Theor Chem Acc (2010) 125:329–334
DOI 10.1007/s00214-009-0698-6

REGULAR ARTICLE

Raman spectroscopy study of breast disease

Marcelo Moreno · Leandro Raniero · Emília Ângelo Loschiavo Arisawa · Ana Maria do Espírito Santo · Edson Aparecido Pereira dos Santos · Renata Andrade Bitar · Airton Abrahão Martin

Received: 30 August 2009 / Accepted: 26 October 2009 / Published online: 26 November 2009

Abstract The aim of this study was to evaluate the vibrational modes of malignant and benign breast tissues with the following diagnosis: fibroadenoma, invasive ductal carcinoma, ductal carcinoma in situ, and fibrocystic condition. Quadratic discriminate analysis, a multivariate statistical method of analysis, showed 98.5% separation between normal and altered tissue. Significant changes were observed at the lower Raman shift for altered tissue. For a better understanding of the spectral differences, a biochemical interpretation was also performed in terms of the reduction and oxidation processes in the cell environment which could be associated with an inflammatory reaction.

Keywords Raman spectroscopy · Breast cancer · Principal component analysis · Biochemical analysis · Linear discriminant analysis · Quadratic correlation analysis

Dedicated to Professor Sandor Suhai on the occasion of his 65th birthday and published as part of the Suhai Festschrift Issue.

M. Moreno · L. Raniero · E. Â. L. Arisawa · A. M. do Espírito Santo · E. A. P. dos Santos · R. A. Bitar · A. A. Martin (✉)
Laboratory of Biomedical Vibrational Spectroscopy, Institute of Research and Development, IP&D, Universidade do Vale do Paraíba, UniVap, Shishima Hifumi Avenue, 2911, São José dos Campos, São Paulo 12244-000, Brazil
e-mail: amartin@univap.br

M. Moreno
Medical School, Universidade Comunitária Regional de Chapecó, Unochapecó, Senador Attílio Fontana Avenue, 591-E, Chapecó, Santa Catarina 89809-000, Brazil
e-mail: moreno@unochapeco.edu.br

Abbreviations

BC	Breast cancer
RS	Raman spectroscopy
FBC	Fibrocystic breast conditions
FD	Fibroadenoma
DCIS	Ductal carcinoma in situ
IDC	Invasive ductal carcinoma
PCA	Principal component analysis
LDA	Linear discriminant analysis
QDA	Quadratic correlation analysis

1 Introduction

Cancer is a leading cause of death worldwide. It has already accounted for 7.9 million deaths in 2007 and, if current trends continue, the toll is projected to be 12 million by 2030 [1]. Breast cancer (BC) is one of the top ten leading causes of death by cancer. Screening programs to identify early cancer or pre-cancer conditions, such as mammography, have increased the rate for early detection, which is the most effective means of achieving better prognosis and lower death rates [1, 2].

The diagnosis of BC is confirmed through biopsy methods, such as fine needle aspiration, surgical or core incision biopsies (with or without imaging guiding), radio-guided occult lesion localization, excision biopsies with wire localization, and mammotomy [2]. All techniques are done with direct access to the suspect breast lesion and require procedures that range from local anesthesia to a hospital stay with use of general anesthesia, which may be traumatic to the patients.

Raman spectroscopy (RS) has been studied as a promising new tool for noninvasive, real-time diagnoses of

benign and malignant lesions in human breast tissue [3–14]. Many efforts have been done to classify the Raman spectra of normal and abnormal tissues, which lead to the correlation between the pathological situation of the tissue and their chemical compositions [9, 15–17]. Statistical methods such as principal component analysis (PCA) have been used to automate the classification of the spectra according to histo-pathological analysis. Indeed many articles have been published based on the classification of different types of cancer with high specificity and sensitivity. For breast malignancy, these values are around 96 and 92%, respectively [18–21]. However, more research still needs to be done on the biochemical interpretation of the spectra based on chemical changes from one individual to another.

