Metabolomics (2009) 5:292–301
DOI 10.1007/s11306-008-0151-1

ORIGINAL ARTICLE

Metabolic characterization of primary human colorectal cancers using high resolution magic angle spinning ^{1}H magnetic resonance spectroscopy

M. Piotto · F.-M. Moussallieh · B. Dillmann · A. Imperiale · A. Neuville · C. Brigand · J.-P. Bellocq · K. Elbayed · I. J. Namer

Received: 21 October 2008 / Accepted: 10 December 2008 / Published online: 24 December 2008

Abstract Colorectal cancer is one of the most frequent and most lethal forms of cancer in the western world. The aim of this study is to characterize by ^{1}H high resolution magic angle spinning NMR spectroscopy (HRMAS) the metabolic fingerprint of both tumoral and healthy tissue samples obtained from a cohort of patients affected by primary colorectal adenocarcinoma. By analyzing HRMAS data using multivariate statistical analysis (PLS-DA), the two types of tissues could be discriminated with a high level of confidence. The identification of the metabolites at the origin of this discrimination revealed that adenocarcinomas are richer in taurine, glutamate, aspartate, and lactate whereas healthy tissues contain a higher amount of myo-inositol and β-glucose. The statistical model resulting from the PLS-DA analysis was subsequently used to perform a blind test on tumoral and healthy colon biopsies. The results of the classification showed that the HRMAS analysis has very high sensitivity and specificity.

Keywords Colorectal cancer · Biopsies · HRMAS · Statistical analysis · Taurine · Myo-inositol

1 Introduction

Colorectal cancer is one of the most frequent cancers in the western world (Jemal et al. (2007); Libutti et al. (2008)). The most common colon cancer is adenocarcinoma which is a malignant epithelial tumor that originates from the glandular epithelium of the colorectal mucosa. Patient prognosis is based on several parameters like the clinical stage of the disease, the histological type and grade of the tumor (Cusack et al. (1996)) and the analysis of oncogenes and molecular markers (Fearon et al. (1990); Leichman et al. (1997)). In order to develop more efficient diagnostic and prognostic tools, current research focuses on both the genetic and the molecular basis of the disease.

^{1}H high resolution magic angle spinning (HRMAS) is a nuclear magnetic resonance (NMR) technique (Andrew et al. (1958); Lippens et al. (1999); Lowe (1959)) that allows the characterization of the metabolic phenotypes of intact cells, tissues and organs, under both normal and pathological conditions (Cheng et al. (1997); Cheng et al. (1998); Cheng et al. (2000); Martínez-Bisbal et al. (2004); Martínez-Granados et al. (2006); Sitter et al. (2002); Sjøbakk et al. (2008); Tzika et al. (2002)). HRMAS allows the assessment of cellular metabolic networks that are directly related to specific genetic information and to the regulation of specific gene transcripts. Several human gynecological, neurological, and urological malignancies have been studied by HRMAS with promising results.

M. Piotto (✉)
Bruker BioSpin, 34 rue de l'industrie, 67166 Wissembourg, France
e-mail: martial.piotto@bruker.fr

M. Piotto · F.-M. Moussallieh · B. Dillmann · K. Elbayed
Institut de Chimie, Université Louis Pasteur, 4 rue Blaise Pascal, 67000 Strasbourg, France

F.-M. Moussallieh · B. Dillmann · A. Imperiale · I. J. Namer (✉)
Department of Biophysics and Nuclear Medicine, University Hospitals of Strasbourg, 67098 Strasbourg, France
e-mail: Izzie.Jacques.NAMER@chru-strasbourg.fr

A. Neuville · J.-P. Bellocq
Department of Pathology, University Hospitals of Strasbourg, 67098 Strasbourg, France

C. Brigand
Department of Digestive Surgery, University Hospitals of Strasbourg, 67098 Strasbourg, France

The aim of the present study is to characterize by ^{1}H HRMAS the metabolic fingerprint of both tumoral and healthy tissue biopsy samples obtained from a cohort of patients affected by primary colorectal adenocarcinoma. One of the objectives of the project is to develop and validate in a clinical set-up a statistical metabolic model capable of achieving a reliable classification of cancerous versus healthy biopsies. In order to obtain reliable statistical results, the study was performed on a relatively large cohort of patients ($n = 44$).

The study of the metabolic content of colon biopsies by magnetic resonance has been the topic of only a few papers in the literature. In fact, the study presented in this manuscript is, to our knowledge, the first one involving the direct comparison of the metabolic content of healthy and cancerous human colon biopsies using intact biopsy samples.

2 Materials and methods

2.1 Patient population

This study involves the use of human colon biopsies obtained from the Tumor bank of the University Hospitals of Strasbourg, France. Between March 2007 and December 2007, 44 patients (26 men, 18 women; mean age: 68 ± 12 years, age range: 45–90 years) with histologically proven colorectal adenocarcinoma were prospectively selected. All patients underwent a surgical resection of the primary lesion and a radical lymphadenectomy. Neither neo-adjuvant chemotherapy nor radiotherapy were performed prior to the operation. For the HRMAS analysis, two specimens containing respectively tumoral and healthy tissue were obtained from each patient. For four patients, only the healthy or the tumoral biopsy was available. The total number of biopsies was therefore equal to 84. Tissue specimens were collected with minimum ischemic delays after resection (average time 50 ± 25 min) and snap-frozen in liquid nitrogen before being stored at −80°C. All tissue samples exhibited a viable tumor/necrosis ratio and were quantitatively and qualitatively adequate to perform a correct HRMAS analysis.

2.2 HRMAS analysis

2.2.1 Rotor preparation

Each biopsy sample was prepared at −10°C by introducing 15 to 20 mg of biopsy into a disposable 30 µl KelF insert. To provide a lock frequency for the NMR spectrometer, 10 µl of D_2O containing 1% w/w TSP were also added to the insert. The insert was then sealed tightly with a conical plug and stored at −80°C until the HRMAS analysis. The insert insures that the entire biopsy sample is detected by the radio-frequency coil of the probe and that no leaks occur during the HRMAS analysis. Shortly before the HRMAS analysis, the insert was placed into a standard 4 mm ZrO_2 rotor and closed with a cap. The ensemble was then inserted into a HRMAS probe pre-cooled at 3°C. All HRMAS experiments were performed at 3°C and were started immediately after the temperature inside the probe had reached the equilibrium condition (5 min). Upon completion of the HRMAS analysis (20 min), the insert was taken out of the rotor and stored back at −80°C. This methodology prevents sample cross-contamination and allows storing all the biopsies in a reliable matter. A considerable amount of time is also saved by eliminating the cleaning process of the rotor. Furthermore, complementary NMR analysis can be performed at a later stage on any of the stored biopsy inserts.

