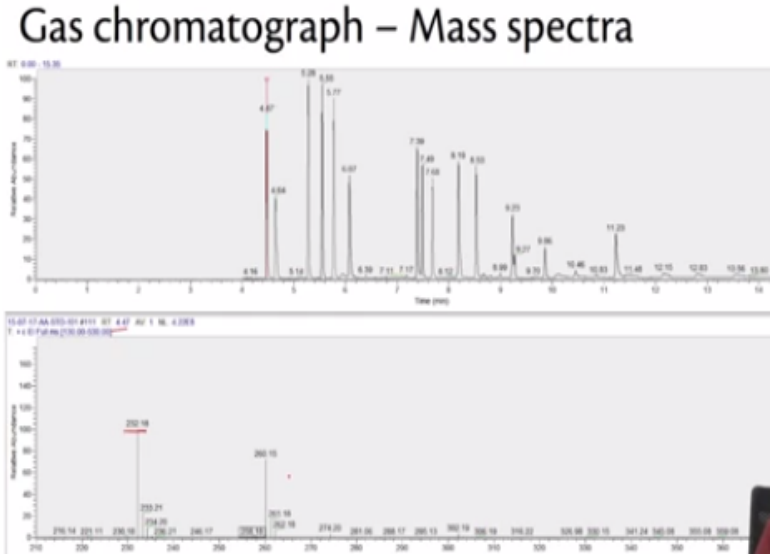
Computational Systems Biology
Aarthi Ravikrishnan
Department of Biotechnology
Indian Institute of Technology – Madras

Lecture - 85
Lab: 13C-Metabolic Flux Analysis using Mass Spectrometry

In this video, we will have a small recap of what has been done previously related to the GCMS data analysis. We will also solve simple exercises, which deals with the data from a GCMS chromatogram and a mass spectrum and see how we can derive useful information from the data that we got. Welcome to this session on C13-Metabolic Flux Analysis using Mass Spectrometry.

In this session, we would be handling a few examples, so this would be a lab session where you would get flavour of what exactly is the kind of data that is being handled in metabolic flux analysis using mass spectrometry and also have few ideas on what kind of problems can be solved using C13-MFA using mass spectrometry.

(Refer Slide Time: 01:13)

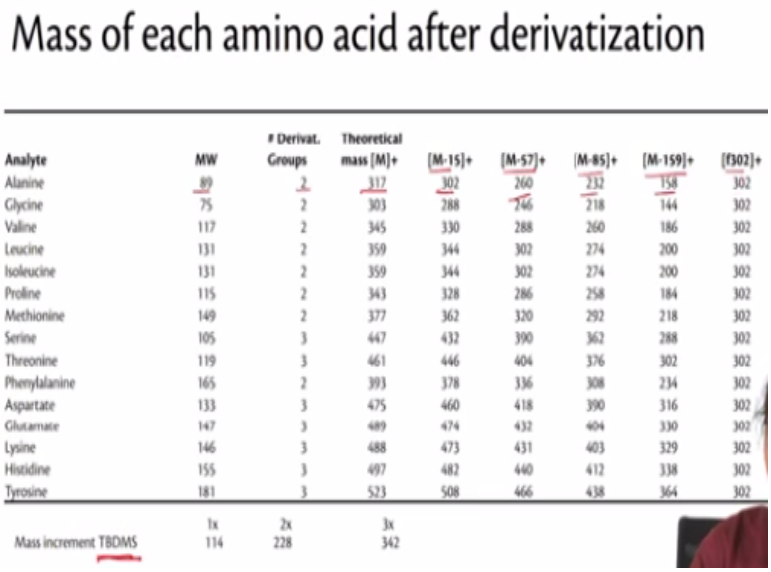


Just to give you an introduction on what exactly is gas chromatography mass spectra. In the earlier lecture, we saw that we have different amino acids, which get hydrolyzed into its salile derivatives for getting analyzed in a gas chromatography. While you pass these analytes in a gas chromatography machine, you would get a chromatogram that is shown in the above panel and also a mass spectrum which is shown in the below panel.

If you look at the top panel you can see different peaks at different retention times say for example 4.47, 4.64, 5.28 and so on and also in the bottom panel you can find out the relative appendences of each mass ion. For example, if you take 232.18, this corresponds to a particular mass and this can be observed at the time 4.47.