In this article the spectral differences between normal breast tissue (NB), fibrocystic breast conditions (FBC), fibroadenoma (FD), ductal carcinoma in situ (DCIS or intraductal carcinoma) and invasive ductal carcinoma (IDC or infiltrating carcinoma) are analyzed. The changes in the spectra were explained from the biochemical point of view with some speculation on the chemical reactions occurring in the tissue.

2 Experimental details

A total of 37 mammary tissue samples were collected from 37 patients. Thirty-one had different kinds of breast disease, and six were normal tissues used as controls. The diseased tissues were gathered either through modified radical mastectomy or conservative breast surgery with the following diagnoses: FD (1 patient), IDC (22 patients), DCIS (2 patients) and FBC (6 patients). The six normal breast tissues were acquired from reductive aesthetic mastectomy. A total of 74 spectra (two to three measurements each) were collected from the samples. Informed consent was obtained from all patients. The study followed the guidelines of the institutional ethical committee (017/2000/CEP).

The samples, after the surgical procedure, were identified, snap frozen and stored in liquid nitrogen (77 K) in cryogenic vials (Nalgene®) until the FT-Raman spectra recording. For FT-Raman data collection, the samples were brought to room temperature and kept moist in 0.9% physiological solution to preserve their structural characteristics, and placed in a windowless aluminum holder for the Raman spectra collection. To avoid the border of the lesion, a center portion of about 1 mm^3 was removed from each pathological breast sample. Soon after Raman analysis, the samples were placed in a 10% formaldehyde solution for further pathological analysis.

To reduce the fluorescence effect (FE), a FT-Raman spectrometer (Bruker RFS 100/S) was used with an Nd:YAG laser at 1,064 nm as the excitation light source. The laser spot size was 200 μm in diameter and the power was kept below 110 mW to better preserve the integrity of the samples and prevent photodecomposition caused by the laser beam irradiation; the spectrometer resolution was set to 4 cm^{-1}. The spectra of pathological breast tissues were recorded with 150 scans and gathered from approximately 1 mm^3 of tissue volume probing a large number of cells ($\sim 10^6$), at least 80% of which were neoplastic cells. The polynomial baseline fitting was done by a Matlab routine with a polynomial order of 5 to remove the FE. All data were normalized by dividing the spectrum by its highest value and mean centering, before statistical analysis. The evaluation of spectra data was done by clustering analysis and a statistical multivariate test based on PCA utilizing MINITAB Release® 14.20 software. The information from this analysis was obtained through four principal components (PC1, PC2, PC3, and PC4). Three pathologists reviewed the diagnosis following criteria from the Brazilian Pathology Society.

3 Results and discussion

3.1 Statistical analysis

The principal components were calculated using a full range of the Raman spectra (IDC, FBC and NB), between 300 and 2,000 cm^{-1}, and a covariance matrix. Figure 1 displays the loading plot for the first four PCs, which correspond to the variation of PCs as a function of wavenumber.

PC1 represents 62% of the data variance. The main contribution of vibration modes and their respective assignments are shown in Table 1 [10, 15, 18, 22–29].

The second PC had 13.4% of the data variance and the major difference was related to changes in the vibrational modes of lipids, phenylalanine, collagens and amino acids (cytosine, guanine, and cysteine at lower wavenumber). PC3 with 4.4% of the data variance also showed change in vibration modes of lipids, amide III and collagens. Besides this, PC4 with 2% of the data variance showed positive and negative contribution for almost all constituents in the range, but there is a significant contribution in the range between 1,000 and 1,500 cm^{-1}. In conclusion, the major changes observed in the spectra were related to amides I, III, lipids, and the aforementioned amino acids.

