SUPPLEMENTARY MATERIAL

NMR based metabolomics approach to elucidate the differential cellular responses during mitigation of arsenic (III, V) in a green microalga

Neha Arora[1#], Durgesh Dubey[3#], Meenakshi Sharma[1], Alok Patel[1], Anupam Guleria[3], Parul A. Pruthi[1], Dinesh Kumar[3]*, Vikas Pruthi[1,2]*, Krishna Mohan Poluri[1,2]*

[1]Department of Biotechnology, [2]Centre for transportation systems (CTRANS), Indian Institute of Technology Roorkee, Roorkee - 247667, Uttarakhand, India
[3]Centre of Biomedical Research, SGPGIMS, Lucknow – 226014, Uttar Pradesh, India

***Corresponding Authors**

Dr. Dinesh Kumar

Email: dineshcmbr@gmail.com

Prof. Vikas Pruthi

Email: vikasfbs@iitr.ac.in

Dr. Krishna Mohan Poluri

Email: mohanpmk@gmail.com; krishfbt@iitr.ac.in

Ph: +91-1332-284779

Fax: +91-1332-286151

[#]**Authors contributed equally**

Figure S1: (A) IC_{50} value and, **(B)** arsenic removal efficiency of *Scenedesmus* sp. IITRIND2 different concentrations (0-500 mg/L) As (III) and As (V)

