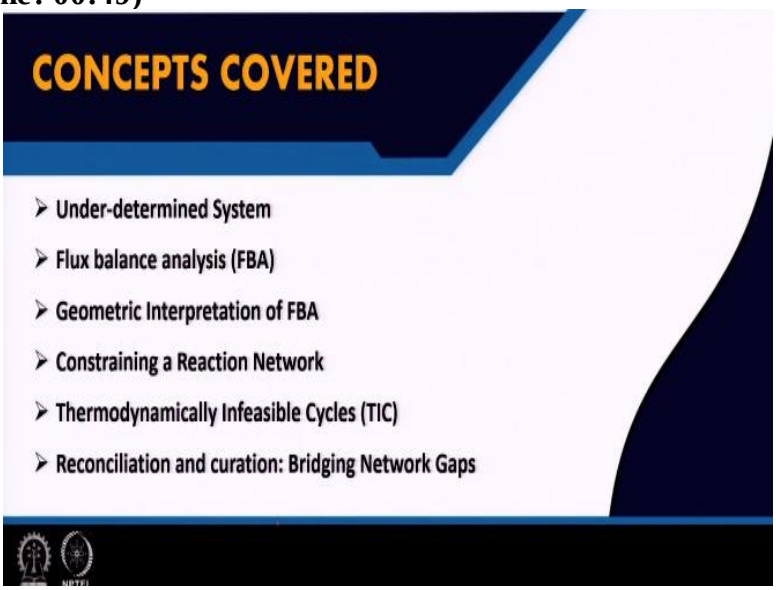
Metabolic Engineering
Prof. Amit Ghosh
School of Energy Science and Engineering
Indian Institute of Technology – Kharagpur

Lecture – 17
Flux Balance Analysis (FBA)

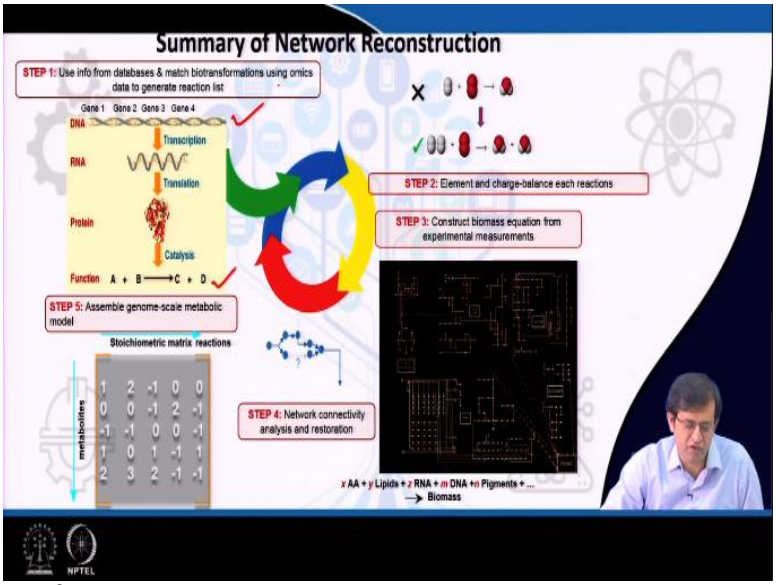
Welcome to metabolic engineering course, so today the topic is flux balance analysis, flux balance analysis in very important concept which involve metabolic network. So, this is the optimization scheme through which you get the fluxes.

(Refer Slide Time: 00:49)



So, we will just go through the concept which we will cover today, we will discuss about under-determined system and then followed by flux balance analysis and then geometric interpretation of FBA and how you can constraint the reaction network? And then the thermodynamically infeasible cycle that exists in the metabolic network and how do you reconciliation and curate the network that is bridging the gap and that will learn from this class.

(Refer Slide Time: 1:22)



So, the summary of the reconstruction process which you already know, so I will just summarize that what you have to do for constructing the metabolic network, you have to start with the genome. So, we can see that genome the entire genome of the organism you can take and then annotate the genome based on the open reading frame, then we have gene 1 gene 2 gene 3 and so on. So, this give rise to protein and then protein give rise to reaction, so from DNA to reaction.

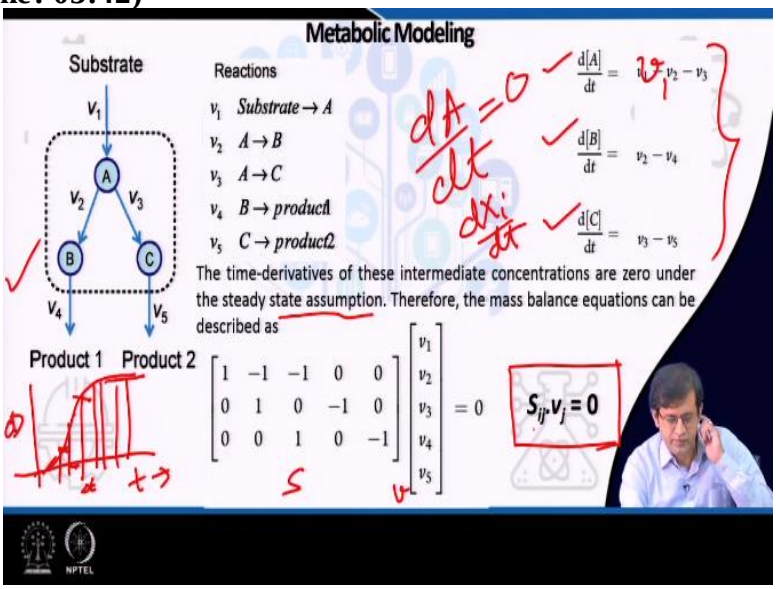
So that mapping have to do that is the first step. And then what do you have to do? You have to balance the charge and element that we learned in the last class where you can do charge balance and the mass balance because the reaction that you take up from the database may not be charged balance or mass balance that is the step number 2. And then in step 3 you construct the biomass equation.

The biomass equation which I discussed last time is basically the cell component, the cell is made up of amino acid, lipid, RNA, DNA, pigments and that gives to the total weight of the cell or that become the biomass equation because each of the component is actually in synthesize inside the cell and as the component as we get more flux through biomass equation, then the cell grows more.

So, this is the important reaction that is the only action which we have to measured experimentally in the metabolic modelling, otherwise all the reaction you can actually get it from the literature or biochemistry by the biomass equation if it is not known then you have to measure experimentally. And then the next step is the network connectivity and all the

nodes are connected and there is no orphan in the network. Then get the stoichiometric matrix which stoichiometric is a mathematical representation of the network, where all the reaction or metabolites are actually compressed in a form of a matrix.

(Refer Slide Time: 03:42)



This matrix you can actually do optimization and so on. So, in the last class, we have done the metabolic modelling of this small network, this one over in the left hand side and in this side only 5 reaction and 3 metabolites. So, I told you to draw the equation for the time derivative of the concentration A, B, C and that you will get the once you write the concentration what is the time derivative of the concentration A it will be v1 minus so already there know, so v1 – v2 - v3.

