

## 5. Results and discussion

### 5.1 Results

By  $^1\text{H}$ -NMR spectroscopy on Bligh and Dyer extracts [126] metabolites profile is obtained, in which low molecular weight compounds are present. This profile is described and analysed here.

#### 5.1.1 $^1\text{H}$ -NMR spectra of hydro-alcoholic extracts

In  $^1\text{H}$ -NMR spectra it is possible to observe signals corresponding to different groups of protons belonging to each molecule present in the polar extracts. Molecules show signals in  $^1\text{H}$  spectrum zones characteristic of specific kind of  $^1\text{H}$  nuclei: for this reason, it is possible to describe each of these spectrum zones separately. A tentative assignment has been performed on the basis of the  $^1\text{H}$  chemical shift values and multiplicity of the signals.

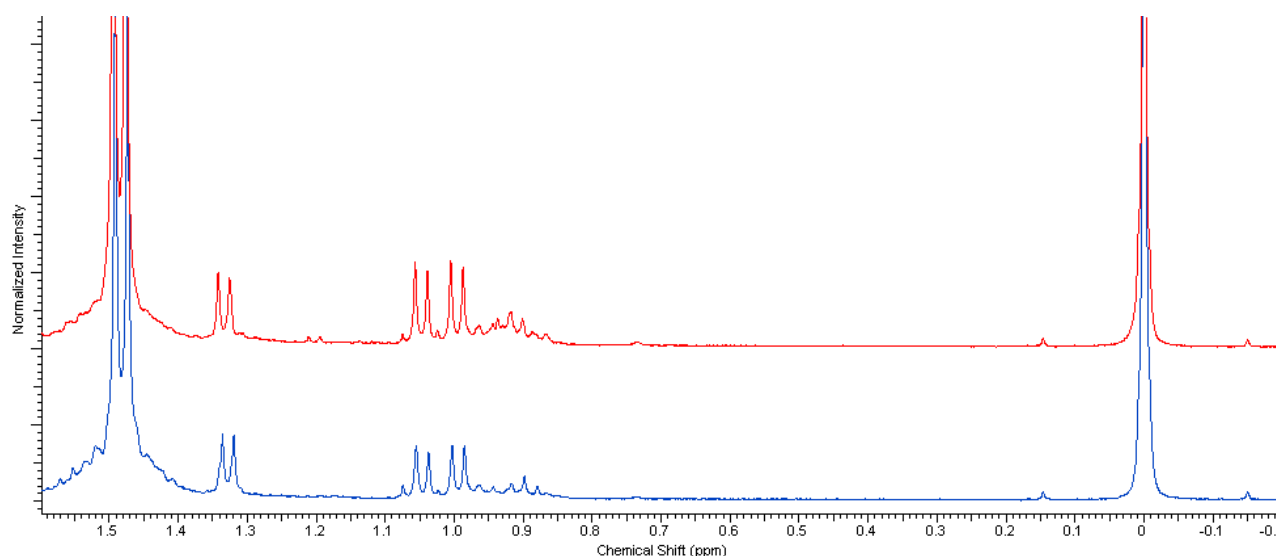


Figure 5.1: Mutant (in red) and wild type (in blue) NMR spectra of hydro-alcoholic extracts at high field.

- 1) At high field there were observed signals relative to the aliphatic  $\text{CH}_3$  protons belonging to aminoacids and small organic acids: fig. 5.1 shows doublets at 1.00 and 1.05 ppm (valine), 1.20 ppm (allothreonine), 1.33 ppm (threonine), 1.48 ppm (alanine), 0.90 ppm (butyrate), 1.08 ppm (isobutyrate) and 1.33 ppm (lactate, superimposed to threonine). The multiplicity of all these signals indicates that these  $\text{CH}_3$  groups are near to only one proton (CH group).

Between 1.6 and 3.3 ppm there are observable a series of multiplets relative to other similar molecules (fig 5.2): lysine, arginine glutamate, glutamine, glutathione, aspartate and malate (more details are available in table 5.1 and 5.2). Other organic acids are evident by the singlet at 1.92 ppm (acetate, more visible in mutant samples, overlapped to the multiplet relative to lysine around 1.9 ppm) and by the singlet at 2.41 ppm (succinate).

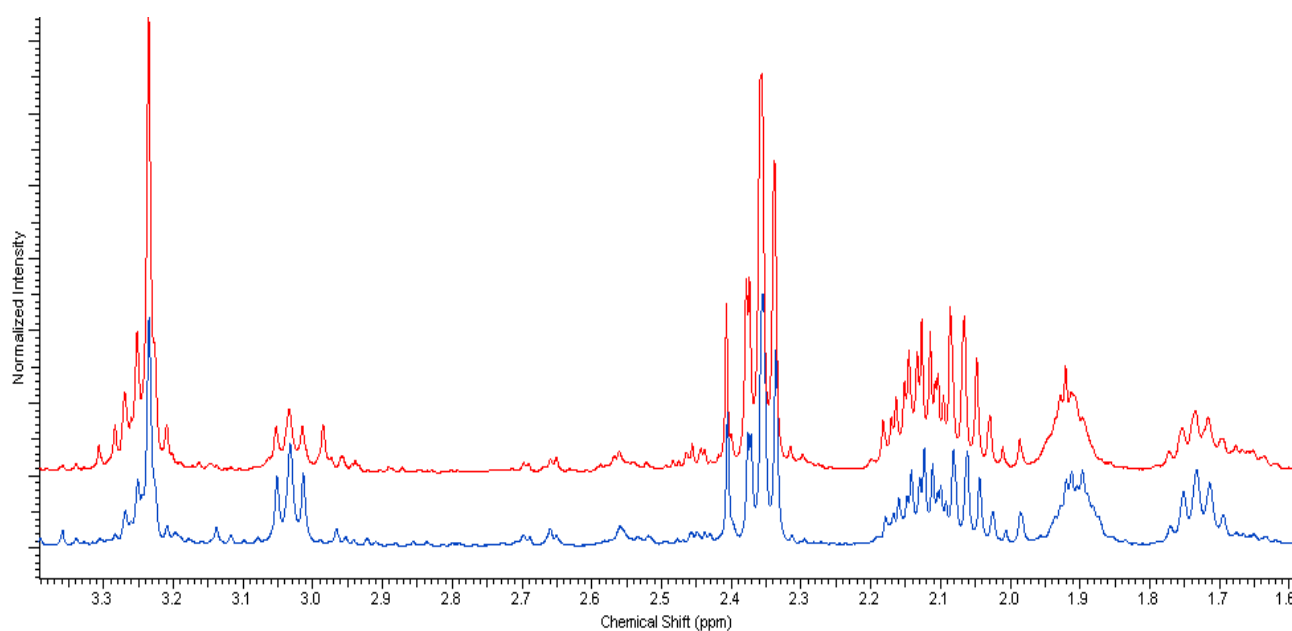


Figure 5.2: Mutant (in red) and wild type (in blue) NMR spectra of hydro-alcoholic extracts between 1.6 and 3.3 ppm.

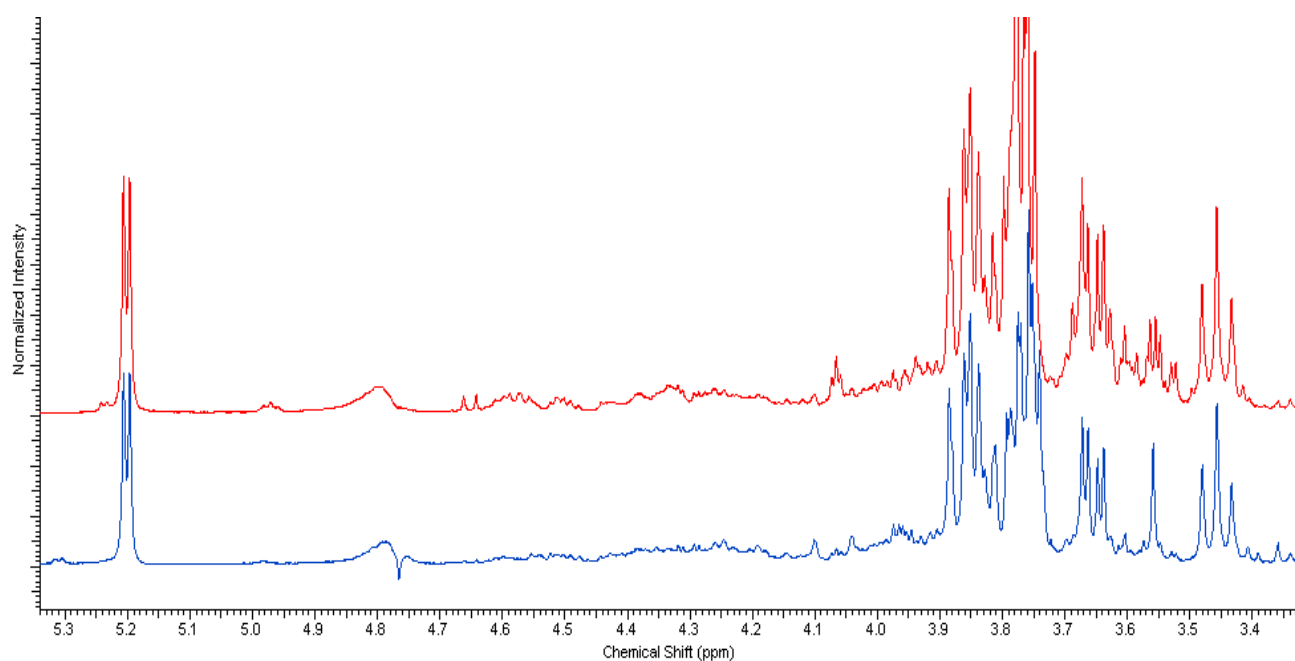


Figure 5.3: Mutant (in red) and wild type (in blue) NMR spectra of hydro-alcoholic extracts in the carbohydrate zone.

2) Between 3.4 and 5.3 ppm there are the resonances of the protons belonging to carbohydrates (fig 5.3): trehalose appear clearly by the triplet at 3.46 and the double-doublet at 3.66. Myo-inositol is more evident in mutant samples, at 4.06 ppm (triplet). However, there were been diagnostic for the assignment of such carbohydrates, doublets relative to the anomeric protons observed at 4.65 ppm ( $\alpha$ -glucose), 5.20 ppm (trehalose) and 5.24 ppm ( $\beta$ -glucose). It is also notable that water residual signal around 4.8 ppm (after the presaturation of the solvent) do not disturb neither the observability nor the quantification of these anomeric protons.

There is also observable a singlet relative to another aminoacid, glycine, at 3.56 ppm, more intense in wild type samples.

3) Finally, at low field (fig 5.4), spectra of the hydro-alcoholic samples present resonances relative to aromatic aminoacids, like histidine (singlets at 7.06 and 7.78 ppm) and tryptophane (singlet at 7.32 ppm), and other short-chain organic acids like fumarate (singlet at 6.52 ppm) and formate (singlet at 8.46 ppm).

There were also been observed signals relative to nucleotides, like orotidine monophosphate (doublet at 5.55 ppm and singlet at 5.77 ppm), and other molecules not yet unambiguously identified, named U, like U3 (doublet at 5.52 ppm) and U4 (triplet at 4.97 ppm and doublet at 6.12 ppm).

In particular, nicotinamide adenine dinucleotide moiety was realized through the doublets at 6.04, 6.09 and 9.15 ppm, the double-doublet at 8.84 ppm, the triplet at 8.20 ppm and the singlets at 8.18 – 8.43 and 9.33 ppm. All these signals are common to the oxidised form ( $\text{NAD}^+$  or  $\text{NADP}^+$ ) of this coenzyme and therefore they are indicated as  $\text{NAD(P)}^+$ . The corresponding reduced form (NADH or NADPH) should have different  $^1\text{H}$  signals relative to the nicotinamide unit.

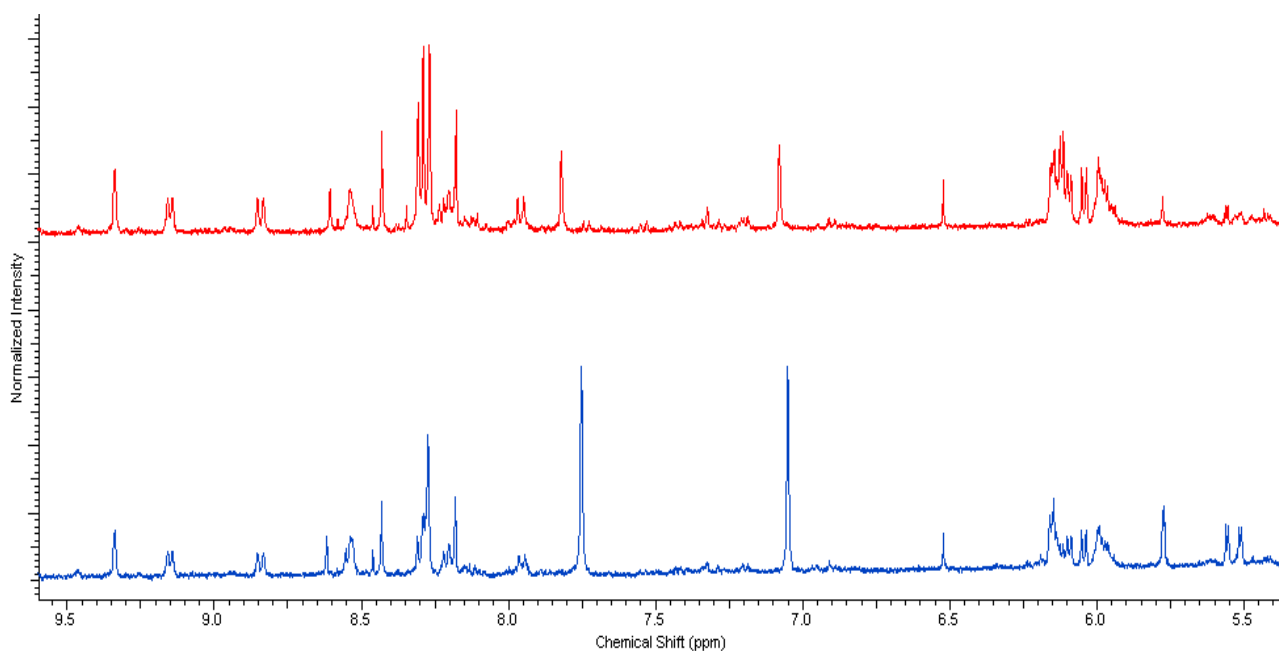


Figure 5.4: Mutant (in red) and wild type (in blue) NMR spectra of hydro-alcoholic extracts at low field.

### 5.1.2 $^1\text{H}$ -NMR spectra of chloroformic extracts

Compared to hydro-alcoholic extracts, the organic phase shows a less number of compounds, but each of them is characterized by a more complex signal pattern (fig. 5.5 and 5.6). Analogously to the hydro-alcoholic extracts, a preliminar tentative assignment on the basis of the  $^1\text{H}$  chemical shift value, multiplicity and corresponding coupling constants is performed. Three classes of prominent molecules can be identified:

- 1) Sterols, predominantly ergosterol (Erg): at high field there are the singlets (0.62 and 0.94 ppm) and the doublets (0.81, 0.83, 0.91 and 1.03 ppm) relative to the terminal  $\text{CH}_3$  of the aliphatic chain and those ones budded to steroidal backbone. Around 5.20 and 5.55 ppm the complex doublets relative to the vinylic protons are observable. In particular, there are two doublets of doublets at 5.17 ( $J = 15.7$  and  $7.85$  Hz) and 5.22 ppm ( $J = 15.7$  and  $7.25$  Hz) relative to the unsaturated carbon belonging to the steroidal backbone. Finally, the signal at 5.55 ppm appears as a pseudo-double-doublet because of an additional long range coupling of  $\sim 2.1$  Hz with an allylic proton [153,154].

2) Acyl-glycerols: there are clearly observable the double-doublets of 1-monoacyl glycerols (1-MAG) at 3.59, 3.69, 4.14 and 4.20 ppm, relative to the diastereotopic protons in C1 (*sn*-1) and C3 (*sn*-3), and the multiplet at 3.92 ppm relative to the *sn*-2 CH. The double doublets are defined by a common geminal coupling constant (about 12 Hz, relative to the H1-H1' and H3-H3' coupling) and by the  $^3J_{H-H}$  relative to the coupling between the *sn*-1 (~6 Hz) and *sn*-3 (4-5 Hz) protons with the *sn*-2 CH [155]. Less intense result the signals relative to the diacylglycerols (DAG) at 3.72, 4.24 and 5.06 ppm [156,157]. Broader are the signals of the tri-acylglycerols (TAG) at 4.13, 4.36 and 5.24 ppm [134,158]: this is due to the major presence of fatty acid chains esterified to the glycerol backbone, as explained below.

3) Fatty acids (FA): the chloroformic extracts are rich of these chemically analogous compounds, characterized by very similar chemical shift values; so all the resonances observed are very near each other and appear as a whole broad signal [143,159,160]. There are observable a series of signals common to all kind of fatty acids at: 0.88 ppm ( $\omega$ CH<sub>3</sub>), 1.29 ppm (CH<sub>2</sub> of the aliphatic chain), 1.59 ppm ( $\beta$ -CH<sub>2</sub> to the carboxyl group) and 2.35 ppm ( $\alpha$ -CH<sub>2</sub> to the carboxyl group) [161]. Moreover other signals are observable, relative to vinylic and allylic protons of unsaturated fatty acids (UFA): around 5.33 ppm there are the signals of the vinylic protons and at 2.02 ppm the ones relative to the allylic protons.

The diallylic protons of poly-unsaturated fatty acids (PUFA) resonate at lower field than the allylic ones; in addition the spectral resolution allows to identified and quantified separately two kind of PUFAs:  $\omega_3$  (indicated as linoleic acid, LLA) that gives a triplet at 2.76 ppm and  $\omega_6$  (indicated as linolenic acid, LNA) that gives a triplet at 2.79 ppm.

The total content of  $\omega_3$  and  $\omega_6$  is used to quantify PUFA amount in the samples, whereas others integrals are used to quantify or calculate UFA, MUFA (mono-unsaturated fatty acids) and SFA (saturated fatty acids) content, as described in 4.2.5 Materials and methods.

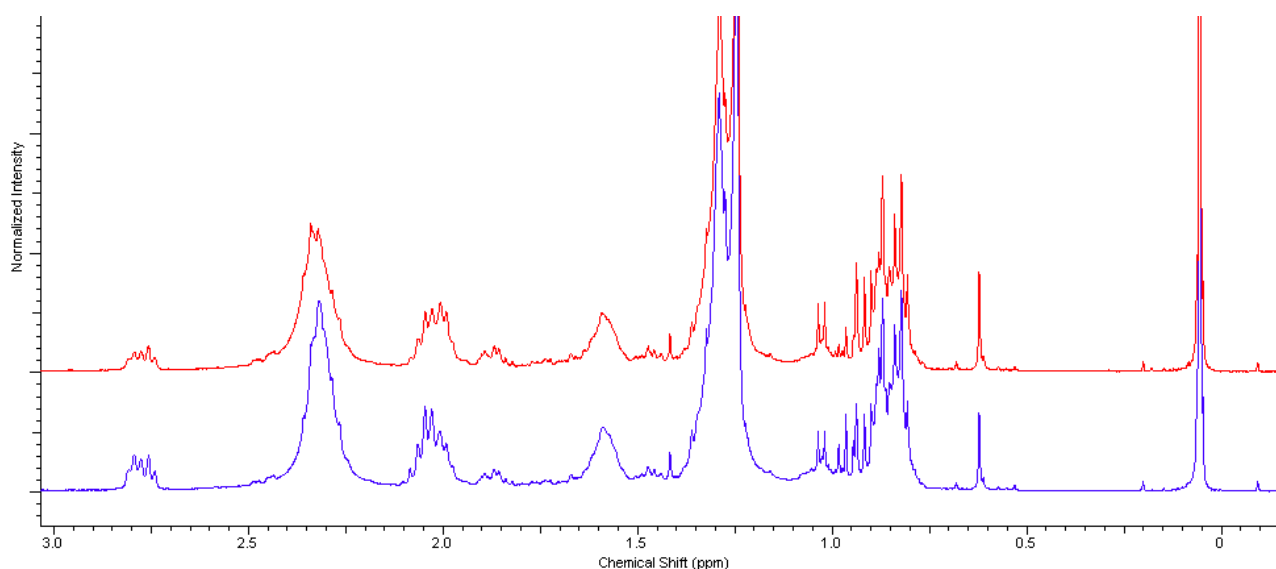


Figure 5.5: Mutant (in red) and wild type (in blue) NMR spectra of organic extracts at high field.

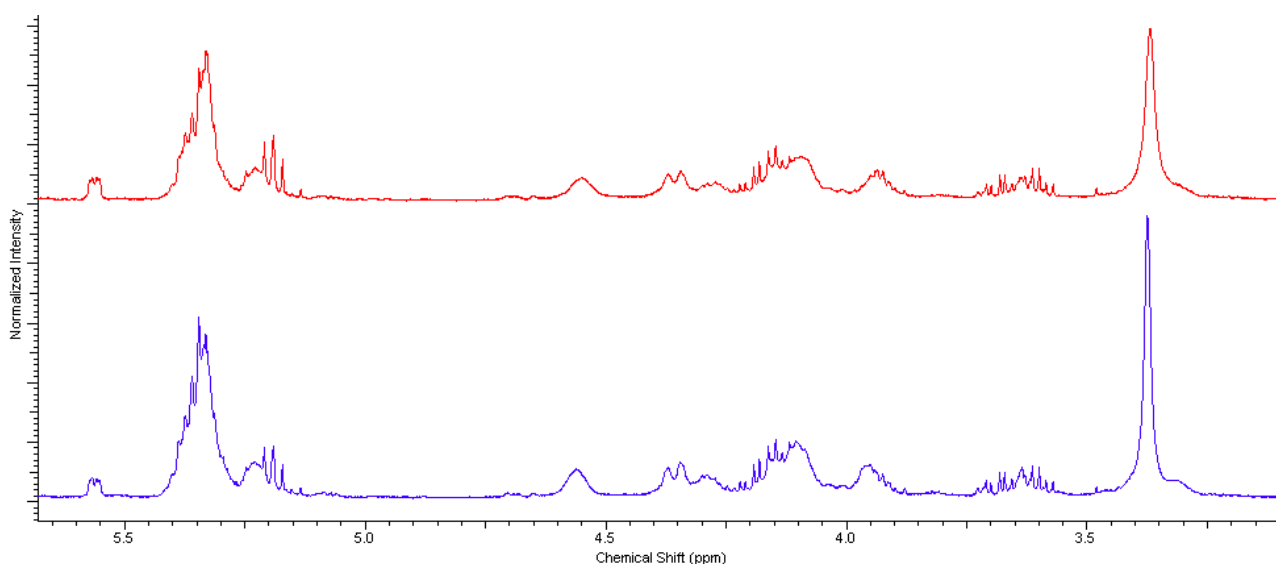


Figure 5.6: Mutant (in red) and wild type (in blue) NMR spectra of organic extracts between 3 and 6 ppm. The signal around 3.37 ppm is residual methanol from the extraction procedure [162].

The spectrum of these compounds in yeasts is rather simple, consisting mostly of C16 and C18 fatty acids [15]. To demonstrate this, the average length of the fatty acids chains ( $x$ ) for all SFAs ( $s$ ), MUFAs ( $\omega_9$ ) and PUFAs ( $\omega_6$  e  $\omega_3$ ) it was expressed, leaving as unknown only the number of the  $(CH_2)_n$  ( $n$ ,  $m$ ,  $o$ ,  $p$ , respectively), as:

$s:$	$x = 1+1+1+n+1 = 4+n$	$\rightarrow$	$n = x-4$
$\omega_9:$	$x = 1+1+1+m+2+2+1 = 8+m$	$\rightarrow$	$m = x-8$
$\omega_6:$	$x = 1+1+1+o+2+4+1+1 = 11+o$	$\rightarrow$	$o = x-11$
$\omega_3:$	$x = 1+1+1+p+2+6+2+1 = 14+p$	$\rightarrow$	$p = x-14$

For all this kind of FA it is computed as 1 the carbonylic carbon (COOH), the one in  $\alpha$ -position, the one in  $\beta$ -position and finally the  $\omega\text{CH}_3$ . All the unsaturated fatty acids have also 2 allylic carbon atoms. Moreover, MUFAs ( $\omega_9$ ),  $\omega_6$  and  $\omega_3$  will have, respectively, 2, 4 and 6 vinylic carbons, relative to their number of unsaturations (1, 2 and 3). Lastly, PUFAs have also 1 ( $\omega_6$ ) or 2 ( $\omega_3$ ) diallylic carbons.

The integral of the signal relative to the  $(\text{CH}_2)_n$  protons of the FA chains (at 1.29 ppm) divided by 2 (addressed as y) could be set equal to the sum of n, m, o and p, weighed for the relative content of each kind of FA revealed.

Operatively, all the integrals relative to FAs have been normalised by the integral of the  $\beta\text{-CH}_2$  (set to 2), representing the total amount of FAs. Then, all the normalised integrals have been divided for the number of protons generated each signal, to have an index of the number of carbon atom corresponding to each group of protons. So, from the normalised integrals, there are calculated the relative amount of each kind of FAs, in the same way described in 4.2.5 Materials and methods:

$$s = 1 - \text{allylic CH}_2 \text{ (at 2.02 ppm)}$$

$$\omega_9 = \text{allylic CH}_2 - (\omega_6 + \omega_3)\text{diallylic CH}_2$$

$$\omega_6 = \omega_6 \text{ diallylic CH}_2 \text{ (at 2.76 ppm)}$$

$$\omega_3 = \omega_3 \text{ diallylic CH}_2 \text{ (at 2.79 ppm)}$$

So, finally:

$$y = s*n + \omega_9*m + \omega_6*o + \omega_3*p = s*(x-4) + \omega_9*(x-8) + \omega_6*(x-11) + \omega_3*(x-14)$$

where \* indicates multiplication.

