Bull. Korean Chem. Soc. 2010, Vol. 31, No. 2 379
DOI 10.5012/bkcs.2010.31.02.379

Combining Information of Common Metabolites Reveals Global Differences between Colorectal Cancerous and Normal Tissues

Young Kee Chae,[*] Woo-Young Kang, Seong Hwan Kim,[†] Jong Eun Joo,[‡] Joon Kil Han,[§] and Boo Whan Hong[#]

Department of Chemistry and Institute for Chemical Biology, Sejong University, Seoul 143-747, Korea
[*]E-mail: ykchae@sejong.ac.kr
[†]Department of Internal Medicine, [‡]Department of Pathology, and [§]Department of Surgery,
Eulji University College of Medicine, Seoul 139-872, Korea
[#]Department of Surgery, Gwangju Samsung Hospital, Kyunggi-Do 464-801, Korea
Received December 2, 2009, Accepted December 24, 2009

Metabolites of colorectal cancer tissues from 12 patients were analyzed and compared with those of the normal tissues by two-dimensional NMR spectroscopy. NMR data were analyzed with the help of the metabolome database and the statistics software. Cancerous tissues showed significantly altered metabolic profiles as compared to the normal tissues. Among such metabolites, the concentrations of taurine, glutamate, choline were notably increased in the cancerous tissues of most patients, and those of glucose, malate, and glycerol were decreased. Changes in individual metabolites varied significantly from patient to patient, but the combination of such changes could be used to distinguish cancerous tissues from normal ones, which could be done by PCA analysis. The traditional chemometric analysis was also performed using AMIX software. By comparing those two results, the analysis *via* ^{1}H-^{13}C HSQC spectra proved to be more robust and effective in assessing and classifying global metabolic profiles of the colorectal tissues.

Key Words: NMR, Metabolomics, Profiling, Colorectal cancer, PCA

Introduction

Colorectal cancer refers to the growth of cancer in the colon, rectum and appendix. 639,000 people die of this kind of cancer worldwide each year, making it the third leading cause of cancer-related deaths (World Health Organization, http://www.who.int/mediacentre/factsheets/fs297/en/). Colorectal cancers are thought to arise from adenomatous polyps in the colon. These mushroom-shaped benign tumors can make transition into cancer over time. The diagnosis of localized colon cancer is mostly through colonoscopy, and the stage of the cancer determines the treatment. When colorectal cancer is found at early stages, it can be cured. However, if detected when distant metastases are present, it becomes less likely to be curable. Therapy is usually through surgery followed by chemotherapy.

The TNM or Duke system is currently used for staging and prognostication of colorectal cancer. This method is mainly based on the histological assessment of tumor invasion and lymph nodal spread.[1] For further understanding at the molecular level, genomics, proteomics, or metabolomics approach has been used recently. Among these, metabolomics can profile metabolic changes that occur in living systems in response to various factors.[2] Metabolomics has already proven its potential in identifying metabolite-based biomarkers in ovarian, brain and liver cancers.[3-5] Significant increases in taurine, choline-containing compounds and lipid resonances were observed in malignant colon mucosa by one-dimensional ^{1}H NMR spectral data.[6] Altered metabolic profiles may provide potential biomarkers for detection, staging, prognostication, and treatment of colorectal cancer.[7]

In this present study, we hypothesized that the metabolic profile in colorectal tissue would provide fingerprints of cancerous tissues that clearly differ from normal tissues. To test the hypothesis, we used two-dimensional ^{1}H-^{13}C HSQC technique,[8,14] and profiled normal and cancerous tissue samples from 12 patients. We also compared our result with the traditional chemometric method based on one-dimensional ^{1}H spectra. This report would demonstrate the robustness of the two-dimensional NMR technique, making it the method of choice for metabolic profiling, and this analysis would be extended to the diagnosis of colorectal cancer based on its global profile of "common" metabolites without the need for a specific biomarker.

Materials and Methods

Patient Population. This study involves the use of human colon samples obtained from 12 patients (4 men, 8 women, mean age: 67.75 years, age range: 41 - 85 years) with histologically proven colorectal cancers. Detailed clinical analysis will be published elsewhere (S.H. Kim, unpublished data).