Then time derivative of concentration B is basically v2 - v4 and then time derivative of concentration C is v3 - v5, this 3 equation you can actually write in front of a stoichiometric matrix. So, this matrix is basically the S matrix that is 1 - 1 - 1 0 0 0 1 0 – 1 0 0 0 1 0 – 1 that is basically we have v1 positive that is why it is 1 and v 2 negative that is why it is minus 1 and v 3 negative that is why it is minus 1 and remaining v3, v4 and v5 are 0 for metabolite A.

Similarly for metabolite B and C, you can write the component 1 - 1 1 – 1 and then you can multiply it with the flux vector this equation the time derivative of the concentration of A B C is actually the intermediate concentration under steady state assumption; we assume that the derivative is 0, so it is not changing with the time. So, this is the assumption that is in flux balance analysis the metabolite concentration does not change with time.

This is a for experimental biology may be difficult to understand, what if you see the metabolites are actually produced by the enzyme and the protein and the proteins are regulated. So, the turnover times are very long, so in this small time, we assume that the concentrations are not changing that is why dA / dt = 0. So, the concentration of the metabolites are not changing with the time and we are assuming a pseudo steady state assumption.

So, this is the assumption which is using the flux balance analysis such that, so is this is S and this is v, so this is a flux vector and this is stoichiometric matrix if you multiply stoichiometric matrix with flux vector what you get? S.v = 0 this is very important equation and this is based on steady state assumption has to be said I suppose only assume that the concentration of the metabolite A, B, C whatever metabolite present inside the intercellular metabolite does not change with time.

So, this is an assumption also determine that the production of the metabolite whatever metabolites produced this is used to have by the other cell the production and consumptions are actually equal. So, there is no accumulation. So, this assumption that is dA / dt is almost constant and this happens mostly in the exponential phase this is this cell is growing this is the cell this is during the exponential phase that is the time.

Where you see whatever metabolize form is equal to that is why you get the maximum growth rate, the growth rate is maximum and the exponential phase and then also you determine the flux at this point only this cell is growing not in a stationary phase and not in a lag phase only in the exponential phase. We assume that the time derivative of the concentration is not changing with time.

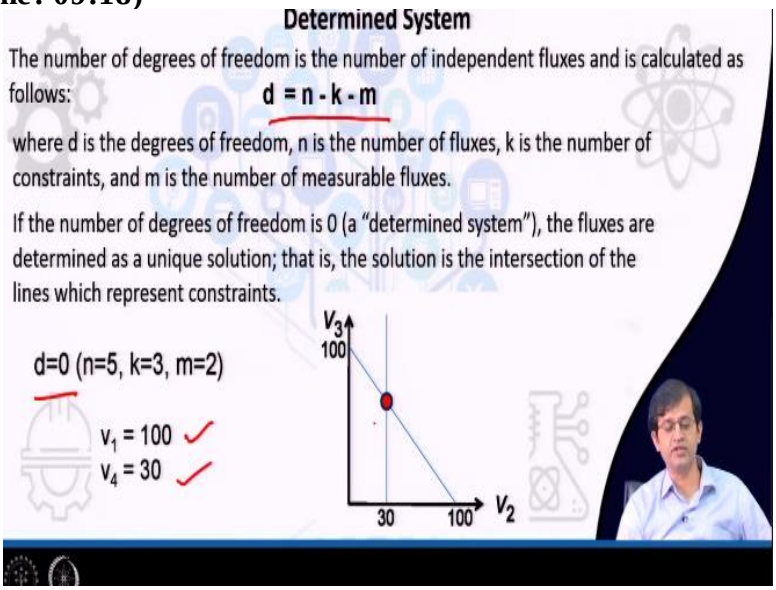
And also we maximize the growth rate that is why we choose the exponential phase if you break this growth curve into small small interval really small small interval, in this small interval of time that the concentration is not changing, during that time the small delta t this is the growth curve is as a function of t and this is OD then what do you see at a function of time delta t is very very small in this small interval dA / dt is 0.

So, this assumption is used to actually solve this equation otherwise, the differential equation we cannot solve this equation but if you apply the steady state assumption then dA / dt

becomes 0 and dA/dt dX/dt, X is the concentration of any metabolite. So, then your stoichiometric equation that is S . v = 0 which is nothing but the linear equation we have to solve the linear equation, now we come to an important system, how do we solve those equations?

Now, we have a linear equation there, so how many questions are here? So, we have 1 2 3, so we have 3 equations if you multiply these matrix, what do you will get 3 equation as a function v1, v2, v3, v5 and you will see that you have more number of variables then the number of equation.

(Refer Slide Time: 09:18)



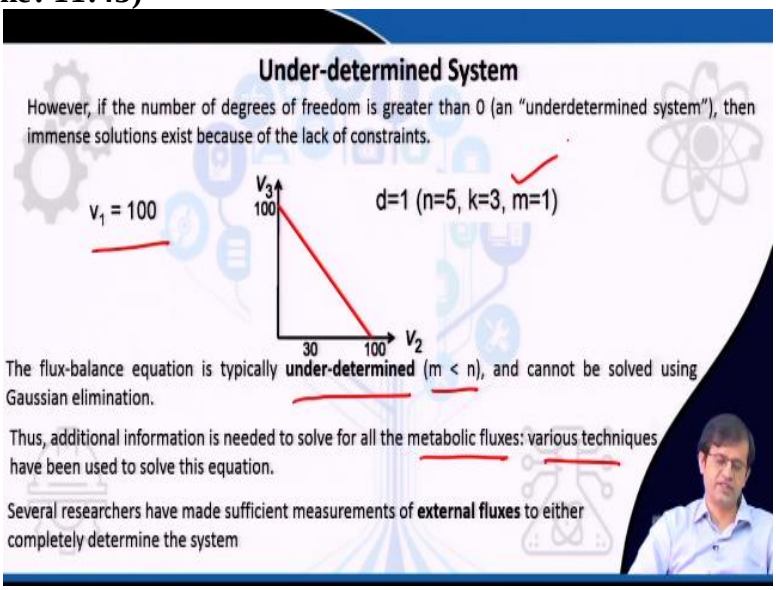
So, this is a problem, first we define what is a determining system where you get a unique solution, the number of degrees of freedom is the number of independent fluxes that is calculated as follows. So, you have an equation that the degree of freedom is equal to n - k - m where n is the number of fluxes and k is the number of constraint and m is the number of measured fluxes.

So, your degree of freedom is basically given by n is the number of fluxes that is in the previous case how many fluxes were there? There were 5 fluxes and then k and how many constraints were there? Constraint we do not have any constraint in the previous example but you can put some constraint as well like environmental constraint and then m is the number of measure fluxes.

So, in the previous case, we did not have any measured fluxes but if you consider this example, why I have only 2 reactions where we have v3 v2 and v1 and v4 is already measured. So, in the previous example, we have 5 reaction out of that v1 and v4 is already defined is measured fluxes. So, your number of fluxes n = 5 and then number of constraint is equal to k = 3 and number of measure fluxes, so these are the measure fluxes.

So, the measure fluxes m = 2 and constraint you apply, suppose you have a given constraint of k = 3 then what is your d discuss in this case? So d = 5 - 3 - 2, so it becomes 0. So, d = 0 if the number of degrees of freedom is 0 then it is a determined system then the fluxes can be measured you will get a unique solution and the solution is intersected by the lines which presented in the constraint. So, the constraint, so if you determine the flux is a determine system because the degree of freedom is 0. So, you can get a unique solution but this is not the case in metabolic network in the metabolic network, it is not a determined system.