From this equation it was calculated the average length of the fatty acids chains (x) present in wild type and mutant samples. In either case, x resulted equal to  $17 \pm 1$  carbon atoms, suggesting that there are no differences in the elongation process between the two strains and confirming the fact that the most abundant FA in the samples are C16 and C18. So, our choice to address  $\omega_6$  and  $\omega_3$  as linoleic and linolenic acid, respectively, is essentially legitimated.

### 5.1.3 Univocal assignment by 2D NMR experiments

The high sensitivity of  $^1\text{H}$  NMR, which stems from the favorable magnetic properties of the  $^1\text{H}$  isotope and its 99.985% natural abundance, is a considerable advantage, but the usefulness of the technique is somewhat reduced by the overlapping of the signals from the detectable metabolites. This is particularly true in such a complex mixtures as cellular extracts, even if our extraction procedure (Bligh and Dyer modified [126]) simplify the system already, allowing the separation of hydro-soluble molecules from lipidic ones.

For this reason the use of two-dimensional NMR experiments was fundamental because these techniques provide valuable informations for metabolite identification by correlating different signals from the same molecule [57]. So, the univocal assignment of resonances observed in 1D  $^1\text{H}$ -NMR spectra is done accurately with the contribute of the diagnostic crosspeaks revealed in 2D-NMR omonuclear spectra (TOCSY, fig. 5.7): in this way it was possible to identify the complete correlation patterns of each metabolite and so univocally assign them.

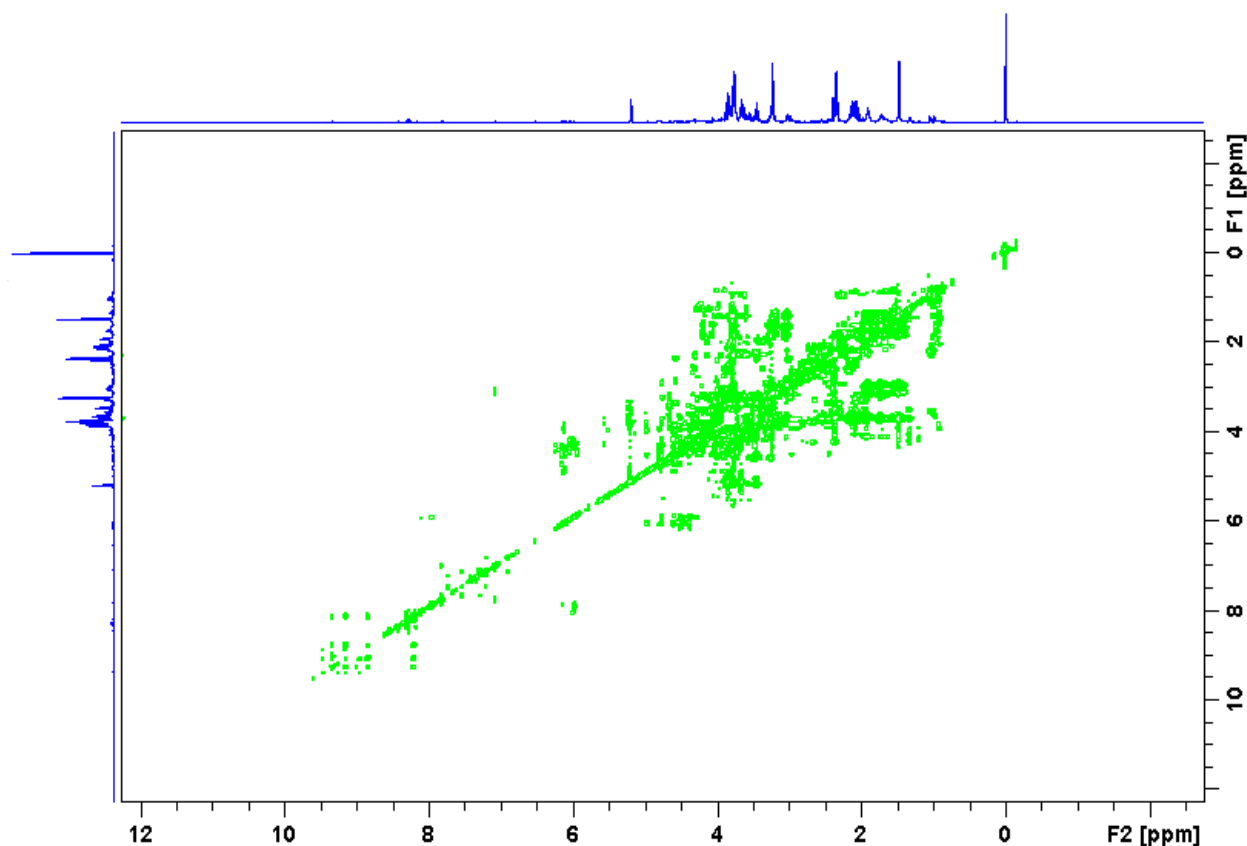


Figure 5.7: 2D-NMR  $^1\text{H}$ - $^1\text{H}$  TOCSY spectrum of an hydro-alcoholic sample.



Figure 5.8 shows how this assignment was performed thanks to the aid of the 2D omonuclear TOCSY crosspeaks; as an example protons belonging to the nicotinamide spin system of NAD(P)<sup>+</sup> are taken. The first signal on the right, at 8.20 ppm, is the triplet relative to the CH in N5 (fig. 5.9): hatched lines indicate the correlation, symmetric respect to the diagonal (continuous line), of this proton with the other <sup>1</sup>H nuclei of the same spin system, i. e., the CH proton in N4 (8.84 ppm), the CH proton in N6 (9.15 ppm) and the CH proton in N2 (9.33 ppm). In the same way it is possible to correlate each other all the proton of the nicotinamide moiety.

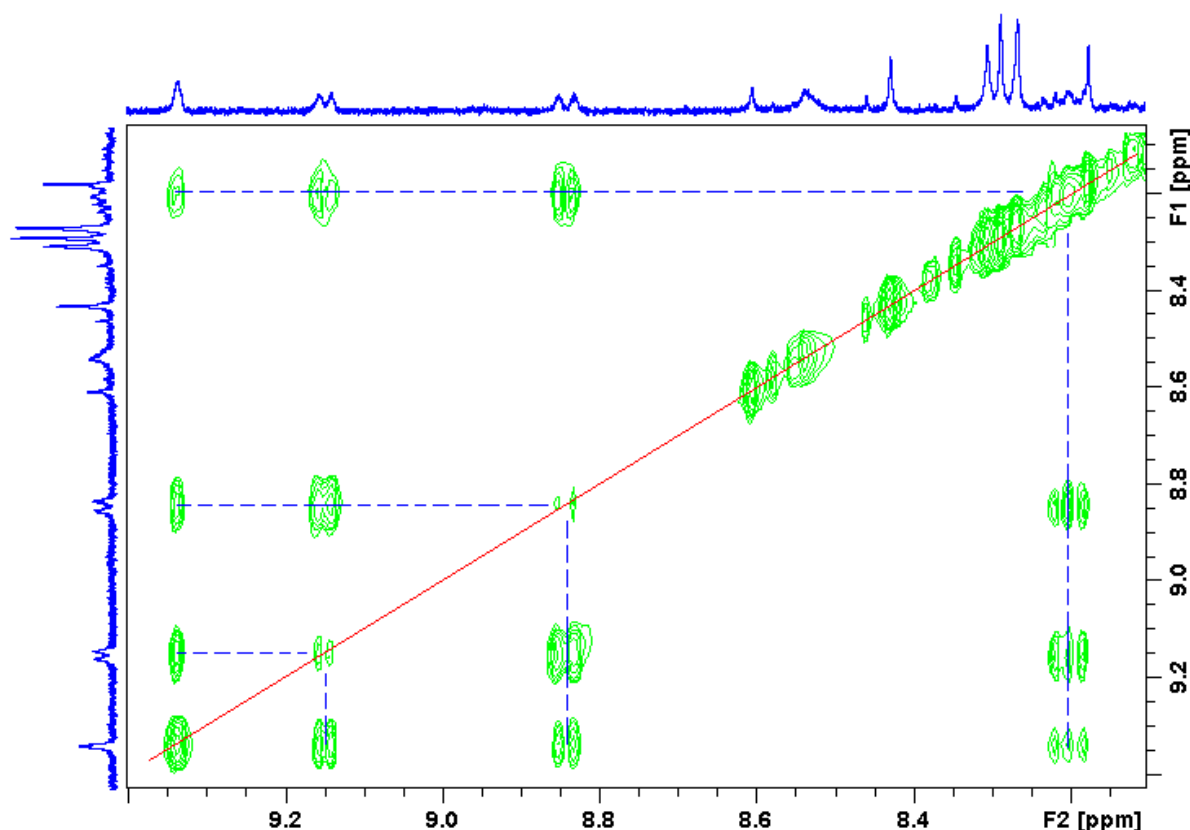


Figure 5.8: 2D-NMR <sup>1</sup>H-<sup>1</sup>H TOCSY spectrum detail.

There are clearly visible the correlations of NAD(P)<sup>+</sup> protons relative to the nicotinamide spin system. Hatched lines help to identify crosspeaks symmetric respect to the diagonal (continuous line).

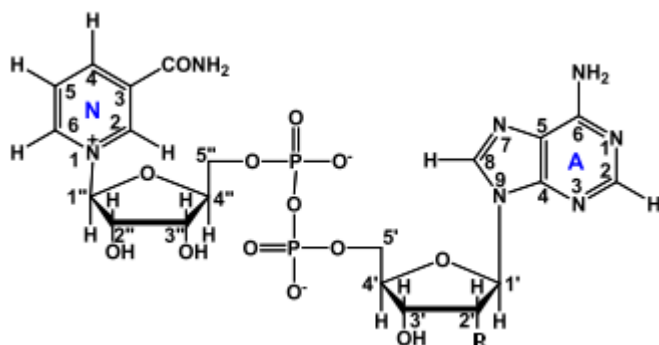


Figure 5.9: NAD(P)<sup>+</sup> structure.  
R could be an -OH or a P<sub>i</sub> group.

Moreover, the correlation identified in eterocorrelated spectra (HSQC and HMBC, fig. 5.10 and 5.11) allow the assignment, respectively, of the direct-bounded and long range-bounded carbon atom of each proton nuclei observed.

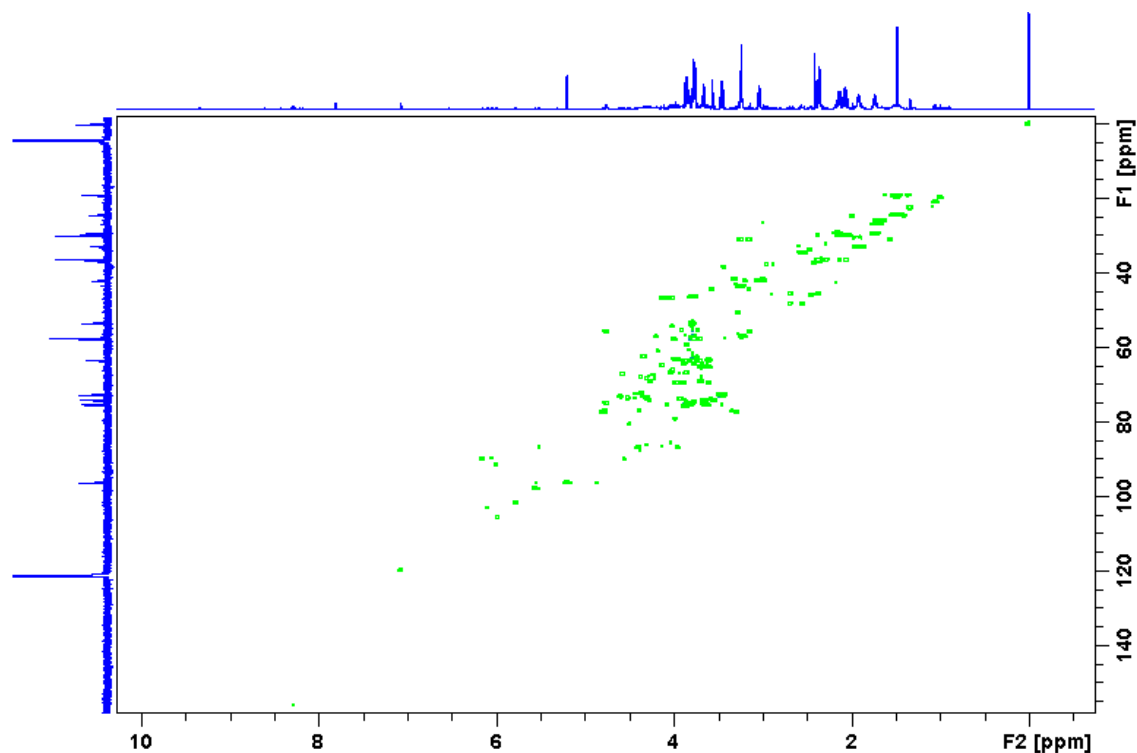


Figure 5.10: 2D-NMR  $^1\text{H}$ - $^{13}\text{C}$  HSQC spectrum of an hydro-alcoholic sample.

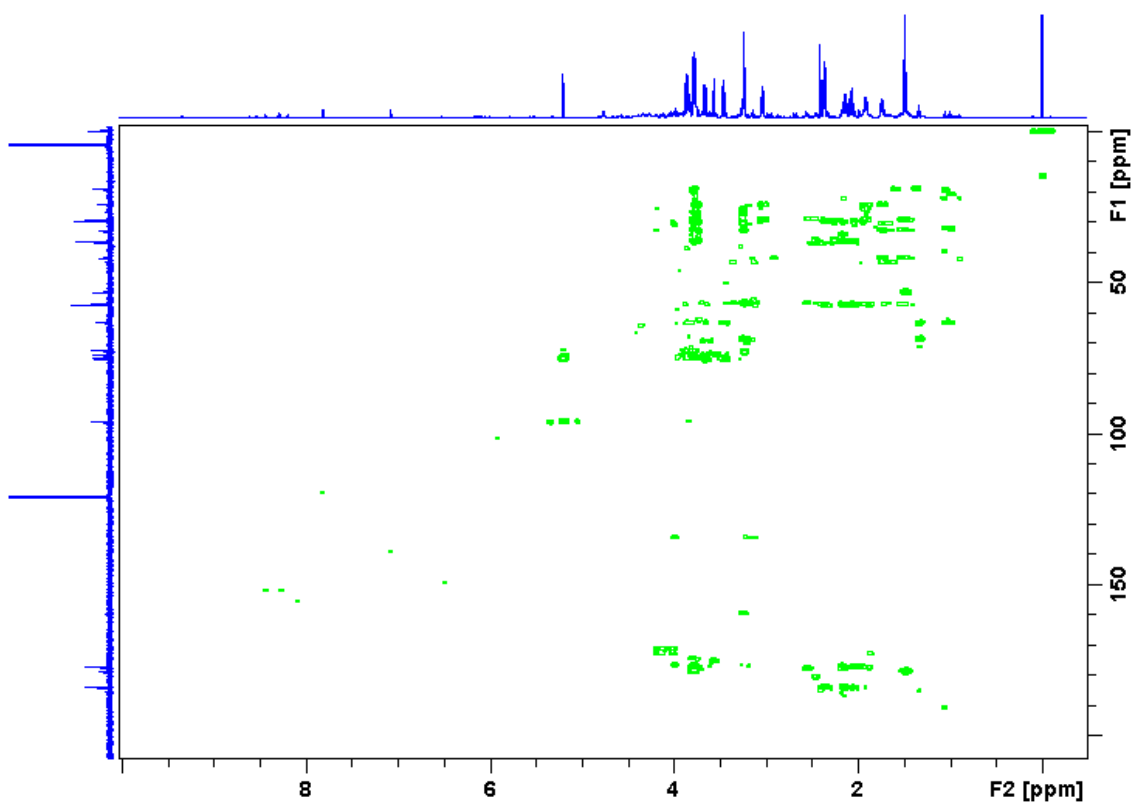


Figure 5.11: 2D-NMR  $^1\text{H}$ - $^{13}\text{C}$  HMBC spectrum of an hydro-alcoholic sample.

Figure 5.12 shows how metabolites univocal assignment was performed thanks to the aid of the 2D heteronuclear crosspeaks, taking glutamate (Glu) as an example.

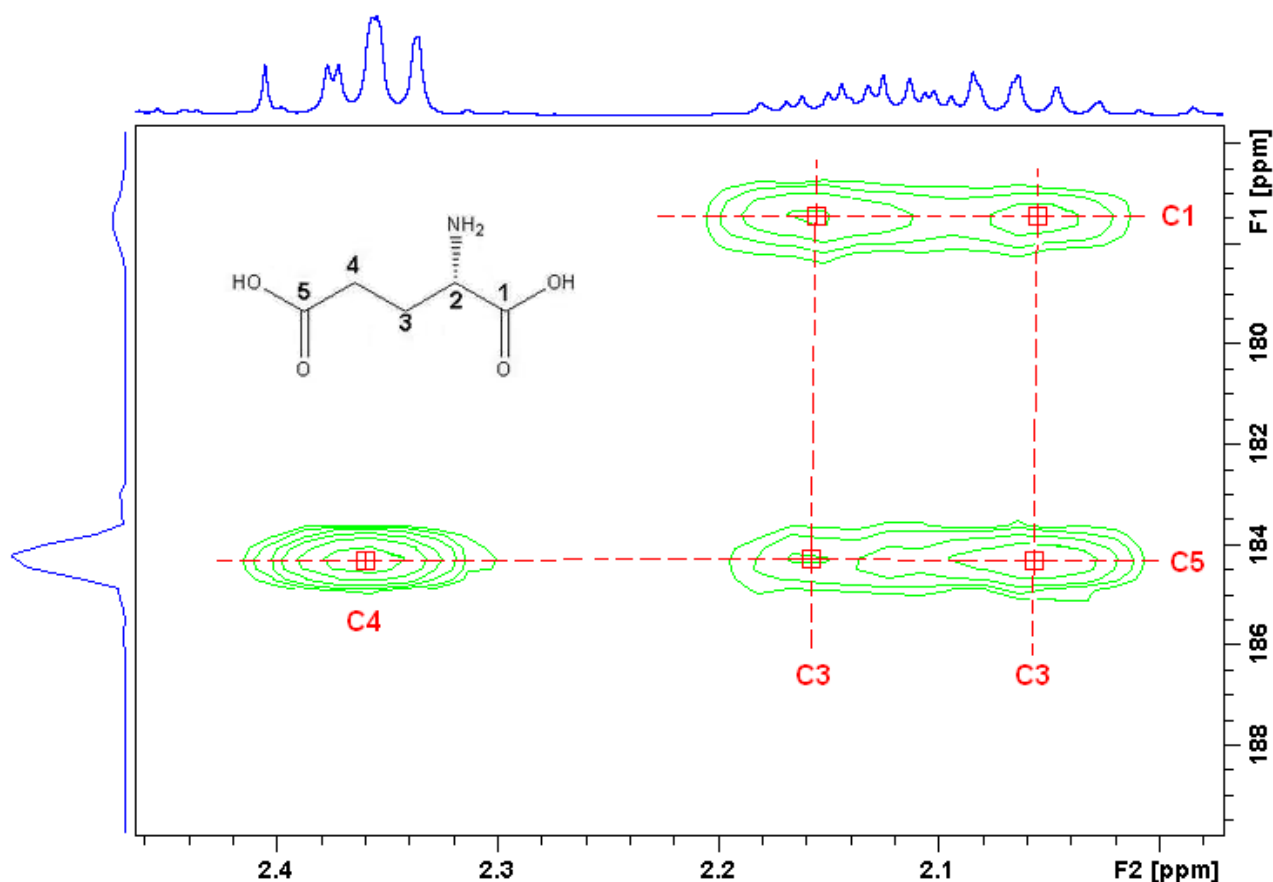


Figure 5.12: 2D-NMR  $^1\text{H}$ - $^{13}\text{C}$  HMBC spectrum detail of an hydro-alcoholic sample. There are reported the carbon assignment of the crosspeaks observed relative to Glu.

In the direct dimension (F2)  $^1\text{H}$  chemical shift is reported, whereas the indirect dimension (F1) is a  $^{13}\text{C}$  chemical shift axe. At 2.07 and 2.12 ppm there are the double-triplets of the protons in C3. These protons are diasterotopic: indeed, they show crosspeaks at the same value of  $^{13}\text{C}$  chemical shift: one correlation at 177.4 ppm, relative to the long range coupled C1, and another crosspeak, at 184.3 ppm, relative to the long range coupled C5. These  $^{13}\text{C}$  assignments are confirmed by the third crosspeak shown in figure 5.12, relative to the  $\text{CH}_2$  in C4 (resonating at 2.36 ppm): also these protons give a correlation at 184.3 ppm, possible only with the long range coupled C5.

All the assignments are also confirmed by comparison with literature data [68,132-144].

The heteronuclear 2D experiments resulted particularly useful for the assignment of the signals observed in  $^{13}\text{C}$  1D NMR spectra, acquired at 600 MHz (see 5.1.4 paragraph).

Moreover, this high magnetic field allows a better resolution of the  $^1\text{H}$  signals as compared to the spectra acquired at 400 MHz. This can be seen in figure 5.13 where is shown, as an example, the spectral zone between 0.80 and 1.1 ppm in the 1D  $^1\text{H}$  spectra acquired at 400 MHz and 600 MHz, respectively. It is possible to note as the spectral profile is more resolved in the spectrum acquired at 600 MHz (grey): there are better defined the triplet at 0.90 ppm (butyrate), at 0.94 ppm (isoleucine) and the doublet at 1.02 ppm (isoleucine), under the higher doublet of valine at 1.00 ppm. Above all, under the other doublet of valine at 1.05 ppm, is more visible the signal at 1.08 ppm (isobutyrate), clearly a doublet: so, the better resolution of the spectra acquired at 600 MHz, highlights the multiplicity of the signals, allowing an assignment improvement.

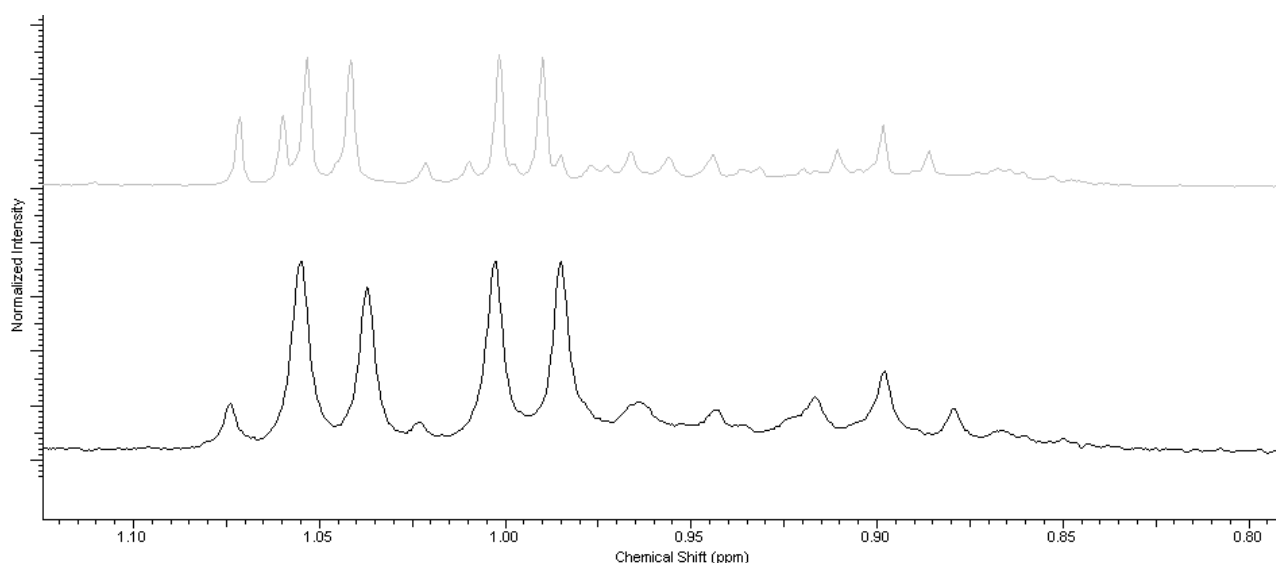


Figure 5.13: 1D  $^1\text{H}$  of samples acquired at 400 MHz (black) and at 600 MHz (grey), respectively.

This can be seen also in 2D spectra: TOCSY spectrum acquired at 400 MHz seemed to suggest that the signal at 1.08 was a singlet, because no crosspeaks were observed (fig. 5.14). At 600 MHz, instead (fig. 5.15), the clear correlation at 1.08-2.39 ppm, allowed to assign the signal (a doublet) to Iso-Btr.

