectivity property of this stoichiometric matrix. So, given a metabolite you can sum all the elements you can sum all the element by taking mods, modulus of the stoichiometric coefficient. So, this is a formula the number of metabolite that participate in a reaction or the number of reaction in which a metabolite participate.

So,  $J_i$  gives you the number of reaction in which metabolite participate that is a connectivity for piruvate you can see that there are 2 reaction for production and 5 reaction for consumption. So, the piruvate is formed from 2 reaction that is ANS1, POR it is produced and it is consumed by another 5 reactions. So, this  $J$  you can calculate for each row  $j$  for each metabolite 1, 2, 3 and  $n$  number of metabolites are there.

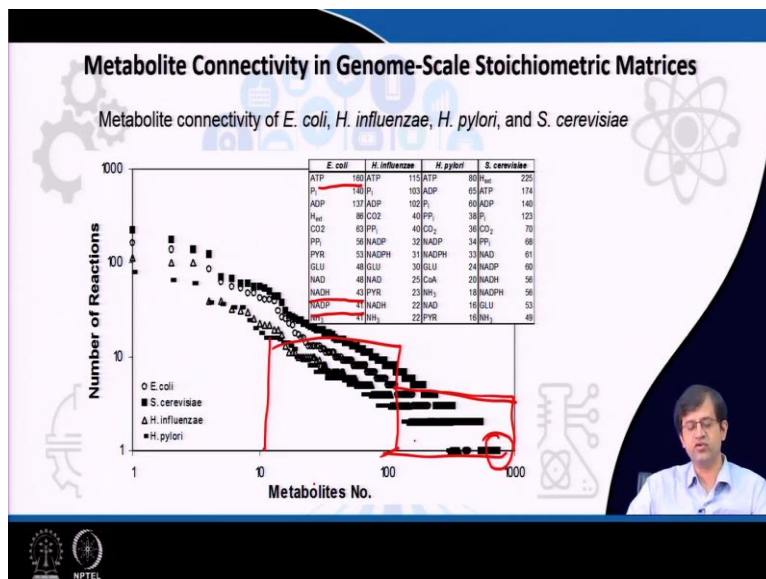
And you calculate and then you know that how many the number of reaction in which a metabolite participant this way can calculate how many reaction a given metabolite participant you can calculate by calculating the  $J_i$  and that is the taking the modules for each coefficient for

sum all the rows taking modules do not take this sign and then you will get the number of reactions a given metabolite participate, similarly for the reaction also you can do the same thing in number of metabolites that participate in a reaction.

So, given a reaction how many metabolites are involved in a reaction that also you can calculate by summing the modulus of the coefficient. So, the modulus of the coefficient you can sum it up along the column then you will get how many metabolites it is involved in a given reaction. This way you can get the connectivity of the network how many links are there for a given node that even evaluate by using this by calculating  $J$  and  $i$ .

So,  $j$  is for the metabolite that is  $i$  goes from 1 to  $m$  and  $j$  goes from 1 to  $n$  that is the number of reactions for every reaction you can calculate how many metabolites are involved and also for every metabolite you can calculate how many reactions it is participating. So, this connectivity property you can evaluate for each and every metabolite and each and every reaction and which will give an idea about the connectivity of the network.

**(Refer Slide Time: 38:35)**



The metabolite connectivity in genome scale stoichiometric matrix is given over here you can see that given a metabolite how many reactions it participant, for example *E. coli* the ATP is involving 160 reaction then we have NADH is involved in 43 action, NADP is involving 41

reaction and it is changing with different model it is not constant interestingly you can see that E. coli, ATP involvement is 160 reaction and then in H. influenzae it is 115 is H. pylori it is 80.

And when it comes to *saccharomyces cerevisiae* the ATP involvement is around 174 was much larger. So, higher the network and the higher the complexity you will have the ATP involvement is more that means many reaction it is involved. And this the number of metabolite and the number of reaction and the number of metabolite also you can see for different organism I have plotted or E. coli, *saccharomyces cerevisiae*, *influenzae* and *pylori*.