To classify the PC1, PC2, PC3 and PC4 data from the IDC, FBC, and NB groups, they were analyzed by discriminant analysis, according to pathological classification. The discriminant analysis can be done by linear (LDA) or quadratic (QDA). In linear analysis, the spectra are classified into a group if the squared distance (also called the

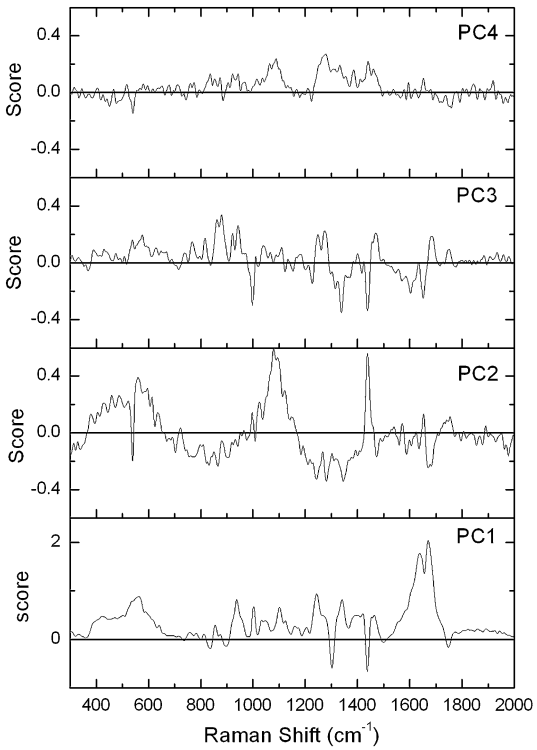

Fig. 1 Loading plot for IDC, FBC and NB group

Table 1 Peak assignments for Raman tissue spectra

Peak position (cm^{-1})	Major assignments
1,730–1,740	Collagen III
1,654–1,655	Amide I (C=O stretching mode of proteins, a-helix conformation)/C=C lipid stretch
1,450–1,500	Stretching (CH_2)—lipids, lycosaminoglycans, metalloproteinases, collagens, and residues
1,245–1,345	Amide III—collagen
1,034–1,175	Lipids and nucleic acids (cytosine, guanine, adenine)
1,083	C–N stretching mode of proteins (and lipid mode to lesser degree)
1,078	C–C or C–O stretch (lipid), C–C or PO_2 stretch (nucleic acids)
1,064	Skeletal C–C stretch lipids
1,031	C–H in-plane bending mode of phenylalanine
1,001	Symmetric ring breathing mode of phenylalanine
935	C–C stretching mode of proline and valine and protein backbone (a-helix conformation)/glycogen
717–719	C–N (membrane phospholipid head)/adenine
679	Guanine ring breathing
669	C–S stretching mode of cystine
643	C–C twisting mode of tyrosine
621	C–C twisting mode of phenylalanine
540	S–S disulfide bridges in cysteine
484–490	Glycogen

Table 2 Summary classification of QDA

Classified into	True group		
Group	IDC	FBC	NB
IDC	43	0	0
FBC	1	11	0
NB	0	1	12
Total number	43	12	12
Number correct	43	11	12
Proportion	0.99	0.92	1.00

Mahalanobis distance) of observation to the group center (mean) is the minimum, but the assumption is made that covariance matrices are equal for all groups. For quadratic analysis, there is no assumption that the groups have equal covariance matrices and an observation is also classified into the group that has the smallest squared distance. However, the squared distance does not simplify into a linear function, and for this reason the analyses in this work were done by QDA and the classification results are shown on Table 2.

The summary of the classification table shows that discriminant analysis identified 98.5% of 34 patients, correctly. This result is extremely satisfactory, since there was no false positive result. Figure 2 displays an ROC curve where the dashed line represents two indistinguishable populations (random data) and the solid line is the combination of PC1, PC2, PC3, and PC4 (calculated using the squared distance of observation to the group center given by quadratic discriminate analysis), which shows 99% sensitivity and 98% specificity. The results from discriminant analysis and the ROC curve showed excellent separation between the normal and abnormal groups with high sensitivity and specificity.