2.2.2 Data acquisition

HRMAS spectra were recorded on a Bruker Avance III 500 spectrometer operating at a proton frequency of 500.13 MHz. This instrument is installed at the Hautepierre university hospital in Strasbourg and is dedicated to the analysis of biopsies by HRMAS. It is operated by qualified scientific and medical personnel in the context of the CARMeN project which aims at the creation of an extensive metabolic database covering most current human tumors. The spectrometer is equipped with a 4 mm double resonance (^{1}H, ^{13}C) gradient HRMAS probe. A Bruker Cooling Unit (BCU) is used to regulate the temperature at 3°C by cooling down the bearing air flowing into the probe. All NMR experiments were conducted on samples spinning at 3,502 Hz in order to keep the rotation sidebands out of the spectral region of interest. For each biopsy sample, a one-dimensional proton spectrum using a Carr–Purcell–Meiboom–Gill (CPMG) pulse sequence was acquired (Bruker *cpmgpr1d* pulse sequence). This sequence allows a better observation of the metabolites by reducing the intensity of lipid signals in the spectrum. The inter-pulse delay between the 180° pulses of the CPMG pulse train was synchronized with the sample rotation and set to 285 µs ($1/\omega_r = 1/3502 = 285$ µs) in order to eliminate signal losses to B_1 field inhomogeneities (Elbayed et al. (2005); Piotto et al. (2001)). The number of loops was set to 328 giving the CPMG pulse train a total length of 93 ms. In order to quantify, in an absolute manner, the amount of metabolites present in each sample, a calibrated Eretic signal (Akoka et al. (1999)) was included in all the 1D experiments. The Lorentzian Eretic signal was added digitally in the spectrum after data acquisition and processing. This procedure allows obtaining an Eretic signal which is perfectly in phase with the NMR signals and

which does not compromise the quality of the baseline. Using a standard solution of known concentration of lactate, the amplitude of the Eretic signal was calibrated to correspond to 0.62 μmol of protons. In order to accurately calculate the metabolite concentration in each biopsy, a correction factor taking into account the difference in pulse widths between the standard lactate sample and the different biopsies was also applied (Wider and Dreier (2006)). The CPMG experiment was acquired with the following parameters: sweep width 14.2 ppm, number of points 32 k, relaxation delay 2 s and acquisition time 2.3 s. A total of 128 FID were acquired resulting in an acquisition time of 10 min. All spectra were recorded in such a manner that only a zero phase order correction was necessary to properly phase the spectrum. The FID was multiplied by an exponential weighing function corresponding to a line broadening of 0.3 Hz prior to Fourier transformation. All spectra were processed using automatic base line correction routines. ^{1}H spectra were referenced by setting the lactate doublet chemical shift to 1.33 ppm.

In order to confirm resonance assignments, two-dimensional homonuclear and heteronuclear experiments were also recorded on eight samples. 2D DIPSI2 spectra (Shaka et al. (1988)) were acquired with a 170 ms acquisition time, a 60 ms mixing time, a 14.2 ppm spectral width and a 1.5 s relaxation delay. Thirty-two transients were averaged for each of the 512 increments during t1, corresponding to a total acquisition time of 8 h. Data were zero filled to a 2 k * 1 k matrix and weighted with a shifted square sine bell function prior to Fourier transformation. 2D ^{1}H-^{13}C HSQC experiments using echo–antiecho gradient selection for phase-sensitive detection (Davis et al. (1992)) were acquired using a 73 ms acquisition time with GARP ^{13}C decoupling and a 1.5 s relaxation delay. A total of 116 transients were averaged for each of 256 t1 increments, corresponding to a total acquisition time of 15 h. Two 1 ms sine-shaped gradient pulses of strength 40 and 10.05 G/cm were used in the experiment. Data were zero-filled to a 2 k * 1 k matrix and weighted with a shifted square sine bell function before Fourier transformation.

2.2.3 *Multivariate statistical analysis*

The spectral region between 4.7 and 0.5 ppm of each 1D CPMG NMR spectrum was automatically binned into regions of 0.01 ppm using the AMIX 3.8 software (Bruker GmbH, Germany). This procedure minimizes the effect of peak shifts due to pH variations. The peak integral within each 0.01 ppm region was computed and normalized with respect to the total integral of the spectrum in the 4.7–0.5 ppm region. This process generated an X data matrix containing 421 columns (chemical shifts) and 84 rows (corresponding to healthy and tumoral biopsies of 44 patients). Data sets were then imported into the SIMCA P 11.0 software (Umetrics AB, Umeå, Sweden) and pre-processed using unit variance scaling of the X columns by weighing each integral region by $1/SD_k$, where SD_k represents the standard deviation of the kth column in the X matrix. This procedure gives equal weight to metabolites present in high and low concentrations while minimizing the contribution from noise.

The X matrix was analyzed using principal component analysis (PCA) within the SIMCA P 11.0 software package. This procedure allows to quickly evaluate the quality of the data and to identify possible outliers. After the PCA analysis, Partial Least Square Discriminant Analysis (PLS-DA) (Tenenhaus (1998)) was conducted in order to build a statistical model that optimizes the separation between the two classes of patients. The number of components of the PLS-DA model was determined by cross-validation. The class membership of each sample was iteratively predicted, using the results to generate a goodness of fit measure ($Q^2 = 1$-PRESS/SS) for the overall model. PRESS is the predicted squared sum of error and represents the squared differences between observed and predicted Y values when each sample is kept out of model development and SS is the residual sum of squares of the previous dimension. Since PLS-DA models are usually built using a relatively small number of samples compared to the number of variables, an extensive cross-validation of the model is mandatory to avoid overfitting of the data (Westerhuis et al. (2008)). The maximum theoretical value for Q^2 is equal to 1 for a perfect prediction. However, a Q^2 value superior to 0.5 is generally considered to be a decent predictor. In order to validate the model, a circular validation procedure was performed. The class labels of the biopsies (cancerous versus healthy) in the Y column matrix were randomly permuted and for each permutation, a new statistical model was computed along with its Q^2 value. This procedure was repeated for a large number of permutations and the Q^2 value were plotted as a function of the number of permutations in the Y matrix. A model was considered as valid if the regression line obtained by circular permutation showed a negative slope and if all the estimated Q^2 were smaller than the original one. This feature proves that none of the models computed on the permuted Y matrix are better than the model computed using the original Y matrix.