(Refer Slide Time: 11:43)



So, what kind of system it is under determined system that is d is not 0, whenever d is not 0 then it is an under-determined system, if the number of degrees of freedom is greater than 0. Then we call it is an under-determined system that means, there is an infinite number of solution exists because of the lack of constraints. So, we will get many solutions, so that when you want to determine the fluxes, so it may there will be many possibilities, you will get many solution to the fluxes v1 v2 v3 v4 v5 that is why there is no unique solution.

So, because the d is not 0, why d is not 0 because for n = 5 that is the number of fluxes and then the constraint = 3 and the number of measure fluxes is only 1. In the previous case we

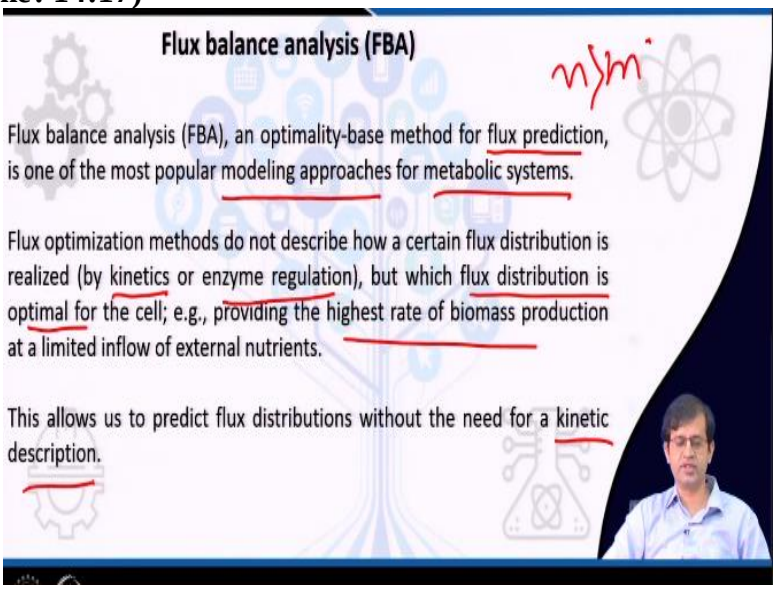
have seen that m is the number of measure fluxes are more but here are the measure fluxes all said that is why it became d = 1. So, the flux balance equation is typically under determined problem where n is greater than m that is number of variables are more than the number of measured fluxes.

So, m is the measure fluxes, n is the number of flux variable and that is flux variable is more than the measure and cannot solve this problem using Gaussian elimination or any other method. Thus, additional information is needed to solve the metabolic fluxes; various techniques have been applied to solve the equation where several researchers have sufficient measurement of external fluxes to either completely determine the system.

So, in this way, you can actually increase the value of n so if you can measure more number of fluxes, external fluxes, the internal fluxes you cannot measure, so it is measure it is very difficult to measure internal fluxes. So, if you can measure as many external fluxes, then it will be possible for making AB as a determine system but right now, it is not possible because the number of variables are more because we have 1000s of variables in the metabolic network that is 1000s of fluxes we have to determine.

So, it is a basically under-determined system where you do not have proper constraint and also we do not have much measure fluxes.

(Refer Slide Time: 14:17)



So, this problem, how to solve the problem, then how we can measure the flux? So, then FBA, flux balance analysis is actually deal with that how we can actually get the flux value

for each and every reaction given the variable more than the measured fluxes. So, n is greater than m. So, numbers of fluxes are more than the number of measure fluxes. So, this problem can be handled using flux balance analysis protocol. So, what is that flux balance analysis an optimality based method flux for flux prediction?

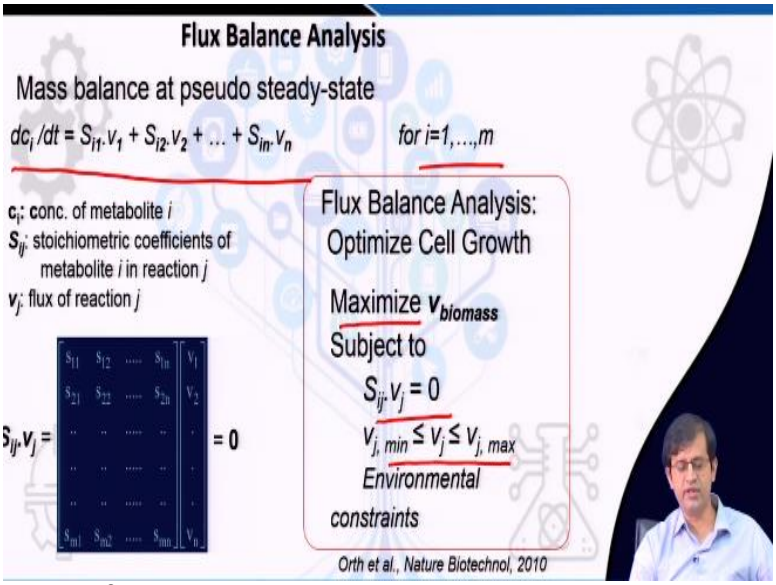
So, you have to predict the flux and how it is done is one of the most popular modelling approaches. So, this is a modelling approach where infer the fluxes for metabolic system where metabolic system you able to infer the fluxes or predict the fluxes and it is one of the most popular method and the flux optimization method do not desire how a certain flux distribution is realized that is by kinetics or enzyme regulation.

So, we do not need kinetic information nor you need the enzyme regulation information, so this method is independent of kinetic information or enzymatic regulation but you can always add the kinetic information there is no harm but if you do not have still you can predict the flux distribution that is what we said but which flux distribution is optimal for the cell that is providing the highest biomass.

So, as I told there are many flux distribution, because there is no unique solution and which of the flux distribution is optimal or which is feasible is based on the highest biomass production. So, they use another technique or another policy where they say that if you can maximize the biomass then you can get the flux value at a limited inflow external nutrient. So, the flux distribution is they say that the flux distribution is optimal or the best available for the cell when you maximize the biomass production.

These allow us to predict the flux distribution without the need of kinetics description, so we do not need a kinetic parameter already I have told that kinetic parameter k have to measure for genome cell metabolic network measuring k value for all the reaction is not possible because we have 1000s of reactions for equalised we have almost 1700 or 1800 reactions. So, this is not possible that is why flux balance analysis is very popular modelling approach where you maximize the biomass in order to get the flux distribution which is optimal for the cell and this does not required kinetic description.

(Refer Slide Time: 17:11)



So, let us come to the flux balance analysis is that the mass balance equation you already know that is dc / dt the concentration of all the metabolites you can write using this equation. So, this equation you can dc / dt is basically you have the stoichiometric component and the flux vector v1, v2, v3 and it goes from i = 1 to m that is multiplied 1, up to m metabolites are the concentration of the metabolite is c_i and this stoichiometric coefficient S_{ij} for metabolite i in reaction j and v_j is the flux for the reaction j.