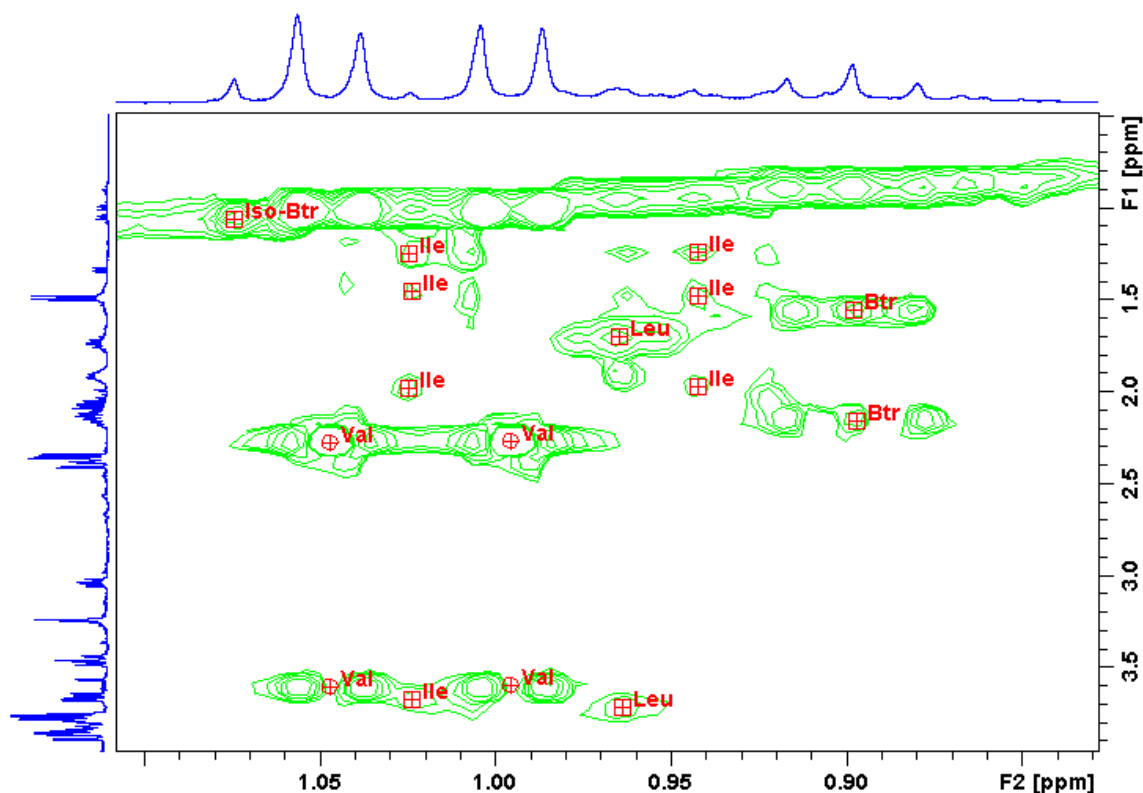


Figure 5.14: TOCSY spectrum acquired at 400 MHz, relative to the spectral zone between 0.8 and 1.1 ppm in F2. There are reported the assignment of the crosspeak observed.  
No crosspeak was observed for the Iso-btr signal at 1.08 ppm, suggesting it is a singlet.

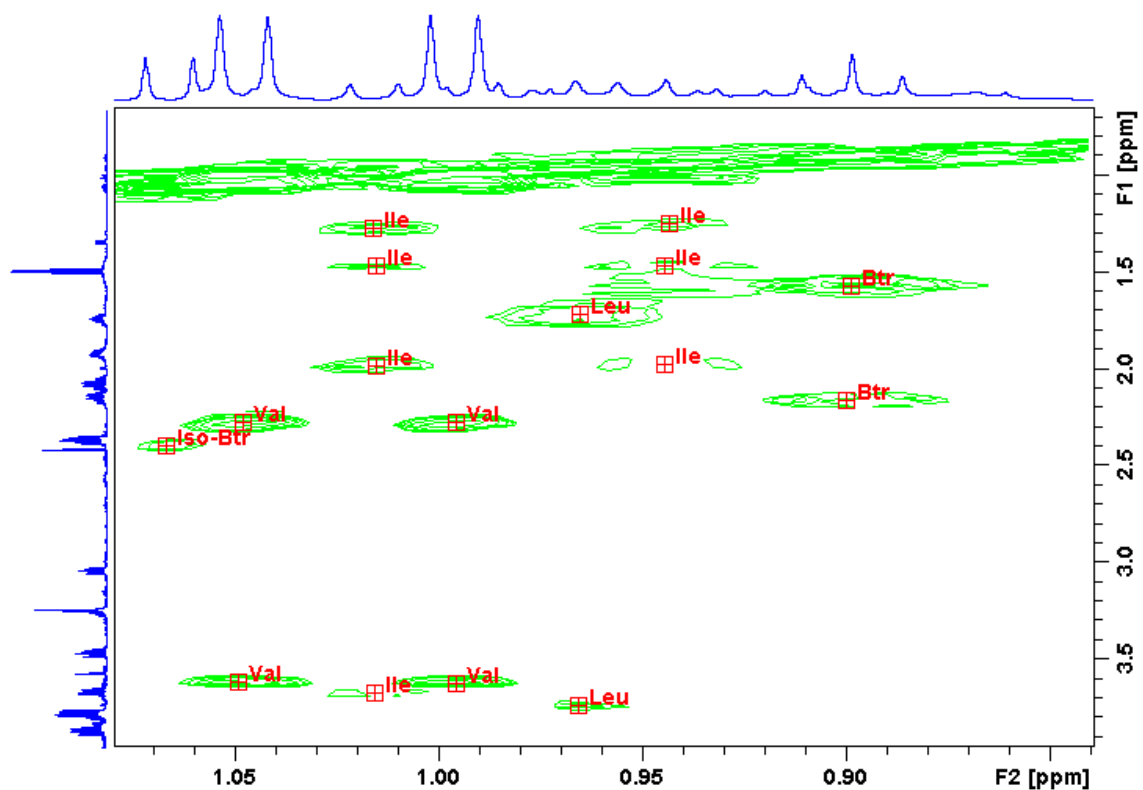


Figure 5.15: TOCSY spectrum acquired at 600 MHz, relative to the same spectral zone displayed in figure 5.14. There are reported the assignment of the crosspeak observed.  
The correlation 1.08-2.39 ppm is clearly visible thanks to the crosspeak named Iso-Btr.

The  $^{13}\text{C}$  chemical shift values assigned thanks to the 2D heteronuclear spectra, together with the resonances observed in the protonic spectra, are reported in the assignment table 5.1, for all the quantified metabolites.

However, the complexity of the mixture constituting each samples allows the quantification of just a subset of the metabolites identified, but the high discriminatory power of the 2D-NMR techniques is able to assign also other metabolites, even if a few only partially (table 5.2).

Table 5.2 - Chemical shift complete pattern of the assigned, but not quantified, metabolites.

Metabolites not integrated	$^1\text{H}$ chemical shifts (ppm)	Observed in
2OH-Butyrate (2OH-Btr)	0.92(t) – 1.54 – 1.66 – 3.99	TOCSY
Leucine (Leu)	0.94(d) – 0.96(d) – 1.72 – 3.72	TOCSY, HSQC
Arginine (Arg)	1.67(m) – 1.90 – 3.24(t) – 3.76(t)	TOCSY, HSQC
Acetate (Ac)	1.92(s)	TOCSY, HSQC
X-Choline (Cho)	1.98(s) – 3.23 – 3.69 – 4.33	TOCSY, HSQC
Proline (Pro)	2.06 – 2.35 – 4.14(t)	TOCSY
Methionine (Met)	2.14 – 2.65(t) – 3.87(t)	TOCSY
Glutamine (Gln)	2.14 – 2.47 – 3.75(t)	TOCSY, HSQC
Glutathione (GSH)	2.17 – 2.56 – 2.95 – 3.78 – 4.00 – 4.56	TOCSY, HSQC
Acetoacetate (AcAc)	2.30(s) – 3.45(s)	TOCSY
$\alpha$ -ketoglutarate ( $\alpha$ -KG)	2.44 – 3.00	TOCSY
Citrate (Cit)	2.45(dd) – 2.66(dd)	TOCSY
Asparagine (Asn)	2.85(dd) – 2.95 – 4.00	TOCSY, HSQC
Ethanolamine (Etn)	3.15 – 3.83	TOCSY
Sucrose (Suc)	3.45 – 3.66 – 3.83 – 5.41(d)	TOCSY
Phenylalanine (Phe)	7.33(d) – 7.38(d) – 7.42(d)	TOCSY

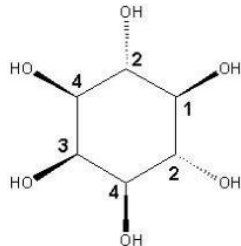
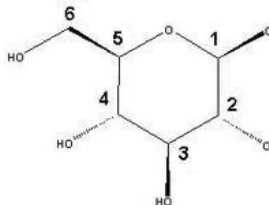
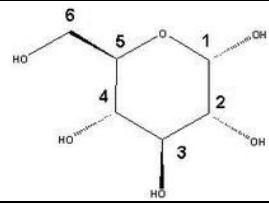
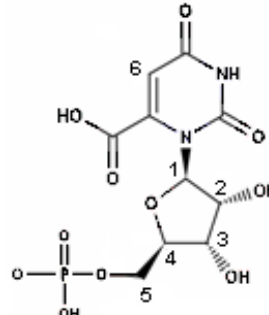
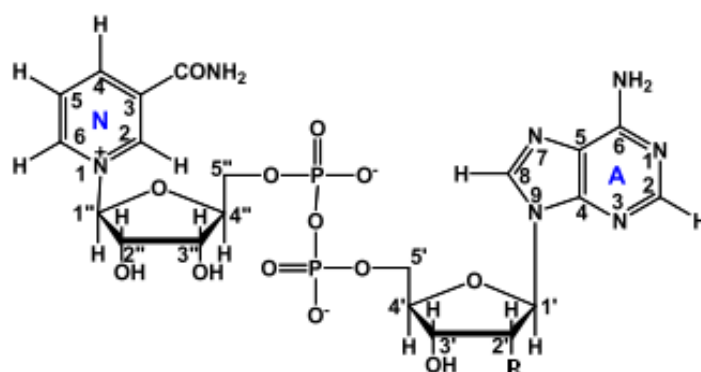
On the other hand, another subset of metabolites are quantified but not univocally assigned yet: they are indicated as “U”. For example, in the hydro-alcoholic extracts there were observed singlets at 2.98 (U1), 8.27 (U5), 8.31 (U6) and 8.62 ppm (U7), all of them more intense in mutant samples, and doublets at 5.31 (U2), 5.52 (U3) and 6.12 ppm (U4); in the chloroformic extracts there was a singlet at 1.42 ppm (U8), instead. Thanks to the 2D homonuclear and heteronuclear experiments there were individuated correlation patterns also for these metabolites, as reported in table 5.1.

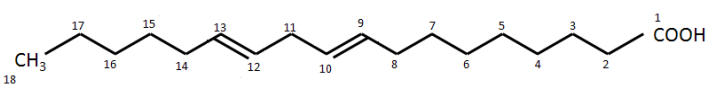
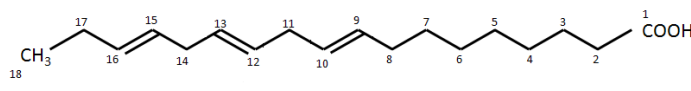
Table 5.1 - Chemical shift complete pattern of the quantified metabolites, assigned through the analysis of 1D and 2D NMR spectra. "U" indicates metabolites not univocally assigned yet.

Metabolites integrated	Structure	Group	$\delta^1\text{H}$ (ppm)	$\delta^{13}\text{C}$ (ppm)
<i>Aminoacids</i>				
Isoleucine (Ile)		6CH <sub>3</sub> 4CH <sub>3</sub> 5-5'CH <sub>2</sub> 3CH 2CH 1COOH	0,94(t) <b>1,02(d)</b> 1,25 - 1,46 1,98 3,69(t)	
Valine (Val)		5CH <sub>3</sub> 4CH <sub>3</sub> 3CH 2CH 1COOH	1,00(d) <b>1,05(d)</b> 2,27 3,60	19.59 20.82 31.97 63.35
Allo-threonine (Allo-Thr)		4CH <sub>3</sub> 2CH 3CH 1COOH	<b>1,20(d)</b> 3,97 4,22	63.03 68.00
Threonine (Thr)		4CH <sub>3</sub> 2CH 3CH 1COOH	<b>1,33(d)</b> 3,58 (d) 4,24 (qui)	22.31 63.35 68.85
Lysine (Lys)		4-4'CH <sub>2</sub> 5CH <sub>2</sub> 3CH <sub>2</sub> 6CH <sub>2</sub> 2CH 1COOH	1,48(m) - 1,51 1,73(m) 1,89 <b>3,04(t)</b> 3,75(t)	24.21 29.27 32.78 41.97 57.49 177.53
Alanine (Ala)		3CH <sub>3</sub> 2CH 1COOH	<b>1,48(d)</b> 3,77(q)	19.05 53.43 178.8
Aspartate (Asp)		3-3'CH <sub>2</sub> 2CH 1COOH 4COOH	2,68(dd) - <b>2,82(dd)</b> 3,88	39.39
Histidine (His)		6-6'CH <sub>2</sub> 5CH 4CH 2CH 1COOH	3,12(dd) - 3,23(dd) 3,97 <b>7,06(s)</b> 7,78(s)	30.90 119.75 138.93
Tryptophane (Trp)		11-11'CH <sub>2</sub> 10CH 6CH 5CH 4CH 8CH 7CH	3,30(dd) - 3,54(dd) 4,12 7,20(dd) 7,28(dd) <b>7,32(s)</b> 7,54(d) 7,74(dd)	57.35

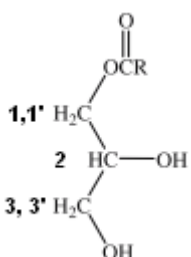
Small organic acids	Glycine (Gly)		2CH <sub>2</sub> 1COOH	<b>3,56(s)</b>	44.28 175.39
	Butyrate (Btr)		4CH <sub>3</sub> 3CH <sub>2</sub> 2CH <sub>2</sub> 1COOH	<b>0,90(t)</b> 1,56 2,16(t)	22.17 42.49 186.97
	Isobutyrate (Iso-Btr)		3CH <sub>3</sub> 2CH 1COOH	<b>1,08(d)</b> 2.39	22.29 39.80 190.80
	Lactate (Lac)		3CH <sub>3</sub> 2CH 1COOH	<b>1,33(d)</b> 4,10 (q)	22.87
	Glutamate (Glu)		3-3'CH <sub>2</sub> 4CH <sub>2</sub> 2CH 1COOH 5COOH	2,07(dt) - 2.12(dt) <b>2,36(dt)</b> 3,76 (t)	28.86 36.39 57.49 177.4 184.3
	Malate (Mal)		3-3'CH <sub>2</sub> 2CH 1COOH 4COOH	2,37(dd) - <b>2,67(dd)</b> 4,30(dd)	45.50 73.28
	Succinate (Succ)		2CH <sub>2</sub>	<b>2,41(s)</b>	37.09
	Fumarate (Fum)		2CH	<b>6,52(s)</b>	138.19
	Formate (For)		1CH	<b>8,46(s)</b>	
	α-Trehalose (Tre)		4CH 2CH 6CH <sub>2</sub> 5CH 3CH 6'CH <sub>2</sub> 1CH	3,46(t) 3,66(dd) 3,75 3,78 3,84 3,88 <b>5,20(d)</b>	72.55 73.90 63.45 74.92 75.39 63.36 96.11

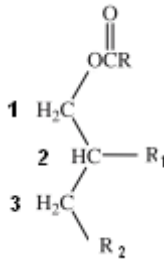
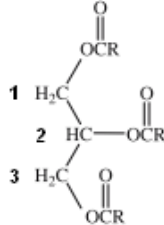
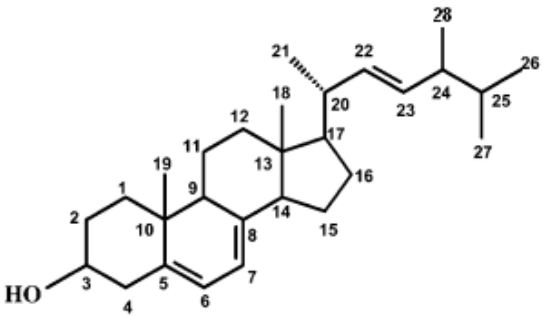


Myo-Inositol (MI)		1CH	3,28(t)	77.20
		4CH	3,52	73.89
		2CH	3,61	75.20
		3CH	<b>4,06(t)</b>	75.05
β-Glucose (β-Glc)		2CH	3,24	78.55
		3CH	3,46	
		6-6'CH <sub>2</sub>	3,73 - 3,88	98.90
		1CH	<b>4,65(d)</b>	
α-Glucose (α-Glc)		4CH	3,41	63.45
		2CH	3,55	
		3CH	3,70	
		6CH <sub>2</sub>	3,84	
		1CH	<b>5,24(d)</b>	95.00
<i>Other hydro-soluble molecules</i>				
Orotidine 5phosphate (OMP)		5-5'CH <sub>2</sub>	3,78 - 3,87	72.03
		4CH	3,95	
		3CH	4,35(t)	
		2CH	4,75	
		1CH	<b>5,55(d)</b>	
		6CH	5,77(s)	101.46
<i>NAD(P)<sup>+</sup></i>				
	R = OH (NAD <sup>+</sup> ) or OPO <sub>3</sub> <sup>=</sup> (NADP <sup>+</sup> )	Rib5'' CH <sub>2</sub>	4,24	68.00
		Rib4' CH	4,38	86.66
		Rib2'', 3' CH	4,51(dd)	89.52
		Rib2' CH	4,81	
		Rib1' CH	6,04 (d)	
		A2 CH	8,18(s)	155.81
		A8 CH	8,43(s)	142.63
		Rib5' CH <sub>2</sub>	4,22	67.49
		Rib3'' CH	4,43	73.36
		Rib4'' CH	4,55	89.77
		Rib1'' CH	6,09(d)	102.80
		N5 CH	8,20(t)	131.52
		N4 CH	8,84(dd)	148.58
		N6 CH	9,15(d)	145.21
		N2 CH	<b>9,33(s)</b>	142.80
<i>Unassigned hydro-soluble molecules</i>				
U1			<b>2,98(s)</b>	
			3,70	
U2			4,11	89.20
			4,25	
			4,31	
			<b>5,31(d)</b>	

		3,88	
		4,01	
		4,43	
U3		4,78	74.70
		<b>5,52(d)</b>	97.78
		3,89	
		4,04	
		4,59	
U4		<b>4,97(t)</b>	75.31
		6,12(d)	91.94
U5		<b>8,27(s)</b>	155.77
U6		<b>8,31(s)</b>	143.65
U7		<b>8,62(s)</b>	143.07
<i>Fatty acids</i>			
Fatty acids (FA)	$\alpha$ CH <sub>2</sub>	2.35	
	$\beta$ CH <sub>2</sub>	<b>1.59</b>	
	CH <sub>2</sub> allylic	<b>2.02</b>	
	CH vinylic	5.33	
	CH <sub>2</sub> diallylic	2.80	
	(CH <sub>2</sub> ) <sub>n</sub>	1.29	
	$\omega$ CH <sub>3</sub>	0.88	
Linoleic acid (LLA)	18CH <sub>3</sub>	0.89	14.44
	4-8,14-17CH <sub>2</sub>	1.29 (m)	23.13, 29.57
	3CH <sub>2</sub>	1.59 (m)	25.27
	8,14CH <sub>2</sub>	2.04	27.75
	2CH <sub>2</sub>	2.27	34.63
	11CH <sub>2</sub>	<b>2.76 (t)</b>	26.05
	9,10,12,13CH	5.32	128.57, 130.43
Linolenic acid (LNA)	18CH <sub>3</sub>	<b>0.96 (t)</b>	14.92
	4-8CH <sub>2</sub>	1.31	29.57
	3CH <sub>2</sub>	1.59 (m)	25.27
	8,17CH <sub>2</sub>	2.04	20.68, 27.75
	2CH <sub>2</sub>	2.32	34.63
	11,14CH <sub>2</sub>	2.79 (t)	26.05
		9,10,12,13CH	5.33
			128.57, 130.43

*Acyl-glycerols*

1-monoacyl glycerols (1-MAG)		1CH <sub>2</sub>	<b>3.59 (dd)</b>
		3CH <sub>2</sub>	3.69 (dd)
		2CH	3.92 (m)
		1'CH <sub>2</sub>	4.14 (dd)
		3'CH <sub>2</sub>	4.20 (dd)

Diacylglycerols (DAG)			3CH <sub>2</sub> 1CH <sub>2</sub> 2CH	<b>3.72 (dd)</b> 4.24 (dd) 5.06	
Triacylglycerols (TAG)			1/3CH <sub>2</sub> 1/'3'CH <sub>2</sub> 2CH	4.13 <b>4.36 (dd)</b> 5.24 (dd)	62.94 62.94 70.01
<i>Sterols</i>	Ergosterol (Erg)		18CH <sub>3</sub>	0.62 (s)	12.75
			27CH <sub>3</sub>	0.81 (d)	23.70
			26CH <sub>3</sub>	0.83 (d)	20.65
			28CH <sub>3</sub>	0.91 (d)	18.3
			19CH <sub>3</sub>	0.94 (s)	16.99
			21CH <sub>3</sub>	1.03 (d)	21.79
			1CH <sub>2</sub>	1.29 (t)	39.1
			2CH <sub>2</sub>	1.47 (m)	32.65
			24CH	1.88 (m)	43.5
			20CH	2.04 (m)	41.10
			4CH <sub>2</sub>	2.46 (t/dd)	41.5
			3CH	3.63 (m)	71.15
			22CH	5.17 (dd)	136.2
			23CH	5.22 (dd)	132.7
			7CH	5.38 (d)	116.99
			6CH	<b>5.56 (d)</b>	120.31
			11-15CH <sub>2</sub>		21.8
			12CH <sub>2</sub>		29.0
			25CH		33.8
			10C		37.75
			16CH <sub>2</sub>		39.8
			13C		43.5
			9CH		46.95
			14CH		55.25
			17CH		56.46
			5C		140.5
			8C		142.05
<i>Unassigned lipo-soluble molecules</i>					
	U8			<b>1.42 (s)</b>	30.40

Multiplicity: (s): singlet; (d): doublet; (dd): double doublet; (t): triplet; (q): quartet; (qui): quintet; (m): multiplet.

About the not univocally assigned molecules, U4 is probably a ribose moiety bounded to a nucleoside: TOCSY correlation at 3.89 ppm suggested the absence of phosphoric groups. Its resonance at 6.12 ppm could be relative to the group involved in the glycosidic bond (C1', as in figure 5.16), supported also by the  $^{13}\text{C}$  chemical shift value found for this signal in the HSQC spectrum.

Analogously, the resonance at 8.31 ppm (U6) should be due to adenosine proton in A8 (fig. 5.16), whereas the signal at 8.27 ppm (U5) could be relative to the proton in A2 of adenosine triphosphate (ATP). Both of these assignments are supported by their  $^{13}\text{C}$  chemical shift value.

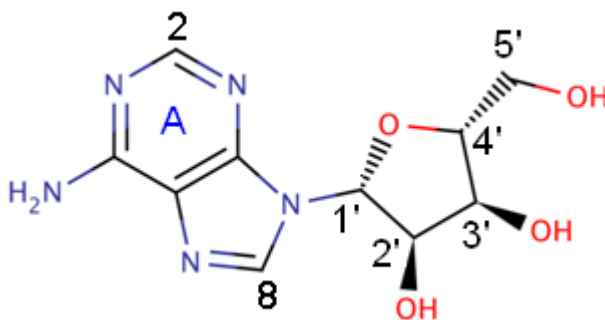


Figure 5.16: Adenosine structure.

Finally, U7 may be assigned as ADP, because 8.62 ppm is a value similar to the proton attached to the A8 carbon atom in adenosine moiety.

### 5.1.4 $^{13}\text{C}$ -NMR spectra from labeling experiment

Despite the enhanced resolution of the  $^{13}\text{C}$  spectra acquired at 600 MHz, they result always in a low signal-to-noise ratio, especially the spectra acquired at natural abundance, because of the intrinsic low sensitivity of the  $^{13}\text{C}$  nucleus.

Figure 5.17 and 5.18 show, for wild type and mutant hydro-alcoholic extracts, respectively, a spectrum acquired at natural abundance (*control sample*) respect to a spectrum relative to sample derived from labeling experiment (*enriched sample*). Figure 5.19 and 5.20 shows the same, but for organic extracts. Signals around 120 ppm and 4 ppm are relative to the cyanide and methyl carbon, respectively, of the  $\text{CH}_3\text{CN}$  quantitative  $^{13}\text{C}$ -reference.

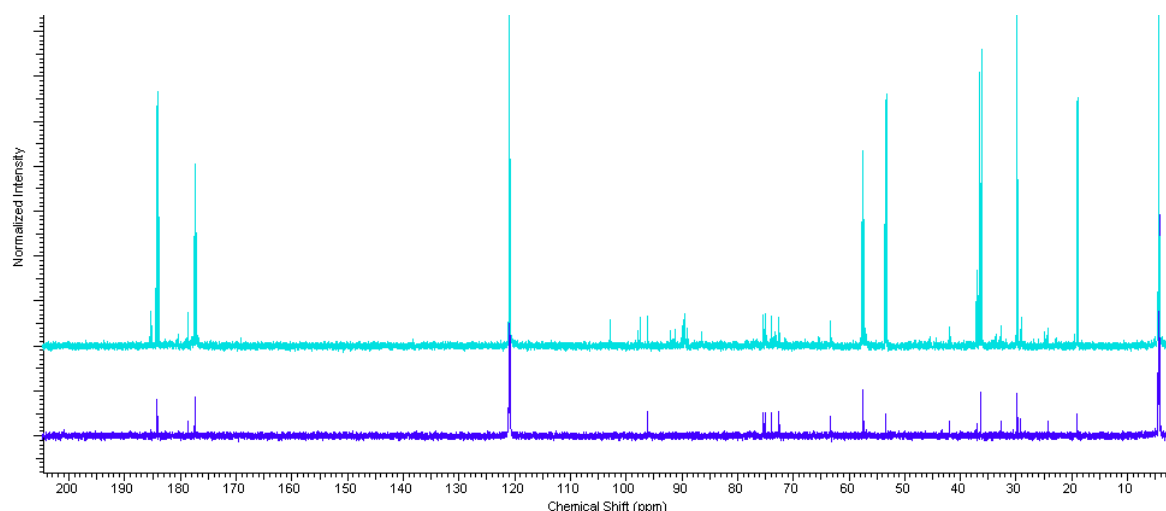


Figure 5.17: 1D  $^{13}\text{C}$  spectra of a control (blue) and an enriched (light blue) wild type hydro-alcoholic sample.