It has been seen using different standards and methods that 4.47 corresponds to alanine when you analyze amino acids in gas chromatography and with this these are few of the derivatized forms and the molecular masses here represent the derivatized weights of different amino acids and here in particular it is alanine.

(Refer Slide Time: 02:48)

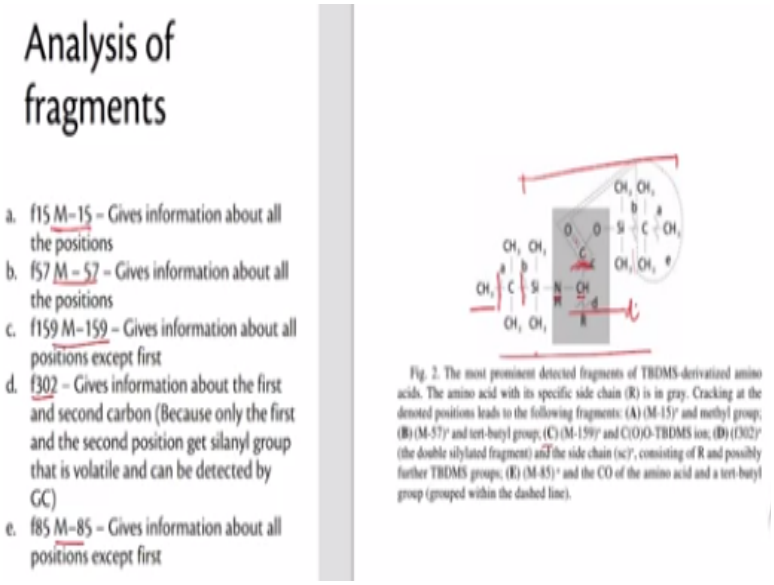


Analyte	MW	# Derivat. Groups	Theoretical mass [M]+	[M-15]+	[M-57]+	[M-85]+	[M-159]+	[f302]+
Alanine	89	2	317	302	260	232	158	302
Glycine	75	2	303	288	246	218	144	302
Valine	117	2	345	330	288	260	186	302
Leucine	131	2	359	344	302	274	200	302
Isoleucine	131	2	359	344	302	274	200	302
Proline	115	2	343	328	286	258	184	302
Methionine	149	2	377	362	320	292	218	302
Serine	105	3	447	432	390	362	288	302
Threonine	119	3	461	446	404	376	302	302
Phenylalanine	165	2	393	378	336	308	234	302
Aspartate	133	3	475	460	418	390	316	302
Glutamate	147	3	489	474	432	404	330	302
Lysine	146	3	488	473	431	403	329	302
Histidine	155	3	497	482	440	412	338	302
Tyrosine	181	3	523	508	466	438	364	302
	1x	2x	3x					
Mass increment TBDMS	114	228	342					

This is the table that depicts the mass of each amino acid after derivatizing. For instance, if you look at alanine, its molecular weight is 89 and the number of groups at which the salile molecules can attach themselves is 2 and so the theoretical molecular mass after derivatizing using this reagent called TBDMS, gets incremented by 228 and so you have a theoretical mass of 317 and as you might have already seen in the previous lectures, we have different fragments that is M-57, M-15, M-85, 159, 301 and so on based on the fragmentations that happen.

Because of this there can be a variation in the molecular weights. This table basically tells you that alanine when derivatized using this reagent called TBDMS gets incremented in its molecular weight and based on the different types of fragmentation, you have different molecular masses.

(Refer Slide Time: 03:58)



In the next slide, I would just tell you about the fragments that are observed. So this is just again a recap of what we saw in the last lecture. M-15, M-57, M-159, f302, and M-85 and these are because the derivatized form of the amino acid gets cleaved at different positions. For example, you have A, in this A you have the loss of CH3 moiety and the molecular weight of CH3 is 15 and so this particular fragment gets detected in the detector and so you have the information about all the positions.