So, this way you can actually get the right the mass balance equation and apply and then you apply the steady state condition that is the dc / dt = 0. So, in the previous slide I told that under a pseudo steady state approximation that the concentration of the metabolite is not changing with time and they are almost constant. So, if you take a time derivative, then it becomes 0 because the derivative of a constant term is 0.

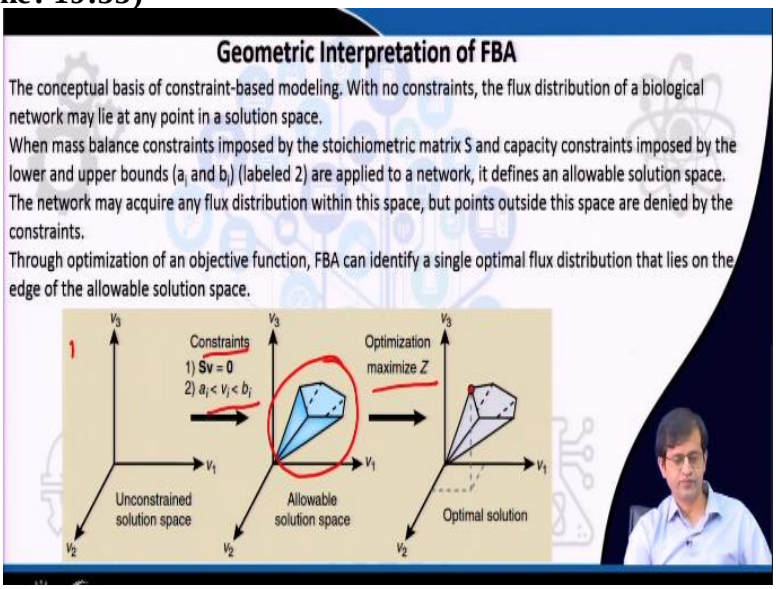
So that is why you can write S . v = 0 or S.v is basically S multiplied by the flux vector and this S . v is a constraint again the flux balance analysis that v is a constraint and also we have another constraint that is a lower bound and upper bound of the fluxes. So, this is another constraint and this is the constraint because due to steady state approximation, so these are the 2 constraint applied in FBA problem.

So, FBA problem and we have many other constraints like environment in that you can put and then you maximize the biomass. So, you maximize the biomass provided that S.v = 0 and the upper bound and lower bound are actually given. So, the fluxes basically were given the

boundary condition. So, if the boundary condition for all the fluxes so that it does not exceeds the boundary value.

So, using this optimization problem the maximization problem you would be able to get the fluxes for all the reaction without applying any kinetic detail. So, this is a very well-known method for modelling metabolic fluxes where you can get the fluxes for all the reaction in few seconds provided you have the metabolic network and all other detail available.

(Refer Slide Time: 19:55)



So, what is the geometric interpretation of FBA? The geometric interpretation we will say that the conceptual basis of constraint based modelling with no constraint the flux distribution of a biological network may lie at any point in the space. So, initially in the figure one, you can see that there is no constraint. So, if you do not have any constraint, then the solution space is very big the solution space it is not bounded.

You can see in the first figure the unconstrained solution space, because there is no constraint the solutions of all the fluxes can lie in any region. So, I am assuming that I have only 3 fluxes, the v 1, v 2, v 3. So, it can be complicated if I take all the reactions. Now, when you in the second case you apply the constraint as whatever you apply the constraint that is S.v = 0.

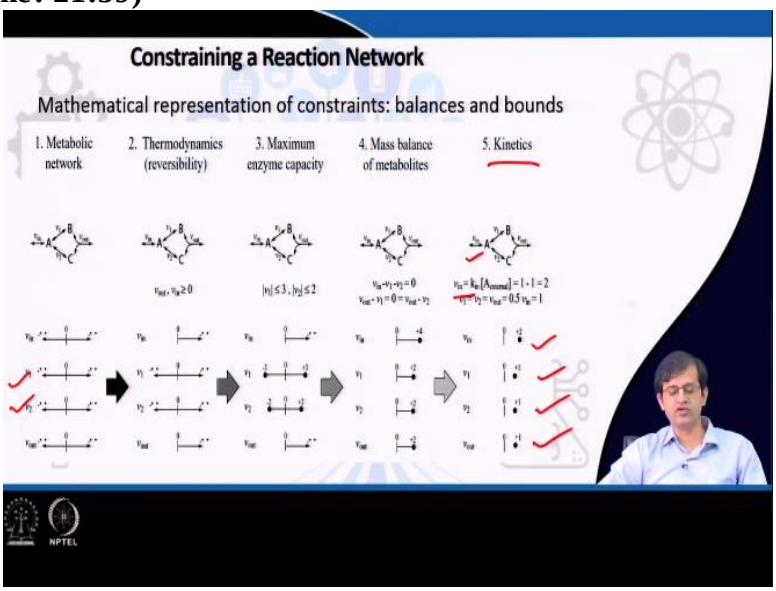
Then you have the boundary condition and that is the range of fluxes A to B if you do not know the flux rate, you can put minus infinity to plus infinity and you know the fluxes from experimental data that the fluxes. So, for example, the extra cellular fluxes you know the

bounds and you can put the bound and then your solution space become constraint. So, you have a cone kind of structure and this is the allowed solution space.

So, because of the constraint your solution was now in shrinking it become smaller and then because you have applied the constraint that is S.v = 0 and also the boundary condition and now, you maximize the objective of the growth and maximize the growth that is Z is the objective function and what you get an optimal solution. So, this is the optimal solution, shown in red dot at this point the growth is maximum.

So, the solution lies in anywhere in the cone but only one point that is at the edge where the biomass is maximum. So, you get optimal solution where the fluxes can be determined provided you are maximizing the growth rate.

(Refer Slide Time: 21:59)



So, the constraint as I discussed that you can apply constraint likes which is S.v = 0 and then boundary condition. So, what is a constraint? So, make it more clear, I give one example why choose a small network which has multiplied A, B, and C and then flux is entering the v as v_{in} and it is going out as v_{out}. Suppose glucose is entering; now I do not know the flux value for each of the reactions then what I just put minus infinity to plus infinity.

So, I do not know whenever it is minus infinity means it is going in the back direction and it is plus infinity is going in the forward direction. So, v_{in} v 1 v 2 v_{out} all of them I have plotted and I have shown that since I do not know the flux for each of the reaction I just put minus

infinity to plus infinity and then apply thermodynamics through thermodynamics you know some of the reactions which are actually reversible or irreversible.

So, whether it is forward going or reverse going that I can know from thermodynamics, if some thermodynamics data are available. So, for v_{in} v_{out} I know from thermodynamics that it is forward by the reaction and v_{out} is also forward by the reaction. So, you are bound automatically change it becomes 0 to some positive value and v_{out} is also 0 to positive value and v1 and v2 options I do not know the thermodynamic parameter.

From thermodynamic parameter I know that the v1 and v2 are actually reversible reactions, so is forward and reverse both are present. So, the range remains the same, so whatever range I assume in the beginning it remained the same. So, it becomes minus infinity to plus infinity for v1 and minus infinity to plus infinity for v2, now we through enzyme capacity or putting data enzyme data into some enzyme data is available.