Metabolite Extraction. Metabolites from the tissues were extracted by a modified version of the hot water extraction method.[9] The tissue sample was put into a 50 mL conical tube, and 16 mL of boiling water was poured in. The resulting mixture was incubated at 121 °C for 15 min, and insoluble remnants of tissues were removed by centrifugation at 4000 g for 15 min. The supernatant was incubated at 4 °C overnight. A 0.45 μm syringe filter was used to remove fine debris from the supernatant. The resulting clear solution was further filtered through a membrane of molecular weight cutoff of 5000 Da (Vivaspin 20, Sartorius Stedim Biotech, Bohemian, NY, USA). The filtered solution was freeze-dried and the mass of the dried extract was measured and dissolved in 5 mM HEPES solution in D_2O with 0.2 mM DSS, 0.5 mM NaN_3 to the final volume-mass ratio

of 17.5 μL per mg of dried extract. However, for some samples as will be mentioned in the result section, the volume-mass ratio was 35 μL or higher depending on the availability of dried extracts. The pH was adjusted to 7.4 with NaOD or DCl (Cambridge Isotope Laboratories, Andover, MA, USA).

NMR Experiments and Data Processing. NMR experiments were performed on Bruker Avance II 500 MHz (Bruker, Germany). The spectrometer was equipped with a triple-resonance (^{1}H, ^{13}C, ^{15}N, ^{2}H lock) probe. Sensitivity enhanced ^{1}H-^{13}C HSQC spectra were collected with 112 scans, 256 increments (TPPI), and 3sine decoupling. The spectral widths were 20 ppm for ^{1}H and 100 ppm for ^{13}C. The carbon carrier frequency was set at 55 ppm.

All spectra were processed and visualized using TopSpin 2.1 (Bruker, Germany) and Sparky software,[10] respectively. Picked peaks were converted to a proper format for MMCD (http://mmcd.nmrfam.wisc.edu)[11] to identify the metabolites using FMQ module (Ian Lewis, personal communication) written in R, a free statistics software package (http://www.r-project.org). FMQ provided a bridge between MMCD and Sparky. FMQ also generated a tailored project file for Sparky, and spectral comparison was facilitated by overlap function. The intensities of resonances of metabolites were measured inside Sparky. The intensity data were standardized using Microsoft Excel, and the resulting data was analyzed by PCA in R. For comparison, AMIX software (Bruker, Germany) was used to analyze one-dimensional ^{1}H spectra. One-dimensional ^{1}H spectrum of each sample was analyzed after simple bucketing.

Results and Discussion

Sample Preparation. The tissue samples were boiled, and only the soluble metabolites coming out of the tissue were used for NMR analysis. Unlike the plant samples which are readily freeze-dried and powderized, the intestinal tissue samples were much greasier, and we decided to boil the whole excised tissues. The initial addition of 16 mL of boiling water was employed in an attempt to denature endogenous enzymes which might degrade or synthesize metabolites during the extraction procedure. When we tried to remove smaller debris from the soluble portion after centrifugation, the syringe filter was clogged completely due to the excess fat suspended inside. After the incubation at 4 °C, fat was solidified at the top of the solution, and we could easily collect the aqueous portion for syringe-filtration. We tried to bypass the ultrafiltration step because of its time-consuming nature, but this step was absolutely necessary on account of the existence of the larger molecules hindered observation of NMR resonances by raising the noise level.

The dried extracts of some tissue samples were so little that we had to prepare more diluted NMR samples. As shown in Table 1, most tissue samples yielded more than 10 mg, but some samples produced as few as 6 mg. We speculate that this is partly due to the experimental variation and also due to the fat content of the excised specimen: more fat, less extract. Even with the Shigemi NMR tubes, the volume should be at least 260 μL, which meant that we had to have at least 15 mg of dried extracts to reach the desired concentration. In case of smaller amount of extracts, we diluted doubly or triply to meet the volume requirement for NMR samples. The intensities of resonances of such samples were multiplied as much as the dilution factors when we analyzed the data. The average mass of the dried extracts from 1 g of wet tissue sample was 14.9 ± 4 mg. This value is only the half of the mouse liver case where about 30 mg of extracts were prepared per g of wet liver (Y. K. Chae, unpublished data). The yield is only around 1.5%, which may reflect the nature of the greasy sample.

Table 1. Extraction yield from wet tissues.