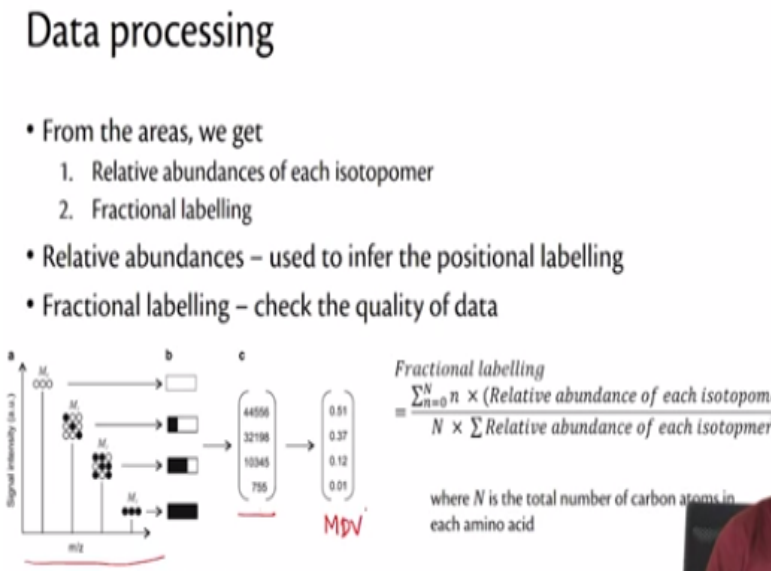
The gray panel here represents the central core of amino acid, so CH, COOH, NH, NR and these are basically the salile moieties that get attached to the active paths of the amino acid. Similarly, if you have the fragmentation at B, you have the tertiary butyl group that gets removed from the entire fragment and you have this portion which gets detected by the detector and so you have again the information about all the positions in the amino acid.

If you look at M-85, for instance M-85 is E and E is the portion in the dotted region, this gets removed and if you observe you can see that the first carbon of the amino acid gets removed and this is not being detected in the detector because of which you have the other parts being

detected and so if you have the fragment M-85 you get information about all the positions but for the first. Similarly, if you have M-159, M-159 is C, and C if it gets cleaved here, the first position goes away and you have the information about rest of the carbon atoms in the amino acid.

Again, if you have f302, 302 is another fragment and 302 gives again the information about the first and the second position, so 302 here is the D part, so D is this. This gets cleaved off and you have the information about first and the second position. With this information you can get a fair idea of the labels at every position in an amino acid and with this you will be able to backtrace to the different fluxes and the pathways that eventually gave rise to these labelings.

(Refer Slide Time: 06:54)



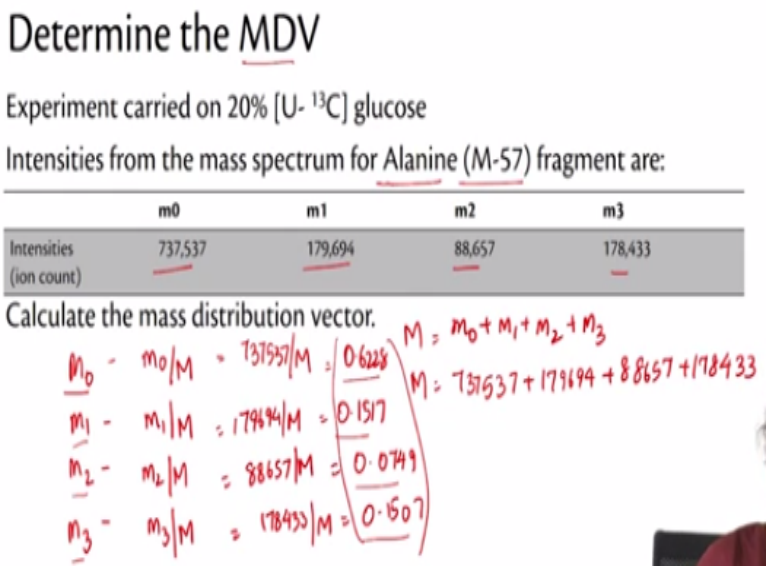
With respect to the data processing, as I already said you have the gas chromatogram and the mass spectra and from the mass spectra you derive different areas and from the areas you tried to find the relative abundances of each isotopomer and the fractional labeling. This is a simple graph which tells you about the signal intensity to M/Z ratio. Here you can find there are M1, M2, M3. M0 has only the non-labeled carbon atoms, whereas M1 has carbon atoms which are labeled at one position, M2 which is labeled at 2 and M3 which is labeled at 3.