Where we will have measured the maximum capacity the maximum value flux value of that enzyme is already known then you can put that value, so that is ±3 and ±2, now you can see that you are range or the bounds is actually much more constraint earlier it was minus infinity to plus infinity. Now, we have become minus 3 to plus 3 and similarly, for v 2 also it become minus 2 plus 2. So, this way if you have enzyme data then you can incorporate and then you can apply the mass balance equation that is $S.v = 0$.

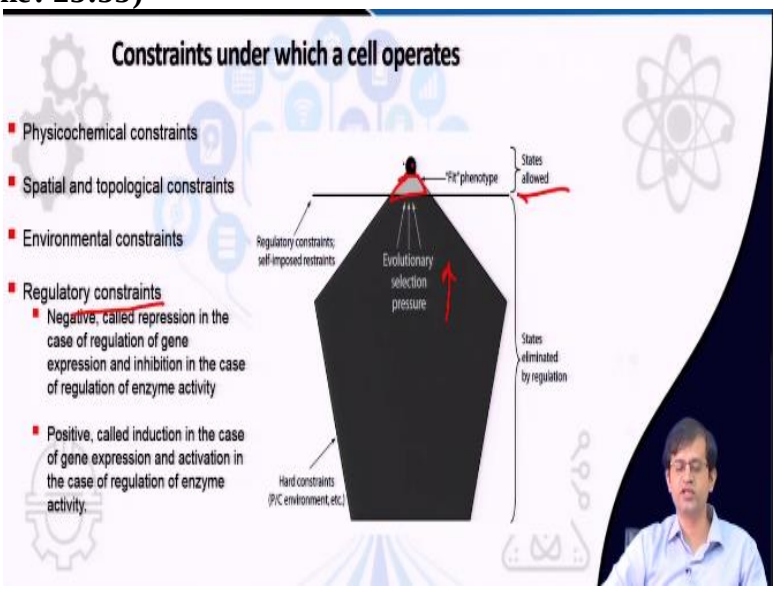
Then also you can see that v_{in} , v1, v2, v_{out} and become much more constraint now, with all the reactions are actually followed going and then the next step what do you see have the kinetic data if you have the kinetic data available, so you can incorporate the kinetic data which is basically the concentration. So, concentration of the metabolite A which is coming in the form of glucose if you know the concentration of the glucose then you multiplied with the k and you can calculates the value of v_{in}.

So, if you know v_{in} then all other values are actually determined, so you get a determined solution. So, unique solution you get, so how you can see that from minus infinity to plus infinity of the fluxes, where you keep on applying constraint based on the available data and you get a unique solution. So, this is what you actually mean to constraint a network, the

constraining network you are putting more constraint or more value or more parameters in the network.

So that your solutions become narrowed down and you get a unique solution and more you would constraint the more accurate your flux prediction will be.

(Refer Slide Time: 25:55)

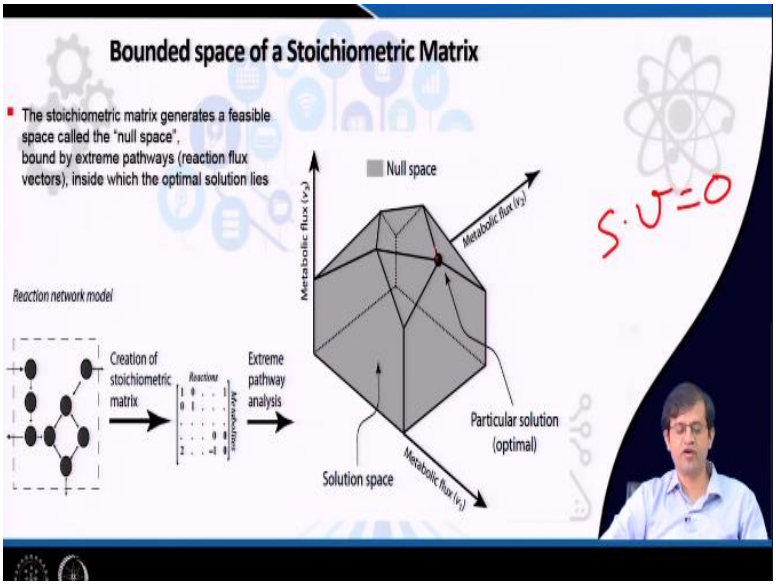


So, this is another example where you can see that initially we have the solution space which is covered in the shaded region in pentagon and then you keep on applying the constraint. So, the constraint may be regulatory maybe it is a regulatory constraint that is the negative repression or inhibition of certain enzymes that you want to incorporate in the model or induction in case of gene expression or activation in case of regulation of enzyme activity.

So, this regulatory constraint you can put in the model and you can see that you are allowed solution space is shrinking. So, because of the evolution pressure, your solution space, this solution space is not allowed. So, only this mass solution space is allowed because of the regulation. So, remaining other solution is not possible, because of the regulation and out of that you see only a dot which is the optimal solution for the network and that is where you have the maximum biomass.

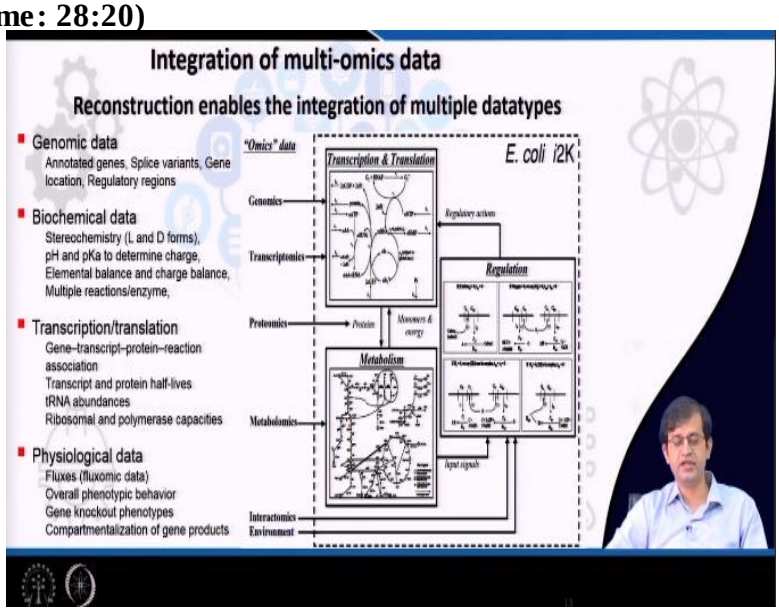
So, this point you eliminate all by regulation and still you get a solution, small solutions space the allowed state and among the allow state you reach a point which is the optimal solution where the biomass is maximum.

(Refer Slide Time: 27:26)



The entire solution space when you apply the constraint S . v = 0, then what happens? It becomes the null space. So, the null space is shown over here in this shaded region which is given by S . v = 0 at any point inside the null space is basically a solution to your FBA problem but not all the solutions are actually will give you the maximum biomass. So, only a point where the particular solution or the optimal solution where you will get the maximum biomass. This way you can reject all other solution only one solution you can get where the biomass is maximum and this you can get it from applying FBA.