3 Results and discussions

3.1 Spectral assignment

Representative HRMAS CPMG magnetic resonance spectra of adenocarcinomas and healthy tissues originating from the

same patient are presented in Fig. 1. Metabolites were assigned using standard metabolite chemical shift tables available in the literature (Martínez-Bisbal et al. (2004)). 2D DIPSI2 and HSQC spectra were also used to confirm some assignments. A total of about 30 different metabolites were identified from these spectra (Table 1). The respective metabolic content of healthy and cancerous biopsies can be directly inferred from Fig. 1 since the intensity of each spectrum was normalized with respect to the amplitude of the Eretic signal and the weight of biopsy present in each sample. Figure 1 tends to show that adenocarcinomas are characterized by a higher level of taurine while healthy tissues contain a larger amount of myo-inositol. It is important to emphasize that since both biopsies were obtained simultaneously from the same patient, the direct comparison of their metabolic content is extremely reliable since the time of

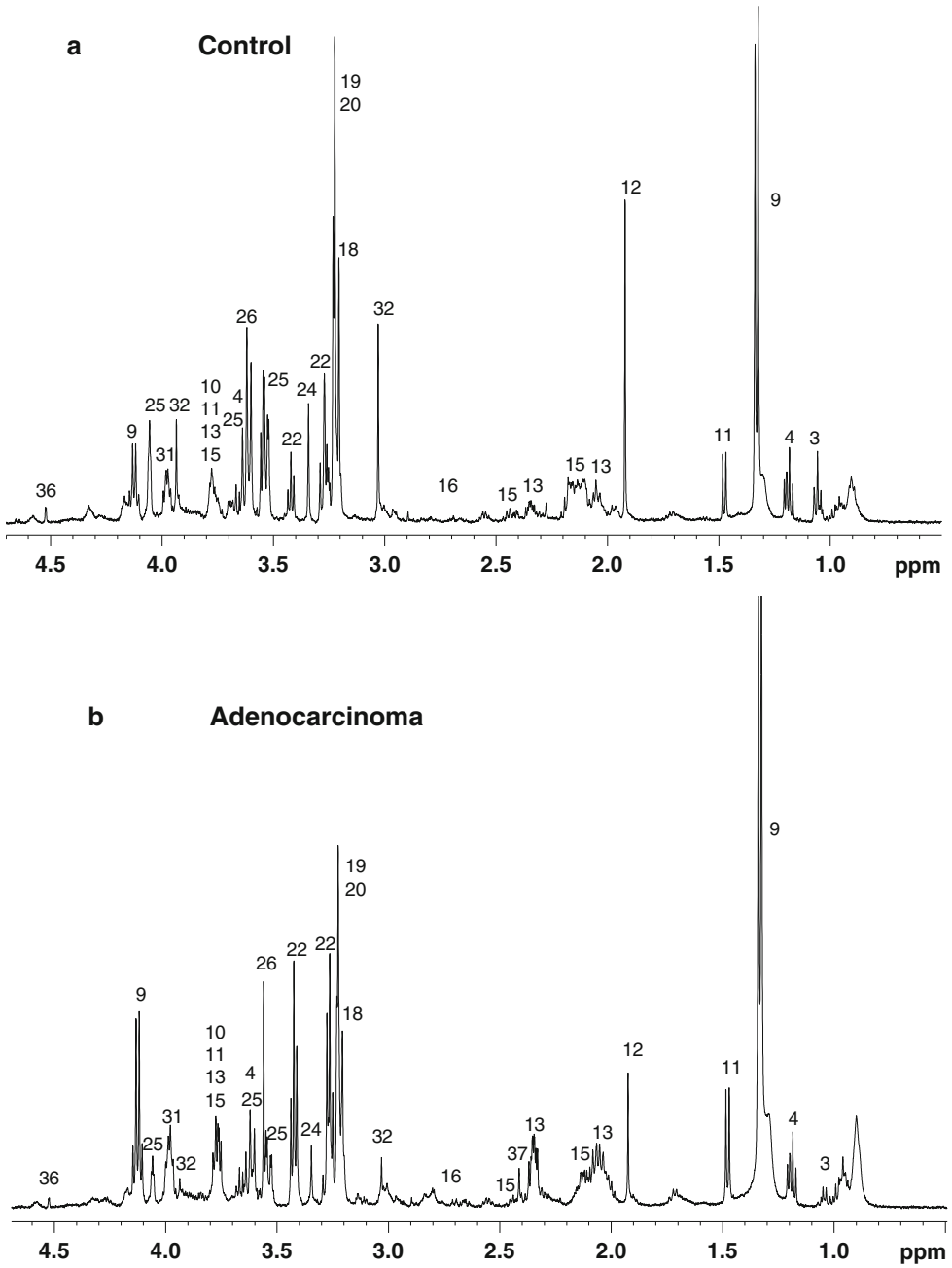

Fig. 1 Representative 1D ^{1}H CPMG HRMAS spectra of healthy (**a**) and adenocarcinoma (**b**) colorectal tissues originating from the same patient. For display purposes, the amplitude of the lactate peak at 1.33 ppm has been cut. Partial metabolite assignment in the 4.7–0.5 ppm region is indicated. The numbers refer to the metabolites listed in Table 1. The Eretic peak at 10.3 ppm, which is not shown in this plot, is a synthetic Eretic peak that is used for the absolute quantification of the metabolites