This matrix basically tells you the signal intensity and from the signal intensity you derive something called a mass distribution vector, that is called MDV. This MDV is the fractional abundance of each isotopologue normalized to the sum of all the isotopologues. To get a clearer

picture, we would be solving a small example in the upcoming slides, but for now this is your mass distribution vector that you get from gas chromatogram and the mass spectra. With this you would go ahead calculating for few other things.

The second thing is fractional labeling. Fractional labeling basically tells you about the labels at different positions in a given amino acid. So, the formula is this and we would again be solving few examples for you to have a better understanding of what exactly is fractional labeling.

(Refer Slide Time: 08:49)

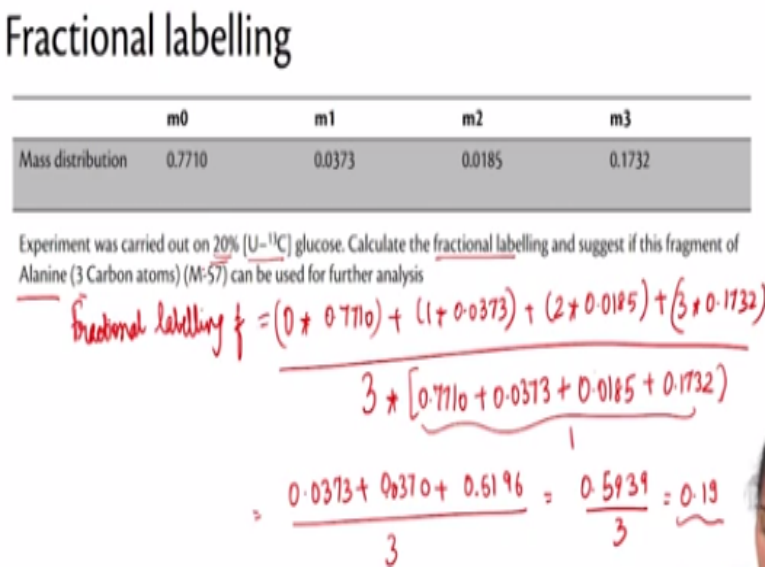


So, this is a small example which I want you to try out. This is a very simple example where you want to find out the MDV. You have been given the intensities of masses, so here you have M0, M1, M2, M3 and you have the intensities, that is the ion count that you get from the mass spectra, so this 7,37,537 and 1,79,000 and so on and so forth. To get the mass distribution vector, I simply do a sum of M0+M1+M2+M3, this I call it as M and for finding out this, I just do a normalization. This is how a simple mass distribution vector looks like.

When you calculate for this, this would just mean 737537/M, whatever the sum might be. Whatever fraction you find here, this roughly comes to 0.6228, this comes to 0.1517, this comes to 0.0749, and this comes to 0.1507. This would basically give you an idea of the mass distributions. This is the first basic information that you would get from any GCMS graph that is obtained. This for alanine, the amino acid with 3 carbon atoms and the fragment is M-57. Like

this you have data for all the fragments and using this you trend to calculate the mass distribution vector.

(Refer Slide Time: 11:36)

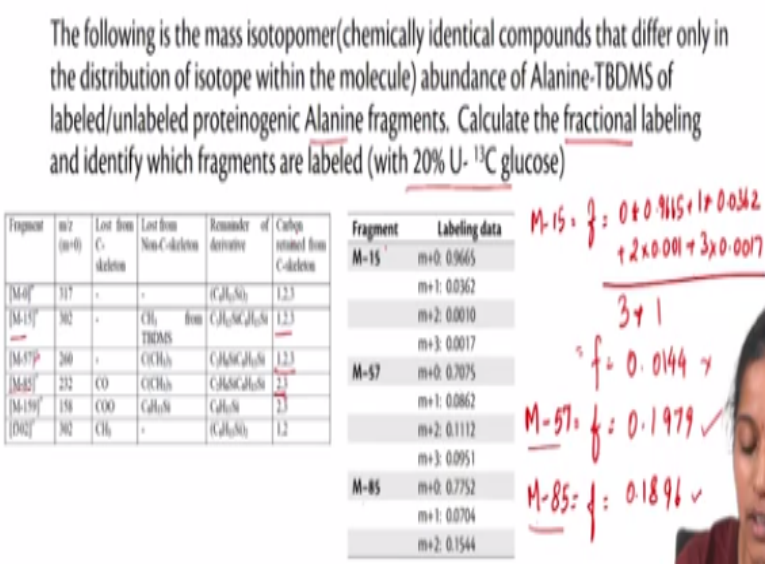


In the next example, we will be talking about fractional labeling. In this case, we have the mass distribution vector, so we already are given with the numbers here, which we earlier calculated, but this is another example here. This experiment was carried out on 20% UC-13 glucose and here we need to first calculate the fractional labeling. If you may recall from the previous lecture, fractional labeling is one of the means by which you can check the quality of your data.