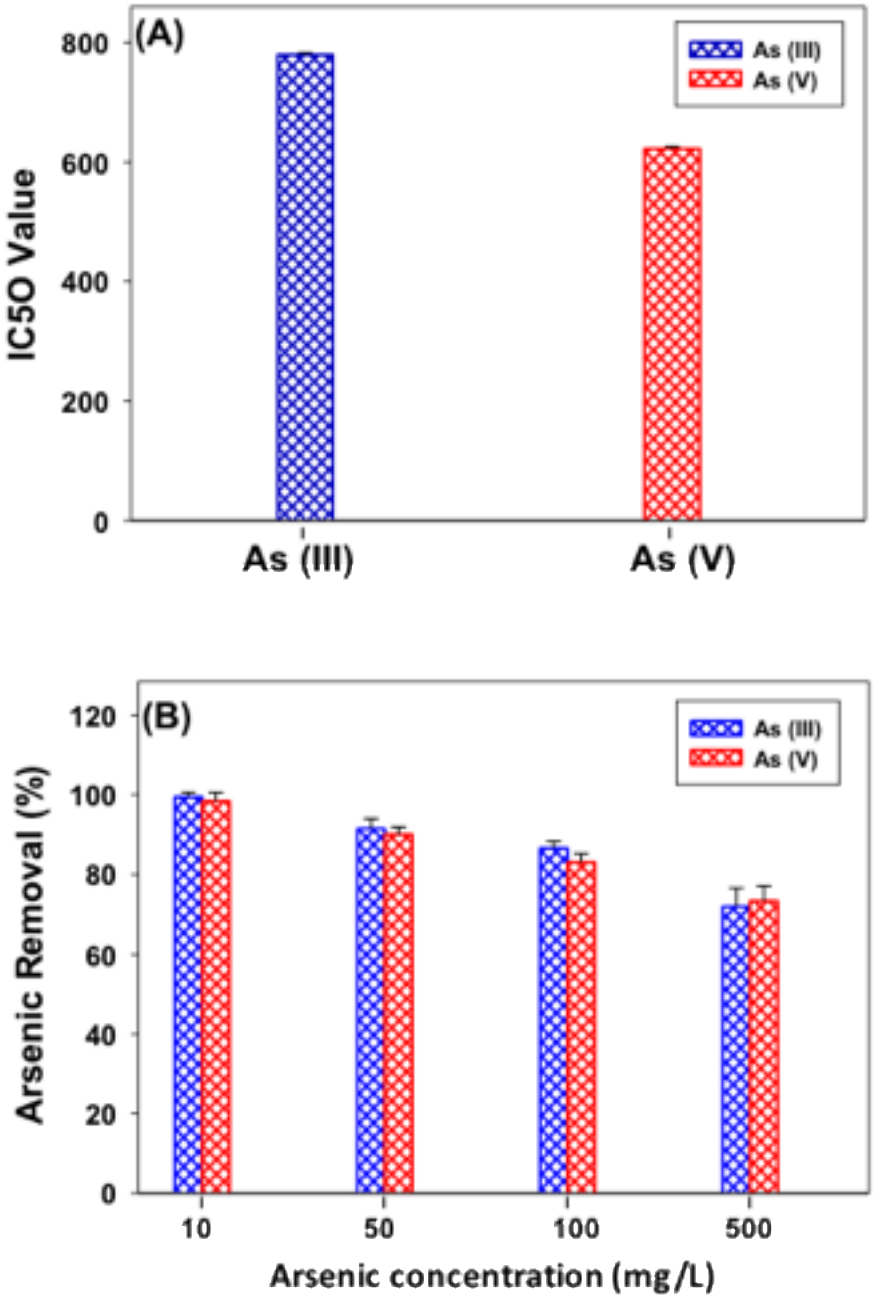


Figure S2: The representative 800 MHz ^{1}H-^{1}H TOCSY spectrum of *Scenedesmus* sp. IITRIND2 showing CH cross-peak assignments for various metabolites.

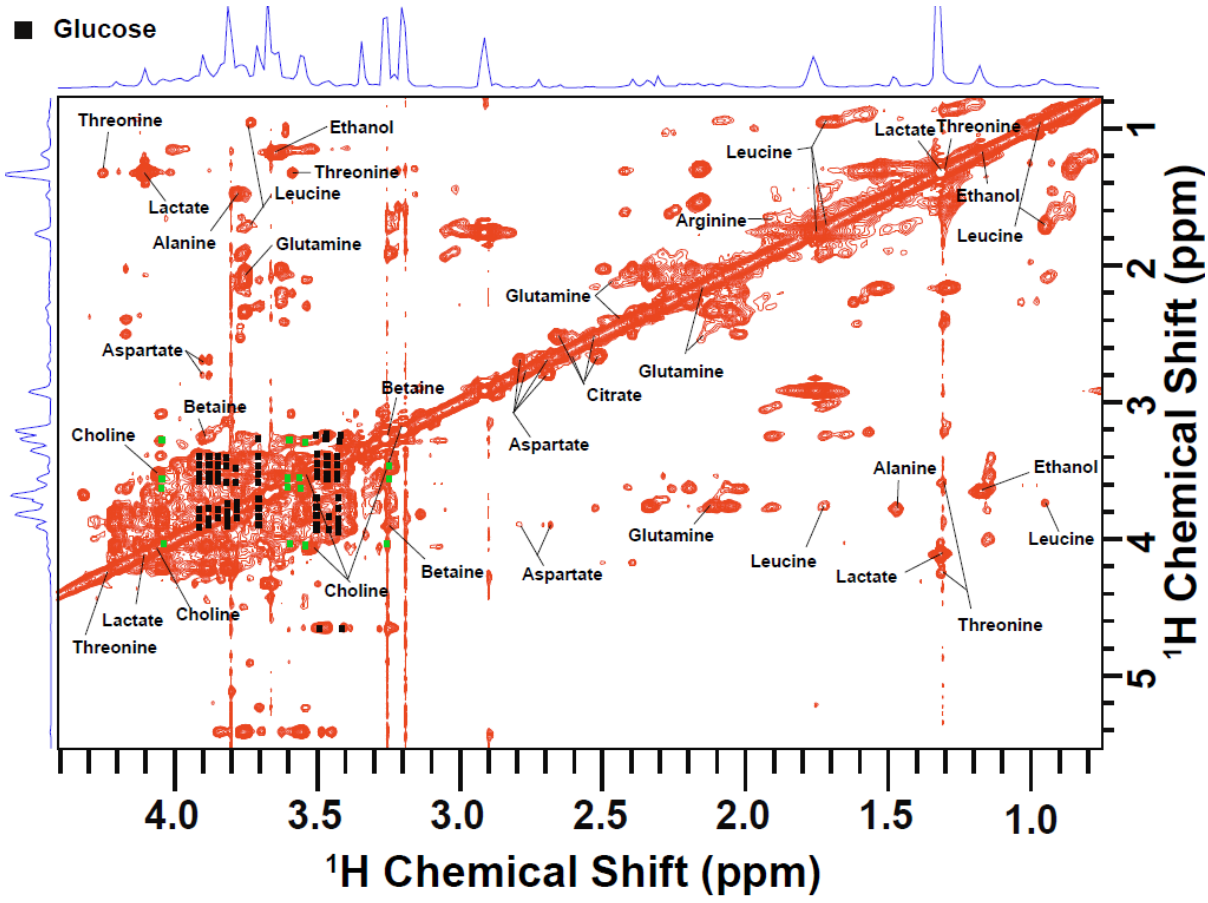
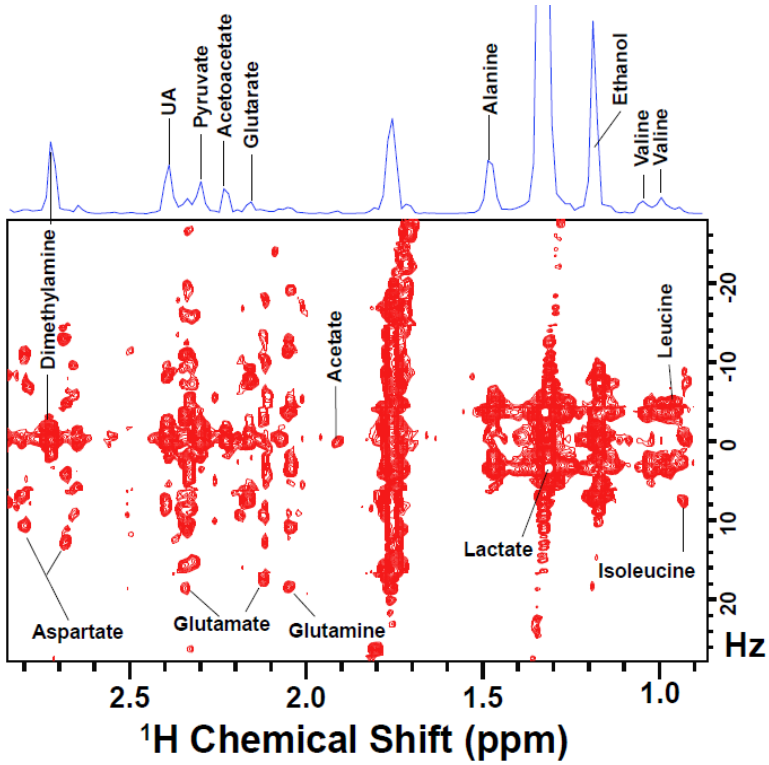