Table 1 ^{1}H resonance assignments of the metabolites present in cancerous and healthy human colons. The nomenclature used for fatty acids is defined in the publication of Martinez-Bisbal (Martínez-Bisbal et al. (2004))

	Metabolites	Group	^{1}H chemical shift (ppm)	^{13}C chemical shift
1	Isoleucine	δCH_3	0.93	13.77
		γCH_3	1.00	17.37
		γCH_2	1.48	27.13
		αCH	3.67	62.25
2	Leucine	δCH_3	0.95	23.36
		$\delta' CH_3$	0.96	24.61
		γCH	1.71	26.69
		βCH_2	1.71	42.40
		αCH	3.73	56.06
3	Valine	γCH_3	0.98	19.14
		$\gamma' CH_3$	1.04	20.56
		βCH	2.27	31.77
4	Ethanol	CH_3	1.17	19.49
		CH_2OH	3.65	60.15
5	Fatty acids (a)	$(2)CH_2$	1.28	34.46
		$(1)CH_2$	1.29	25.35
6	Fatty acids (b)	$(2)CH_2$	2.04	27.39
		CH_2	2.80	28.05
7	Fatty acids (a) (b)	$(n)CH_2$	1.28	32.24
8	Fatty acids (c)	$(2)CH_2$	1.58	27.31
		$(1)CH_2$	2.25	36.06
9	Lactate	CH_3	1.33	22.65
		CH	4.12	71.05
10	Lysine	γCH_2	1.46	24.19
		δCH_2	1.71	29.14
		βCH_2	1.90	32.50
		εCH_2	3.01	41.84
11	Alanine	βCH_3	1.47	18.80
		αCH	3.78	53.18
12	Acetate	CH_3	1.92	25.97
13	Glutamate	βCH_2	2.06	29.69
		γCH_2	2.34	35.93
		αCH	3.76	57.19
14	Methionine	εCH_2	2.12	16.50
15	Glutamine	βCH_2	2.14	NA
		γCH_2	2.44	33.37
		αCH_2	3.77	NA
16	Aspartic acid	$\beta CH_2(u)$	2.70	39.15
		$\beta CH_2(d)$	2.80	39.15
		αCH	3.90	54.81
17	Phosphoethanolamine	$-CH_2-NH_3^+$	3.22	43.21
		$-CH_2-O-$	3.98	62.94
18	Choline	$-N^+-(CH_3)_3$	3.21	NA
		βCH_2	3.52	69.84
19	Phosphorylcholine	$-N^+-(CH_3)_3$	3.22	56.49
		βCH_2	3.61	69.03
		αCH	4.18	60.59
20	Glycerophosphocholine	$-CH_2-NH_3^+$	3.23	NA
		αCH_2	4.29	56.36
21	Arginine	βCH_2	1.91	30.17
		δCH_2	3.22	43.23
22	Taurine	$-CH_2-NH_3^+$	3.26	50.07
		$-CH_2-SO_3^-$	3.42	38.07
23	Proline	γCH_2	2.00	26.45
		$\beta CH_2(u)$	2.07	31.70
		$\beta CH_2(d)$	2.35	31.70
		$\delta CH_2(u)$	3.33	48.75
		$\delta CH_2(d)$	3.41	48.75
		αCH	4.13	63.81
24	Scyllo-Inositol	All Hs	3.34	76.25
25	Myo-Inositol	C5H	3.27	76.96
		C1H, C3H	3.54	73.80
		C4H, C6H	3.62	75.05
		C2H	4.05	74.90
26	Glycine	αCH	3.56	44.10
27	Threonine	αCH	3.59	63.14
		βCH	4.26	68.55
28	Glycerol	1,3 $CH_2OH(u)$	3.56	65.10
		1,3 $CH_2OH(d)$	3.65	65.10
		–CH(OH)–	3.78	74.82
29	β-Glucose	C6H(u)	3.73	63.53
		C6H(d)	3.89	63.53
		C1H	4.65	NA
30	α-Glucose	C1H	5.23	NA
31	Serine	αCH	3.84	59.08
		βCH	3.97	62.90
32	Creatine	CH_3	3.03	39.67
		CH_2	3.93	56.32
33	Asparagine	αCH	4.00	53.94
34	Tyrosine	CH 3,5	6.88	118.34
		CH 2,6	7.18	133.32
35	Phenylalanine	CH 2,6	7.32	131.82
		C4	7.40	131.69
36	Ascorbic acid	C4H	4.53	NA
37	Succinic acid	(α,βCH2)	2.40	NA

ischemia is identical for both samples. The nature and the concentration of the metabolites present in each biopsy can therefore be reliably compared.

3.2 Statistical analysis

PCA analysis of the CPMG data applied to the 4.7–0.5 ppm chemical shift range and using the entire patient population was not able to discriminate adenocarcinomas from healthy tissues in a statistically significant manner. A few outliers were however identified by this analysis. In this type of studies, outliers can occur because of erroneous spectrum referencing, sample pollution due to the fixing medium (o.c.t. compound, Tissue-Tek, Sakura) employed by the pathologists or an unusual amount of lipids. In our study, 12 biopsies out of a total of 84 were removed, leaving a cohort of 72 biopsies. PLS-DA analysis of the CPMG data was initially performed on the same population of 72 biopsies. As for the PCA analysis, buckets within the 4.7–0.5 ppm range were initially used. Biopsies were classified in a Y column matrix as being either adenocarcinomas (value of 0) or healthy tissues (value of 1). The analysis generated a two component PLS-DA model characterized by a faithful representation of the Y data ($R^2Y = 0.80$) and, more importantly, by a very good cumulative confidence criterion of prediction ($Q^2 = 0.68$). The score plot of the PLS-DA model (Fig. 2a) showed a very clear separation of the two sets of biopsies. The loading plot (Fig. 2b) of the PLS-DA model revealed a statistically significant elevation of taurine, glutamate, aspartate, and lactate in adenocarcinomas compared to healthy tissues. Conversely, the metabolic signature of normal colon tissue consisted of a higher concentration of myo-inositol and β-glucose. The assessment of the metabolites responsible for this classification can lead to the detection of pathological biomarkers of tumor metabolism. It is necessary to stress that the discriminant power of each metabolite is linked to its relative difference in concentration between each groups (tumor versus healthy tissue) for each clinical variable considered (i.e., histological classification). Therefore, a metabolite present in subtle concentration in one group but completely absent in the second group, will play a major role in discriminant analysis. Conversely, a metabolite detected in high concentration in both groups will be less significant in the clustering process (inconsistent relative difference).