By that I mean if you get a fractional labeling from this fragment as 0.2 and when you have used 20% UC-13 glucose as your substate, then you could go ahead with confidence to use this fragment for your further analysis. In this case, if you take a look, if I call fractional labeling as f, then with the formula earlier shown, it is nothing but (0*0.7710) + (1*0.0373) + (2*0.0185) + (3*0.1732)/3 * (0.7710 + 0.0373 + 0.0185 + 0.1732).

The sum of the denominator is going to be 1, so this is (0.0373+0.370+0.5196)/3, this roughly comes to 0.5939/3 is around 0.19. If you observe this data, you get a fractional labeling of 0.19, when you have used 20% UC-13 glucose, which means I cannot go ahead with this fragment because this is the level of confidence I have. I have 0.19 is the labeling that I observe when I give 20% that is 0.2 glucose. This is how you calculate fractional labeling.

(Refer Slide Time: 14:56)



In the next example, it is again to identify the fractional labeling and identify which fragments are labeled. As I was already telling you, that alanine is a 3 carbon amino acid and there are different fragments which you can observe when you pass them through a GCMS equipment. Here you have the information of every fragment and the mass distribution vector. You have M+0 that is basically M0, M1, M2, M3 and here you have again for mass M-57, you have masses. M-85, you have two of them, simply because for the fragment M-85 you just get the information for 2 carbon atoms and not all of them.

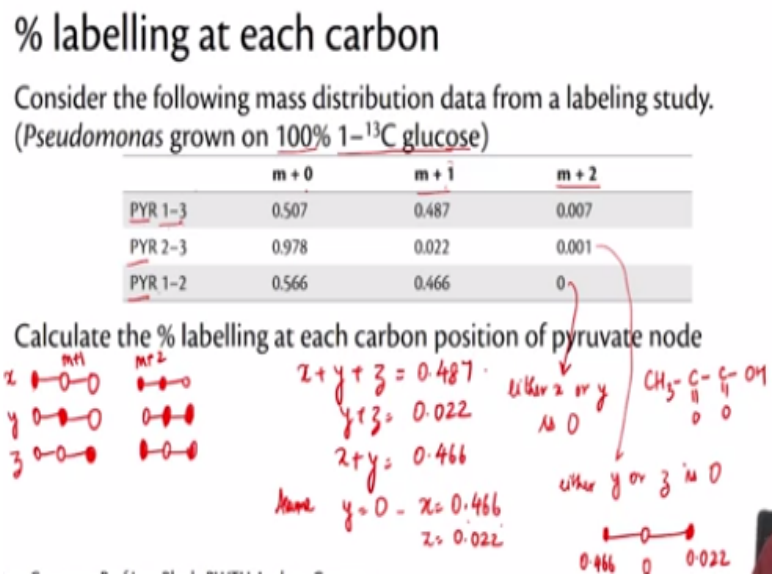
With this again you tend to calculate the fractional labeling as it might be called. Again you apply the same set of formula so that is for M-15. I will again have f is the same as whatever we calculated before (0*0.9665) + (1*0.0362) + (2*0.001) + (3*0.0017)/the number of carbon atoms in alanine is 3 and also the fragment M-15 can capture the information about all the 3carbon in the amino acid, so you have 3*1. In this case, f is going to be 0.0144.

Similarly, if you have M-57, you do the same set of calculations and in this case for M-57, using these values you find that the f is 0.1979 and similarly if you do it with M-85 fragment, you find that f is 0.1896. These values are simple to calculate if you just use the formula which is shown, it is very simple to get these values. Now comes the question of which fragment should I use for my further analysis.

If you observe the values of fractional labeling of M-15, it is very less that is 0.014, when I had 20% of uniformly labeled 13C glucose as my sub-state. I would rule out this fragment from my further analysis and if I take a look at M-57, I already have the information from this table that M-57 is going to give me the information about all the 3 carbon in my amino acid and also if I look at the fractional labeling, I can see that it is very close to 0.2 and so I will go with this fragment.