Figure S3: The representative 800 MHz ^{1}H-^{1}H J-Resolved (JRES) spectrum of *Scenedesmus* sp. IITRIND2 showing CH cross-peak assignments for various metabolites. Two different spectral region of the JRES spectrum are expanded for improved visualization: **Top Panel** - δ(0.9-2.8) ppm and, **Bottom Panel** - δ(3.1-4.6) ppm. Assigned peaks are annotated, whereas unassigned (UA) metabolite signals are labelled as UA.



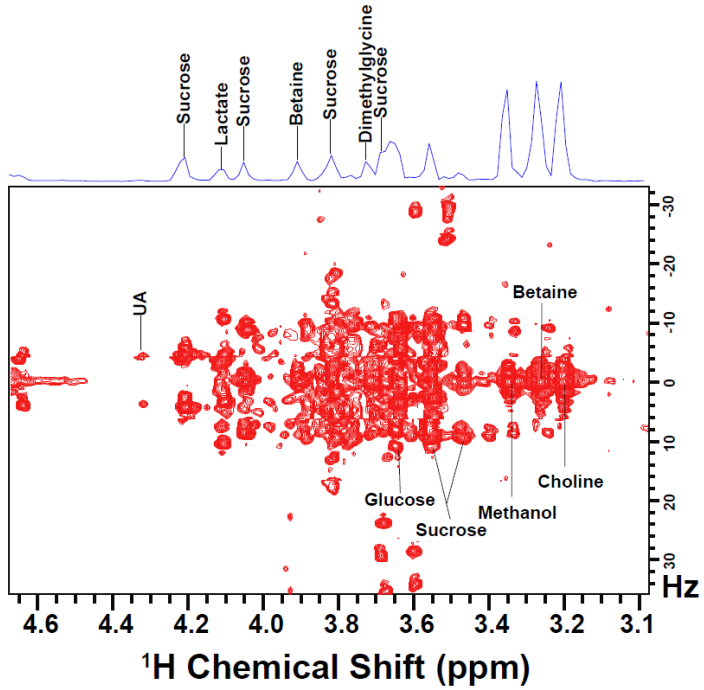


Fig. S4: The representative 800 MHz ^{1}H-^{13}C HSQC spectrum of *Scenedesmus* sp. IITRIND2 showing CH cross-peak assignments for various metabolites. GPC – Glycerophosphocholine.