Patient #	normal tissue		cancerous tissue	
	wet weight (g)	dried extract (mg)	wet weight (g)	dried extract (mg)
3	1.049	17	1.306	23
6	1.851	25	4.601	47
7	2.396	28	5.027	40
9	1.204	21	1.583	25
10	0.418	8	0.558	11
11	1.405	18	2.504	36
12	0.885	14	2.175	25
13	1.394	20	6.357	49
14	0.571	11	0.889	16
15	0.497	8	1.54	36
16	0.844	13	3.258	60
17	0.714	6	2.77	35

NMR Experiments. Each 2D HSQC experiment took about 12 hours, which was just due to the limited use (only at night) of the NMR spectrometer. We believe a 4 hour experiment would have produced a spectrum with enough sensitivity and resolution. If the spectrometer had been equipped with an autosampler feature, the data collection time could have been reduced to one third. The automatic tuning and matching module would play a great supporting role to the autosampler. The cryoprobe would be another critical feature to reduce the data collection time since it is known to produce at least 10 times as large a signal-to-noise ratio as the room-temperature probe. With these two equipments, the data collection time could drop to less than an hour. Compared to the traditional profiling method based on one-dimensional NMR data, these two-dimensional NMR experiments take at least 10 times as much time, which is why the former is still the method of choice when there are a large number of samples to be analyzed. Nonetheless, as will be mentioned in the next section, if the accuracy and robustness are concerned, the latter can be considered as a potent alternative. In fact, a recent report showed the quantitative NMR data could be extracted from the 8 minute ^{1}H-^{13}C HSQC experiment.[9] Fig. 1 shows one of the spectra collected in this report along with the names of the identified metabolites.

NMR Data Processing and Analysis. NMR data were processed using TopSpin 2.0. The processed data could directly be read into Sparky where the peak list was generated. The peak list was properly formatted and sent to MMCD to identify the metabolites in the samples. The identities of metabolites were confirmed in Sparky using the overlay feature. In practice, the most important step was referencing the spectra since the slight

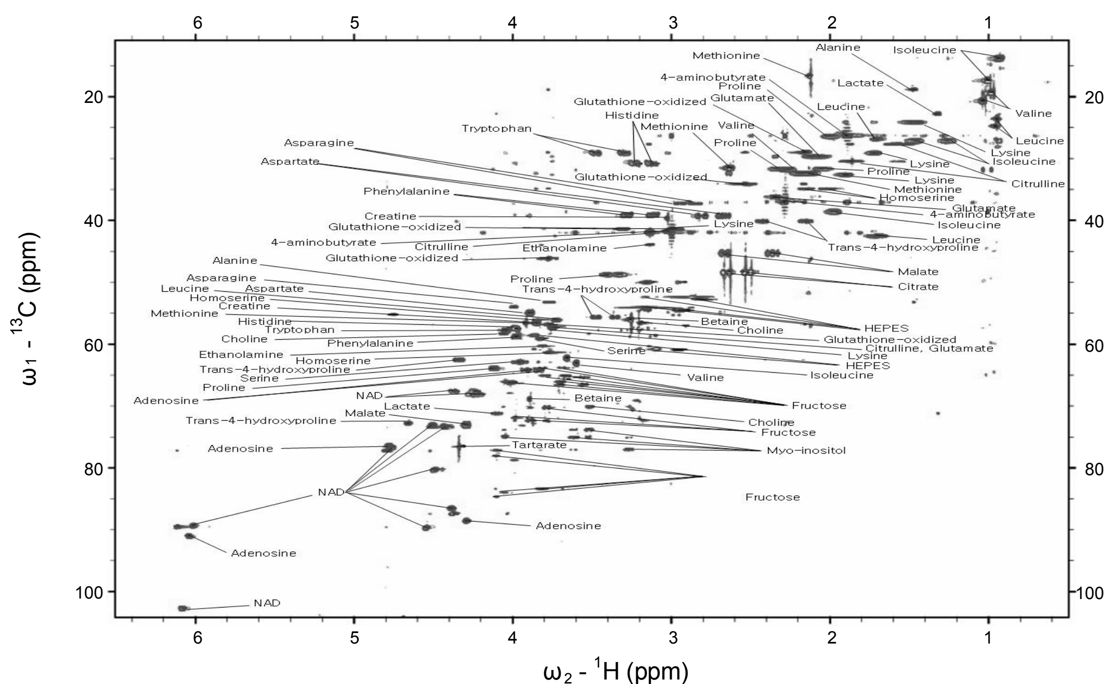

Figure 1. Two-dimensional ^{1}H-^{13}C HSQC spectrum of metabolites from colorectal cancer tissue. The assigned resonances are labeled with the names of the metabolites.

variation in chemical shift could generate a list of improper metabolites from MMCD. Fortunately, DSS showed up in the spectrum, though with a low intensity providing the reference point. Since we constructed a table of representative resonances of major metabolites,[12] the intensities of those peaks were quickly measured and used for semi-quantitative analysis.