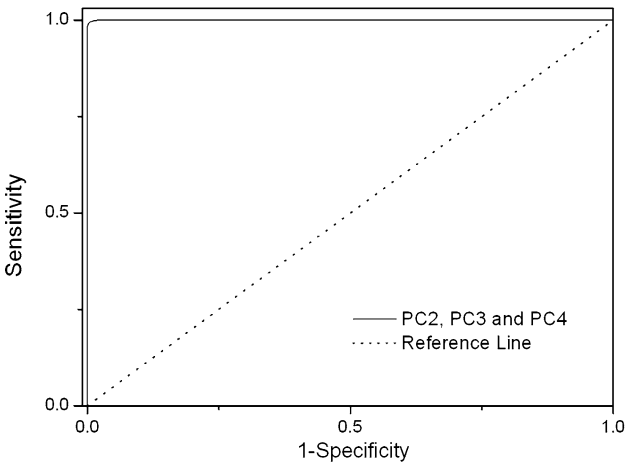

Fig. 2 ROC curve, a *dashed line* represents two indistinguishable populations (random data) and the *solid line* is the combination of PC2, PC3, and PC4

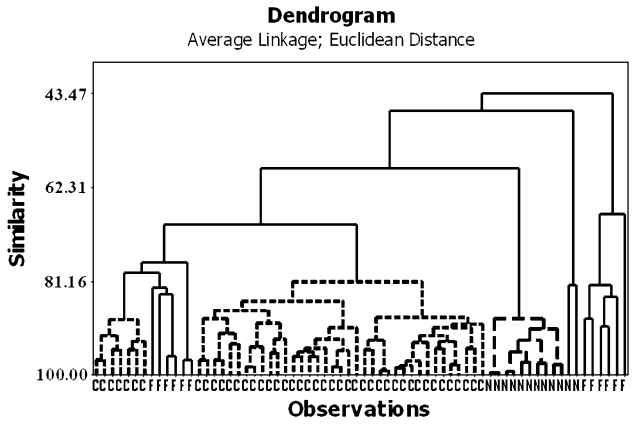

Fig. 3 Dendogram of IDC (*C*), FBC (*F*) and NB (*N*) groups

Figure 3 shows the tree-like diagram (dendrogram, using an average linkage and Euclidean distance of PC1, PC2 and PC3), which groups the spectra into clusters according to similarity level. All groups were correctly separated, but they were divided additionally into various subgroups. This result suggests a variation in the chemical composition of the tissue in these groups, which can be expected from different people living in different environments.

Figure 4 shows the box plot of the groups, the black line represents the group average and the gray shadow is the range of variation.

The Raman data from the NB group showed a typical spectrum for a normal tissue. The peaks were sharp with low values of full width at half maximum (FWHM) and there is a peak around 1,740 cm^{-1}. Nevertheless, the IDC group spectra had broad bands and the peak around 1,740 cm^{-1} almost disappeared.

The major changes on the spectra of the FBC group were at a lower wavenumber range. The presence of cysteine amino acid (a sharp peak around 540 cm^{-1}) and a small peak at the right side, which is assigned to tryptophan/cytosine, guanine modes (573 cm^{-1}) as shown in the Fig. 4 [11, 21]. The s(C–C) aromatic ring phenylalanine at 1,001 cm^{-1}, appeared stronger, and the bands among 1,200–1,400 cm^{-1} became broader and lost resolution.

4 Biochemical interpretation of the spectra

Among the benign diseases, FD is one of the most common in the female breast that is composed of both fibrous and granular tissue [24]. Unfortunately, mammography or sonography cannot reliably diagnosis the difference between small noncalcified FDs and small carcinomas with regular borders [30]. FBC is also a benign breast disease which manifests morphologically as fibrosis, inflammation, cysts and micro-cysts, but the clinical and radiologic findings are not specific. This condition, many times, causes doubt on the part of the patient and the medical evaluator. DCIS is a disease condition that may manifest clinically or radiologically either as a malignant or a benign disease. This disease has the same histological