(Refer Slide Time: 28:20)

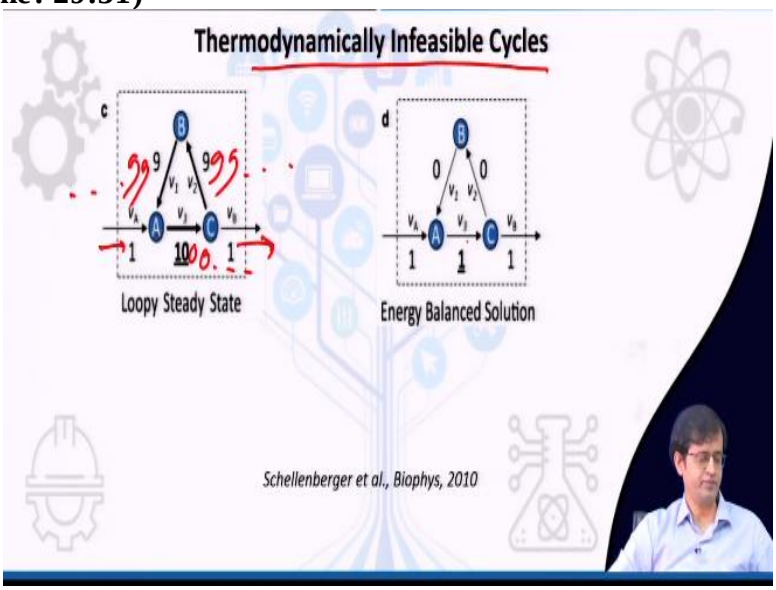


So, this is a crowded slide where you can see that the metabolism the metabolic network is actually does not have any regulation. So, the FBA formalism you apply, you know that there is no regulation but you can incorporate the regulatory information or transcriptional information in the network in that we will explain later also. So, various data like biochemical data that you can use for metabolic network construction.

But also you can incorporate transcriptomics data, protein-protein interaction data, various data you can apply in the metabolic network which will give you much more constraint network. This kind of genome scale metabolic network actually enables integration of high throughput data, for example the multi-omics data which you can incorporate in the model to make the model more constrained.

So, as in the previous slide also I told that you are putting constraint that the regulatory constraint or the boundary condition that actually allowing the network to shrink in space, so that you get more accurate solution.

(Refer Slide Time: 29:31)



Another important concept that we will learn today is it thermodynamically infeasible cycle what is a thermodynamically infeasible cycle? Suppose I have a 3 reaction network. So, we can see that the flux v_a is entering the network and then flux v_b which is going out. So, it is entering with a flux value of 1 and is going out with plus value of 1 and then the there is a loop which is A going to B and B going to C, you can see that v1, v2, v3 the value taken by is around 9.

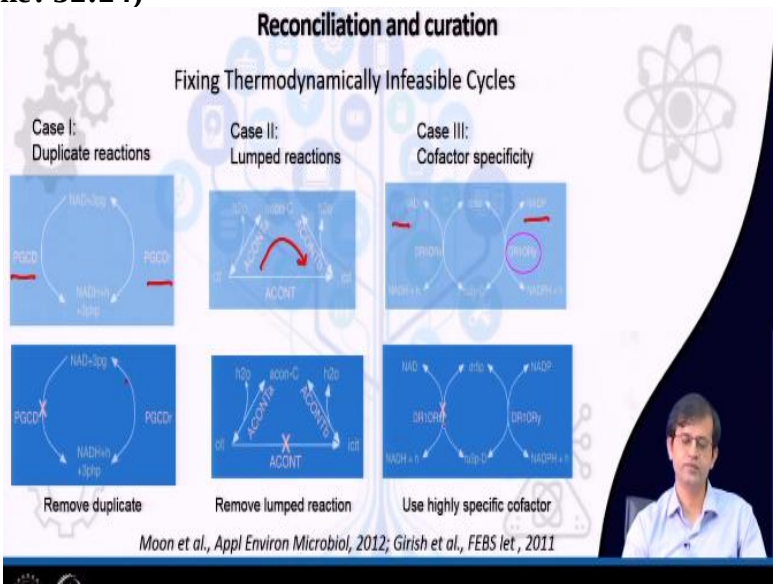
So, A to B you have 9 and B to C we have 9 and A to C we have 10 but your input flux is 1, so input and output flux is 1. So, this is will give rise to a biologically infeasible solution because your total suppose if the cell is going the uptake rate of the glucose is only 1 but the internal fluxes is 9 or 10, so it is not actually a feasible solution. So, the optimal solution that you get from FBA may give you a higher value than the input flux.

Then you can say that something is wrong in the network and this is based on the creature of law. So, if you apply creature of law at every node you will see that the summation is 0. So, the amount of flux entering and the amount of flux entering going out is balanced. So, there is a mass balance there is no problem, so mass balance is holding true here but still you get biologically infeasible flux why? So, this is because of the thermodynamically infeasible cycle.

And if you increase the number like if I put 9 9 9 9, 9 9 9 and 0 0 0 this is also true, if put any number it will hold true. Increase as many number of 0 as many number of 9 we will see that the solution exists and it is the mass balance is true. So, this kind of infeasible loop exists in the metabolic network and how do you eliminate the infeasible loop and these infeasible thermodynamically infeasible cycle can be removed by removing these reactions.

So, these in order to remove these reactor you have to run the loop less FBA, so loop less FBA these TIC, thermodynamically infeasible cycles are removed. So, you allow the flux to go through A2 through reaction v3, this is the concept which is used because when you construct the metabolic network, we will see that TIC exist and that give rise to a thermodynamically infeasible the flux values are very high which is biologically not feasible.

(Refer Slide Time: 32:24)



So, we will learn about how to reconcile these how to fix this thermodynamically infeasible cycle and in the case 1 you can see the duplicate reaction. So, this is a reaction where you can see PGCD and PGCDr which is this is a forwarding reaction and it is a reversible reaction.

And this kind of duplicate reaction exists in the network because you keep on adding reaction you will see that you might have the same there are 2 reaction for this conversion from metabolite A to B.

And how do you remove these infeasible loop the better is to remove the duplicate reaction, so if you remove PGCD, then you are retaining you are not losing any information in the network that means you are keeping the reaction reversible that is forward and reverse going just you are crossing 1 component. So, removing the duplicate reaction will fix the infeasible cycle. So, you can remove fixing infeasible cycle by just by removing the duplicate reaction.

Then I come to the lumped reaction, lumped reactions are the reaction where many reactions are lumped together, for example citrate to isocitrate. So, citrate to isocitrate you can either you can go from citrate to isocitrate directly through ACONT or you can go through two reaction. So, if you want to take the other path and that is from ACONTa to ACONTb and then you are covering 2 reaction but if you get both the reactions together then it will be infeasible cycle.