On account of each sample was unique (one cancer and one normal sample from one patient, no replicates) and experimental variation existed as always, the intensities of resonances had to be normalized carefully for proper semi-quantitative comparison. The internal standard, HEPES, played as an excellent reference to normalize intensity data since all NMR samples contained the same concentration of HEPES, 5 mM. After normalization to HEPES, we multiplied the dilution factor of the sample to the measured intensities since some samples were less concentrated due to the smaller amounts of dried extracts. This is

Table 2. Comparison of metabolite levels in normal and cancerous tissues.

metabolite	average increase (fold)	standard deviation	relative error (%)
Alanine	1.229137	0.670701	54.56678
Asparagine	1.478282	0.823448	55.70303
Aspartate	1.824012	1.245487	68.2828
Choline	1.369423	1.005727	73.44167
Creatine	1.599603	3.004061	187.8005
Citrulline	1.298738	0.884338	68.09212
Ethanolamine	0.995173	0.521629	52.41591
Glutamate	2.012502	1.373001	68.22355
Histidine	1.289351	0.75002	58.17032
Isoleucine	1.463596	0.776274	53.03879
Lactate	2.095451	1.612676	76.96078
Leucine	1.282245	0.627715	48.9544
Lyssine	1.20491	0.765548	63.53572
Malate	0.832433	0.312644	37.55791
Methionine	1.429754	0.747843	52.30574
Myoinositol	1.257173	1.87351	149.0256
Phenylalanine	1.257093	0.675449	53.73104
Proline	1.75535	0.977948	55.71241
Serine	1.300596	0.691601	53.17569
Valine	1.360372	0.794825	58.42703
Arginine	1.207031	0.998068	82.6878
Carnitine	1.632586	1.280369	78.42582
Glutathione (oxidized)	1.721084	1.001966	58.21713
Glycine	1.666617	1.170533	70.23407
Succinate	1.556828	1.023015	65.71149
Taurine	4.75352	4.420021	92.98416
Threonine	1.461987	0.874539	59.81853
Acetate	1.214736	0.589445	48.52457
Betaine	1.119952	1.173058	104.7418
Glycerol	0.983628	0.966382	98.24672
Glutamine	1.491061	1.179091	79.07729
Glucose	0.621077	0.676671	108.9513
Putricine	1.449777	0.793168	54.70966
Uridine	1.03167	0.550931	53.40184
Maltose	1.085144	1.39541	128.5921
Ornithine	1.405354	0.708642	50.42444

Figure 2. (a) Increases in metabolite concentrations upon transition to cancer. Note that the y-axis is in logarithmic scale. (b) Average relative changes in metabolite concentrations upon transition to cancer. Standard deviations were denoted with error bars.

Figure 3. (a) PCA analysis of metabolite concentrations extracted 2D HSQC data. (b) PCA analysis of 1D ^{1}H NMR spectra of normal tissues (open circles) and cancerous tissues (filled circles).

based on the assumption that there is a linear relationship between concentration and resonance intensity. According to the previous study,[12] this assumption could hold if the concentration was lower than 10 mM which served as an upper bound of the concentration of most metabolites. From the rough comparison between the resonance intensities of HEPES and other metabolties, we believe that the linearity between concentration and peak intensity should hold. Furthermore, we speculate that this relationship holds at least up to 50 mM, but we need more data to confirm it.

We compared the normal and cancerous tissues of each patient. As shown in Fig. 2a, we can see the "trends" of the changes of metabolites upon the transition of normal tissue to cancer. For example, taurine, glutamate, and lactate showed a strong increase in such a transition while creatine, malate, and glucose showed a sharp decrease. Similar results have recently been published.[1,3,13] Their findings also indicated that levels of taurine, glutamate, aspartate, and lactate were elevated while those of myoinositol and glucose were lowered in the cancerous tissues. We provide in this report more metabolites that showed noticeable changes, which can constitute the basis of diagnosis or stage identification of the cancerous tissues.

As can be seen in Table 2 and Fig. 2b, the amount of changes varied from patient to patient, leading to a large relative variation

among patients. In case of myoinositol, the relative variation was almost 150%. If we should use the individual values of concentration changes of metabolites to characterize cancerous tissues, we would have to deal with large error bars and a small confidence window. In such a case, it would be very difficult to say whether the tissue is normal or cancerous. To better analyze and diagnose the cancer tissues, we needed a better way to draw any meaningful pattern from these widely spread data. Principle component analysis was chosen as a fit to deal with such data.