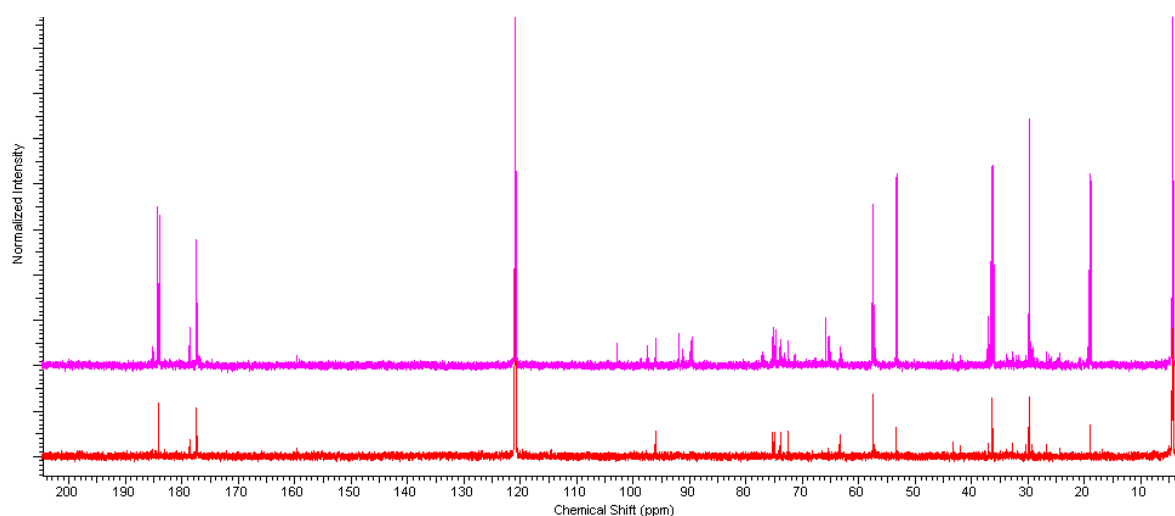


Figure 5.18: 1D  $^{13}\text{C}$  spectra of a control (red) and an enriched (pink) mutant hydro-alcoholic sample.

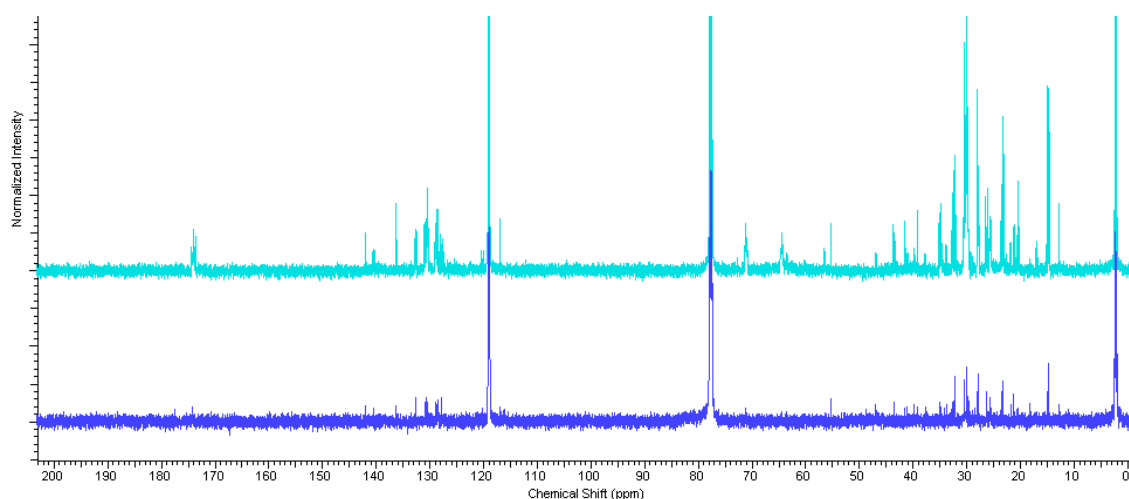


Figure 5.19: 1D  $^{13}\text{C}$  spectra of a control (blue) and an enriched (light blue) wild type organic sample.

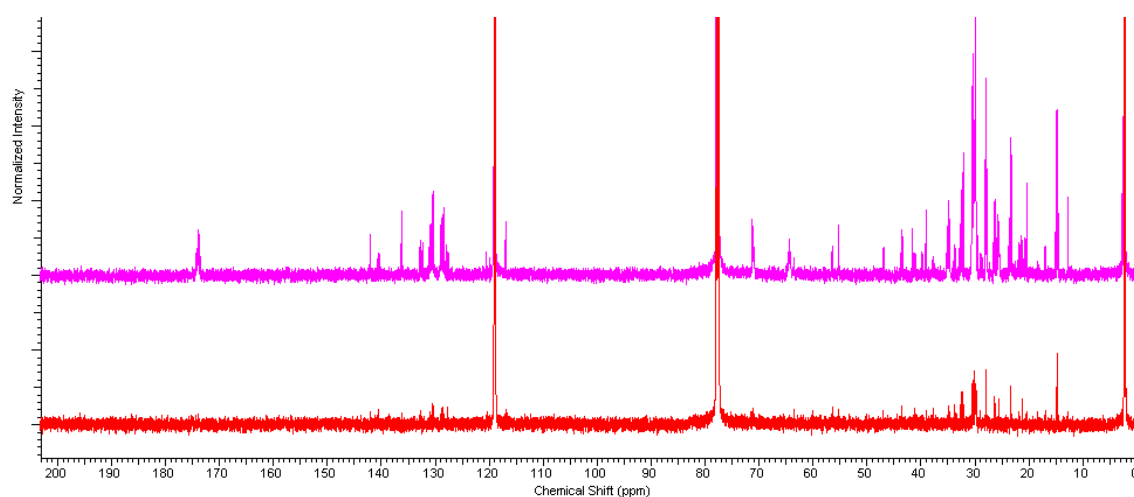


Figure 5.20: 1D  $^{13}\text{C}$  spectra of a control (red) and an enriched (pink) mutant organic sample.

Note as the resonances in the spectra obtained after infusion of  $^{13}\text{C}$ -labeled substrate appear to be enhanced compared to that of natural abundance spectra. These increased intensities indicate that the  $^{13}\text{C}$  label has been incorporated in the metabolites [2,90].

In both hydro-alcoholic and organic control spectra, resonances of the most abundant metabolites found in these extracts were observed at the corresponding  $^{13}\text{C}$  chemical shift value, as reported in table 5.1. Some of these signals are relative to carbon atoms of molecules that do not show relevant incorporation of  $^{13}\text{C}$  in the enriched spectra, for example: valine and trehalose in hydro-alcoholic samples and glycerol backbone of acyl-glycerols in organic extracts. For these metabolites the enhancement of the  $^{13}\text{C}$  resonance is not relevant and so they are simply considered not enriched.

Alanine (Ala), glutamate (Glu), lysine (Lys) and succinate (Succ), in polar extracts, and fatty acids (FA) and ergosterol (Erg) in organic samples, show significant enrichment, instead.

### 5.1.5 Isotopomer analysis

The introduction of  $^{13}\text{C}$ -labeled molecules into a bioreaction network yields non-randomly  $^{13}\text{C}$ -labeled metabolites that can be measured quantitatively by  $^{13}\text{C}$  NMR. The labeling leads to the selective enhancement of signals from nuclei with low natural abundance (1.1%). Moreover, if there is one - or more - adjacent carbon atom labeled, a complex spectral pattern around the central resonance will be observed, corresponding to the multiplicity given by the mononuclear  $^{13}\text{C}$ - $^{13}\text{C}$  direct coupling constant ( $^1J_{\text{CC}}$ ). In a  $^{13}\text{C}$  NMR spectrum, in fact, the isotopomer distribution is resolved in detail because a labeled carbon atom produces different hyperfine splitting signals depending on the labeling state of its direct neighbors [103].

The optimization of NMR measurement parameters and the proper interpretation of all the peak fine structures in these kind of spectra require a skilled NMR specialist. As an example, fig. 5.21 shows the spectral pattern observed around the glutamate resonance relative to its C4 position.

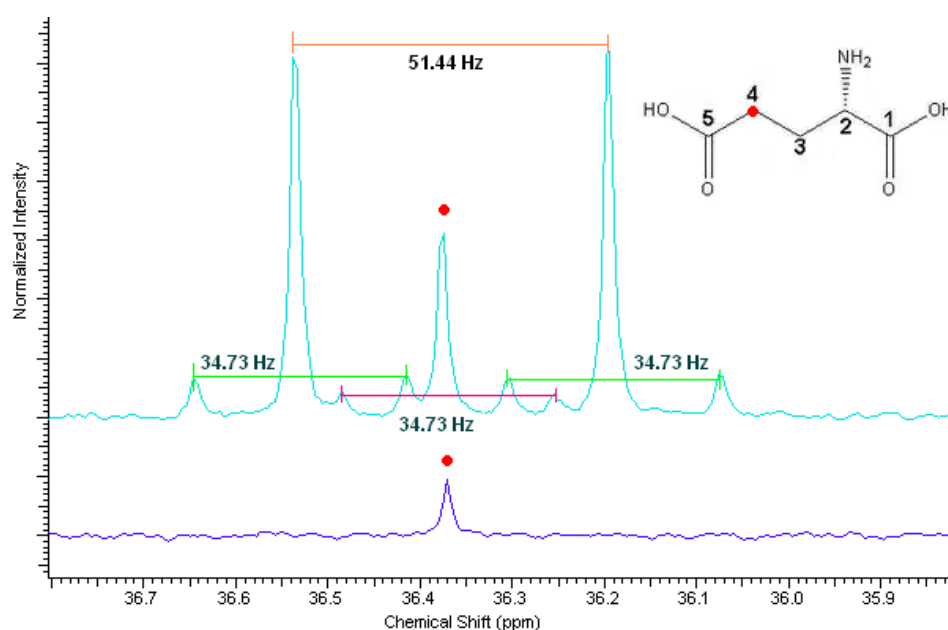


Figure 5.21: Spectral pattern of glutamate C4 in an unlabeled (blue) and an enriched (light blue) WT hydroalcoholic sample. Red dots indicate the enhanced singlet relative to C4. Colored lines connect the doublets (C4-C5, orange; C4-C3, pink) and the double doublet (C5-C4-C3, green) observed. There are also reported the corresponding coupling constants values.

In the control spectrum (blue) there is only the singlet relative to the natural abundant  $^{13}\text{C}$  at the glutamate C4 position (36.39 ppm). The corresponding signal (red dot) in the enriched spectrum (light blue) is increased because of the incorporation of  $^{13}\text{C}$  from the labeled substrate.

Moreover, it is evidently present a doublet, defined by a coupling constant of 51.44 Hz (orange line), characteristic of the coupling between an aliphatic and a carboxylic carbon atom [163,32]: hence this doublet must be relative to molecules of glutamate labeled in both C4 and C5 position. It is also visible a weak doublet, characterised by a  $^1J_{\text{CC}}$  of 34.73 Hz (pink line), representative of the coupling between aliphatic carbons ( $\text{sp}^3\text{-sp}^3$ ) [32,90,164]: this means that, even if in a less amount, the C4-C3 enriched fragment is also present. Finally, there is a double-doublet (green lines) defined by both the coupling constant seen, relative to the C3-C4-C5 enriched fragment.

The same was done for all the other complex patterns observed in both the hydro-alcoholic and organic extracts; then, the fractional enrichment ( $Y_{\text{Ci}}$ ) of each carbon atom, for each fragment, was calculated as described in 4.4.1 Materials and methods.

So, the hydro-alcoholic metabolites resulting enriched are:

- ❖ glutamate (Glu, fig. 5.22): there were observed enhanced singlet signals for C1, C2, C3 and C4 atoms, but not for the C5. Moreover, the doublets observed allow the quantification of the  $\text{C}_2$  fragments C1-C2, C2-C3, C4-C3 and C4-C5. The double doublet observed in C4 allows the quantification of the  $\text{C}_3$  fragment C3-C4-C5 (observed from C4), whereas another  $\text{C}_3$  fragment (C1-C2-C3, observed from C2) was revealed only in WT samples.

Naturally, all of the  $\text{C}_2$  fragments were observed in both the positions involved, for example: there was observed the fragment C1-C2, quantified from the C1 carbon atom, and the same fragment was quantified from the C2 position (C2-C1).



Finally, since  $^1J_{CC}$  relative to the coupling between C3-C2 and C3-C4 atoms are similar, it was not possible to quantify these two fragments separately from the C3 position (Glu C3-C2 or C3-C4). However, no double doublet was observed in C3 position: this means that the C<sub>3</sub> fragment C2-C3-C4 does not occur.

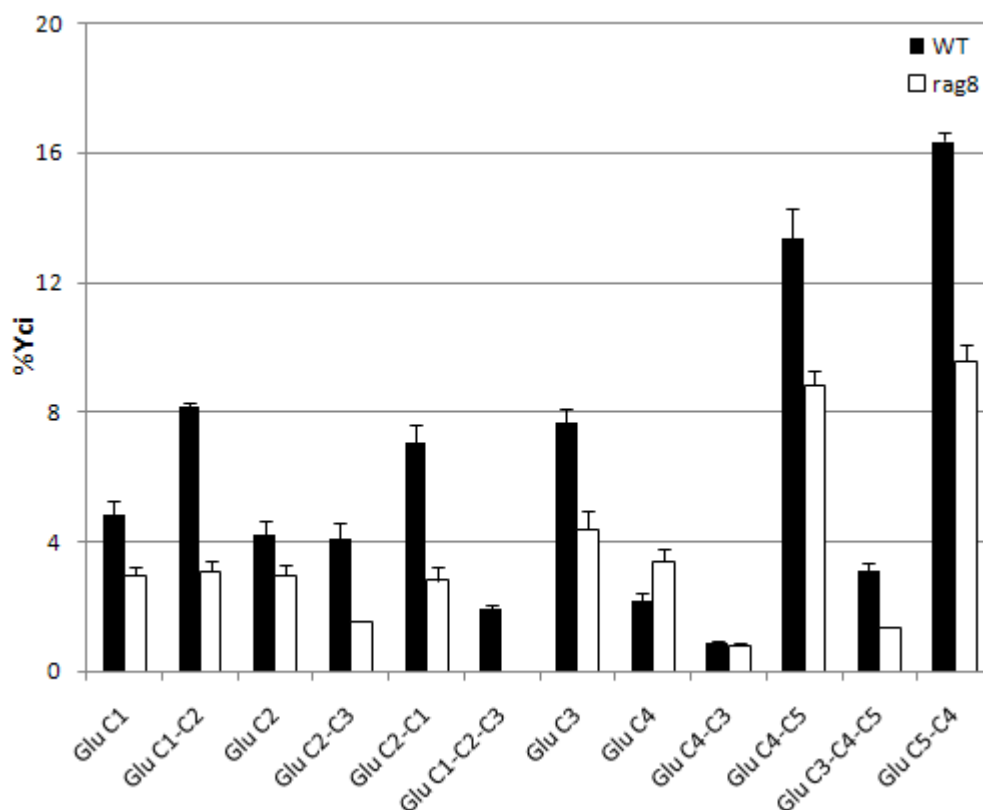


Figure 5.22: Fractional enrichment of the enriched glutamate (Glu) carbon atoms observed. For the C3 position, the coupling constant is unable to discern between C3-C2 and C3-C4 fragments.

- ❖ Alanine (Ala, fig. 5.23): only the singlet in C1 and C3 positions appear enriched. Moreover, the doublet relative to the C2-C3 fragment is observable from both the positions involved (C2-C3 and C3-C2, statistically in agreement each other).
- ❖ Lysine (Lys, fig. 5.24): in this case, both WT and mutant samples show an enrichment relative to the C2 (singlet) and C2-C1 (doublet) fragments exclusively. Enriched singlets were observed in C3 and C5 positions only in the *rag8* mutant samples.

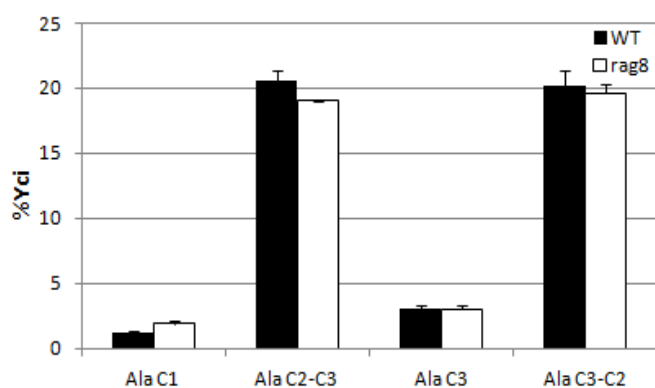


Figure 5.23: Fractional enrichment of the enriched alanine (Ala) carbon atoms observed.

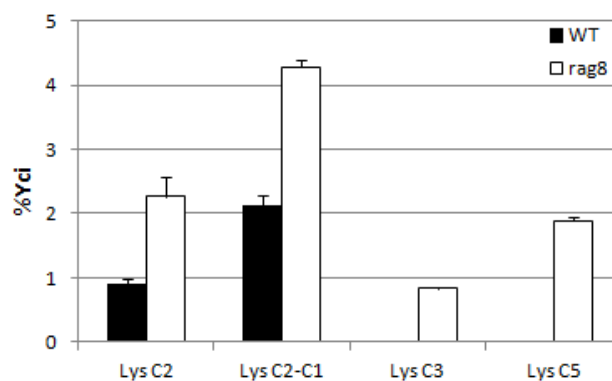


Figure 5.24: Fractional enrichment of the enriched lysine (Lys) carbon atoms observed.

❖ Succinate (Succ, fig. 5.25): this is a symmetric molecule that leads to a scrambling of the labeling during its generation in both TCA and glyoxylate cycle, so that the enrichment in C1 can not be discern from the enrichment in “C4” position (called, in fact, C1, in fig. 5.25). The same is for C2 and “C3” position (called, anyway, C2). In this case, it was chosen to report the quantification of the fragment C1-C2 from the C2 position only.

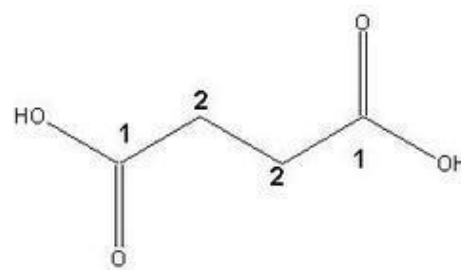
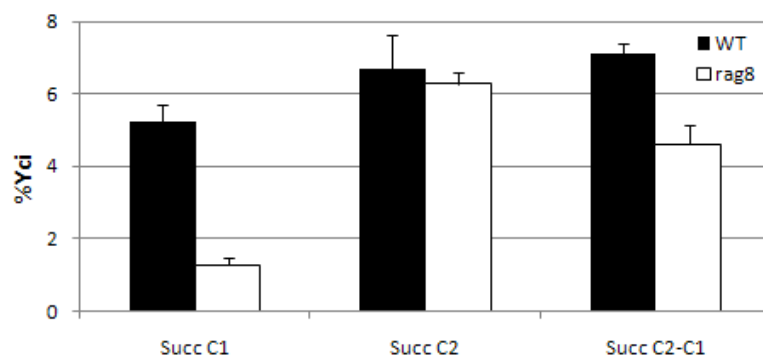


Figure 5.25: Fractional enrichment of the enriched succinate (Succ) carbon atoms observed. It is shown also its molecular structure.

Even if the relative intensities of the multiplet components derived from the  $^{13}\text{C}$  spin–spin scalar coupling fine structure reflects the relative abundance of isotopomers fragments, it must be noted that only the one-bond scalar couplings,  $^1J_{\text{CC}}$ , are usually large enough ( $\sim 30 - 60$  Hz) to be well resolved in  $^{13}\text{C}$ -NMR spectroscopy, while long-range couplings are often too small ( $< 5$  Hz).

Moreover, the dispersion of the  $^1J_{CC}$  values is often insufficient to distinguish the signals from two different spin pairs  $^{13}C_a - ^{13}C$  and  $^{13}C_b - ^{13}C$ , as in the case of C3-C2 and C3-C4 Glu fragments, simply because the corresponding coupling constants are too similar each other [32]. For this reason, a complete isotopomers analysis is often not possible for molecules with more than 3 carbon atoms, so  $^{13}C$ -NMR spectral analysis is limited to revealing the abundance of groups of isotopomers.

About the lipo-soluble metabolites, fatty acids (FA) and ergosterol (Erg) result enriched.

Unfortunately, even if the digital resolution of such spectra is relatively high (0.6 Hz), signals relative to the FA moiety, have too similar  $^{13}C$  chemical shift values: this means that resonances are too near each other, compared to the coupling constants involved, and significant distortions of the line intensities of a given multiplet occur, compromising the integration of the signal for the enrichment quantification because the spectrum become of superior order (strong spin-spin scalar coupling effects, see 4.4 Materials and methods). So the complexity of the mixture is so high that the pattern of multiplets generated by the enrichment around each FA signal do not permit an interpretation through the  $^1J_{CC}$  coupling constants values. However, it was possible to evaluate an averaged value of  $Y_{Ci}$  for the principal FA carbon atoms resulting in an enrichment statistically not different between WT and mutant samples (data not shown).

About ergosterol (Erg), all of its carbon atoms was observed in the  $^{13}C$  NMR control spectra and the enriched positions showed only singlet and/or doublet multiplicity. In particular:

- for the carbon atoms C1, C4, C7, C8, C18, C22 and C27 only singlets were observed at their corresponding  $^{13}C$  chemical shift value.
- The carbon atoms C3, C5, C9, C17, C19 and C26 result enriched as both singlet and doublet, instead.
- No singlets were observed for the carbons C6, C10, C12, C16, C20, C23 and C25; however they gave doublets centered at their  $^{13}C$  chemical shift value.

- The only carbon resulting not enriched was C28. C2 and C14, instead, were not determined because they are overlapped to other signals (C16 of fatty acids moiety and MeOH residue, respectively).
- Finally, C11, C15 and C21 appear together as a complex pattern around 21.8 ppm, too near each other to be considered separately. The same is for C13 and C24, at 43.5 ppm. Their multiplicity (both a singlet and a doublet) will be interpreted later (5.2.2 Isotopomer and flux analysis).

In particular, the complex pattern observed around 132.7 ppm, relative to Erg C23, is due to the overlap with the signals of C16 linolenic acid (LNA). Figure 5.26 shows clearer these signals, allowing to unveil that this complex pattern is defined by:

- a doublet around the enriched Erg C23 singlet (black square), characterized by a  $^1J_{CC}$  of  $\sim 43.5$  Hz, typical of a  $sp^3-sp^2$  carbon coupling [163] and therefore relative to Erg C23-C24 enriched fragment;
- another doublet around the enriched LNA C16 singlet (grey circle), characterized by a  $^1J_{CC}$  of  $\sim 70.5$  Hz, typical of a  $sp^2-sp^2$  carbon coupling [163] and hence relative to the C16-C15 LNA enriched fragment.

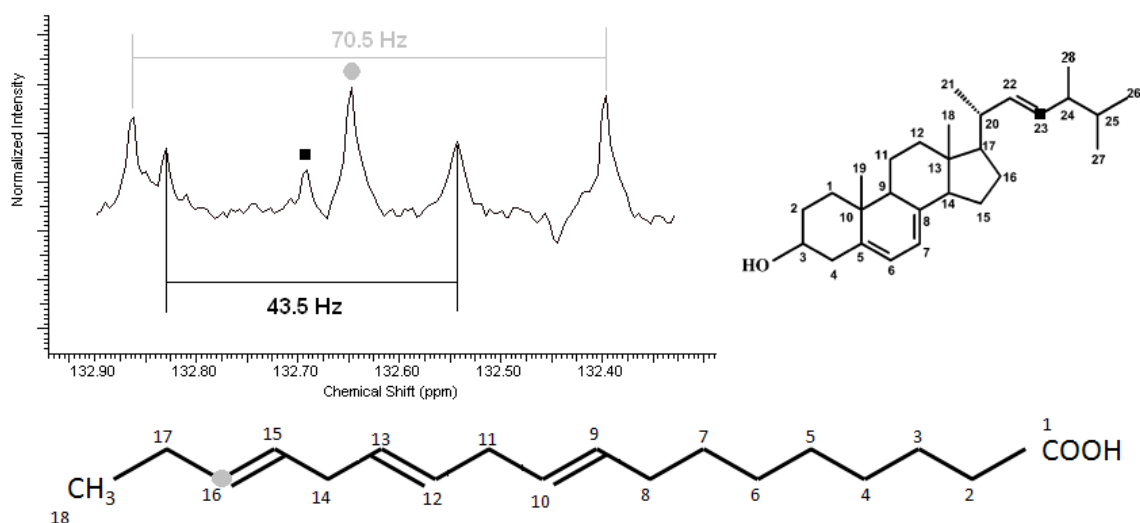


Figure 5.26: Spectral pattern of Erg (structure on the right) C23 (black square) overlapped with LNA (structure at the bottom) C16 (grey circle) signals. There are also indicated the corresponding coupling constants.

This is an example of the fundamental role the coupling constants have in explaining the complex spectral pattern resulted in such spectra.

It could be noted now that these doublets are effectively not centered to the chemical shift value of their corresponding singlet and that they are slightly shifted to higher field, instead: this is because of the  $^{13}\text{C}$  isotope effect (see 4.4 Materials and methods) and this was observed in all the  $^{13}\text{C}$  NMR spectra acquired (see also fig. 5.21).

### 5.1.6 Quantitative analysis

The assignment reported in table 5.1 permits the conceptual passage from a fingerprinting analysis to a metabolic profiling analysis because all changes in the distribution of signals intensities between wild type (WT) and mutant (*rag8*) samples spectra are associated to known metabolites. This allows to understand the most significant chemical-biological differences induced by the genetic manipulation to fluxes distribution and metabolic phenotype of yeast cells.