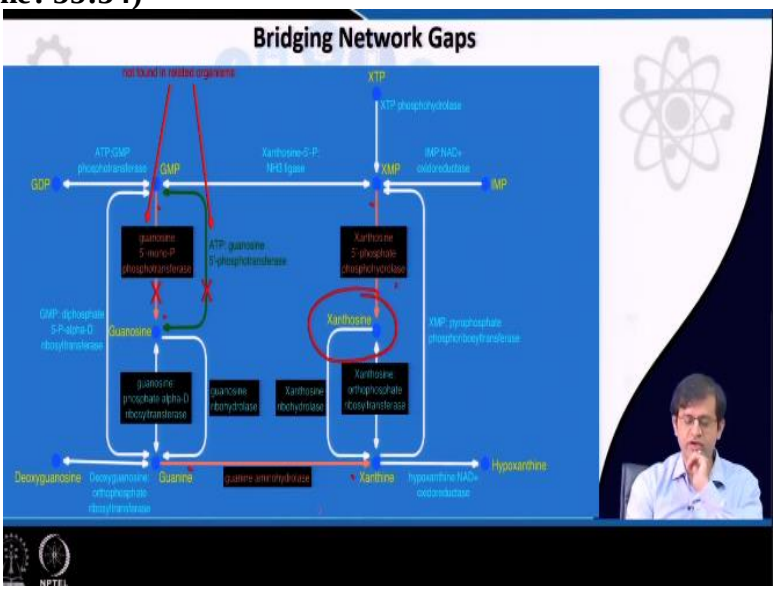
So, how do you remove this infeasible cycle just by removing the reaction that is ACONT, so if you remove ACONT that is citrate to isocitrate then you are not losing any information in the network where you are the reaction network is complete and still you are able to remove the lumped reaction that is from citrate to isocitrate it. So, by removing the lumped reaction you can remove the infeasible cycle and also there are cofactors specificity which exists in the network.

So, sometimes you keep the same 2 reactions for same metabolic conversion, so these 2 reactions one is specific to NAD and another is specific to NADP. So, you have to keep only one reaction if you keep 2 reactions in the network. Then again also it will create an infeasible cycle and this infeasible cycle you can remove by removing one of the reactions. So, you have to see which of the co-factor out of these 2 co-factor enzyme which one is actually much more specific that you keep and remove the other one.

So, in this way you can remove the NAD specific enzyme and keep the NADP specific enzyme that you have to see which one is actually highly specific co-factor. So, these are the techniques, these are the major 3 category where you can remove the thermodynamically

infeasible cycle but there are there exists many infeasible cycle in the network that you have to see individually and you would be able to remove the infeasible cycle.

(Refer Slide Time: 35:34)



So, another example, I will show you are how we can actually bridge the network gap, so while you construct the metabolic network, we will see that you have left out many gaps in the network and how do you remove the gaps? So, here in the example, you can see that, so the Xanthosine is in this network Xanthosine is not produced but it is consumed. So, I want to add a reaction, so that the Xanthosine is produced inside the cell to Xanthosine is not producing the cell but it is consume.

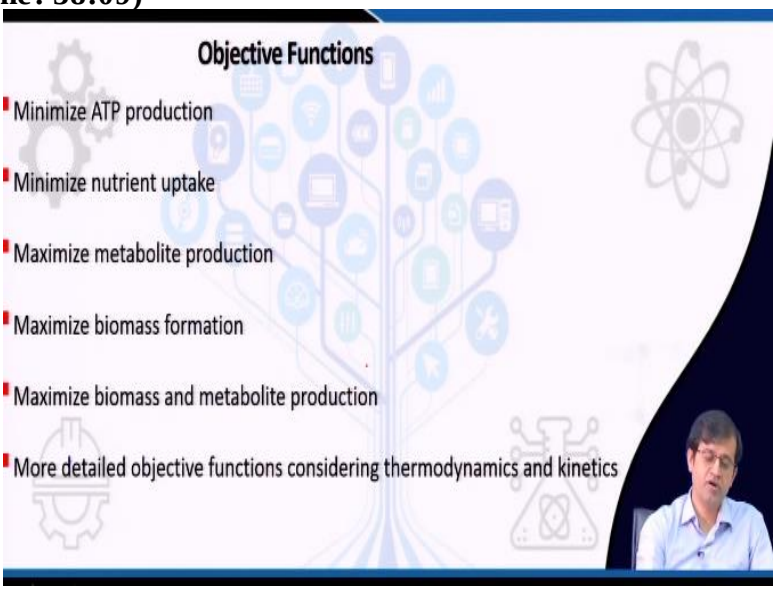
So, to do that, I can actually make a new addition I can look for hypothetical protein through homology charge and try to add a reaction either I can add between these 2 points or these 2 points or these 2 points, this three region I add a reaction and that you can run that I can add a new reaction based on the homology search I can look for hypothetical protein which I can catalyse this reaction.

So, I started with the search and I found that that there is 3 enzymes one is from GMP to guanosine by this reaction, you can bridge the gap or if I see this gene or it is not actually located in the closely neighbour species. So, it is that is why I cannot add this one. So, I cannot add this one because it is not present in the closely related organism for which the network I am getting and then also if I add when an amino hydrolyse then also from when and Xanthine, then also I see that it is forming an infeasible loop.

So, these red dotted line you can see that it is forming an infeasible cycle, so this one is also not possible and then what do I left is basically the other one the Xanthosine-5-phosphate phosphohydrolase. So, this is a much more feasible, so it when you do the bridging network, when you fill the gap then you have to decide logically which one is actually suitable for your network.

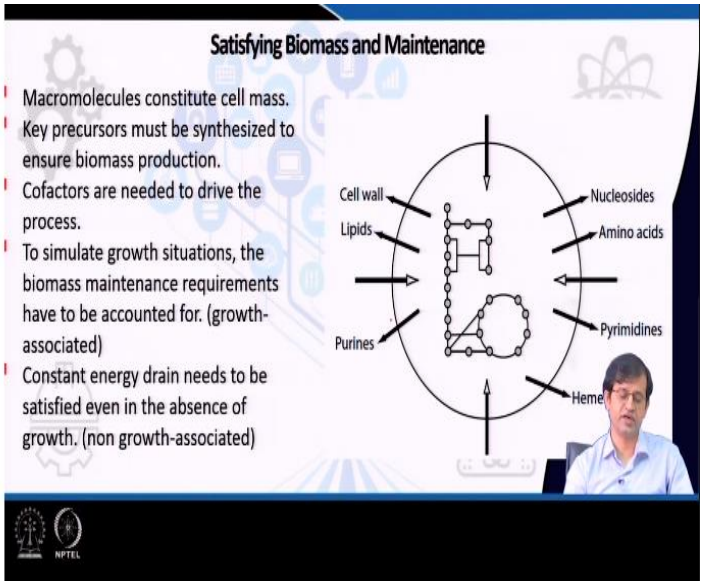
So, out of 3 choices, I found that this is only actually feasible, where if you add these reactions then you can see the Xanthosine is produced inside the cell and you are able to bridging the network we can bridging the network gaps are based on the literature data.

(Refer Slide Time: 38:09)



So, there are many in the metabolic network you can either minimize ATP production or you can minimize nutrient uptake rate or you can maximize certain metabolite production or already I told about maximization of the biomass or you can maximize the biomass and metabolite production both more detail objective function considered thermodynamics and kinetics of the cell.