The intensity data were read into the R software package and the PCA analysis was applied. The script was kindly written and provided by Ian Lewis (Univ. of Wisconsin-Madison, USA). As shown in Fig. 3a, normal and cancerous tissues can be clearly divided into two regions by the diagonal line. Furthermore, it is evident which metabolite contributes more to such a separation. As expected from Fig. 2a, increase of taurine, lactate, and glutamate, and decrease of malate, creatine, and glucose contributed for the separation, and those metabolites were considered the indicators for the transition to cancerous tissues. Fig. 3b shows the widely used figure of PCA analysis based on the one-dimensional spectrum. Here, the separation of normal and cancerous tissues is not so apparent. They could be roughly grouped, but the separation is not enough for diagnosis. This could have been improved by using a very carefully designed bucketing method; however it would also mean that the data could be manipulated until the desired separation was achieved. The PCA result in Fig. 3a is clearly out of this controversy because there is no such bucketing or binning to improve the separation.

Conclusion

We have shown that two dimensional HSQC spectra produced a more robust and reliable result than the method based on one-dimensional spectra. The PCA result from HSQC data showed clearly separated groups corresponding to normal and cancerous tissues. This kind of diagnosis does not depend on finding a unique biomarker, instead uses "common" metabolites and a combination of their concentrations as a whole. We prospect that this method can lead to the diagnosis and classification of not only the colorectal cancers but also other cancer types.

References

1. Chan, E. C.; Koh, P. K.; Mal, M.; Cheah, P. Y.; Eu, K. W.; Backshall, A.; Cavill, R.; Nicholson, J. K.; Keun, H. C. *J. Proteome. Res.* **2009**, *8*, 352-361.
2. Nicholson, J. K.; Lindon, J. C.; Holmes, E. *Xenobiotica.* **1999**, *29*, 1181-1189.
3. Denkert, C.; Budczies, J.; Kind, T.; Weichert, W.; Tablack, P.; Sehouli, J.; Niesporek, S.; Konsgen, D.; Dietel, M.; Fiehn, O. *Cancer Res.* **2006**, *66*, 10795-10804.
4. Petrik, V.; Loosemore, A.; Howe, F. A.; Bell, B. A.; Papadopoulos, M. C. *British Journal of Neurosurgery* **2006**, *20*, 275-280.
5. Yang, J.; Xu, G.; Zheng, Y.; Kong, H.; Pang, T.; Lv, S.; Yang, Q. *Journal of Chromatography B* **2004**, *813*, 59-65.
6. Moreno, A.; Rey, M.; Montane, J. M.; Alonso, J.; Arus, C. *NMR Biomed.* **1993**, *6*, 111-118.
7. Longley, D. B.; Allen, W. L.; Johnston, P. G. *Biochim. Biophys. Acta* **2006**, *1766*, 184-196.
8. Fan, T. W. M. *Progress in Nuclear Magnetic Resonance Spectroscopy* **1996**, *28*, 161-219.
9. Lewis, I. A.; Schommer, S. C.; Hodis, B.; Robb, K. A.; Tonelli, M.; Westler, W. M.; Sussman, M. R.; Markley, J. L. *Analytical Chemistry* **2007**, *79*, 9385-9390.
10. Goddard, T. D.; Kneller, D. G., 3 ed.; University of California, San Francisco.
11. Cui, Q.; Lewis, I. A.; Hegeman, A. D.; Anderson, M. E.; Li, J.; Schulte, C. F.; Westler, W. M.; Eghbalnia, H. R.; Sussman, M. R.; Markley, J. L. *Nat. Biotechnol.* **2008**, *26*, 162-164.
12. Kang, W. Y.; Chae, Y. K. *J. Kor. Magn. Reson.* **2009**, *13*, 7-14.
13. Piotto, M.; Moussallieh, F. M.; Dillmann, B.; Imperiale, A.; Neuville, A.; Brigand, C.; Bellocq, J. P.; Elbayed, K.; Namer, I. J. *Metabolomics* **2009**, *5*, 292-301.
14. Ha, J. H; Won, E. Y.; Yoon, H. S.; Kang, S.; Bae, K. H.; Park, S. G.; Park, B. C.; Choi, B. S.; Lee, J. H.; Chi, S. W. *Bull. Kor. Chem. Soc.* **2009**, *30*(7), 1655-1657.