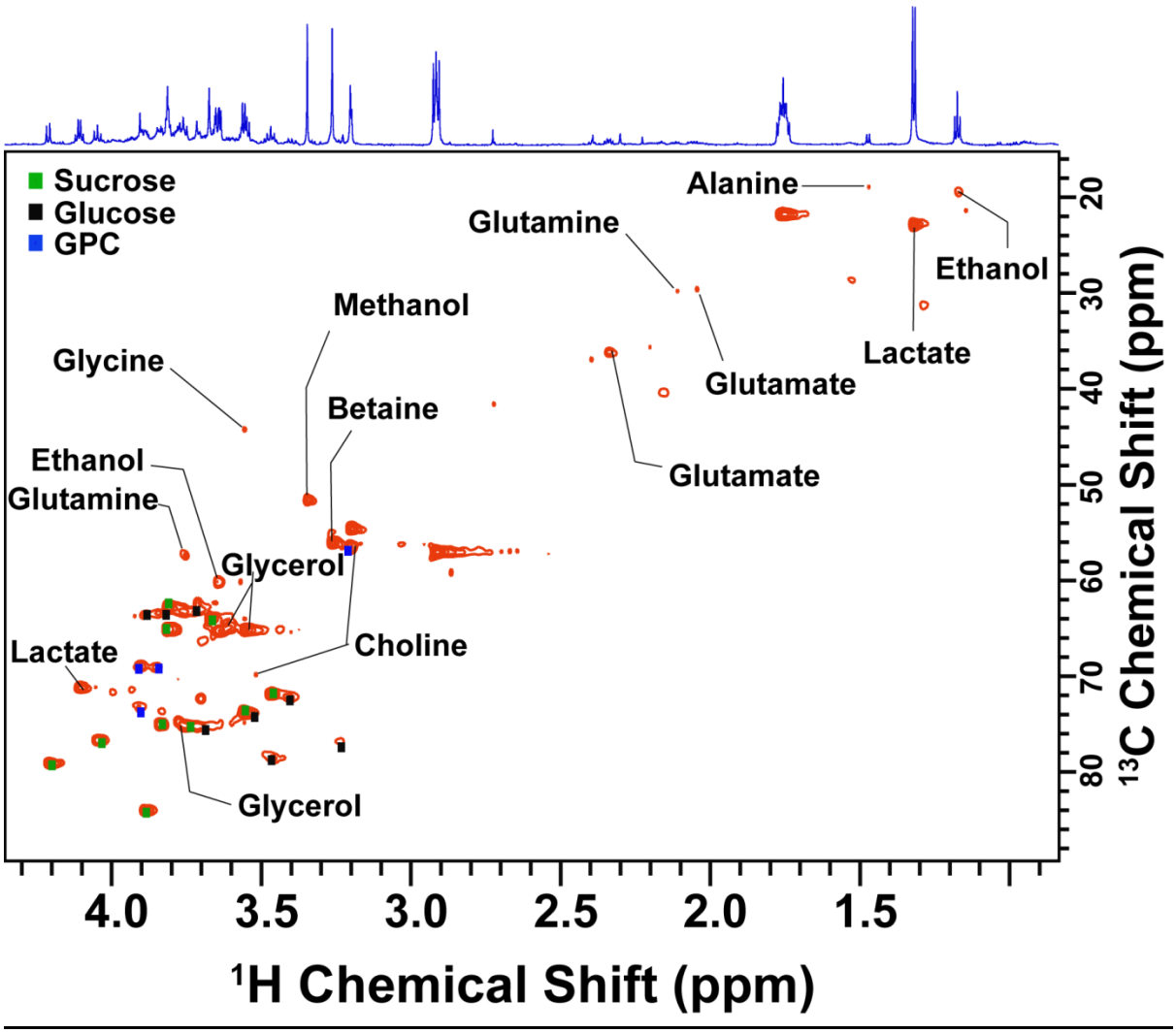


Table S1: List of metabolites along with their respective chemical shifts and their respective metabolic change patterns as a consequence of uptake of Arsenic III and V metal uptake. one-way ANOVA was conducted to determine significant ($p < 0.001$) metabolic changes. The up (↑) and down (↓) arrows represent, respectively, increased and decreased metabolite levels.