Since the initial PLS-DA model using the entire set of biopsies gave extremely promising results, a second PLS-DA model was built using only a subset of the available biopsies. More precisely, only the first 50 biopsies of our sample set, corresponding to 27 tumoral cases and 23 controls, were used to build a new model. Moreover, since the most discriminant metabolites appeared consistently to be, in a decreasing order of importance, taurine, myo-inositol, glutamate, and aspartate only these metabolites were used for model development. This choice resulted in an X matrix containing only 30 columns, thereby decreasing considerably the risk of overfitting the data. The resulting PLS-DA model was characterized by a R^2Y of 0.83 and a Q^2 of 0.80. The score and the loading plot of the new PLS-DA model are shown in Fig. 3a, b respectively. The detailed analysis of these plots shows that the average position of the different patients is similar to the model using the whole set of patients (Fig. 2a). The PLS-DA model was validated by circular validation (Fig. 3c). The remaining 22 colon biopsies (14 tumoral cases and 8 controls) were then used as a test set and subjected to a blind classification process using the previously determined statistical model. The outcome of the PLS-DA classification process is presented in Fig. 4. A visual inspection of the score plot shows that the 14 cancerous patients and the 8 healthy patients are classified in the correct region. This result is very promising and shows that the HRMAS analysis has very good sensitivity and specificity. Following this classification process, a new PCA analysis was performed on the whole biopsy cohort ($n = 72$) using only data points corresponding to taurine, myo-inositol, glutamate, and aspartate. The PCA results showed this time a clear differentiation between the two groups of biopsies (data not shown). This result is extremely important since PCA alone is rarely capable of separating biological samples in an accurate way. This feature proves the high discriminating power of taurine, myo-inositol, glutamate, and aspartate in healthy and tumoral colon tissues. Since the weight of biopsy present in each sample is accurately known, the Eretic peak can be used to quantify the resolved taurine peak at 3.42 ppm by simple integration. The average taurine concentration of cancerous and healthy biopsy samples was found to be 4.5 ± 1.7 mmol/kg and 2.0 ± 0.6 mmol/kg, respectively. This result is clearly in agreement with the conclusions of the statistical analysis.

The results presented in this manuscript were finally compared to related studies published in the literature. In 1996, Moreno and Arús (Moreno and Arús (1996)) compared the metabolic content of tumoral colon biopsies ($n = 16$) with normal mucosa biopsies ($n = 10$) using extraction techniques and high resolution liquid state ^{1}H NMR. They found that colon tumors were characterized by a higher level of taurine, glutamate, aspartate, lactate, spermine, putrescine, glutathione, and glycerophosphoethanolamine. Colon tumors were also associated with a significant decrease in myo- and scyllo-inositol. The results presented in our manuscript agree to a large extend with the results obtained by Moreno and Arús. In our spectra of intact biopsies, taurine, glutamate, aspartate, and lactate are clearly present in higher concentrations in adenocarcinomas while myo-inositol and β-glucose are detected in higher concentrations in healthy tissues. Spermine and putrescine are however not detected in our HSQC spectra. Like Moreno and Arús, we find that, among these metabolites, taurine and myo-inositol are the most promising biomarkers. More recently, Seierstad (Seierstad et al. (2008)) reported a HRMAS study of human