If I look at M-85, from this table I can see that I get the information about the second and the third carbon. I see it is 0.1896, it is not a very bad idea to take up this fragment and conclude more about the second and the third positions in alanine. I would take these two fragments for my further analysis and go ahead with my calculations. This is how you use fractional labeling and you also take the information from different fragments to get an overall idea of how the labels are in your amino acids. These are few examples.

(Refer Slide Time: 19:11)



In the next example, we will try to see how to calculate the percentage labeling at each carbon in a given molecule. If you have say pyruvate, here I take the example of pyruvate. Pyruvate as well all know is a 3 carbon compound and it basically has CH3COCOO. It has 3 carbon atoms and this study was done on 100% 1-C13 glucose. So, this means that it is a positional tracer and if

you recall from the previous lecture, the positional tracer can be used to resolve alternate pathways in a given cell.

Here since I use 13C glucose which is labeled at the first position, I get information this again is the mass distribution vector of pyruvate with respect to different positions labeled. For instance, if you look at pyruvate 1-3, you have the information about all the 3 carbon atoms in pyruvate and here you have M+0, which means this mass talks about all the unlabeled positions in pyruvate. If I look at M+1 and M+2, this M+1 talks about the pyruvate which has 1 carbon labeled and this one could either be in the first position, second position or in the third position.

Similarly, if I look at M+2, it can either be at the first and second, second and the third or the first and the third positions. Since I have this distribution table, I would now like to see how each position in pyruvate is labeled. This is a very simple example. So, first what we try to do is we try to construct this as a small toy. I am just drawing the carbon atoms in the pyruvate with different positions labeled. This is M+1.

M+2 again is this, this and this. So, now I would call this mass pyruvate that is labeled at the first position as X. Pyruvate which is labeled at the second position as Y. Pyruvate which is labeled at third position as Z. If you take a look at this data at M+1, you can see that pyruvate from 1-3 has a mass distribution of 0.487, so I write a simple algebraic equation, where you say X+Y+Z is 0.487. This is because you have M+1, M+1 is the fragment, which has at least one carbon that is labeled.

Here this X is basically the fragment where the first position is labeled. Y is this and Z is this. This is basically the sum of all the isotopomers of M+1. Next one is Y+Z is 0.022 and X+Y is 0.466. Does it make sense. If you now take a look at M+2, M+2 means that you have two positions that are labeled and if you look at 1, 2, you see it is 0, which means that either X or Y is 0 from this data because the first is X and second is Y.

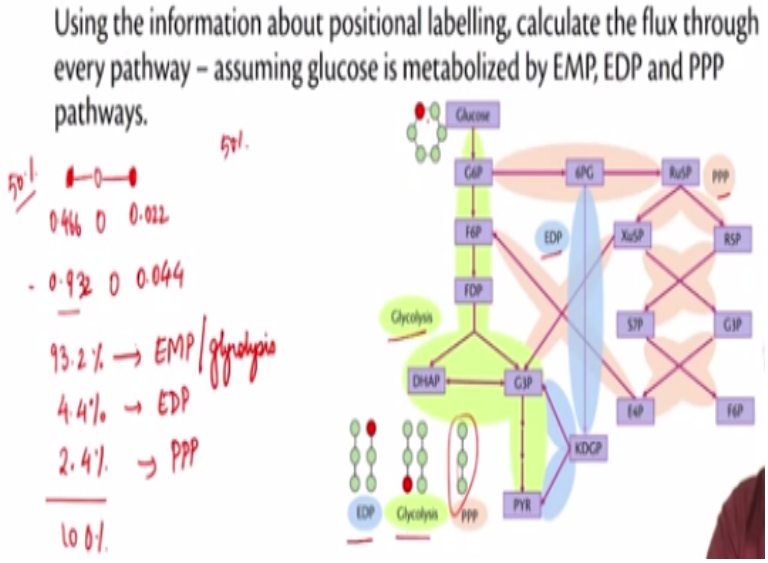
If you look 2 and 3, that is the second and the third position, again you have a value very close to 0. This from this again I can infer that Y or Z is 0. I can make these inferences from this and

from this I can say either X or Y or Z is 0. If you look at the fragments M+2 for pyruvate 2, 3 and 1, 2, you can see that Y has a higher probability of being 0. If I assume now Y = 0, this set of equations become very trivial to solve. X becomes 0.466 and Z becomes 0.022.