Metabolite Name	Assignment	Chemical shifts (δ) in ppm	Relative Change*	
			As (III) Spiking vs Control	As (V) Spiking vs Control
Amino acids				
Leucine	δ-CH3 δ-CH3	0.95 (d) 0.96 (d) $^{\epsilon}$	↑	↑↑
Isoleucine	γ-CH3 δ-CH3	0.93 (t) $^{\epsilon}$ 1.00 (d)	↑	↑↑
Valine	γ-CH3 γ-CH3	0.98 (d) 1.03 (d) $^{\epsilon}$	--	↑↑
Threonine	γ-CH_3	1.29 (d) $^{\epsilon}$	--	↑
Alanine	β-CH_3 α-CH	1.47 (d) $^{\epsilon}$ 3.79 (q)	↑	↑↑
Proline	γ-CH2 ½ β-CH2 ½ β-CH2	2.00 (m) $^{\epsilon}$ 2.06 (m) 2.34 (m)	↑	↑↑
Glutamate	β-CH_2 γ-CH_2	2.11 (m) $^{\epsilon}$ 2.34 (m)	↑	↑↑
Glutamine	β-CH_2 γ-CH_2	2.12 (m) 2.44 (m) $^{\epsilon}$	--	↑
N,N-Dimethylamine (DMA) $^{\#}$	N-CH_3	2.71 (s)	--	↑
Sacrosine $^{\#}$	N-CH_3	2.72 (s)	--	↑
N,NN,-Trimethylamine (TMA) $^{\#}$	N-CH_3	2.86 (s)	↑	↑
Glycine	α-CH_2	3.56 (s)	↑	↑↑

Cysteine	β-CH2 α-CH	3.07 (m) 3.97 (dd) $^{\epsilon}$	↑	↑↑
N,N-dimethylglycine	N-CH3	3.71 (s)	↑	↑↑
Tyrosine	C2H & C6H C3H & C5H	6.88 (d) $^{\epsilon}$ 7.18 (d)	↑↑	↑
Phenylalanine	C2H & C6H C4H C3H & C5H	7.31 (m) 7.37 (m) 7.43 (m) $^{\epsilon}$	↑	↑
γ -glutamyl-phenylalanine	C2H, C6H & C4H C3H & C5H	7.3 (m) $^{\epsilon}$ 7.4 (m)	↑	↑↑
Organic acids				
Lactate	β-CH_3 α-CH	1.32 (d) 4.10 (q) $^{\epsilon}$	↑↑	↑
Acetate	CH_3	1.91 (s)	--	↑
Glutarate	β, δ- CH2	2.16 (t)	--	↑
Succinate	α, β-CH_2	2.39 (s)	--	↑↑
Citrate	½ γ-CH2 ½ γ-CH2	2.52 (d) $^{\epsilon}$ 2.69 (d)	--	↑↑
Fumarate	CH	6.51 (s)	↑↑	↑↑
Formate	CH	8.44 (s)	↑↑	↑↑
Carbohydrates/sugar				
Sucrose	C10H C12H	3.46 (t) 3.55 (dd)	↑	↑↑

	C13H C11H C17H & C19H C5H & C9H C4H C3H C7H	3.66 (s) 3.75 (m) 3.77 (m) 3.81 (dd) 4.04 (t) 4.21(d) [€] 5.40(d)		
α-Glucose	C1H	4.63 (d)	↑	↑↑
β-Glucose	C1H	5.22 (d)	↑	↑↑
Mannose/Trehalose	C1H	5.19 (d)	↑	↑
Glucose-1-phosphate	C1H	5.55 (d,d)	↑↑	↑
Phosphagen				
Choline/PC	$N–(CH_3)_3$	3.20 (s)	↑	↑↑
GPC	$N–(CH_3)_3$	3.22 (s)	↑	↑
Osmolytes				
Glycerol	½ γ-CH2 ½ γ-CH2	3.63 (d) [€] 3.65 (d)	↑	↑
Betaine	$N–(CH_3)_3$ $β-CH_2$	3.26(s) 3.91 (s)	↓↓	↓
TMAO	$N–(CH_3)_3$	3.27(s) [€]	↓↓	↓
Nucleotides				
Adenine	C2H	8.19 (s) [€]	--	↓