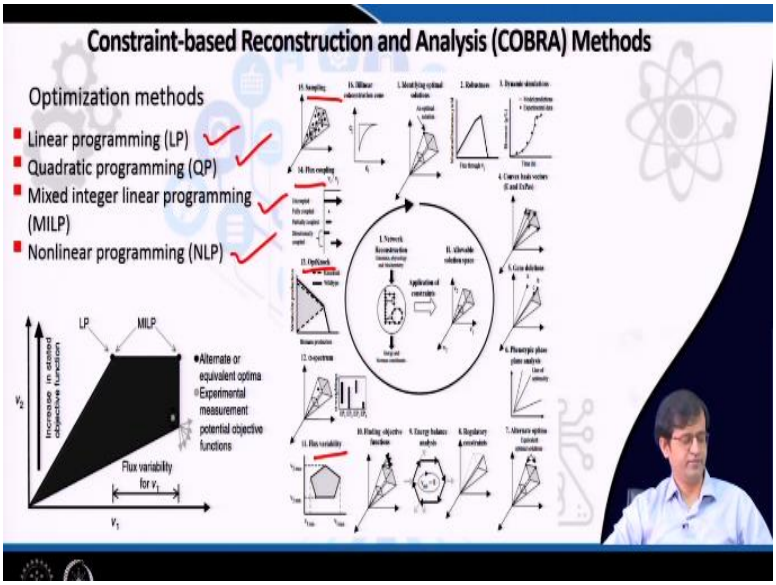
(Refer Slide Time: 38:35)



So, you have to understand that, whether the given reaction you are maximizing whether it is a growth couple or not. So, growth coupled reaction or is it not a growth coupled reaction. So, how do you check that whether the reaction is growth couple or not and the best way to actually maximize that reaction and you see, if you maximize a certain reaction, if you see that the growth is 0, then it is you can say that reaction is not growth coupled.

And if a reaction if you maximize certain reaction and if you see that growth is automatically coming up that is a nonzero growth you are getting then it is a growth couple reaction. This way you can actually measure which of the reaction is actually growth coupled and which of the reaction is not growth coupled to simulate a growth simulation the biomass maintenance recommend have to be account for that is the growth associated component that you have to add in the network constant energy demands needs to be satisfied even in the absence of growth.

(Refer Slide Time: 39:40)

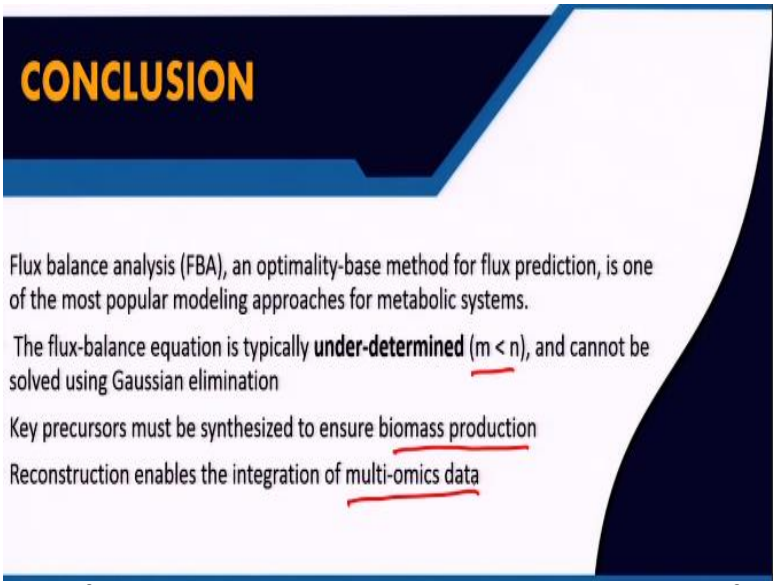


So, these are the constraint based reconstruction analysis method which is known as COBRA. COBRA toolbox in matlab also which is available which involves a lot of constraints. So, each of the methods which are available the optimization method while different constraints are applied and you can see the solution space is shrinking with increasing number of constraints and this method the constraint based reconstruction analysis method involved optimization method.

And the solver you use basically linear programming most of the time we use linear programming and then mixed integer linear programming and sometimes we use nonlinear programming and also quadratic programming. These various algorithms have been developed we will discuss in subsequent classes for understanding the gene deletion or they are understand the regulatory constraint or the flux variability and analysis and OptKnock flux variability, flux coupling, sampling.

So, various toolbox is available that is known as COBRA where you can apply this method and analyse the network and to understand it better where you can infer the flux or new phenotypes you can predict or you can apply different genetic perturbation to improve the production of certain metabolites that also you can do using this method.

(Refer Slide Time: 41:08)

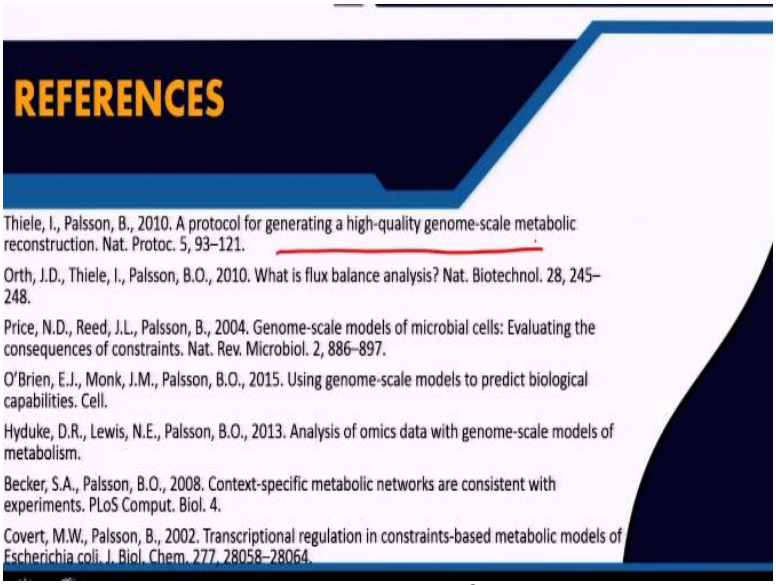


So, in conclusion, the flux balance analysis an optimality based method for flux prediction is one of the most popular modelling approaches for metabolic system. So, using FBA you will be able to predict flux and it is one of the most popular modelling approach and the flux balance analysis is typically under-determined system where the number of fluxes are more than the measure fluxes. So that is n is greater than m and cannot be solved using Gaussian elimination.

That is why you have to use a formalism which is FBA where you maximize the biomass and you get the flux solution and the key precursor must be synthesised to ensure biomass production of the biomass components should be synthesizing inside the cell, otherwise the cell will not be growing. So, you have to make sure that the biomass components are synthesizing inside the cell, how you determine the biomass component? It is through experiment?

So, you have actually run an experiment to determine how much DNA, how much protein, how much RNA is present inside the cell and that must be produced inside the cell to make the cell growing and reconstruction enables integration of multi omics data, today we have lots of high throughput data that you can integrate into the metabolic network and understand the different properties of the network, why do you make the network much more constrained which is very specific to address your problem.

(Refer Slide Time: 42:39)



The reference you can see that there is a protocol for generating high quality genome scale metabolic reconstruction this one you can read, where is a very useful where you will know what are the steps required for genome scale metabolic reconstruction. Then whatever flux balance analysis which you learn today, you can read in more detail in nature biotechnology paper, also the tools the constraint was tool, you can read more in detail in nature review microbiology where it gives details of different methodology used to actually constraint the network.

Moreover, you can go through all the references and get a more idea about flux balance analysis and how we can use to address different metabolic engineering problem. So, I close here thank you for listening.