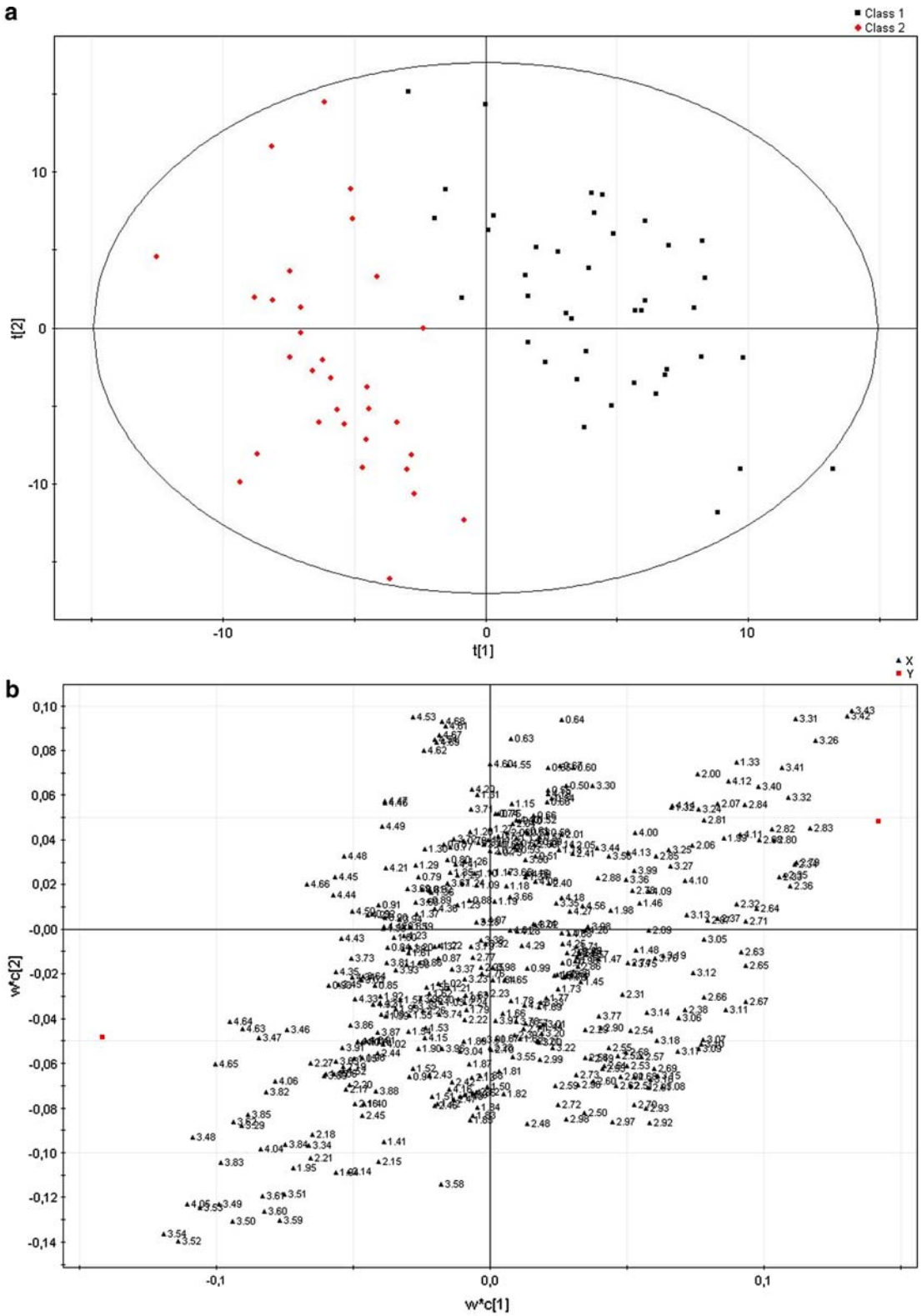

Fig. 2 Two component PLS-DA model obtained when classifying the entire cohort of colorectal biopsies ($n = 72$) according to histopathological criteria and using the 4.7–0.5 ppm spectral region. The PLS-DA model is characterized by the following parameters: $R^2Y = 0.80$ and $Q^2 = 0.68$. The high Q^2 value indicates good predictive capabilities for the model. **a** PLS-DA score plot. Cancerous and healthy colons are represented by black squares and red rhombs, respectively. **b** PLS-DA loading plot. Each point (chemical shift in ppm) represents the center of a single bucket corresponding to an interval of 0.01 ppm in the HRMAS spectrum. These buckets are associated with one or more metabolite resonances

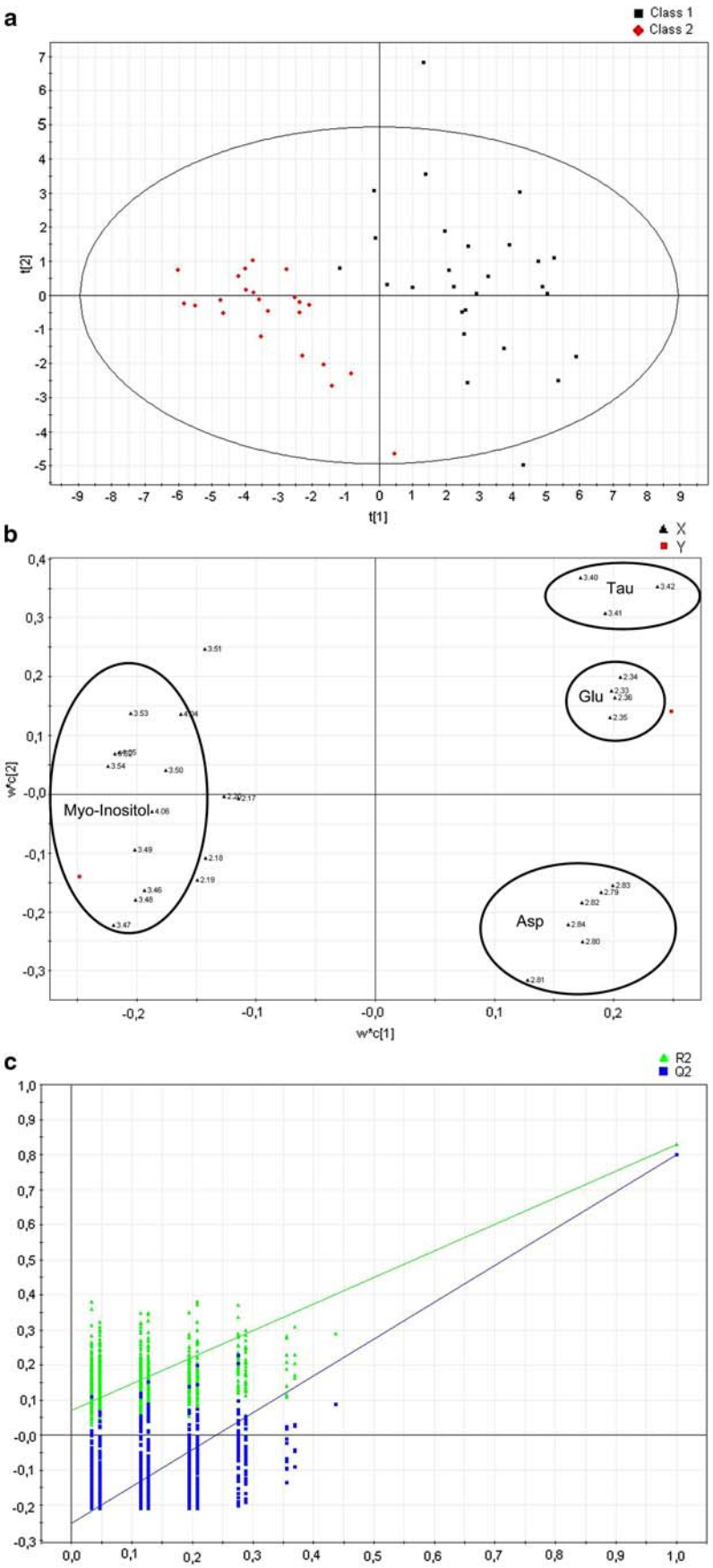

Fig. 3 Two component PLS-DA model obtained when classifying a subset of the cohort of colorectal biopsies (27 tumoral cases and 23 controls) according to histopathological criteria and using only the most important discriminant variables determined in the previous model (taurine, glutamate, aspartate, and myo-inositol regions). The PLS-DA model is characterized by the following parameters: $R^2Y = 0.83$ and $Q^2 = 0.80$. **a** PLS-DA score plot. Cancerous and healthy colons are represented by black squares and red rhombs, respectively. **b** PLS-DA loading plot. Each point (chemical shift in ppm) represents the center of a single bucket corresponding to an interval of 0.01 ppm in the HRMAS spectrum. **c** Model validation using 999 permutations. The regression line represents the correlation coefficient between the original Y (histopathological classification) and the permuted Y versus the cumulative Q^2. A negative slope of regression line suggests that there is no model overfit. Q^2 and R^2Y are represented by blue squares and green squares, respectively