	C6H	8.19 (s)		
ATP	C7H C12H C2H	8.61 (s) € 8.25 (s) 6.13 (d)	↑↑	↑↑
Others				
Isopropanol	β-CH_3	1.16 (d)	--	--
Ethanol	CH3	1.17 (t)	--	↑↑
3-hydroxy-isovalerate	γ-CH3	1.24 (s)	↑	↑↑
Acetone	CH_3	2.22 (s)	--	↑
Aceto-acetate	δ-CH3	2.3 (s)	--	↑
N,N-Dimethyl-formamide (DMF)	N-CH_3	2.85 (s)	--	↑
Ethanolamine	N-CH_2	3.13(m)	↑	↑↑
Methanol	CH_3	3.34 (s)	↑	↑↑
Hydroxy-benzoate derivative	C2H & C6H C3H & C5H	7.02 (d) € 7.79 (d,d)	↓	↓↓

Note: All the values of the metabolites were statistically significant with p value 0.001 respectively. For visualization interpretation, single (↑,↓) and double (↑↑,↓↓) arrows are used to represent relative change in the mean value of metabolite concentration (as evaluated from their respective box-plots). "€" represents he metabolite peak used for evaluating the quantitative difference as represented here using up (↑) and down (↓) arrows in the case of multiple signals/chemical shift values. "#" The metabolites showing discrepancy in the box-plot calculation because of some outlier values.

Appendix S1:

Materials and Methods: NMR Methodology for 2D NMR experiments:

Unambiguous assignments of various metabolites of *Scenedesmus* sp. IITRIND2 were obtained using two-dimensional *J*-resolved (JRES), TOCSY (total proton-proton correlation spectroscopy) and HSQC (heteronuclear single quantum correlation) spectrum. These spectra were recorded for some of the randomly selected normal and arsenic spiked samples.

Homonuclear 2D *J*-resolved spectra (jresgpprqf) were recorded in magnitude mode using quadrature phase (QF) with water pre-saturation during recycle delay (RD) of 2 sec. 16K data points were collected along direct proton (*F*2) dimension with spectral width of 16 ppm, whereas, along indirect J- Couplings (*F*1) dimension, 80 points (increments) corresponding to spectral width of 78 Hz (~0.0976 ppm) were collected and for each *F*1 increment, 32 transients were acquired. Prior to Fourier transformation (FT), free induction decay (FID) signals were weighted in both dimensions by a sine-bell function and zero-filled in the *F*1 dimension to 256 data points. The spectra were tilted by 45° to provide orthogonality of the chemical shift and coupling constant axes and subsequently symmetrized about the *F*1 axis.

Two-dimensional ^{1}H-^{1}H TOCSY (dipsi2esgpph) and sensitivity enhanced ^{1}H–^{13}C HSQC (hsqcetgp) spectra were acquired in phase sensitive mode using time proportional phase incrementation (TPPI). 2D TOCSY spectrum was recorded using 2048 data points along direct dimension (*F*2) and 512 increments along indirect dimension (*F*1) with 32 transients per increment and a spectral width of 12 ppm in both dimensions. The FIDs were weighted using a sine–bell-squared function in both dimensions and zero filled to 2048 and 4096 data points, respectively, in the *F*1 and *F*2 dimensions prior to FT. For TOCSY experiment a mixing time of 80 ms was used. Spin-lock was achieved by a DIPSI2 pulse sequence during the TOCSY mixing time. The relaxation delay (RD) was 3.0 sec.

2D HSQC spectrum was recorded with inverse ^{13}C detection and using ^{13}C decoupling during acquisition using GARP-1.2. A relaxation delay of 2.0 sec was used between successive pulse sequence cycles and a refocusing delay equal to 1/(4*1*J*C–H =1.75 ms) was employed. 2048 x 256 complex *t*2 (^{1}H) and *t*1(^{13}C) data points with 128 scans per increment were acquired with spectral widths of 12 and 165 ppm, respectively, in the ^{1}H and ^{13}C dimensions. The FIDs were weighted using a sine-bell-squared function in both dimensions and zero filled to 4096 and 2048 data points, respectively, in the *F*2 and *F*1 dimensions prior to FT. After FT, the final spectrum was manually phase corrected and ^{1}H and ^{13}C dimensions were referenced to lactate methyl protons and carbon, respectively, at 1.31 and 22.5 ppm.