Where you can see that the number of metabolites is more in case of *saccharomyces cerevisiae* and the number of reaction, so some metabolite have many reactions as well as some of the metabolite as you can see 1 to 10 where they are involved in 100 of reaction almost close to 100. So, here for some metabolite from 10 to 100 metabolite you can see it is involving 10 reaction, this part of the way you can see that these are the metabolite which are involved in almost close to 10 reaction.


And there are 100 to 1000 metabolite which are involved in less than 5 reactions, less than 5 reactions you can see some reactions which are involved in only in one reaction some metabolites which are involved with so these are the metabolites which are involved in only one reaction very few. So, you can see that the network is very highly connected in the sense that most of the metabolites are connected to many reactions.

So, these are highly connected nodes so these metabolites 1 to 10 are very highly connected they have almost close to 100 connection and from 10 to 100 you can see that it has an average connection of around 10 and from 100 to 1000 you see and they are having around 4 to 5 connections. So, this is the network connectivity you can evaluate across organism and dependent across you can know how the network is changing the connectivity is changing in detail.

**(Refer Slide Time: 41:17)**

## CONCLUSION

- Stoichiometric matrix is derived from annotated genomes given knowledge of enzyme stoichiometries and is a mathematically compact description of metabolic maps
- The chemical elements, ionic charge, and biochemical moieties must be balanced in the stoichiometric matrix
- The stoichiometric matrix is 'sparse', i.e. few non-zero elements



In conclusion this stoichiometric is derived from the annotated genome and that is the knowledge of the enzyme, the biochemistry and mathematically you can represent the network in the form of stoichiometric matrix and you can then draw the metabolic maps. Today we learn how to draw a metabolic map from the stoichiometric matrix. The stoichiometric matrix actually gives you the biochemistry of the cell.

And the annotated genome and is stored in the form of a matrix in a mathematical way and that you can be used to draw the metabolic map, the chemical elements ionic charged biochemical moiety must be balanced in this stoichiometric matrix. So, in this stoichiometric matrix you have to make sure that you are balancing the elements, charge and different moieties and the stoichiometric matrix which you will learn today is basically a sparse matrix.

That is it has few nonzero elements. So, most of the elements are 0 only few elements are nonzero. So, today we learned a very important topic that is how to construct this stoichiometric matrix and how you can draw the metabolic map from these stoichiometric matrix.

**(Refer Slide Time: 42:32)**

## REFERENCES

- Ouzounis, C.A. and P.D. Karp, Global properties of the metabolic map of Escherichia coli. *Genome Research*, 2000. **10**(4): p. 568-76.
- Edwards, J.S. and B.O. Palsson, The Escherichia coli MG1655 in silico metabolic genotype: Its definition, characteristics, and capabilities. *Proceedings of the National Academy of Sciences*, 2000. **97**(10): p. 5528-5533.
- Edwards, J.S. and B.O. Palsson, Systems properties of the Haemophilus influenzae Rd metabolic genotype. *Journal of Biological Chemistry*, 1999. **274**(25): p. 17410-6.
- Schuster, S., T. Höfer, Determining all extreme semi-positive conservation relations in chemical reaction networks: a test criterion for conservativity. *J. Chem. Soc. Faraday Trans.* 1991. **87**: p. 2561-2566.
- Meyer, C. D. *Matrix analysis and applied linear algebra* (Society for Industrial and Applied Mathematics, Philadelphia, 2000).
- Strang, G. *Linear Algebra and its Applications* (Saunders College Publishing, Fort Worth, 1988).
- Jeong, H., et al., The large-scale organization of metabolic networks. *Nature*, 2000. 407(6804): p. 651-654.
- Schilling, C.H. and B.O. Palsson, The underlying pathway structure of biochemical reaction networks. *Proceedings of the National Academy of Sciences of the United States of America*, 1998. 95(8): p. 4173-4178.



So, these are the references which you can look at and read more about this topic and thank you.