Integration of metabolite signals reported in **bold** in table 5.1 permits a quantitative analysis of the data. The signal at 1.33 ppm could be relative to both threonine  $\text{CH}_3$  in C4 and lactate  $\text{CH}_3$  in C3 position: they were integrated together because this signal rest the most clear and clean for both the pattern. However, since *K. lactis* does not perform lactic fermentation, it was assumed that lactate could be ignored and that the total amount quantified from this resonance is all relative to threonine.

Each mutant metabolite content reported in figure 5.27 (white bars), is expressed as the mean of the concentration of that metabolite in all mutant samples, compared with the wild type mean value (black bars), with the respective standard deviation, expressed as  $\mu\text{mol}\cdot\text{g}^{-1}$  wet weight. It can be seen that the differences between WT and *rag8* mutant samples are fundamentally quantitative, as the univariate Student's t-test statistically demonstrates.

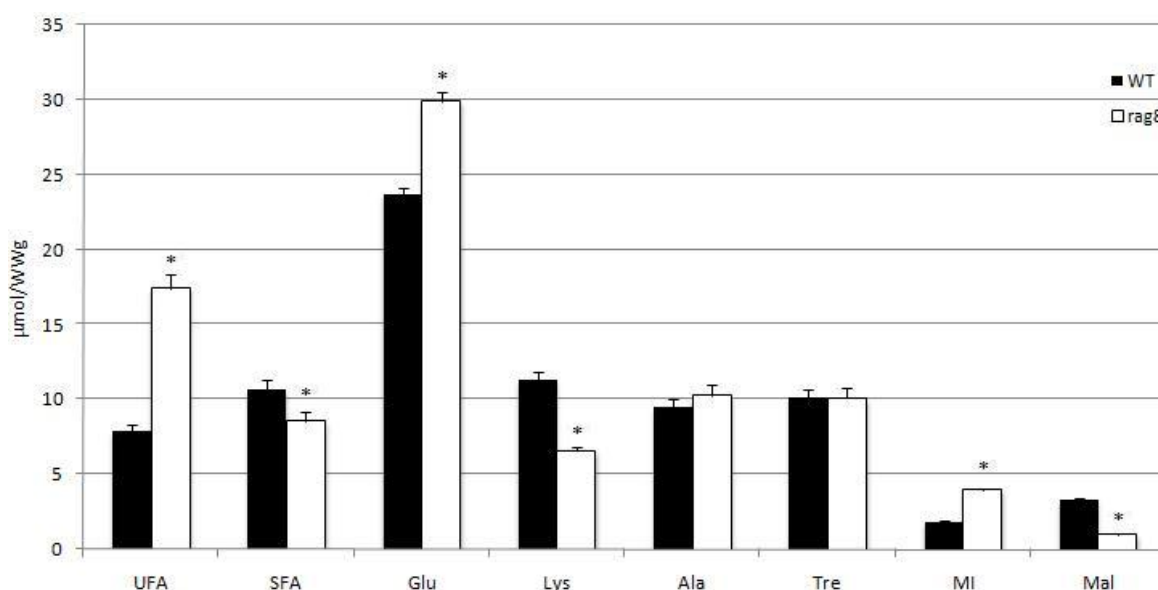


Figure 5.27a: Histograms relative to the concentrations of the metabolites most abundant in hydro-alcoholic and chloroformic extracts. Stars indicate that metabolites whose variation in concentration between wild type (WT) and mutant (*rag8*) samples is statistically significant, at a univariate Student's t-test.

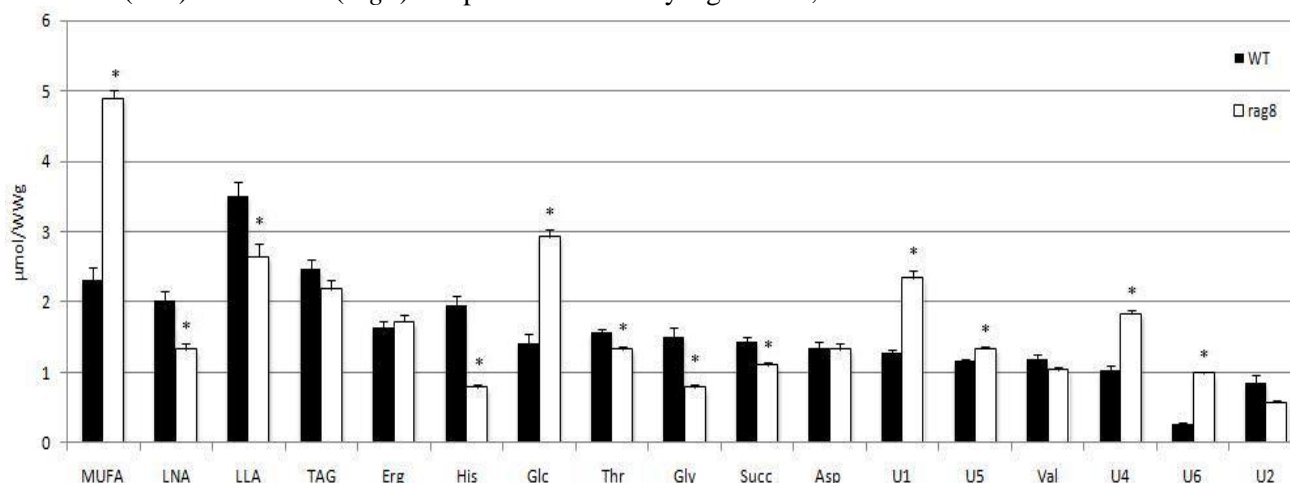


Figure 5.27b: Histograms relative to the concentrations of the metabolites present in intermediate amount in hydro-alcoholic and chloroformic extracts. Stars indicate that metabolites whose variation in concentration between wild type (WT) and mutant (*rag8*) samples is statistically significant, at a univariate Student's t-test.

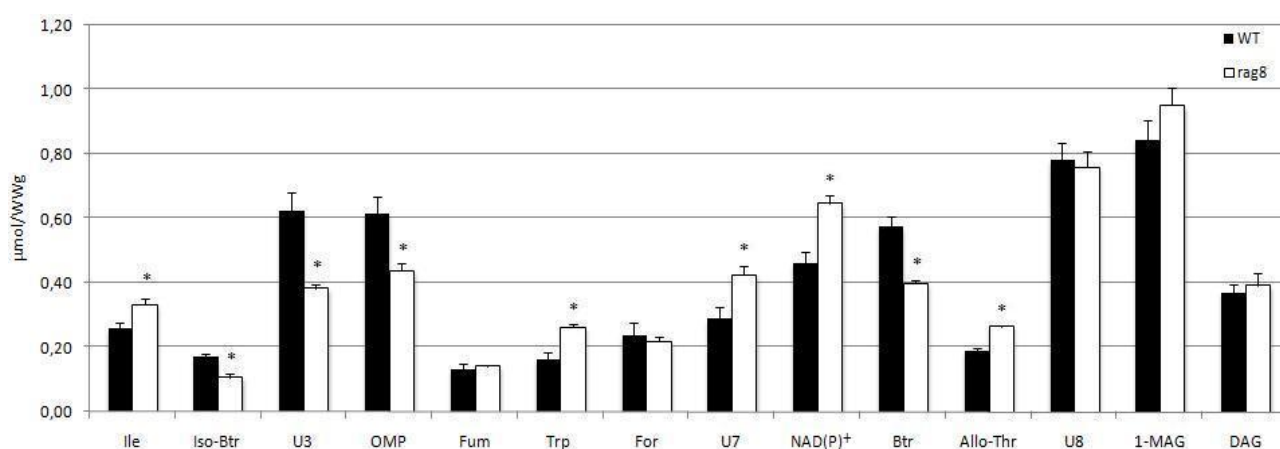


Figure 5.27c: Histograms relative to the concentrations of the metabolites present in trace in hydro-alcoholic and chloroformic extracts. Stars indicate that metabolites whose variation in concentration between wild type (WT) and mutant (*rag8*) samples is statistically significant, at a univariate Student's t-test.

Mutant samples show changes in several metabolite levels: among the ones present at higher concentrations (fig. 5.27a), we observed a significant increase in the unsaturated fatty acid (UFA), glutamate (Glu) and myo-inositol (MI) content and a decrease in the saturated fatty acids (SFA), lysine (Lys) and malate (Mal) levels. Among the metabolites present at lower concentrations (fig. 5.27b and 5.27c), mono-unsaturated fatty acids (MUFA), glucose (Glc), isoleucine (Ile), tryptophane (Trp), allo-threonine (Allo-Thr) and NAD(P)<sup>+</sup> are increased in *rag8*, as compared to the wild type strain, while both poly-unsaturated fatty acids (PUFA) linoleic (LLA) and linolenic (LNA) acids, histidine (His), threonine (Thr), orotidine monophosphate (OMP), glycine (Gly), succinate (Succ), butyrate (Btr) and isobutyrate (Iso-Btr) are reduced.

To visually display the differences induced by the genetic manipulation, two correlation matrixes were built, one for each strain (fig. 5.28), reporting linear correlation Pearson's index values (R) for each couple of metabolites assigned, so matrices appear symmetric. Colors aid us to see how the relationships between couples of metabolites change connected with the genetic mutation: in this sense, these patterns of correlations can be interpreted as a fingerprint of the underlying system.

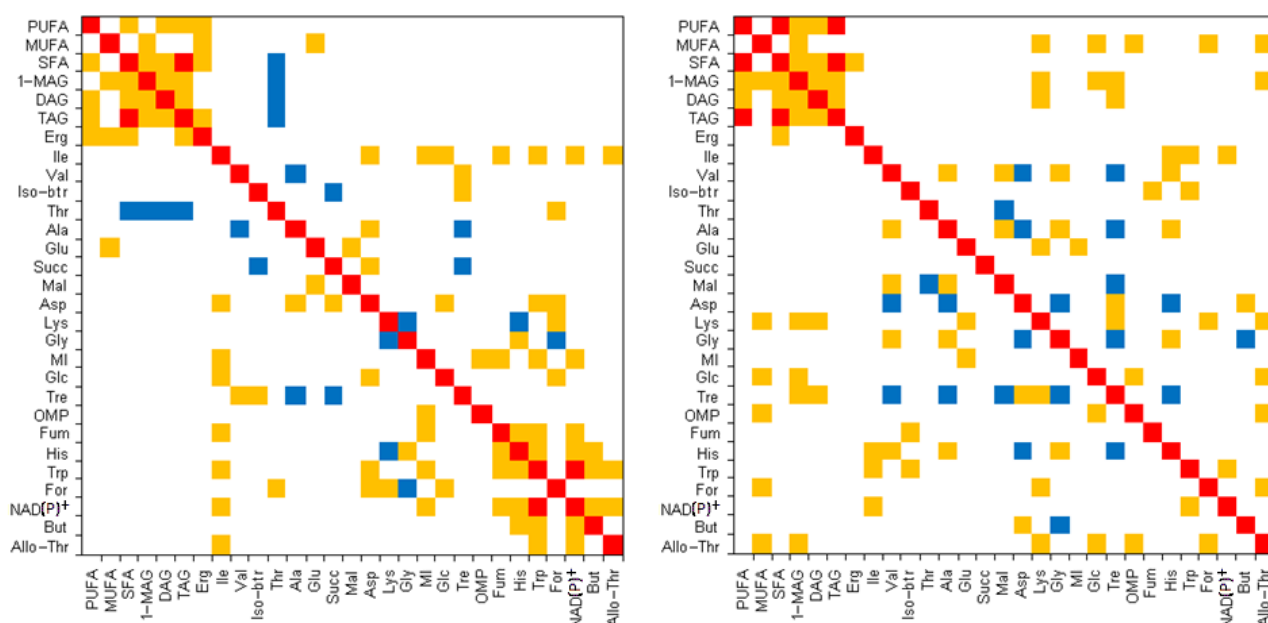


Figure 5.28: Correlation matrix of wild type (left) and mutant (right) samples. Metabolites are reported along the two axes. Red squares indicate a very strong positive correlation ( $R > 0.97$ ), orange squares indicate strong positive correlation ( $0.75 < R < 0.97$ ) whereas blue squares indicate negative correlation ( $R < -0.75$ ).

Detecting and identifying correlation pattern changes is the first and crucial step in understanding the underlying biochemical changes in *K. lactis* cells metabolism (metabolic phenotype). However, while some correlations are in accordance with our intuitive expectations, most others, in particular between seemingly distant metabolites through the metabolic network, cannot be easily understood.

For this reason, Principal Component Analysis (PCA) is performed to discern significant patterns in complex data sets: in fact, this multivariate statistical approach is particularly appropriate in situations in which there are more variables than samples in the data set [62].



## 5.2 Discussion

### 5.2.1 Statistical analysis and metabolic network model

Principal Component Analysis was performed to explore the data field. The new variables (PCs) were displayed in a 2D scores plot, allowing the analysis of the distribution and grouping of the samples in the new variable space and the identification of eventual outliers. Fig. 5.29 shows a clear separation between wild type and mutant samples: the first principal component (PC1), explaining 53% of the total variability, was responsible for this separation among *rag8* and wild type cells. One mutant sample was eliminated because resulted in an outlier at Hotelling's statistic (fig. 5.30).

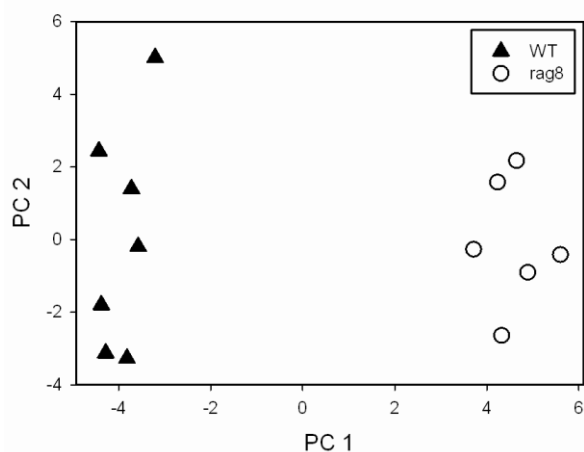


Figure 5.29: PCA scores plot relative to 7 wild type samples (black triangles) and 6 mutant samples (white circle). Total variability represented: 69% (PC1: 53%; PC2: 16%)

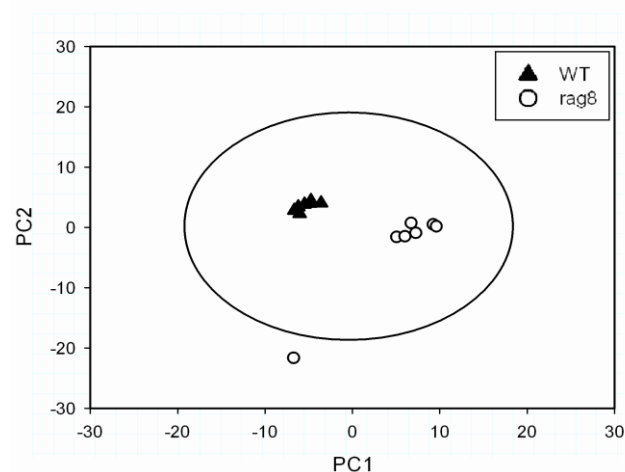


Figure 5.30: Hotelling's plot; one *rag8* sample results in an outlier.

So, in fig. 5.31 there are reported the loadings values of each variable respect to only the first principal component (PC1). In fact, since each PC is a combination of the original variables, each variable has a weight (loading factor) which indicates the strength of influence that variable (metabolite) has on the overall profile for a set of samples. These weights are normalized from 0 to 1 and can be visually inspected in usually one- or two-dimensional loadings plot.

In the 1D loadings plot there can be easily individuated the variables influencing the separation among wild type and *rag8* samples: dashed lines indicate loadings above and under 0.7 and  $-0.7$ , considered statistically significant ( $P < 0.01$ ).

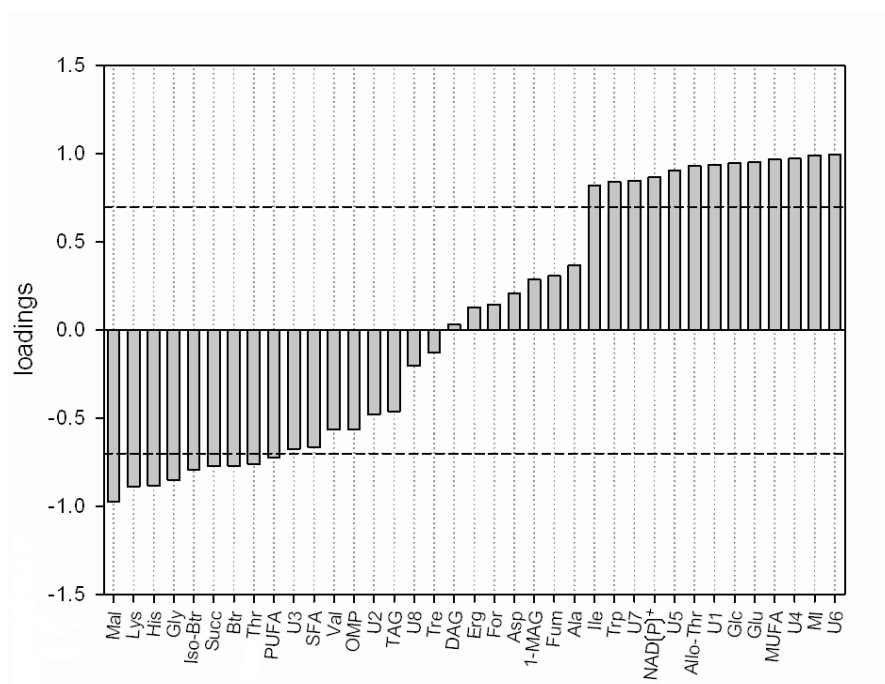


Figure 5.31: 1D PCA loadings plot showing loading values for each variable measured (metabolites), relative to PC1.

In the 2D loadings plot (fig. 5.32) the same variables are grouped together into two clusters, on the right and on the left side of the plot, at the higher positive and negative loading values, respect to PC1.

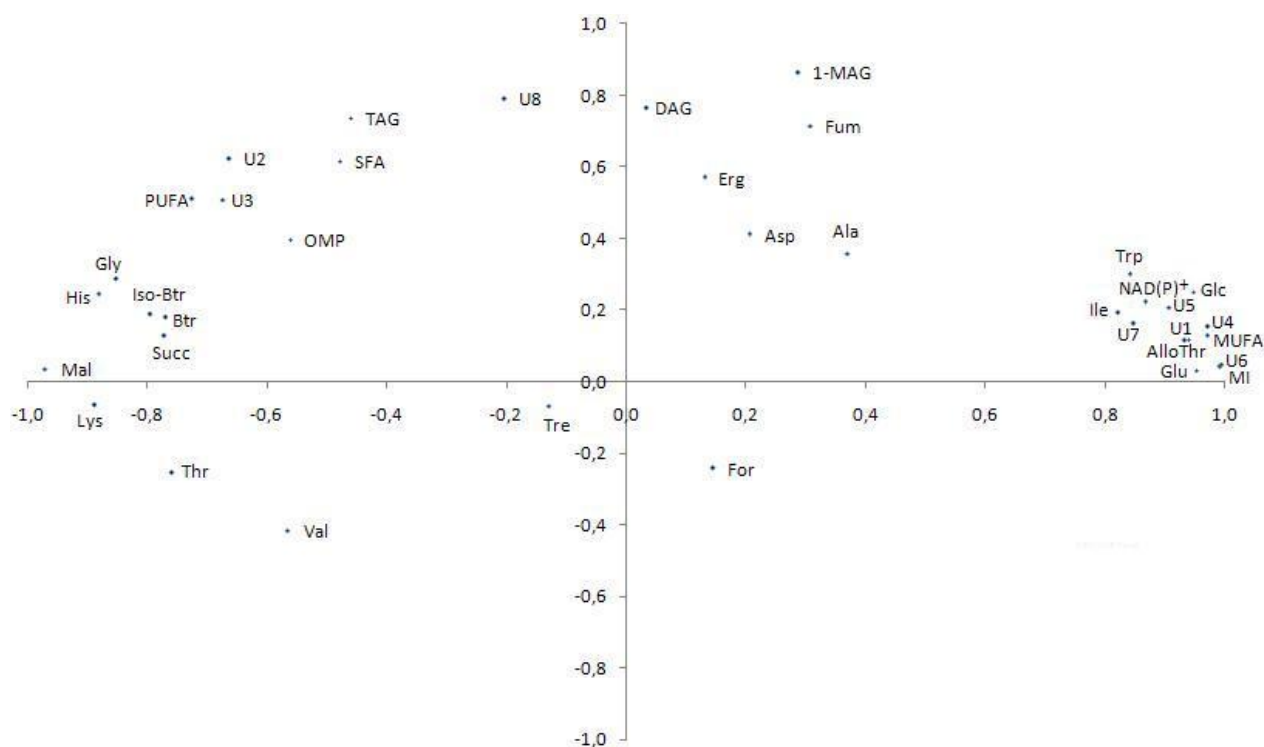


Figure 5.32: 2D PCA loadings plot showing loading values for each variable measured, relative to PC1 and PC2.

In this 2D plot it is possible to observe the correlation degree of the most discriminant variables, in a multivariate sense respect to what was observed in the correlation matrices: more the variables are near each other, indeed, more is similar the kind of information they carry on about the discrimination between WT and mutant samples.

PCA results, naturally, are in agreement with what was evidenced by univariate Student's t-test (see figure 5.27): the variables important in the discrimination, indeed, are the same metabolites present in a statistically higher amount in WT and mutant samples, respectively.

Metabolic networks, which are defined by the chemical nature of the metabolites, the stoichiometry for their interconversion and the biochemical pathways, are well known for many organisms of biotechnological interest, including *Kluyveromyces lactis*. In fact, eucaryotes central carbon metabolism represents settled text-book knowledge and it can be found in detail in many free database available on-line [13].

In contrast, in vivo control and regulation of metabolic networks appears to be a largely unexplored field for experimental research [32]. To reveal cause and effect relationships it is important to gain deeper insight to the complex metabolic responses at the level of intracellular metabolite concentrations and fluxes [55]. To achieve this purpose, metabolites determination in *rag8* mutant and wild type cells was employed to the construction of a metabolic network inter-connecting these compounds in mitochondria, endoplasmic reticulum (ER) and peroxisome organelles (fig. 5.33).

The altered values of the metabolites levels observed in the *rag8* strain could be interpreted as an obligate response of the mutant cells to re-adjust the intra-cellular redox balance.

In fact, all *rag* mutants were characterized by reduced production of ethanol and by the inability to accumulate glycerol [41] through fermentation pathways for redox balancing and osmoregulation [165].

Therefore, these strains are unable to support the cytoplasmic reoxidation of the NAD(P)H redox excess produced during glycolysis and pentose phosphate pathway (PPP) [39,51,166]. *K. lactis*, in particular, displays a oxidative utilization of glucose mainly through the PP pathway with higher production of NADPH [37,39,167,168].

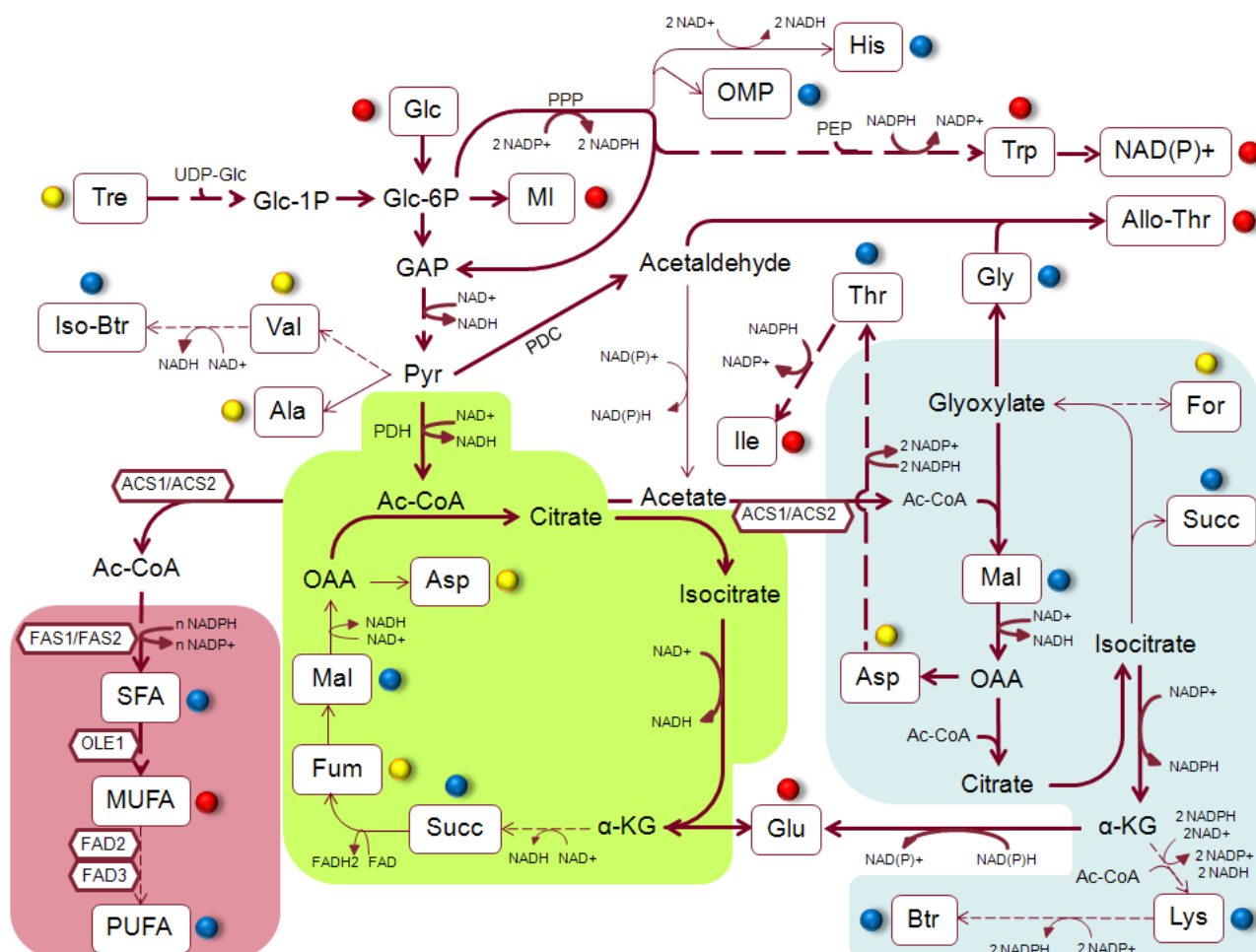


Figure 5.33: Statistical metabolic network. Colored balls denote mutant metabolites content respect to wild type samples: red balls indicate metabolites present in higher amount, blue balls indicate that ones present in lower amount and yellow balls indicate metabolites present in statistically same amount. There are also shown some important cell compartment, like mitochondrion (light green), endoplasmic reticulum (pink) and peroxisome (light blue). Hexagons near the reaction arrows reported the genes discussed.