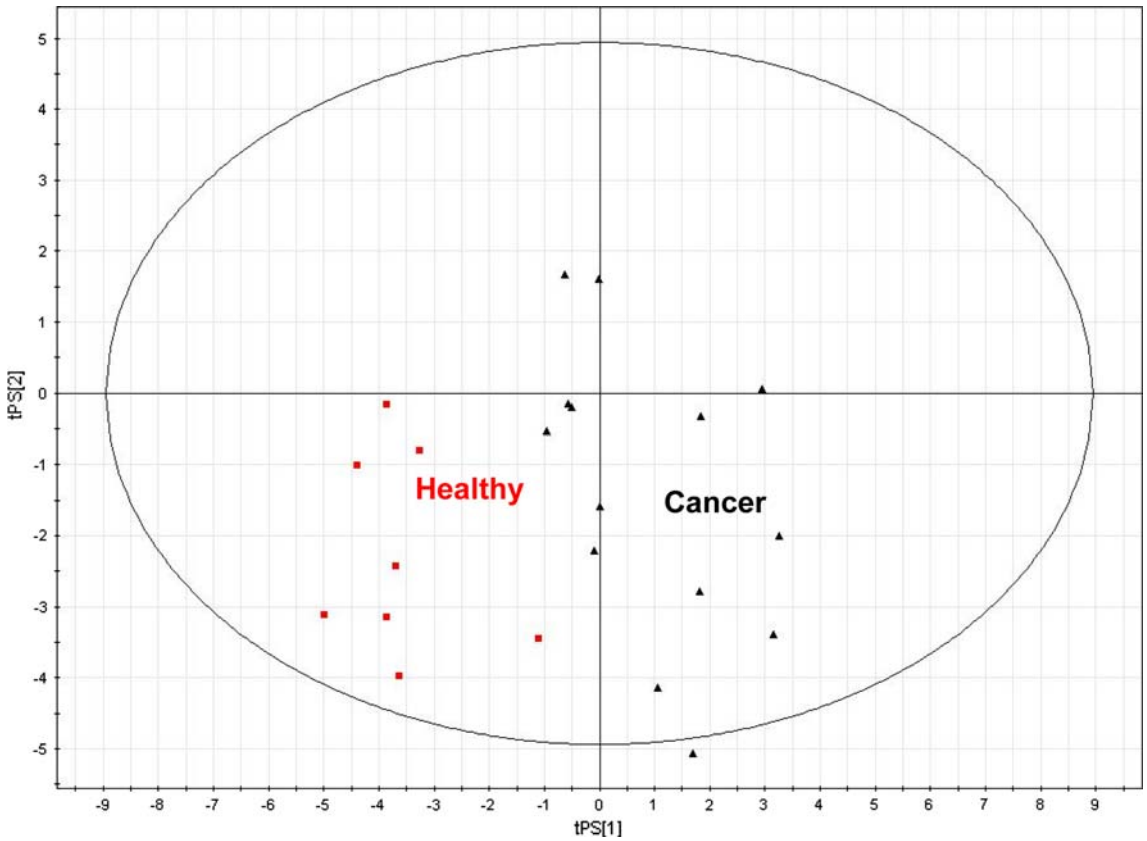

Fig. 4 Blind test classification of colon biopsies (14 tumoral cases and 8 controls) using the metabolic model presented in Fig. 3. Cancerous and healthy colons are represented by black triangles and red squares, respectively

rectal cancer biopsies and colorectal xenografts. The aim of their study was to identify the most appropriate human rectal xenograft model for preclinical magnetic resonance spectroscopy studies. They made no attempts at differentiating tumor from healthy human colon samples. However, the results they published are clearly in agreement with those presented here. The metabolites they identified are similar to the ones reported in Table 1 and colon tumors are characterized by a higher level of taurine.

Our results are therefore in agreement with previous studies already published in the literature. The strength of our approach is clearly to work on intact colon biopsies with an analytical technique (HRMAS) capable of delivering some information as to the nature of the tissue in a reasonable amount of time (ca. 20 min).

4 Concluding remarks

In conclusion, our results show that adenocarcinomas can be separated with a good specificity and sensibility from healthy colorectal biopsies solely on the basis of their metabolic content. An elevated amount of taurine, glutamate aspartate, and lactate in a colorectal biopsy is highly symptomatic of an adenocarcinoma. Conversely, a high amount of myo-inositol and β-glucose is a signature of a healthy tissue. Our HRMAS study is one of the first ones to actually use a PLS-DA metabolic model to classify tumoral and healthy colon biopsies. The next crucial step will be to investigate if a metabolic analysis can discriminate between different patient populations within the cancer group—the purpose being the identification of possible markers of unfavorable patient evolution (i.e., appearance of precocious metastasis or local recurrence). Such monitoring requires a careful longitudinal follow-up of patients and this will be the object of our next study.

Acknowledgments This work is part of the CARMeN project and was supported by grants from Région Alsace, Oséo, Communauté Urbaine de Strasbourg, Conseil Départemental du Bas-Rhin, Bruker BioSpin, University Louis Pasteur and University Hospitals of Strasbourg. The technical assistance of Dr. J. Raya, Mrs. H. Kada, and Mrs. T. Tong is gratefully acknowledged.

References

Akoka, S., Barantin, L., & Trierweiler, M. (1999). Concentration measurement by proton NMR using the ERETIC method. *Analytical Chemistry, 71*(13), 2554–2557. doi:10.1021/ac981422i.

Andrew, E. R., Bradbury, A., & Eades, R. G. (1958). Removal of dipolar broadening of nuclear magnetic resonance spectra of solids by specimen rotation. *Nature, 183*, 1802–1803. doi:10.1038/1831802a0.