If I assume this, I can say that X, where you have the first position labeled is 0.466, so I can now write this X has a higher probability of being labeled, so here it is 0.466 and Y is 0 and Z is 0.022. There could be other case as well in this case, where Z is 0 and you can try and simulate it by yourself, but this is just an example to show you how it works. If I have this distribution, I can say that the first position is labeled with a probability of 0.466 and the third position is labeled with a probability of 0.022 and second position has no label at all.

This could be one of the solutions of the pyruvate being labeled at so and so positions.

(Refer Slide Time: 25:52)



Now if I want to go see how this positional information can be used to calculate the flux. Again, if you recall from the previous lecture, pyruvate can be derived from three different pathways, which is most commonly used and again assuming that this organism follows all these 3 pathways, you can say that pyruvate can either come from glycolysis, EDP or pentose phosphate pathway.

If I were to give glucose with the first carbon labeled as my sub-state, there can be different labeling patterns. In my previous example, here I have used 100% C13, which is labeled at the first position as my source, if I want to resolve and understand from which pathway has these labelings come from, I again go back to this labeling, here I have 0.466 and 0.022. If I look at this information that is M+0 is 0.507, which means this is the information of all the 3 carbon atoms and I can say that only 50% of the whole pyruvate is labeled.

"Professor – student conversation starts" Is it M+0 means not labeled at all. Not labeled at all, which means 50% is not labeled. 50% is labeled and 50% is not labeled. **"Professor – student conversation ends"** I do not use a uniformly labeled sub-state, but rather I just use a positional tracer, which means there can be fluxes through other pathways. I cannot say that PPP was not active simply because I do not observe this kind of pyruvate.

There could have been PPP that was active. It is just a matter of how we calculate it. Now we take this labeling pattern. This was 0.466, this was 0, this was 0.022. This was 50% of the fraction. There was another 50%, which was unlabeled from the data that we had, which means that if I want to calculate for 100% obviously I will multiply it by 2 and so I get this as 0.932, 0, and 0.044.

From the biochemistry and from the metabolic network, we have the evidence that if pyruvate was produced from 1-C13 labeled glucose it should have come from glycolysis because of different enzymes that act in glycolysis. We can say that our flux through glycolysis was 93.2% and second carbon has no labels at all and again in this case there is no pathway that can actually tell you about the second position labeled carbon in pyruvate. We do not have any evidence of a pathway.

Here we call this as EMP or glycolysis and with the third position, third position basically tells you something about the EDP pathway. In this case, we can say that 4.4% was through EDP, but since you have the total to be 100%, the rest whatever you get should have been via pentose phosphate pathway. This basically corresponds to 2.4%. **"Professor – student conversation**

starts" It is just a matter of our own understanding that (()) (29:58) how do you decide which pathway flux (()) (30:03).

That is because of the enzymes that act in glycolysis. The enzymes do not reshuffle the carbon bonds in random orders. They have specific orders in which they scramble the bonds. If glucose comes via glycolysis to pyruvate, only these set of enzymes act over glucose and because of that you have an understanding of where the label had come from. Similarly, for EDP and PPP. **"Professor – student conversation ends"**

If you look at PPP, if you have the first carbon labeled, glucose, right in the first step there is an exit of the first carbon as carbon dioxide, which means that the other metabolites from PPP will not be labeled if you have 1-C13 glucose. This is the overall picture of how you can use C13 labeled data from a mass spectrum.

"Professor – student conversation starts" This is just an illustration. Why are you multiplying the denominator by 3? 3 is because of the number of carbon atoms is alanine. It depends on different amino acids. Also, the fragment that you are analyzing. If at all have I to take say M-85, I get information only about the second and third. I have to multiply it only by the number which is here, that is 2. **"Professor – student conversation ends"**

In this video, we had a small recap of the GCMS techniques that we followed and we also did a couple of small examples that dealt with the data which was obtained from GCMS and looked about how to do calculations using the data that we obtained from the experiments that we carried out. In the next video, Dr. Karthik Raman would be back and he would be talking to you about logical models with special focus on Boolean Networks.