Metabolites data comparison between wild type and mutant strains shows that the levels of metabolic intermediates produced in pathways generating NADH/NADPH - like His and Iso-Btr - were decreased (blue balls in fig. 5.33), while those associated with its reoxidation routes - like Ile, Glu and Trp - were increased (red balls) in *rag8* cell extracts.

As direct consequence of this, other metabolites show reduced level in mutant samples, instead: Thr, for example, from which is generated Ile with  $\text{NADP}^+$  production. Moreover, if the metabolic flux is directed to Glu synthesis to reoxidise NADH to  $\text{NAD}^+$ , it is clear that TCA flux is deviated at the level of  $\alpha$ -ketoglutarate ( $\alpha$ -KG), leading to a reduced amount of the following metabolites of the cycle (Mal and Succ). Lys is synthetised from  $\alpha$ -KG too, so its content result reduced probably to promote NADH reoxidation during Glu synthesis. Btr appears still reduced: this is probably due to the reduced content of its precursor (Lys) in mutant cells.

Finally, the increased content of  $\text{NAD(P)}^+$ , which is generated starting from Trp, found in mutant samples, also suggests the activation of pathways leading to the *de novo* synthesis of this coenzyme, to overcome the excessive generation of cytoplasmic NAD(P)H.

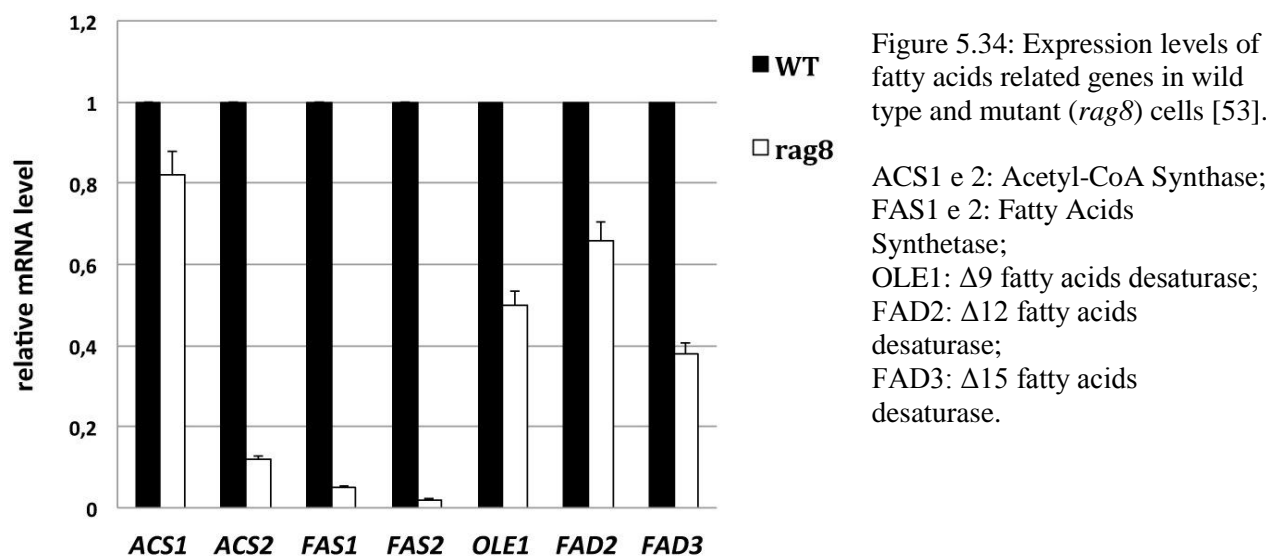
The revealing of OMP, an intermediate in the *de novo* synthesis of pyrimidine, appeared in agreement with the presence in both the strains of the *ura3* mutation, that blocks the biosynthesis of uracyl, leading to OMP accumulation. However, OMP level is lower in mutant samples and this could be explained considering that, in *S. cerevisiae*, serine/threonine protein kinase controls uracyl permease [169], so it was suggested that the depletion of the corresponding CKI gene in this strain (*rag8*), lead to a decreased uptake of uracyl, in the presence of an altered pyrimidine biosynthesis.

The increased level of allo-threonine, generated by the condensation of acetaldehyde with glycine (EC 4.1.2.49), and the consequent reduced level of Gly found in mutant samples, could be explained by a slightly reduced activity of the aldehyde dehydrogenase (EC 1.2.1.5)  $\text{NAD(P)}^+$  dependent: a plain conversion of all the acetaldehyde to acetate, catalysed by this enzyme, would lead to an excessive production of NAD(P)H [168], a condition that mutant cells probably avoid favoring its partial conversion to allo-threonine. Moreover, intracellular accumulation of acetaldehyde generated by PDC (EC 4.1.1.1) or by Thr aldolase (EC 4.1.2.5) [170] should be toxic for the cells [171].

Glucose supplied as substrate is essentially used for acetyl-CoA generation. Ac-CoA, produced in the Krebs cycle from pyruvate or by the condensation of acetate with the coenzyme A by the *KLACS1* and *KLACS2* genes products, is a key intermediate in cellular metabolism for energy generation in the mitochondria, for fatty acids  $\beta$ -oxidation in the peroxisome and for the synthesis of fatty acids in the ER.

In mutant samples acetyl-CoA flux seems to be directed to an higher synthesis of fatty acids (FA). However, it could be noted that the increase in the total FA amount (UFA+SFA) observed in *rag8* cells, was only limited to the mono-unsaturated fatty acids (MUFA) content. In fact, the  $\omega_3$  and  $\omega_6$  polyunsaturated fatty acids (PUFA) showed, by univariate student's t test, a statistically significant decrease (fig. 5.27b).

To determine if the different level of fatty acids observed in mutant samples can be ascribed to mechanisms of regulation controlled at transcriptional level, it was evaluated the expression profile of genes directly involved in their synthesis (fig. 5.34).



About the two acetyl-coA synthase genes, it was shown a higher transcriptional level of *ACS1*, as compared to *ACS2*, in *rag8* cells.

In *Saccharomyces cerevisiae* cultures grown aerobically on glucose, both ACS1 and ACS2 were expressed simultaneously, whereas only the ACS2 transcript was detected in anaerobic cultures [172]. Moreover, Acs1 and Acs2 localization in this closely related yeast is still under debate, even though Acs2 is believed to be mainly cytosolic and Acs1 is believed to be mainly peroxisomal [173,174].

So it was assumed a possible unequal distribution of Ac-CoA between cellular compartments in the mutants, as if in *K. lactis* cells grown under these conditions, ACS1 was responsible for most of the total ac-CoA synthesis capacity [172].

Also all the other genes involved in the FA synthesis show lower transcriptional levels in *rag8* cells. In particular, the less transcription of FAS1 and FAS2 genes is in agreement with the minor amount of SFA revealed in mutant samples (fig. 5.27a); the higher content of the total FA found in *rag8* cells, indeed, is all imputable to the higher level of MUFA present in these samples.

However, also the  $\Delta 9$  desaturase gene (OLE1) shows a reduced transcriptional level in *rag8* cells.

This could be explained bethinking that, in contrast to MUFA content, PUFA amount in mutant samples is decreased. Since the levels of FAD2 and FAD3 mRNA (coding for the  $\Delta 12$  and  $\Delta 15$  FA desaturase [175]) are reduced, a minor activity of these enzymes is assumed, leading to an accumulation of MUFA in mutant samples. Moreover, there must be also a less turnover of MUFA in *rag8* samples to account their massive presence despite the reduced transcription of OLE1.

Glucose assumption from the medium is about 20% for both the strains (data not shown) and its extracellular concentration is more than 10 times over the  $K_m$  of its high affinity transporter for *Kluyveromyces lactis* [44], ensuring the  $v_{max}$  condition of transport is satisfied. So, if glucose assumption from the medium is not enhanced, its intracellular accumulation (and MI higher content too) revealed in mutant samples rests to be explained.

### 5.2.2 Isotopomer analysis and flux analysis

$^{13}\text{C}$ -labeling experiment is an ideally suited technique to determine flux ratios [32]. This approach exploits the underlying relationship between the observed labeling of a metabolite and the multiple pathways that might lead to that labeling state in a complex metabolic network, providing a powerful method for probing metabolic pathways. In this way, differences in the metabolite phenotype between wild type and mutant strains can be easily quantified [93].

It has just seen as the complex pattern of signals observed in the  $^{13}\text{C}$ -NMR spectra indicate the presence of multiply labeled isotopologue species with significant abundance. So, the differential  $^{13}\text{C}$ -labeled isotopomer profiles can serve as a fingerprint of the metabolic networks activity and could reflect both qualitative and quantitative differences in the metabolic pathways that lead to the synthesis of each metabolite [104].

Once all the observed  $^{13}\text{C}$  NMR fine structures have been translated into fragment abundances the breakdown of the glucose six-carbon backbone supplied as substrate can be interpreted.

The more labeled material put into the system, the higher the intracellular carbon labeling and thus all fractional enrichment measurement sensitivities are improved. But, if all positions of the input substrate would have the same  $Y_{Ci}$ , then all intracellular carbon atom pools will finally have the same  $Y_{Ci}$ , whatever the fluxes in the system are. Thus, there is no flux information in such experiments [176]. Choosing as substrate  $[1,2-^{13}\text{C}_2]\text{glucose}$ , instead, it was possible to quantify the relative fluxes through glycolysis and pentose phosphate pathway within the metabolic network.

As fig. 5.35 shows, a molecule of  $[1,2-^{13}\text{C}_2]\text{glucose}$  could be metabolized by the glycolytic pathway or by the pentose phosphate pathway (PPP). The chemical mechanisms of these metabolic reactions are still well known and since our goal is to follow how the carbon backbone labelling state change during such pathways, each metabolite structure will be simply represented with white balls, for  $^{12}\text{C}$  unlabeled carbon atoms, and black balls for  $^{13}\text{C}$  labeled carbon positions.



From  $[1,2-^{13}\text{C}_2]\text{glucose-6phosphate}$  (Glc-6P), a molecule of  $[2,3-^{13}\text{C}_2]\text{pyruvate}$  (Pyr) will be generated through the glycolysis, whereas the PPP will produce a molecule of  $[1,3-^{13}\text{C}_2]\text{pyr}$  or pyruvate labeled only at C3 position, derived from intermediates such as F-6P or glyceraldehyde-3-phosphate (GAP)\*.

The succeeding activity of pyruvate dehydrogenase (PDH) or pyruvate decarboxylase (PDC) enzymes leads to the production of  $[2-^{13}\text{C}]\text{acetyl-CoA}$  (ac-CoA), if pyruvate followed the PPP.

$[1,2-^{13}\text{C}_2]\text{Ac-CoA}$  will be produced from  $[2,3-^{13}\text{C}_2]\text{pyr}$  followed the glycolytic pathway, instead.

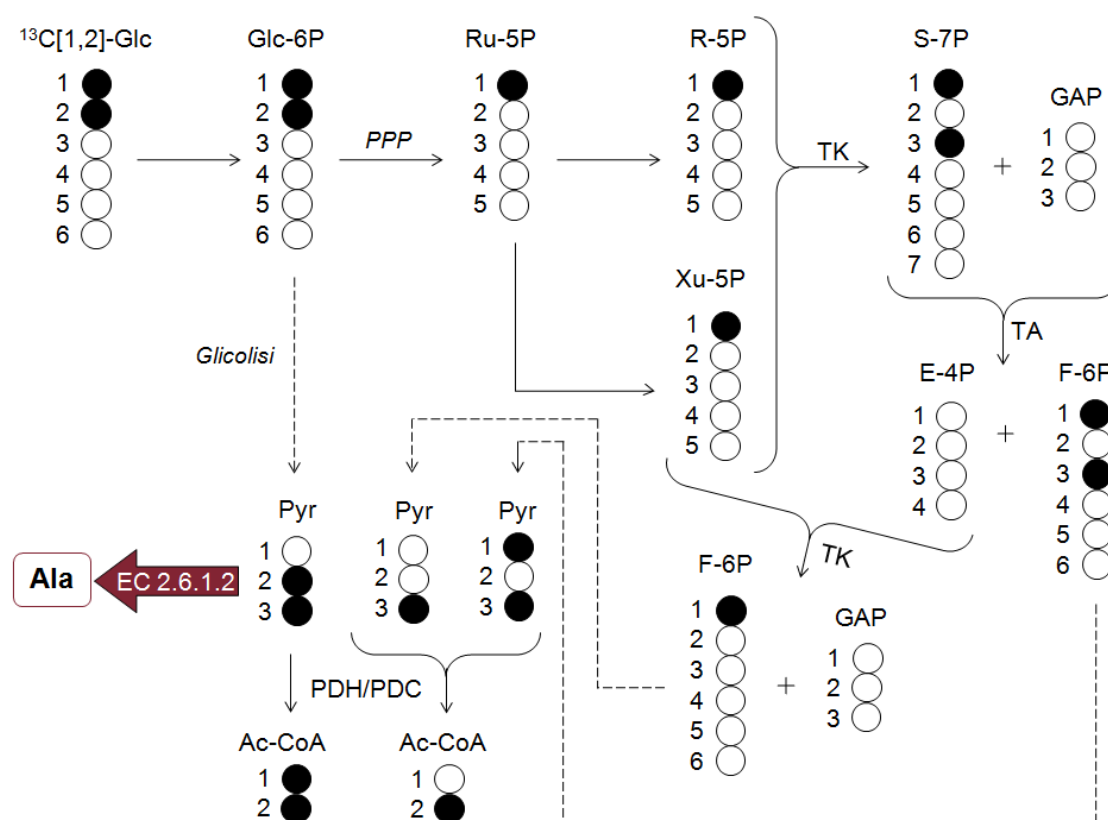


Figure 5.35: Metabolic pathways leading to acetyl-CoA (Ac-CoA) production from  $[1,2-^{13}\text{C}_2]\text{glucose}$  ( $[1,2-^{13}\text{C}_2]\text{Glc}$ ). Glc-6P: glucose-6phosphate; Ru-5P: ribulose-5phosphate; R-5P: ribose-5phosphate; Xu-5P: xylulose-5phosphate; S-7: sedoheptulose-7phosphate; GAP: glyceraldehyde-3phosphate; E-4P: erytrose-4phosphate; F-6P: fructose-6phosphate; Pyr: pyruvate; Ala: alanine. TA: transaldolase; TK: transketolase; EC 2.6.1.2: alanine transaminase.

Alanine (Ala) is one of the enriched hydro-soluble metabolites revealed in the samples. It is produced from pyruvate by a transaminase (EC 2.6.1.2), so its enrichment state directly reflects pyruvate isotopomer distribution.

\* Also one totally unlabeled molecule of pyruvate is always generated, from the C4-C5-C6 fragment of Glc original backbone.

The fragment C2-C3 (or C3-C2), so, is originated from the  $[2,3-^{13}\text{C}_2]$ pyruvate coming from the glycolytic pathway, whereas the singlets observed at C1 and C3 position are referred to both the labeling state generated by the PPP.

The ratio between the enrichment measured as singlet at the C3 position and the enrichment measured as doublet in the same site (C3-C2) could give a quantitative evaluation of the PPP/glycolysis relative flux ratio. The same value ( $\sim 0.15$ ) was obtained for wild type and mutant samples suggesting that there are no differences between the fluxes through these two pathways, until alanine level at least.

In particular, the singlet observed in C3 could be relative to both the  $[1,3-^{13}\text{C}_2]$ Ala and  $[3-^{13}\text{C}]$ Ala isotopomers, and the singlet relative to C1 position could be referred to the  $[1,3-^{13}\text{C}_2]$ Ala again and to  $[1-^{13}\text{C}]$ Ala, instead.

To perform a complete isotopomer analysis for this three carbon atoms metabolite, it must be considered that the number of the possible isotopomers for Ala is  $2^3 = 8$ . Table 5.3 shows all these isotopomers and the corresponding notation used for the discussion below.

Table 5.3 - Alanine (Ala) isotopomers.  
Notation:  $^{12}\text{C}$  = white balls = 0;  $^{13}\text{C}$  = black balls = 1.

Ala isotopomer	Notation	Ala isotopomer	Notation
$\begin{array}{c} 1 \text{ } \bigcirc \\ 2 \text{ } \bigcirc \\ 3 \text{ } \bigcirc \end{array}$	000	$\begin{array}{c} 1 \text{ } \bullet \\ 2 \text{ } \bigcirc \\ 3 \text{ } \bullet \end{array}$	101
$\begin{array}{c} 1 \text{ } \bullet \\ 2 \text{ } \bigcirc \\ 3 \text{ } \bigcirc \end{array}$	100	$\begin{array}{c} 1 \text{ } \bigcirc \\ 2 \text{ } \bullet \\ 3 \text{ } \bullet \end{array}$	011
$\begin{array}{c} 1 \text{ } \bigcirc \\ 2 \text{ } \bullet \\ 3 \text{ } \bigcirc \end{array}$	010	$\begin{array}{c} 1 \text{ } \bullet \\ 2 \text{ } \bullet \\ 3 \text{ } \bigcirc \end{array}$	110
$\begin{array}{c} 1 \text{ } \bigcirc \\ 2 \text{ } \bigcirc \\ 3 \text{ } \bullet \end{array}$	001	$\begin{array}{c} 1 \text{ } \bullet \\ 2 \text{ } \bullet \\ 3 \text{ } \bullet \end{array}$	111

- No singlet was observed in Ala C2 position: this means that the 010 isotopomer could not be considered. The same is for the 110 and 111 isotopomers, because no doublet relative to the C2-C1 coupling was revealed.
- The isotopomer 100 could not exist because there is no metabolic pathway leading the label only in C1 position (fig. 5.35): the enrichment observed as singlet in this position, so, is totally due to the 101 isotopomer.
- The 011 isotopomer is completely represented by the quantification of the enrichment relative to the C3-C2 fragment observed.

Singlets quantified in C1 and C3 position, instead, represent the 10X and X01 *cumomer* respectively. A cumomer is a set of one or more isotopomers that contain a particular labeled fragment: X indicates an atom that is either labeled or unlabeled [83]. In this case, 10X cumomer contain the  $^{13}\text{C}_1$ - $^{12}\text{C}_2$  fragment observed at the C1 position, whereas X01 contain the  $^{12}\text{C}_2$ - $^{13}\text{C}_3$  fragment observed at the C3 position. Two equations could be deduced from this observation:

$$10X = 101$$

$$X01 = 001 + 101$$

There are two unknowns for two equations, so the system is solvable:

$$101 = 10X$$

$$001 = X01 - 101 = X01 - 10X$$

Considering also the 011 isotopomer, it is possible to obtain the relative percentage of each alanine labeled isotopomers revealed by  $^{13}\text{C}$ -NMR spectroscopy. However, to complete Ala isotopomer analysis also the 000 isotopomers must be considered.

This totally unlabeled isotopomer could be calculated from the natural abundant Ala C3 signal in the  $^{13}\text{C}$ -NMR control spectrum. This signal, indeed, could be referred to the total amount of Ala present in the samples, i.e., it corresponds to the sum of all the revealed isotopomers (including the totally unlabeled).

In this way, the relative percentage of *all* the Ala isotopomers are calculated (fig. 5.36).

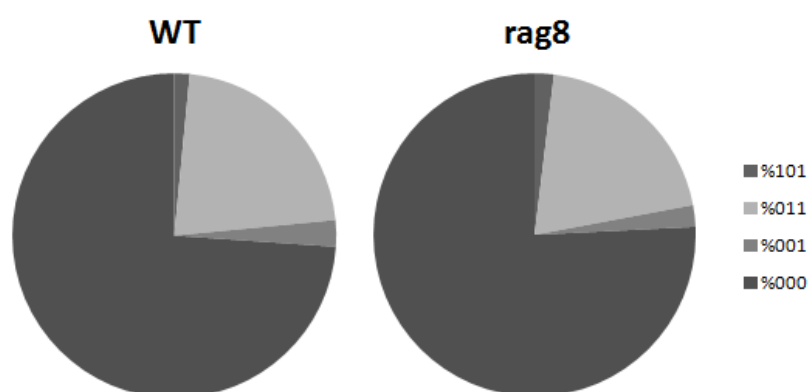


Figure 5.36: Relative percentage of all the alanine isotopomers present in WT and *rag8* samples.

Note as the totally unlabeled isotopomer (000) has the higher abundance suggesting that the most of the material involved in cell metabolism is unlabeled. During glycolysis, indeed, two molecule of pyruvate are generated from [1,2-<sup>13</sup>C<sub>2</sub>]glucose: one labeled (singlet or doublet) and one unlabeled. However, in this way, 50% of both labeled and unlabeled material is expected. We found ~75% (>50%) of unlabeled isotopomers, instead: this means that there could be present unlabeled source of material, as aminoacids or other compounds, coming from the medium itself.

Finally, it is possible to note that, if the singlet measured in Ala C1 position is only due to the 101 isotopomer, it could be considered as an index of the TA activity within the PPP (see fig. 5.35). The ratio between this singlet and the one observed in Ala C3 position (relative to both the 101 and 001 isotopomers, instead) could give a relative quantification of the TA activity, respect to TK activity, within the PPP flux (table 5.4).

Table 5.4 - Transaldolase activity relative to transketolase, within PPP, in WT and mutant (*rag8*) samples.

TA / (TA+TK)	WT	<i>rag8</i>
(s)C1/(s)C3	0.37 ± 0.02	0.64 ± 0.01

<sup>13</sup>C isotopomer analysis is also an important technique for exploring metabolic pathways that intersect in the tricarboxylic acid (TCA) cycle [177].

Pyruvate could be entered the Krebs cycle as acetyl-CoA (ac-CoA), thanks to the pyruvate dehydrogenase complex/pyruvate decarboxylase activity (PDH/PDC) or as oxalacetate (OAA) generated by pyruvate carboxylase (PC).

When the  $^{13}\text{C}$ -labeled glucose supplied as substrate is oxidized it contributes enriched carbons to all intermediates of the TCA cycle and to all metabolites in exchange with cycle intermediates.

Fig. 5.37 shows glutamate (Glu) and lysine (Lys) labeling pattern achievable after the first route of the TCA cycle, with pyruvate coming from glycolysis (fig. 5.37a) and PPP (fig. 5.37b).

Glutamate is another enriched metabolite found in the hydro-alcoholic extracts. It was generated by the transamination of  $\alpha$ -ketoglutarate so its labeling state is the same of its precursor ( $\alpha$ -KG).

Pyruvate coming from glycolysis and entered in the Krebs cycle as  $[1,2-^{13}\text{C}_2]\text{ac-CoA}$  could generate, by the first TCA route, Glu isotopomer labeled in C4-C5 position (fig. 5.37a), which is observable by the enrichments quantified as doublet in both C4 and C5 positions (XX011 and XXX11 cumomers, respectively). However, the corresponding abundances of the  $\text{C}_2$  fragments C4-C5 and C5-C4 seem not in agreement each other (fig. 5.22). This could be completely explained revolving that the XXX11 cumomer comprises both the XX011 and the XX111 cumomers, the latter represented by the  $\text{C}_3$  C3-C4-C5 fragment revealed in C4 position. Glu C5-C4 fragment amount, indeed, results equal to the sum of the C4-C5 and C3-C4-C5 fragments, in both the strains.

The same pyruvate pool passed through the PC activity, instead, could generate Glu isotopomer containing the C2-C3 fragment, observed from both the position involved (011XX and X110X cumomers, respectively). However the second cumomer is not discernable from the other one (X011X) observed in the same position (Glu C3-C2 or C3-C4), so no more considerations are possible.