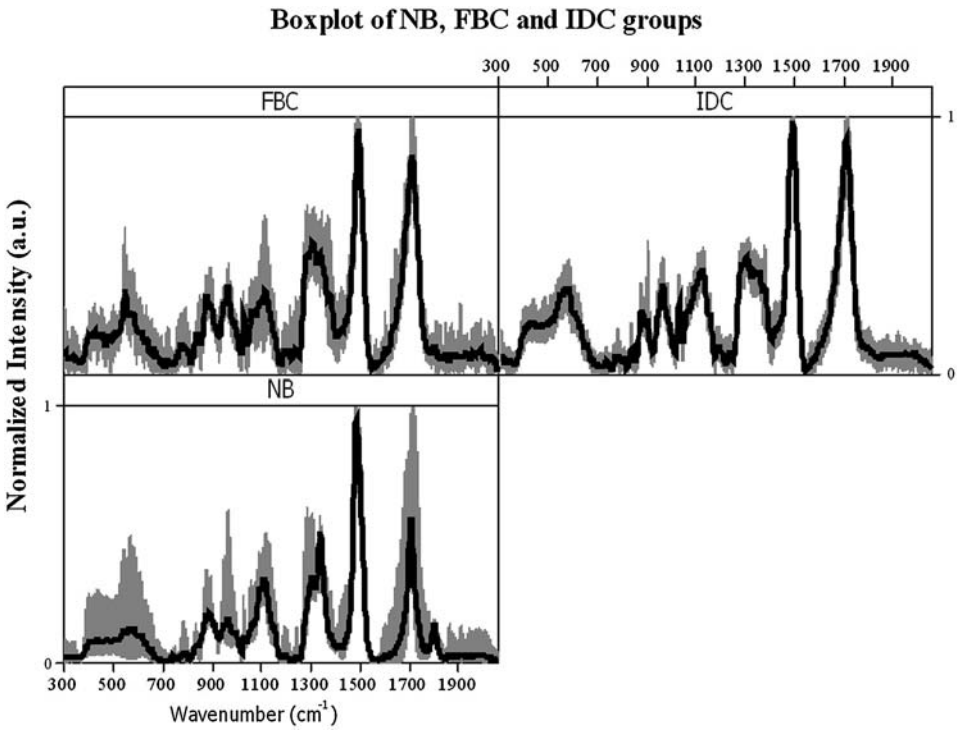

Fig. 4 Boxplot for NB, FBC, and IDC groups

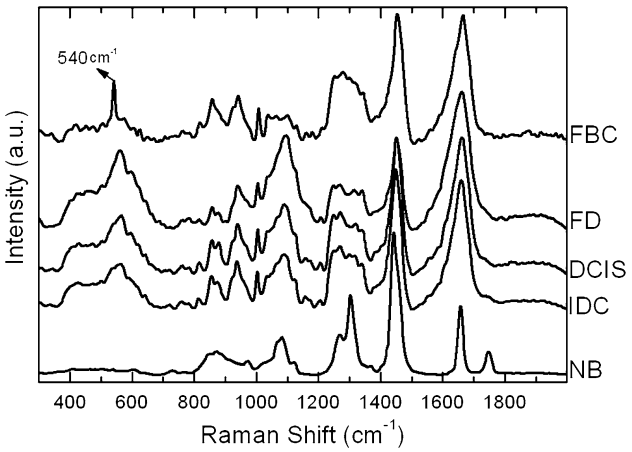

Fig. 5 Average spectra for NB, FBC, and IDC clusters

alteration characteristics as IDC, but with no ductal membrane invasion [2, 31].

Figure 5 shows the average spectra of all groups. The DCIS and FD spectra were added to the graph, where the diagnostic was done by histopathological analysis. These two spectra were not included in the statistical analysis preview due to the small number of patients, but it was possible to comment on the biochemical changes as a function of a patient's diagnosis.

For a better understanding of the spectra, the biochemistry involved in the reactions of reactive oxygen and nitrogen species must be considered. Some of these forms are free radicals that can damage lipids, proteins, and nucleic acids. Free radicals are chemical species that have a single unpaired electron in an outer orbital. Energy created by this unstable configuration is released through reactions with adjacent molecules, such as inorganic or organic chemicals—proteins, lipids, carbohydrates—particularly with key molecules in membranes and nucleic acids [32]. Biological systems are protected against oxidative species by mechanisms designed to suppress potentially harmful oxidative processes. Indeed, oxygen is involved in many metabolic reactions and many molecules like superoxide and hydrogen peroxide in addition to their relative, nitric oxide (NO) can be produced. These species have been implicated in both physiologically helpful and harmful reactions [33, 34].