Cheng, L. L., Anthony, D. C., Comite, A. R., et al. (2000). Quantification of microheterogeneity in glioblastoma multiforme with ex vivo high-resolution magic-angle spinning (HRMAS) proton magnetic resonance spectroscopy. *Neuro-Oncology, 2*(2), 87–95. doi:10.1215/15228517-2-2-87.

Cheng, L. L., Chang, I. W., Louis, D. N., & Gonzalez, R. G. (1998). Correlation of high-resolution magic angle spinning proton magnetic resonance spectroscopy with histopathology of intact human brain tumor specimens. *Cancer Research, 58*(9), 1825–1832.

Cheng, L. L., Ma, M. J., Becerra, L., et al. (1997). Quantitative neuropathology by high resolution magic angle spinning proton magnetic resonance spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America, 94*(12), 6408–6413. doi:10.1073/pnas.94.12.6408.

Cusack, J. R., Giacco, G. G., Cleary, K., et al. (1996). Survival factors in 186 patients younger than 40 years old with coloractal adenocarcinoma. *Journal of the American College of Surgeons, 183*, 105–112.

Davis, A. L., Keeler, J., Laue, E. D., & Moskau, D. (1992). Experiments for recording pure-absorption heteronuclear correlation spectra using pulsed field gradients. *Journal of Magnetic Resonance (San Diego Calif), 98*(1), 207–216.

Elbayed, K., Dillmann, B., Raya, J., Piotto, M., & Engelke, F. (2005). Field modulation effects induced by sample spinning: application to high-resolution magic angle spinning NMR. *Advances in Magnetic Resonance, 174*, 2–26. doi:10.1016/j.jmr.2004.11.017.

Fearon, E., Cho, K. R., Nigro, J., et al. (1990). Identification of a chromosome 18q gene that is altered in colorectal cancers. *Science, 247*, 49–56. doi:10.1126/science.2294591.

Jemal, A., Siegel, R., Ward, E., et al. (2007). Cancer Statistics. *CA: A Cancer Journal for Clinicians, 57*, 43–66.

Leichman, C. G., Lentz, H. J., Danenberg, K., et al. (1997). Quantitation of intratumoral thymidylate synthase expression predicted for disseminated colorectal cancer response and resistance to protracted infusion of 5-fluorouracil and weekly leucovorin. *Journal of Clinical Oncology, 15*, 3223–3229.

Libutti, S. K., Saltz, L. B., & Tepper, J. E. (2008). Principles and practice in oncology. In V. T. De Vita, T. S. Lawrence, & S. A. Rosemberg (Eds.), *Colon cancer in cancer* (pp. 1232–1285). Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins.

Lippens, G., Bourdonneau, M., Dhalluin, C., et al. (1999). Study of compounds attached to solid supports using high resolution magic angle spinning NMR. *Current Organic Chemistry, 3*, 147–169.

Lowe, I. J. (1959). Free induction decays of rotating solids. *Physical Review Letters, 2*, 285–287. doi:10.1103/PhysRevLett.2.285.

Martínez-Bisbal, M. C., Martí-Bonmatí, L., Piquer, J., et al. (2004). 1H and 13C HR-MAS spectroscopy of intact biopsy samples ex vivo and in vivo 1H MRS study of human high grade gliomas. *NMR in Biomedicine, 17*(4), 191–205. doi:10.1002/nbm.888.

Martínez-Granados, B., Monleon, D., Martínez-Bisbal, M. C., et al. (2006). Metabolite identification in human liver needle biopsies by high-resolution magic angle spinning 1H NMR spectroscopy. *NMR in Biomedicine, 19*(1), 90–100. doi:10.1002/nbm.1005.

Moreno, A., & Arús, C. (1996). Quantitative and qualitative characterization of 1H NMR spectra of colon tumors, normal mucosa and their perchloric acid extracts: Decreased levels of myo-inositol in tumours can be detected in intact biopsies. *NMR in Biomedicine, 9*(1), 33–45. doi:10.1002/(SICI)1099-1492(199602)9:1<33::AID-NBM391>3.0.CO;2-G.

Piotto, M., Bourdonneau, M., Furrer, J., et al. (2001). Destruction of magnetization during TOCSY experiments performed under magic angle spinning: Effects of Radial B1 inhomogeneities. *Journal of Magnetic Resonance (San Diego, Calif.), 149*, 114–118. doi:10.1006/jmre.2001.2287.

Seierstad, T., Røe, K., Sitter, B., et al. (2008). Principal component analysis for the comparison of metabolic profiles from human rectal cancer biopsies and colorectal xenografts using high-resolution magic angle spinning 1H magnetic resonance spectroscopy. *Molecular Cancer, 7*, 1–13. doi:10.1186/1476-4598-7-33.

Shaka, A. J., Lee, C. J., & Pines, A. (1988). Iterative schemes for bilinear operators: Application to spin decoupling. *Journal of Magnetic Resonance (San Diego, Calif.), 77*, 274–293.

Sitter, B., Sonnewald, U., Spraul, M., Fjosne, H. E., & Gribbestad, I. S. (2002). High-resolution magic angle spinning MRS of breast cancer tissue. *NMR in Biomedicine, 15*(5), 327–337. doi:10.1002/nbm.775.

Sjøbakk, T. E., Johansen, R., Bathen, T. F., et al. (2008). Characterization of brain metastases using high-resolution magic angle spinning MRS. *NMR in Biomedicine, 21*(2), 175–185. doi:10.1002/nbm.1180.

Tenenhaus, M. (1998). La régression PLS: Théorie et Pratique. Paris.

Tzika, A. A., Cheng, L. L., Goumnerova, L., et al. (2002). Biochemical characterization of pediatric brain tumors by using in vivo and ex vivo magnetic resonance spectroscopy. *Journal of Neurosurgery, 96*(6), 1023–1031.

Westerhuis, J. A., Hoefsloot, H. C. J., Smit, S., et al. (2008). Assessment of PLSDA cross validation. *Metabolomics, 4*(1), 81–89. doi:10.1007/s11306-007-0099-6.

Wider, G., & Dreier, L. (2006). Measuring protein concentrations by NMR spectroscopy. *Journal of the American Chemical Society, 128*(8), 2571–2576. doi:10.1021/ja055336t.