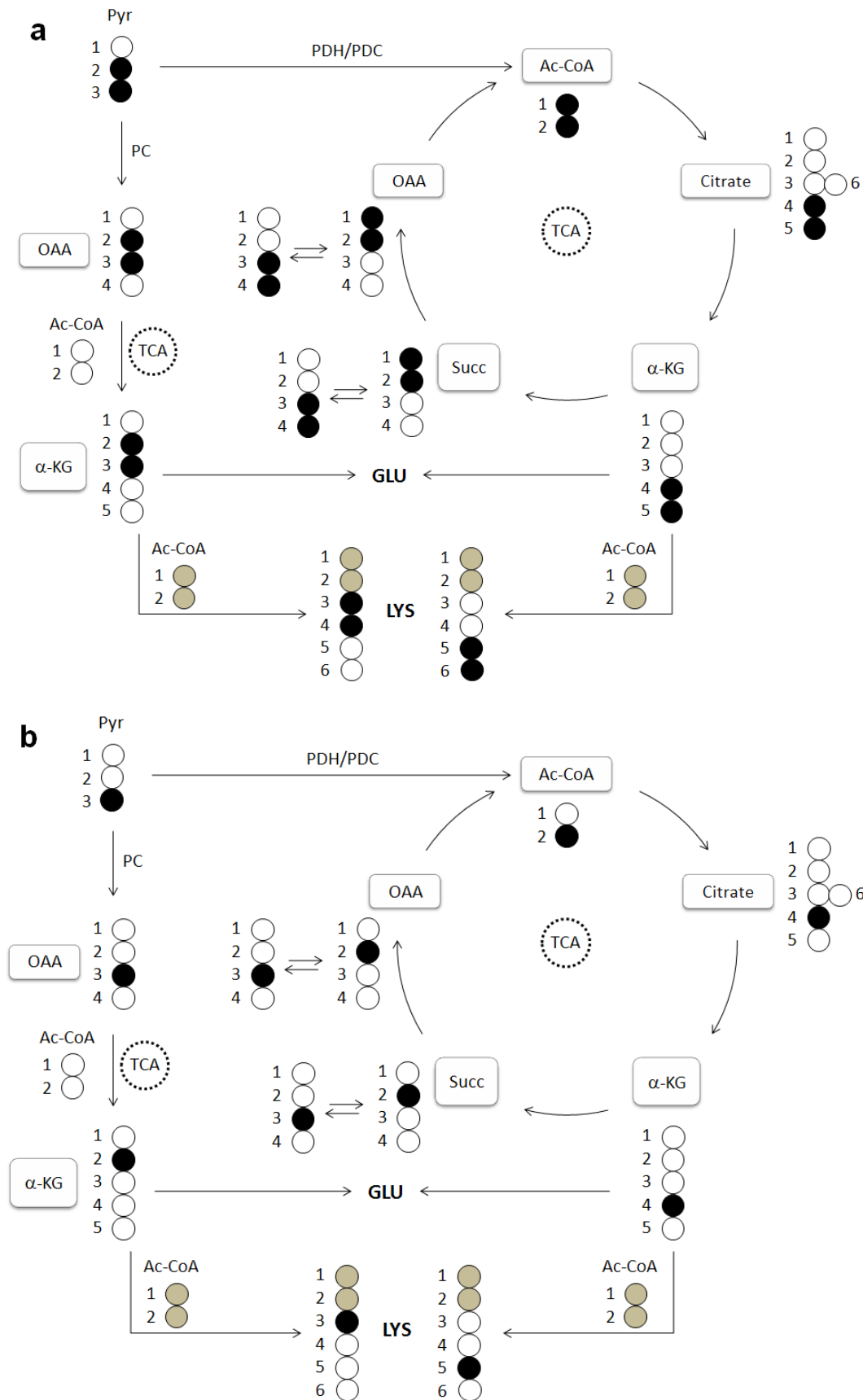


Figure 5.37: Glutamate (Glu) and Lysine (Lys) labeling state achievable after the first TCA round. Pyruvate (Pyr) is coming from glycolysis (a) or PPP (b). Glu labeling state is the same of  $\alpha$ -ketoglutarate ( $\alpha$ -KG). The ac-CoA unit employed in Lys synthesis could be labeled or unlabeled (grey balls).

Analogously, pyruvate coming from PPP and entered in the Krebs cycle as ac-CoA or OAA (by PDH/PDC or PC activity, respectively) leads to Glu isotopomers labeled as singlet in C4 or C2 positions (fig. 5.37b).

No enrichment was observed as singlet in Glu C5 position because it should be due to an ac-CoA pool labeled only in C1 and this labeling state is impossible to obtain.

Moreover, note that in the TCA cycle,  $\alpha$ -KG is then converted to succinate (Succ) with a scrambling of the labeling because the molecule become symmetric (fig. 5.38). This leads to the formation of two Succ pools: 50% with a certain labeling state and 50% with the symmetric ones.

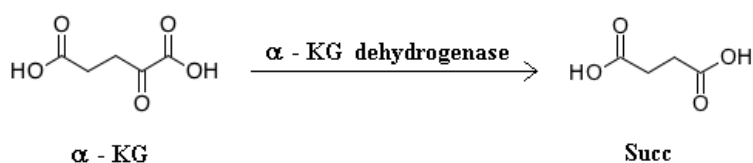


Figure 5.38: Conversion of  $\alpha$ -ketoglutarate ( $\alpha$ -KG) to succinate (Succ). This reaction, catalysed by the  $\alpha$ -KG dehydrogenase, leads to the scrambling of the labeling in Succ because this molecule is symmetric.

So, going on in the TCA routing, the OAA obtained after the first TCA route, in the second one could condensate with differently labeled ac-CoA (2<sup>nd</sup> route, in fig. 5.39).

In each case it can be seen that Glu C1-C2 fragment is formed. Moreover, there could be also originated new fragments, as the one singly labeled at the C3 position and the C3-C4 and C3-C4-C5 fragments. However, it is clear that the relative abundance of the C1-C2 and C3 fragments are very similar each other whereas C4-C3 and C3-C4-C5 fragments are lower. This suggests that the second route of TCA most probably involved  $^{13}\text{C}$ -unlabeled ac-CoA, in agreement with the fact that, in both the strains, most of the material involved in cells metabolism is unlabeled, as the prominent relative percentage of the totally unlabeled Ala isotopomer has confirmed yet (000 ~ 75%).

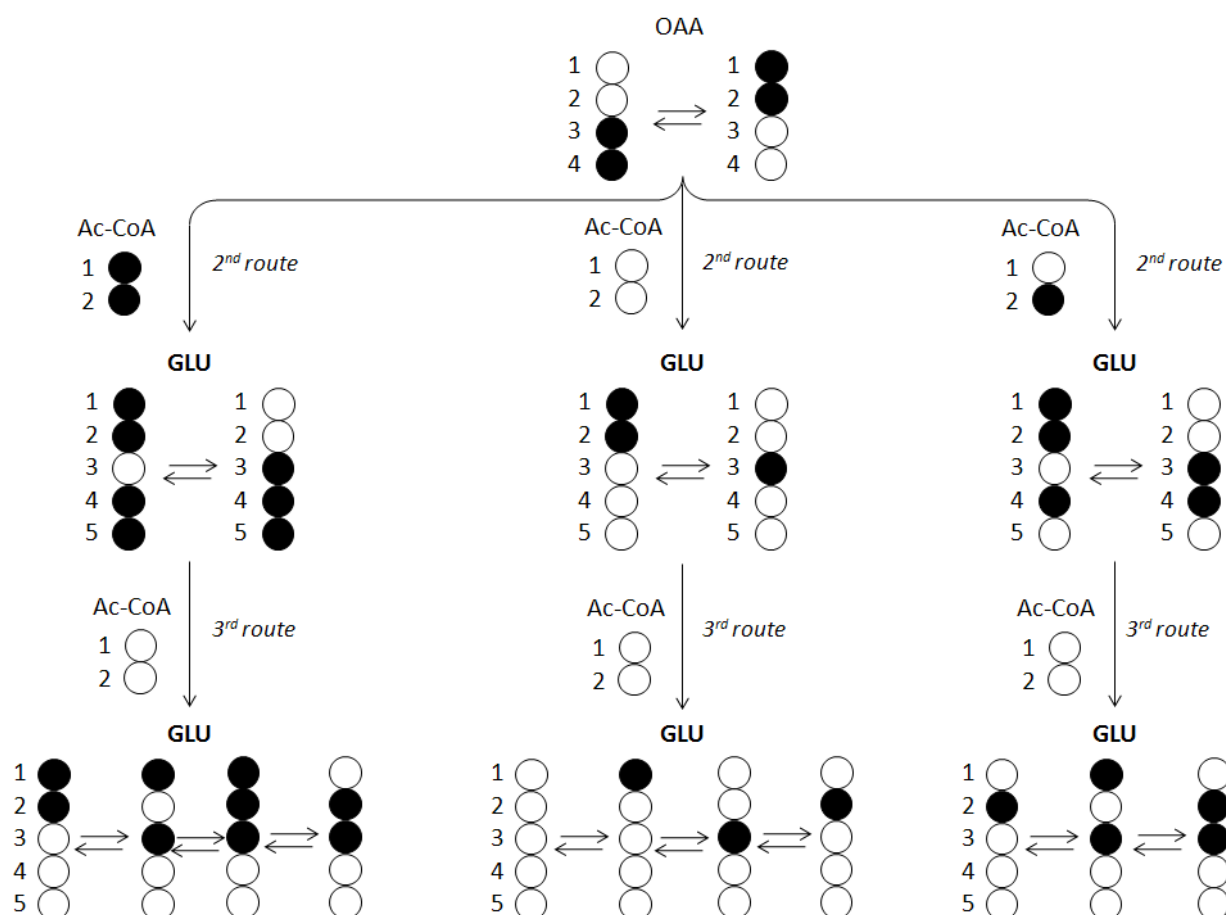


Figure 5.39: Possible glutamate (Glu) isotopomers generated by TCA routing.

Going on through the third TCA route (3<sup>rd</sup> route, in fig. 5.39), the condensation with differently labeled ac-CoA it is again possible: however, it was chosen to show only the most probably results obtained by the condensation with unlabeled ac-CoA. This leads to the formation of other isotopomers, including those ones containing C1 and C1-C2-C3 labeled fragments.

Although the C1 fragment is relatively high (it is generated in each case), it results lower in *rag8* samples (fig. 5.22): for this reason it was supposed a minor TCA activity in the mutants.

C1-C2-C3 fragment, instead, was observed only in WT samples (fig. 5.22): this is in agreement with the fact that the quantifications of the C<sub>2</sub> fragments C1-C2 and C2-C1 are in statistical agreement in mutant samples whereas in WT samples C1-C2 fragment seems to be a little greater. Glu C1-C2 fragment amount in WT samples, indeed, is nearish to the sum of Glu C2-C1 and C1-C2-C3 fragments.



Moreover, the presence of the C1-C2-C3 fragment only in the WT samples could not be ascribed to the simply TCA routing. For this reason, it was supposed a glyoxylate shunt, active only in this strain, giving OAA totally labeled, from which Glu C1-C2-C3 fragment could be generated by TCA activity (fig. 5.40).

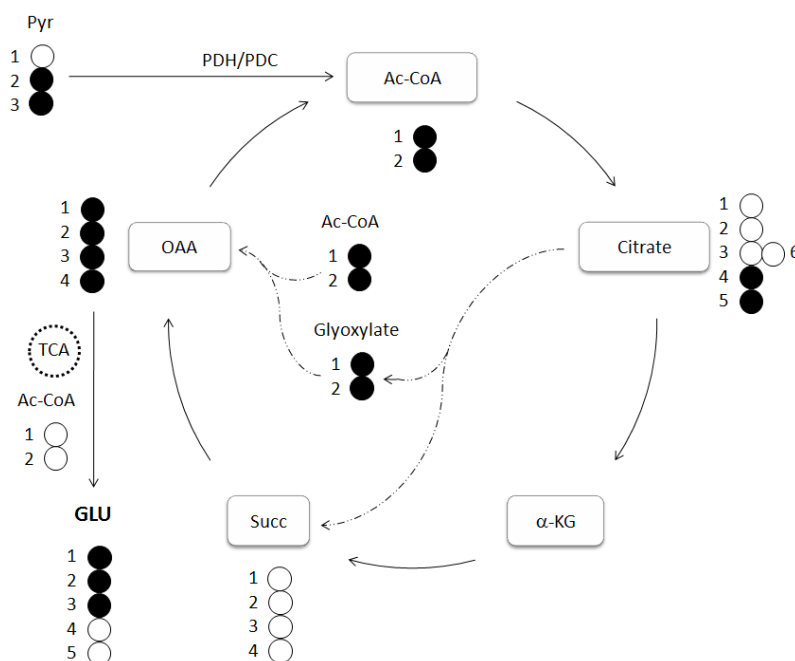


Figure 5.40: Glyoxylate shunt (dashed lines) generating glutamate (Glu) C1-C2-C3 fragment.

Lysine (Lys) is another hydro-alcoholic metabolite resulted enriched; in yeasts it is synthesized in peroxisomes starting from acetyl-CoA and  $\alpha$ -ketoglutarate [178]. The ac-CoA unit corresponds to the C1-C2 fragment in Lys backbone; both WT and mutant samples display an enrichment observed as singlet in C2, relative to the ac-CoA moiety coming from the PPP, and a labeling state doublet (C2-C1) relative to the ac-CoA pool passed through the glycolytic pathway (grey balls in fig. 5.37).

An interesting difference is about the labeling state of the  $\alpha$ -KG moiety building Lys backbone: the fact that it results totally unlabeled in WT samples suggests a different source (unlabeled) of this precursor for lysine biosynthesis in this strain. We have just suggested that a possible unlabeled source could be the medium (*peptone*) that is evidently less metabolized by mutant cells, in agreement with the minor total content of Lys quantified in *rag8* samples (fig. 5.27a). As a result, mutant cells must satisfy Lys biosynthesis in a different way.

Moreover, for both the strains, this  $\alpha$ -KG pool must be compartmentalised because its labeling state is not mixed with the other labeled compounds derived from the substrate. As the enrichments observed in mutant samples on Lys C3 and C5 positions are only singlets, data suggest that this labeling state comes from PPP intermediates. Finally, since Lys biosynthesis occurs in peroxisomes, it was supposed that this alternative pathway is a PPP compartmentalised in these organelles.

It has just seen that [1,2- $^{13}\text{C}_2$ ]glucose could be metabolized by the PPP and [1,3- $^{13}\text{C}_2$ ]fructose-6-phosphate could be generated by transaldolase (TA) activity. This metabolite could be converted to fructose-1,6-bisphosphate (EC 3.1.3.11 or EC 2.7.1.11) first and then to glyceralone-3-phosphate (EC 4.1.2.13), from which it is generated glyceraldehyde-3-phosphate (GAP, EC 5.3.1.1). Fig. 5.41 shows as GAP could be transformed to serine (Ser) and so to pyruvate (Pyr), maintaining the labeled carbon in C1 and C3 position. Moreover,  $^{13}\text{C}$ -unlabeled pyruvate could be derived from a pool of  $^{13}\text{C}$ -unlabeled aminoacids.

Pyruvate generated as described, could enter the glyoxylate cycle as OAA, thanks to pyruvate carboxylase (PC) activity, reacting with an ac-CoA pool derived from  $^{13}\text{C}$ -unlabeled aminoacids again or from PDH/PDC activity. This leads to the generation of an  $\alpha$ -KG pool characterized by C2 and/or C4 labeled positions, corresponding to the Lys C3 and C5 singlets observed, respectively.

Moreover, Lys biosynthesis involved another molecule of ac-CoA (indicated with grey balls in fig. 5.41) that could be labeled or not. Since no doublet was observed in Lys C3 position (and a single label in Lys C1 position is impossible to obtain), isotopomers containing the  $\text{C}_3$  C1-C2-C3 fragment could not exist. There are 64 possible isotopomers for Lys: fig. 5.42 shows those one revealed in mutant samples.

This is the first time this supposed compartmentalized pathway is observed in *K. lactis* cells: for the closely related yeast *S. cerevisiae*, literature reported that the carbon backbone of lysine coming from  $\alpha$ -ketoglutarate were synthesized in the mitochondria [179].

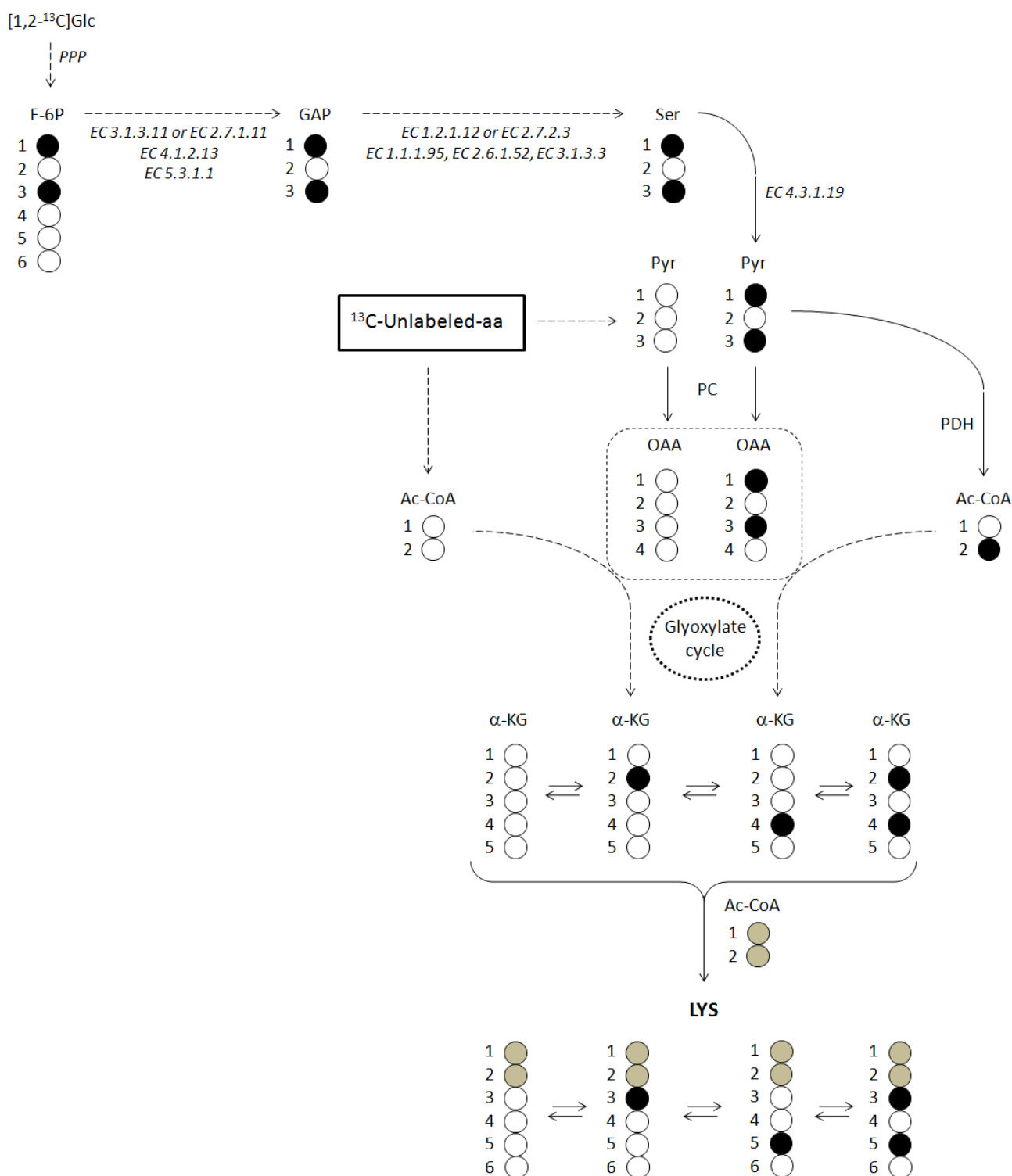


Figure 5.41: Lysine (Lys) biosynthesis pathway supposed for mutant cells.

EC 3.1.3.11: fructose-1,6-bisphosphatase; EC 2.7.1.11: 6-phosphofructokinase; EC 4.1.2.13: fructose-bisphosphate aldolase; EC 5.3.1.1: triose-phosphate isomerase; EC 1.2.1.12: glyceraldehyde-3-phosphate dehydrogenase; EC 2.7.2.3: phosphoglycerate kinase; EC 1.1.1.95: phosphoglycerate dehydrogenase; EC 2.6.1.52: phosphoserine aminotransferase; EC 3.1.3.3: phosphoserine phosphatase; EC 4.3.1.19: serine ammonia-lyase.  $^{13}\text{C}$ -Unlabeled aa: pool of  $^{13}\text{C}$ -unlabeled aminoacids. PC: pyruvate carboxylase; PDH: pyruvate dehydrogenase complex.

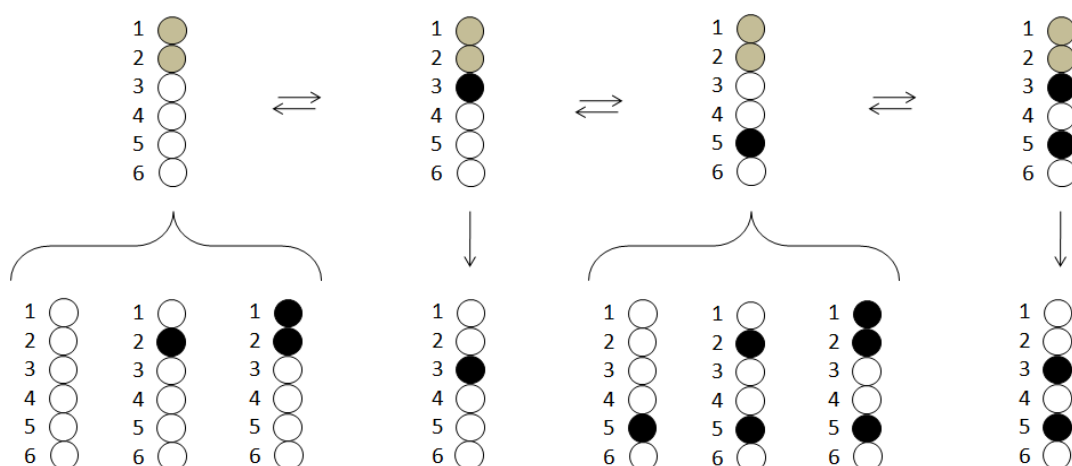


Figure 5.42: Lysine cumomers generated by the supposed pathway (first row): ac-CoA units are indicated with grey balls. Lys isotopomers are represented in the second row, respect to the possible labeling state of the C1 and C2 carbons, in agreement with the labeling pattern observed in mutant samples.

However, even if it is evident that no other  $\alpha$ -KG molecules differently labeled concur to Lys biosynthesis, it must be verified if this compartmentalized  $\alpha$ -KG pool, generated from PPP intermediates, could exit from its compartment: for example, it could concur to Glu biosynthesis (fig. 5.33). Since  $\alpha$ -KG is a precursor of both lysine and glutamate, Glu C2 and C4 positions labeled as singlet correspond to the Lys C3 and C5 singlets observed (fig. 5.37b).

Since the ratio between the enrichment measured as singlet in Glu C4 and that one measured as doublet in the same position (C4-C5 fragment) quantify the relative ratio between PPP and glycolysis, it was seen that in WT samples this value ( $0.16 \pm 0.01$ ) is very close to that one obtained analogously for Ala (C3/C3-C2 =  $0.15 \pm 0.01$ ), meaning that such relative fluxes remain constant in the biosynthetic pathways from Ala to Glu, confirming that the contribution of the supposed peroxisomal flux is absent in WT samples.

In *rag8* samples, instead, the ratio Glu C4/C4-C5 is about double ( $0.38 \pm 0.02$ ) as compared to that one obtained from Ala: this demonstrates that the compartmentalized  $\alpha$ -KG pool (whose labeling state comes from PPP, indeed) could exit from its compartment and that it concurs to Glu biosynthesis.

About succinate, the ratio between enrichments observed as singlet in C1 (or C4) and C2 (or C3) quantify the relative fluxes of ac-CoA coming from the PPP and passed through PC or PDH/PDC, respectively (fig. 5.43). The values obtained for this ratio result to be lower in *rag8* samples, as compared to WT, suggesting a reduced flux of PC for the TCA in mutant samples.

This explains the lower enrichment observed as singlet in *rag8* Glu C2, as compared to WT (fig. 5.22), even if, also for this position, the peroxisomal contriubute exists, as in Glu C4. Indeed, also the label occuring in Glu C2 is due to the PC activity, whereas the labeling of Glu C4 site is due to the PDH/PDC activity, within the PPP flux (fig. 5.37b).

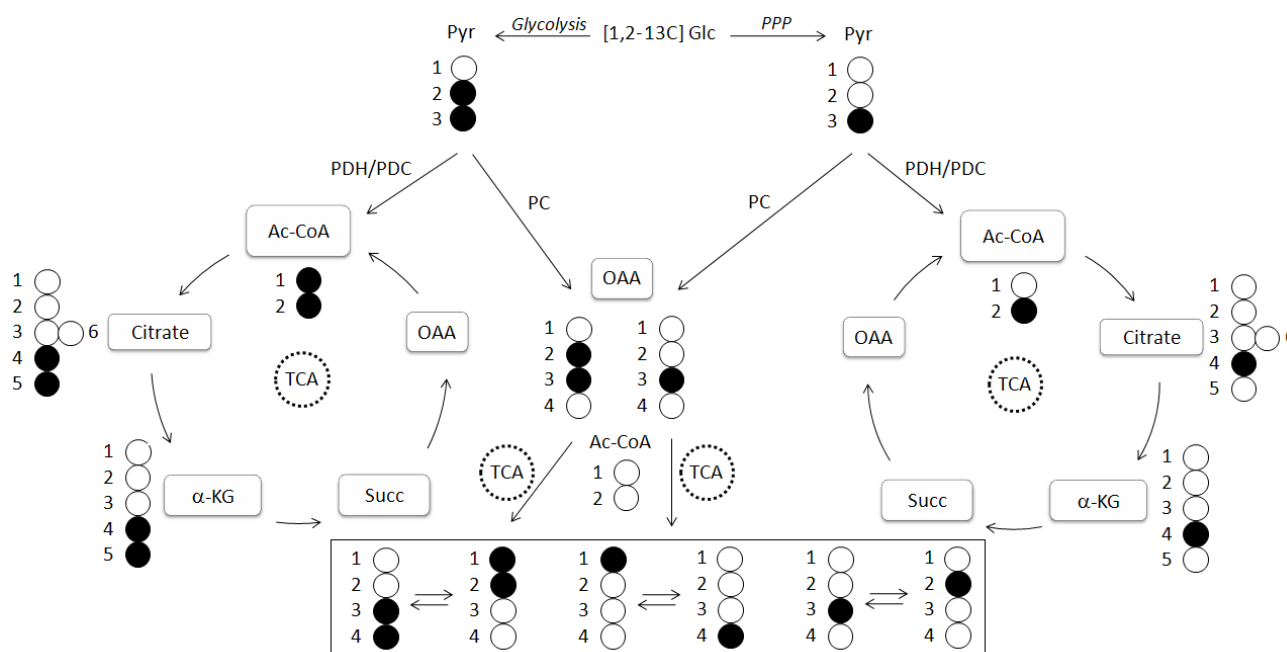


Figure 5.43: Succinate (Succ) isotopomers deriving from fluxes through PC or PDH/PDC and starting from both [2,3-<sup>13</sup>C<sub>2</sub>]pyruvate (glycolysis) and [3-<sup>13</sup>C]pyruvate (PPP) are shown in the box.