However, a high concentration of oxidative species promotes the conversion of NO to higher oxide forms, such as nitrogen dioxide and peroxynitrite. Peroxynitrite is a highly reactive molecule, which could induce changes in proteins by oxidizing the sulfhydryl groups of cysteine and other amino acids [35].

The production of such oxidative species leads to a formation of nitrotyrosine, which has been identified as an indicator of cell damage. Besides this, inflammation and the aging process are recognized as a peroxynitrite-triggered mechanism of nitrosative injury. An increase in cysteine band intensities for FBC and IDC groups may suggest an organism defense mechanism, considering that cysteine is a precursor of glutathione and a powerful intracellular antioxidant with free SH groups, which are scavengers of peroxynitrite, and also of oxyradicals [36].

The phenylalanine peaks around 1,001 and 1,209 cm^{-1} are also more intense for the FBC, FD, DCIS, and IDC groups. In fact, phenylalanine is present in various processes that involve both the formation of collagen fibers, observed in FBC and FD, and in neoplasm processes, DCIS and IDC, that are characterized by uncontrolled cell growth. This could be an explanation for the increased peak intensity observed in abnormal tissues. The increase in intensity of the phenylalanine peaks for the DCIS and IDC provides useful information in differentiating between the DCIS and IDC grades [37].

Normal breast tissue presents a delicate stroma, consisting of loose connective tissue (collagen III) and large quantities of lipids, as observed in the NB spectra (1,740 cm^{-1}). Under pathological conditions, like cystic fibrosis or cancer, this delicate stroma is replaced by more resistant fibrous connective tissue, which contains mostly proteins [38]. This would explain the lower intensity of the peak at 1,740 cm^{-1} and widening of bands for the FBC, FD, DCIS, and IDC groups.

5 Conclusions

A reasonable separation between normal and modified tissue was obtained using principal component analyses. The comparison between all spectra studied showed more important contribution at lower Raman shift for altered tissue, which could be associated with the reduction and oxidation process in the cell, which resulted in an inflammatory reaction. The phenylalanine also shows an increased contribution in all groups, except in the normal tissue group. This could be correlated with the presence of collagen fibers. Finally, a decrease of peak at 1,740 cm^{-1} and the widening of bands could be related to the formation of fibrous connective tissue.

Acknowledgments The author would like to thank the pathologists Cíntia Lopes, Jerso Menegasi, and Rosane Aguiar for supporting the histopathologic analysis. The author also thanks the patients for permitting this research. A word of gratitude is due to FAPESP and CNPq for their financial support on Projects 0114384-8 and 301362/2006-8, respectively.