Note as these labeling state correspond to the fragments observed by <sup>13</sup>C-NMR spectroscopy.

The ratio between Glu C2-C3 and C4-C5 fragments could give quantitative informations about the relative contribute of PC and PDH/PDC but, this time, within the glycolytic flux (fig. 5.37a). These ratios result again lower in mutant samples, confirming a reduced PC activity in *rag8* samples also within the glycolytic flux.

Finally, the ratio between Lys enrichments measured as singlet (at the C2 position) and the enrichments measured as doublet in the same site (C2-C1 fragment) will give a quantitative evaluation of the PPP/glycolysis relative flux ratio. Once again, the values obtained for this ratio is about the same for both the strains ( $\sim 0.5$ ), indicating that the nature of the ac-CoA pool employed in Lys synthesis is the same in WT and mutant cells. However it results to be different from the value obtained from Ala ( $\sim 0.15$ ), suggesting that the pool of PPP intermediates recycling in glycolysis is not homogeneous, but compartmentalized instead.

Labeling experiments give information also about the unresolved question around the intracellular accumulation of glucose (Glc) and myo-inositol (MI) observed in mutant cells extracts.

It was just seen that glucose assumption from the medium is the same from both the strains, so its accumulation can not be ascribed to an higher substrate uptake. Since no enrichment was revealed for its  $^{13}\text{C}$  resonances, glucose high intracellular presence can not even be due to an eventual reduced substrate utilization. The answer is hidden in the high intracellular amount of trehalose (Tre) found in both the samples: also this metabolite results not enriched, so it is not a storage product generated from the labeled Glc supplied as substrate; it must be assumed from the medium, instead, in which it was unexpectedly revealed. So trehalose, inside the mutant cells, could be disrupted into two molecules of glucose by trehalase enzyme (EC 3.2.1.28), leading to its high intracellular accumulation.

Moreover, from Glc, MI could be synthesised through hexokinase (EC 2.7.1.1), synthase (EC 5.5.1.4) and phosphatase (EC 3.1.3.25) activity.

So, the high intracellular concentration of these metabolites found in mutant cells extracts could has an osmoregulatory function, according to the *rag8* strains inability to accumulate glycerol [41] through fermentation pathways for redox balancing and osmoregulation [165]. MI has also fundamental role against cells wall alterations [180], characteristic of *rag* mutants [41].

About the lipo-soluble metabolites the most attractive is ergosterol (Erg).

To explain and correctly evaluate the multiplicity observed in its enriched  $^{13}\text{C}$ -NMR signals, it must be considered Erg biosynthetic pathway. As described in 1.3.7 Introduction, Erg is synthesized by mevalonate pathway starting from two molecule of acetyl-CoA. Then the pathway continues with a series of reactions always involving these condensation products: fig 5.44a and 5.44b aid to follow how the original ac-CoA units are rearranged until the formation of Erg structure.

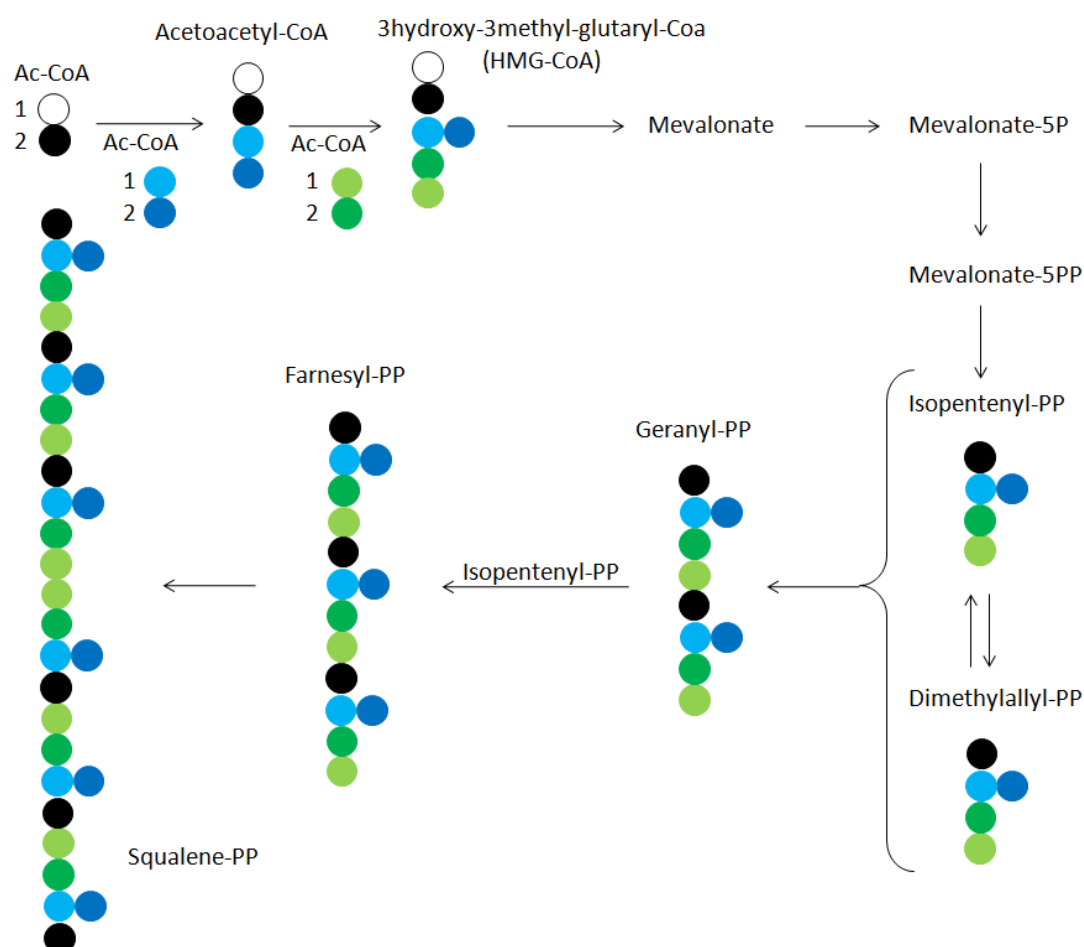


Figure 5.44a: Erg biosynthetic pathway. Different acetyl-CoA (Ac-CoA) units are represented with different colors: the lighter is always referred to the C1 position whereas the darker is always referred to the C2 one.

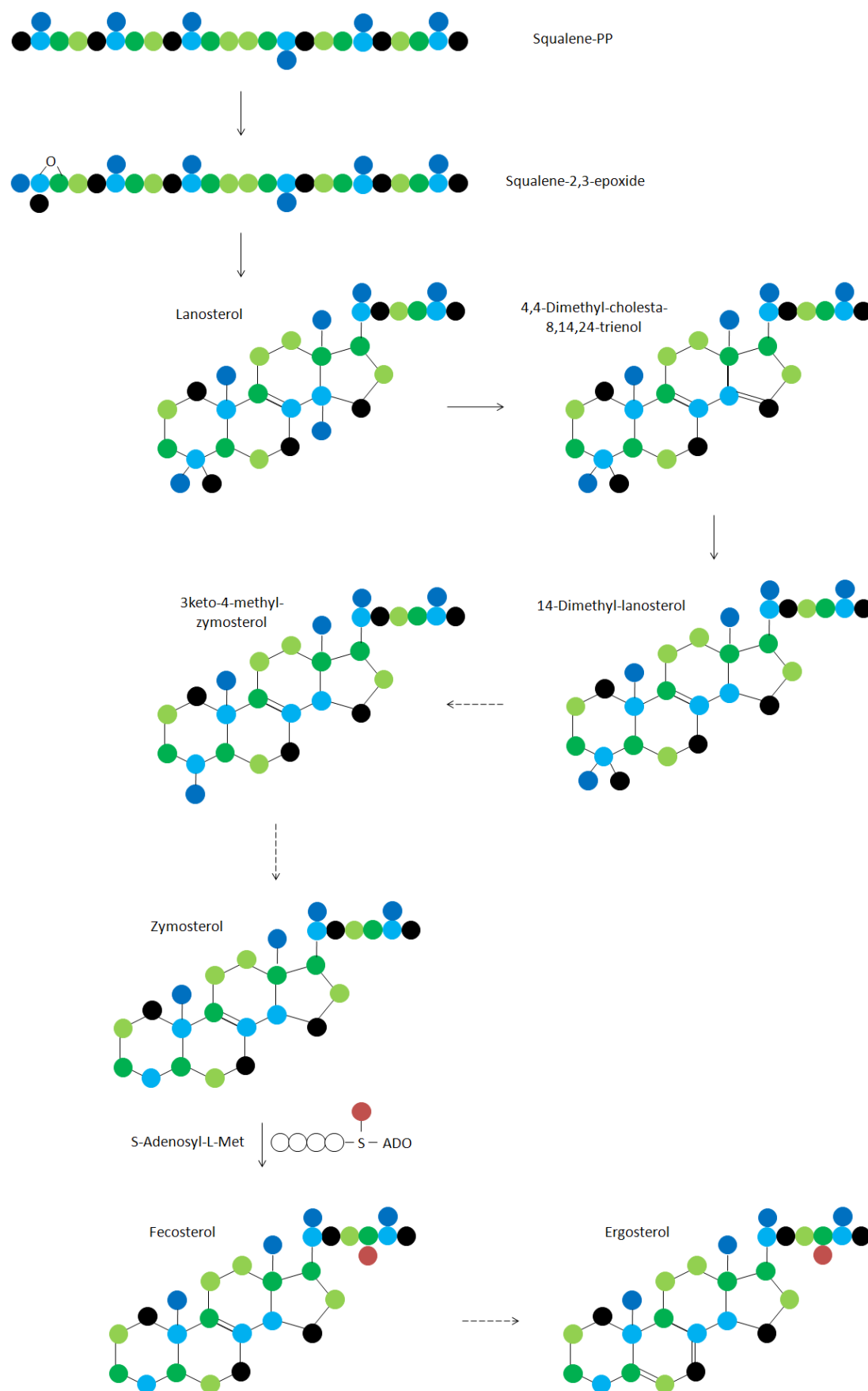


Figure 5.44b: Erg biosynthetic pathway. Red ball indicates an external source of carbon.



Now it is clear that Erg final structure is composed by different ac-CoA units:

a) Ac-CoA units that have conserved both the original carbon atoms (indicated with a red circle in tab. 5.5). The labeling state of the corresponding Erg positions depend on the labeling state of these ac-CoA units:

- Erg carbon atoms corresponding to C2 ac-CoA position will be labeled as both singlet and doublet, depending on the metabolic pathway followed for the ac-CoA unit generation (PPP or glycolysis, respectively). This is in agreement with the labeling state observed in C3, C5, C9, C17, C19 and C26 positions.
- Erg carbon atoms corresponding to C1 ac-CoA position, instead, will be labeled only as doublet, because there are no metabolic pathways leading to a single label in ac-CoA C1 position. This is in agreement with the labeling state observed in Erg C6, C10, C12, C16, C20, C23 and C25 positions.

b) Ac-CoA units that have lost their counterpart leading to corresponding Erg carbon atoms labeled only as singlet. This is in agreement with what was observed in C1, C4, C7, C8, C18, C22 and C27;

The only carbon resulted not enriched was C28: this is because sterol C24-methyl transferase (EC 2.1.1.41) converting zymosterol to fecosterol, inserts a CH<sub>3</sub> group in this position involving an “external” source of carbon, not labeled (S-Adenosyl methionine).

Bearing in mind these considerations it was possible to resolve the multiplicity of the other overlapped signals observed.

For examples, C2 should be only a doublet, because it corresponds to an ac-CoA C1 position: however it was not possible to observe its multiplicity because it is overlapped to the complex pattern of FA C16 signals.

Analogously, Erg C14 should be only a singlet, because it corresponds to an ac-CoA unit that has lost its counterpart. Unfortunately, its singlet is overlapped with MeOH signal; however, no doublet was observed around these signals, so this condition could be considered verified.

Erg C13 and C24 signals, both resonating around 43.5 ppm, correspond to ac-CoA C2 position, so they lead to a complex spectral pattern resulting from the overlap of their two singlets and doublets. Known this, it was possible to understand how to measure their relative coupling constants (fig. 5.45):  $^1J_{CC}$  relative to C24-C23 coupling ( $\sim 43$  Hz) was well in agreement with the corresponding coupling constant measured in C23 ( $\sim 44$  Hz, table 5.5). The same is for the other  $^1J_{CC}$  relative to the C13-C12 fragment ( $\sim 35$  Hz), in agreement with the corresponding coupling constant measured in C12 ( $\sim 34$  Hz).

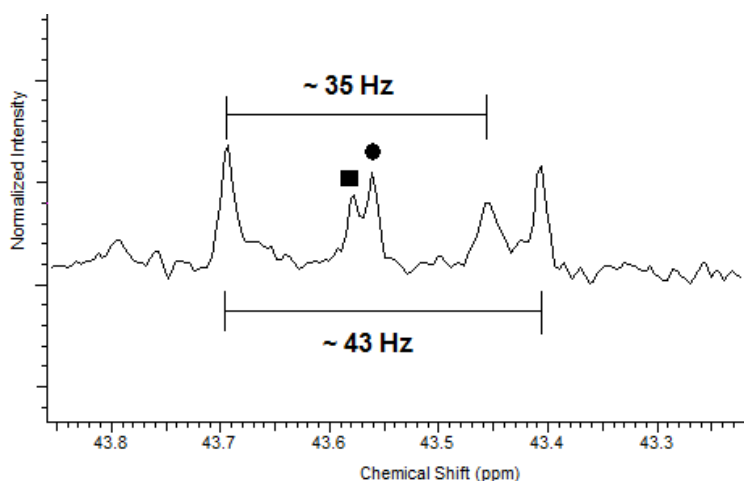


Figure 5.45: Erg C13 (square) and C24 (circle)  $^{13}\text{C}$  resonances and their coupling constants relative to C13-C12 and C24-C23 fragments, respectively.

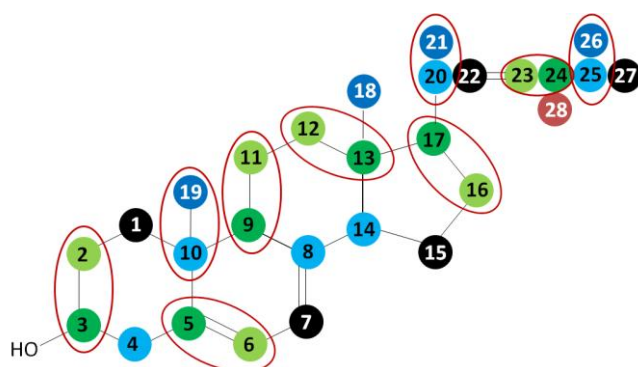
Finally, the complex pattern observed around 21.8 ppm is relative to C11, C15 and C21 altogether. Since C11 corresponds to an ac-CoA C1 carbon atom, it will give only a doublet. C15 is relative to an ac-CoA unit that has lost its counterpart, so it gives only a singlet, instead. C21 correspond to an ac-CoA C2 carbon atom and so it will show both a singlet and a doublet.

This means that the broad singlet observed in  $^{13}\text{C}$  NMR spectra acquired is due to the overlap of C15 and C21 singlets, whereas C11 and C21 determine two overlapped doublets.

Table. 5.5 summarized these considerations.

Table 5.5 - Assignment of Erg carbons in terms of Ac-CoA fragments, multiplicity and coupling constants.

Erg C position	Ac-CoA position	Multiplicity		$^1J_{13C-13C}$ [Hz]
		singlet	doublet	
1	2	√	-	-
2	1	n.d.	n.d.	n.d.
3	2	√	√	~ 37
4	1	√	-	-
5	1	√	√	~ 70
6	1	-	√	~ 70
7	2	√	-	-
8	2	√	-	-
9	2	√	√	~ 35.5
10	1	-	√	~ 35
11	1	-	√	~ 35.5
12	1	-	√	~ 34
13	2	√	√	~ 35
14	2	n.d.	n.d.	n.d.
15	2	√	-	-
16	1	-	√	~ 35
17	1	√	√	~ 35
18	2	√	-	-
19	2	√	√	~ 34
20	1	-	√	~ 34.3
21	2	√	√	~ 34.4
22	2	√	-	-
23	1	-	√	~ 44
24	2	√	√	~ 43
25	1	-	√	~ 36
26	2	√	√	~ 36
27	2	√	-	-
28	Met	not enriched		



Ergosterol structure and legend of table 5.5:

√ = observed  
 - = not observed  
 n.d. = not detected

Fig. 5.46 shows the fractional enrichments measured in each position of Erg carbon structure. It can be seen a clear trend indicating an higher enrichment level in mutant samples, as compared to WT (~ 50%): this means that, even if Erg concentration results statistically the same in both the strains (fig. 5.27b), *rag8* cells evidently used mostly, for Erg biosynthesis, the labeled substrate supplied.

Moreover, the observation of all these enriched sites is an advantage because, in this way, there are many measurements useful to calculate ratios between fluxes. Only those positions showing a labeling state both singlet (s) and doublet (d) (C3, C5, C9, C13, C17, C19, C21, C24 and C26), indeed, reflect the labeling state of the corresponding ac-CoA units belonging to the pool used for Erg biosynthesis (originated by PPP or by glycolysis, respectively). So, the ratio between these enrichments (s/d) could give quantitative information about the relative flux through these two pathways, as just seen before.

The averaged values obtained for this flux ratios result similar once again in both the strains (0.4 – 0.5): this suggests that the nature of the ac-CoA pool utilized for Erg synthesis is the same in WT and mutant cells.

Finally, for the FA moiety it was only possible to evaluate an averaged value of  $Y_{Ci}$  for the principal FA carbon atoms resulting in an enrichment statistically not different between WT and mutant samples. This supports the supposition that MUFA accumulation revealed in mutant samples should be due simply to their less turnover and not to an higher biosynthetic flux from ac-CoA.

Acetyl-CoA is the C2-carbon donor also for FA synthesis and elongation (see 1.3.6 Introduction) so, analogously to Erg, the relative ratio between the averaged enrichment as singlet (s) and the averaged enrichment as doublet (d) is useful to obtain quantitative information about the relative flux through PPP and glycolysis. This s/d ratio results in a similar value for both the strains (~ 0.20), suggesting again a similar nature of the ac-CoA pool used for FA synthesis in both WT and mutant samples. However, this value is different from the one obtained from Erg data.

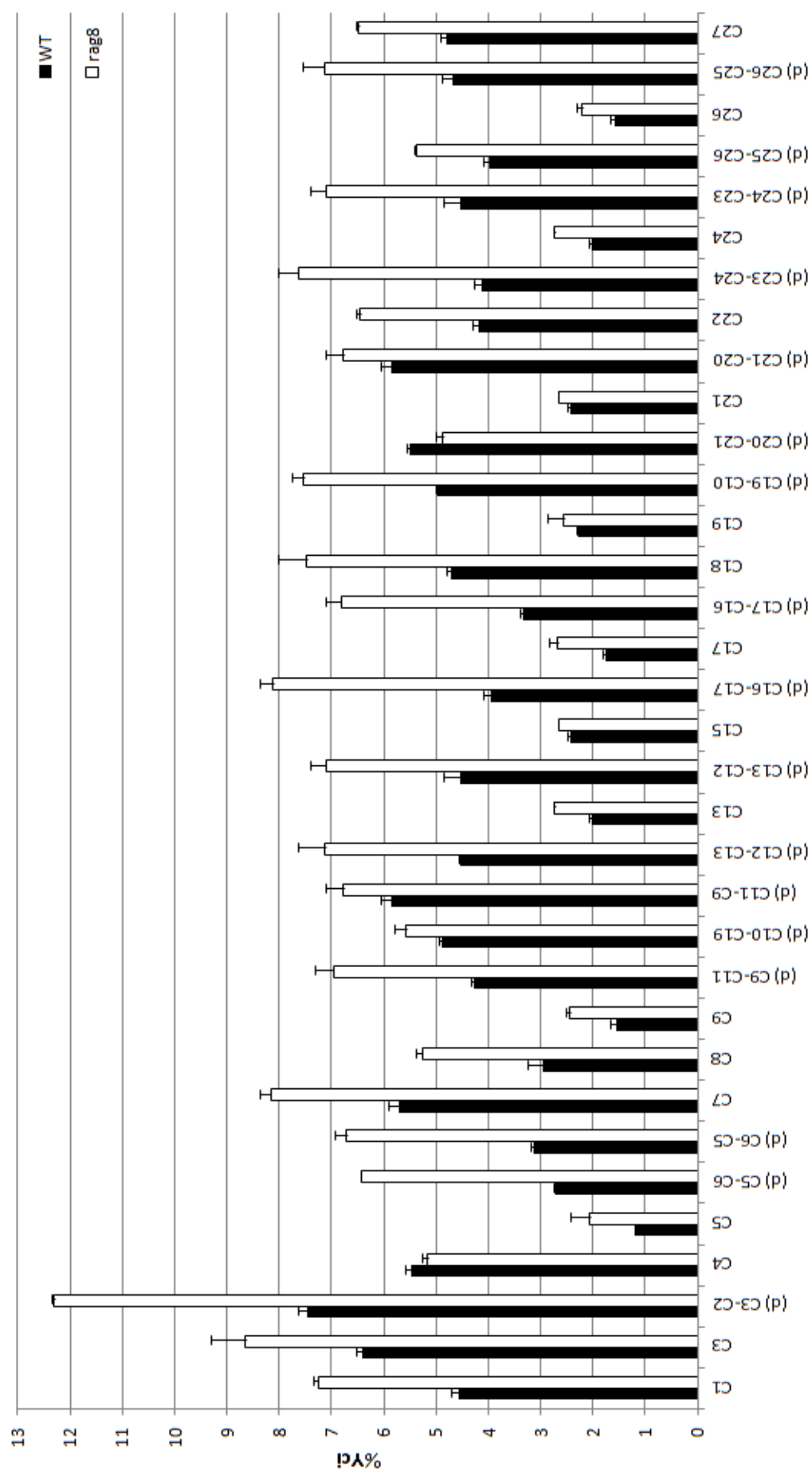


Figure 5.46: Fractional enrichments of the enriched ergosterol (Erg) carbon atoms observed by  $^{13}\text{C}$  NMR spectroscopy.

Table 5.6 summarize the values obtained in WT and mutant samples for the relative ratio between the PPP and the glycolytic fluxes (PPP/glycolysis) calculated respect to the ac-CoA moiety of the metabolites resulted enriched in both the hydro-alcoholic and organic cells extracts.

As just seen, Glu and Succ will be not considered now, because we are interested with evaluate “pure” ac-CoA pools, whereas, in such metabolites, the contribution of the peroxisomial flux could confuse our considerations. For this reason they are discussed separately, before.

Table 5.6 - WT and *rag8* relative ratio between PPP and glycolytic fluxes, calculated in both hydro-alcoholic (alanine and lysine) and organic (fatty acids and ergosterol) metabolites.

	PPP/glycolysis	WT	<i>rag8</i>
<b>Ala</b>	(s)C3/(d)C3-C2	$0.15 \pm 0.01$	$0.15 \pm 0.01$
<b>FA</b>	(s)/(d)	$0.22 \pm 0.01$	$0.18 \pm 0.01$
<b>Lys</b>	(s)C2/(d)C2-C1	$0.42 \pm 0.07$	$0.53 \pm 0.06$
<b>Erg</b>	(s)/(d)	$0.5 \pm 0.2$	$0.4 \pm 0.1$

From the values reported in table 5.6 it seems that there are at least two different ac-CoA pools: one relative to the ac-CoA contributing to Ala and FA synthesis (PPP/glycolysis  $\sim 0.2$ ) and another one relative to an ac-CoA pool generating Lys and Erg, characterized by an higher contribution of the flux through the PPP (PPP/glycolysis  $\sim 0.5$ ).

In the first case, ac-CoA could be a mitochondrial pool (the same of the mit-pyruvate generating Ala in *K. lactis* cells [179]), common to that one used for FA synthesis. This is in agreement with the existence of a specialized fraction of the endoplasmic reticulum (ER) called the MAM (Mitochondria Associated Membrane) fraction, in which probably FA synthesis occurs, and with the fact that mitochondria do not only adapt their activity to different physiological demands, but they are in addition able to modulate their sub-cellular distribution to local cellular requirements.

This is accomplished by the formation of a reticulated mitochondrial network which responds to cellular needs by constant fission and fusion events, thereby optimizing mitochondrial performance [23].

On the other hand, we have the evidence that the  $\alpha$ -KG pool involved in Lys biosynthesis is compartmentalized (in peroxisome, we supposed) and that it could exit from this organelle (it concurs to Glu synthesis, indeed). So we could suppose that this  $\alpha$ -KG pool, once outside, reacts with an ac-CoA pool, common to the one used in Erg synthesis, that is characterized by an higher contribution of the flux through the PPP and that could be at least localized nearby peroxisomes.

Moreover, since the endoplasmic reticulum (ER) is not a compact and well localized organelle within the cells, but an interconnected labyrinth of channels distributed within all the cytoplasm, is it possible that it establishes close contacts with other organelles [181]: so, we could reasonably suppose that Erg is synthesized in an ER zone connected to peroxisomes.