References

1. World Health Statistics (2008) WHO Library Cataloguing-in-Publication Data. ISBN:978-92-4-0682740

2. Hayat MA (2008) Methods of cancer diagnosis, therapy, and prognosis—breast carcinoma. Springer, Berlin. ISBN:978-1-4020-8368-6
3. Haka AS, Shafer-Peltier KE, Fitzmaurice M, Crowe J, Dasari RR, Feld MS (2005) Proc Natl Acad Sci USA 102:12371–12376
4. Stone N, Matousek P (2008) Cancer Res 68(11):4424–4430
5. Krishna CM, Kurien J, Mathew S, Rao L, Maheedhar K, Kumar KK, Chowdary MV (2008) Expert Rev Mol Diagn 8(2):149–166
6. Miranda Marzullo AC, Neto OP, Bitar RA, da Silva Martinho H, Martin AA (2007) Photomed Laser Surg 25(5):455–460
7. Matousek P, Stone N (2007) J Biomed Opt 12(2):024008
8. Haka AS, Volynskaya Z, Gardecki JA, Nazemi J, Lyons J, Hicks D, Fitzmaurice M, Dasari RR, Crowe JP, Feld MS (2006) Cancer Res 66(6):3317–3322
9. Utzinger U, Heintzelman DL, Mahadevan-Jansen A, Malpica A, Follen M, Richards-Kortum R (2001) Appl Spectrosc 55:260A–294A
10. Bitar RA, Tierra-Criollo C J, Ramalho LNZ, Netto MM, Martin AA (2006) J Biomed Opt 11(5), 054001-1/8
11. Shafer-Peltier KE, Haka AS, Fitzmaurice M, Crowe J, Myles J, Dasari RR, Feld MS (2002) J Raman Spectrosc 33:552–563
12. Smith J, Kendall C, Sammon A, Christie-Brown J, Stone N (2003) Technol Cancer Res Treat 2(4):327–332
13. Matousek P, Stone N (2007) J Biomed Opt. doi:10.1117/1.2718934
14. Kneipp J, Schut TB, Kliffen M, Menke-Pluijmers M, Puppels G (2003) Vib Spectrosc 32:67–74
15. Mahadevan-Jansen A, Richards-Kortum RR (1996) J Biomed Opt 1:31–70
16. Lieber CA, Majumder SK, Billheimer D, Ellis DL, Mahadevan-Jansen A (2008) J Biomed Opt. doi:10.1117/1.2899155
17. Lieber CA, Mahadevan-Jansen A (2003) Appl Spectrosc 57(11):1363–1367
18. Stone N, Kendall C, Smith J, Crow P, Barr H (2004) Faraday Discuss 126:141–157
19. Clark RJH, Heiser RE (1993) Advances in spectroscopy. Wiley, New York
20. Majumder SK, Kanter E, Robichaux-Viehoever A, Jones H III, Mahadevan-Jansen A (2007) In: Vo-Dinh T, Grundfest WS, Benaron DA, Cohn GE, Raghavachari R (eds) Advanced biomedical and clinical diagnostic systems V. Proc SPIE 6430:64300Q. doi:10.1117/12.724873
21. Lauridsen RK, Everland H, Nielsen LF, Engelsen SB, Norgaard L (2003) Skin Res Technol 9:137–146
22. Puppels GL, Del Mul FF, Otto C M, Greve J, Robert-Nicoud M, Arndt-Jovin DJ (1990) Nature 347:301–302
23. Stone N, Kendall C, Shepherd N, Crow P, Barr H (2002) J Raman Spectrosc 33:564–573
24. Deng H, Bloomfield VA, Benevides JM, Thomas G (1999) J Biopolymers 50:656–666
25. Ikoma T, Kobayashia H, Tanaka J, Walsh D, Mann S (2003) J Biol Macromol 32:199
26. Pawlukojc A, Leciejewicz J, Ramirez-Cuesta AJ, Nowicka-Scheibe J (2005) Spectrochim Acta A 61:2474–2481
27. Naumann D (2001) Appl Spectrosc Rev 36:239–298
28. Shi Y, Wang L (2005) J Phys D 38:3741–3745
29. Li T, Yaman H, Arakawa T, Narhi LO, Philo J (2002) Protein Eng 15:59–64
30. Della Rovere GQ, Warren R, Benson JR (2006) In: Taylor & Francis, pp 242–251. ISBN:978-1-84184-384-1
31. Hall AJ, Knaus JV (2005) In: An atlas of breast disease. CRC Press, p 50. ISBN:1-58070-533X
32. Costas M, Mehn MP, Jensen MP, Que L Jr (2004) Chem Rev 104(2):939–986
33. Lehninger AL, Nelson David L, Cox Michael M (2004) Lehninger principles of biochemistry. W.H. Freeman, San Francisco, pp 842–862. ISBN:13:978-0716743392
34. Valacchi G, Davis PA (2008) Oxidants in biology. Springer, Berlin, pp 74–106. ISBN:978-1-4020-8398-3
35. Zhang SM, Willett WC, Selhub J, Manson JE, Colditz GA, Hankinson SE (2003) Cancer epidemiology. Biomarkers Prev 12:1188–1193
36. Gilad E, Salvatore C, Zingarelli B, Salzman AL, Szabó C (1997) Life sciences 60(10):169–174
37. Rehman S, Movasaghi Z, Tucker AT, Joel SP, Darr JA, Ruban AV, Rehman IU (2007) Raman Spectrosc 38:1345–1351
38. Penteado SCG, Fogazza BP, Carvalho CS, Arisawa EAL, Martins MA, Martin AA, Martinho HS (2008) J Biomed Opt 13(1):014018