Surface-enhanced Raman spectroscopy of urine for prostate cancer detection: a preliminary study

Greta Del Mistro · Silvia Cervo · Elena Mansutti · Riccardo Spizzo · Alfonso Colombatti · Pietro Belmonte · Renzo Zucconelli · Agostino Steffan · Valter Sergo · Alois Bonifacio

Abstract Surface-enhanced Raman scattering (SERS) spectra were obtained from urine samples from subjects diagnosed with prostate cancer as well as from healthy controls, using Au nanoparticles as substrates. Principal component analysis (PCA) of the spectral data, followed by linear discriminant analysis (LDA), leads to a classification model with a sensitivity of 100 %, a specificity of 89 %, and an overall diagnostic accuracy of 95 %. Even considering the very limited number of samples involved in this report, preliminary results from this approach are extremely promising, encouraging further investigation.

Keywords SERS · Raman · Prostate cancer · Urine · Nanoparticles

G. Del Mistro · E. Mansutti · V. Sergo · A. Bonifacio (✉)
Department of Engineering and Architecture, University of Trieste, Via Valerio 6a, 34127 Trieste, TS, Italy
e-mail: abonifacio@units.it

S. Cervo · E. Mansutti · A. Steffan
CRO-Biobank, CRO Aviano, National Cancer Institute, Via Franco Gallini 2, 33081 Aviano, PN, Italy

S. Cervo · A. Steffan
Clinical Cancer Pathology, CRO Aviano, National Cancer Institute, Via Franco Gallini 2, 33081 Aviano, PN, Italy

R. Spizzo · A. Colombatti
Division of Experimental Oncology 2, CRO Aviano, National Cancer Institute, Via Franco Gallini 2, 33081 Aviano, PN, Italy

A. Colombatti
Centre of Excellence MAT Department of Medical and Biological Science, University of Udine, P.le M. Kolbe 4, 33010 Udine, Italy

P. Belmonte · R. Zucconelli
Urology Department, Policlinico San Giorgio, Via Agostino Gemelli 10, 33170 Pordenone, PN, Italy

Introduction

The burden imposed by prostate cancer on public health is very relevant: it is the most frequent non-dermatologic cancer and the second leading cause of cancer death in men. Screening for prostate cancer has become a main controversial topic. The currently used screening tools—serum PSA (prostate-specific antigen) and DRE (digital rectal examination)—have a rather low predictive accuracy (reviewed in [1]), and consequently, the real benefit of these screening methods in reducing the mortality due to prostate cancer was not agreed upon [1]. In fact, there was a great concern about the risks related to overdiagnosis and overtreatment of "insignificant" prostate cancers (i.e., those cancers not threatening patient survival). In support of these criticisms, while prostate cancer detection increased linearly with the use of PSA, the survival rate did not significantly change. This implies that a greater number of indolent tumors are diagnosed whereas real tumors are not identified and thus not properly treated. For these reasons, more effective diagnostic tools are needed for an efficient screening of prostate cancer.

Surface-enhanced Raman scattering (SERS) is a vibrational spectroscopy technique, which has recently gained much interest for the diagnosis of cancer, in particular for prostate cancer [2]. In a very recent work, prostate cancer is reported to have been detected from blood serum SERS spectra with a diagnostic accuracy as high as 98 % [3]. SERS is rapid, portable, and relatively inexpensive, and thus, it is an excellent candidate for screening. The aim of this preliminary study is to report the use of SERS spectra of urine, another body fluid widely used in clinical analysis, to diagnose prostate cancer by directly comparing samples from healthy donors with those collected from patients with prostate cancer. Since urine comes in contact with the prostate, there might be changes in its biochemical composition, which could be detected by

SERS. The testing of urine has an additional advantage, that the sampling is not invasive. Hereafter, the diagnostic potential of SERS spectra of urine has been assessed using both principal component analysis (PCA) and linear discriminant analysis (LDA), two widely used chemometric techniques, often combined together in an approach called PCA-LDA [4, 5]. Despite the limited number of samples, results clearly show the occurrence of marked spectral differences, suggesting that diagnostics based on urine SERS is worthy of further investigation.

Materials and methods

Collection, processing, and storage of urine samples

Urine samples of all participants were provided by the biobank of the CRO National Cancer Institute (Aviano, Italy). Nine prostate cancer patients undergoing radical prostatectomy (age range 56±4) were enrolled by the Urology Department of Policlinico San Giorgio (Pordenone, Italy), while nine healthy subject (age range 64±5) were directly enrolled by CRO-Biobank. Samples were collected at diagnosis, so that no therapies could affect the results of our analysis. All patients underwent blood tests (PSA), urological examination, and transrectal ultrasound-guided biopsy of the prostate. The healthy controls were asymptomatic and negative for blood tests (PSA). All controls have been re-contacted 1 year after the urine collection, resulting still asymptomatic. All participants provided written informed consent to participate in this study. The CRO-Biobank project has been approved by the CRO Institutional Ethics Committee. Morning urine samples were collected from fasting subjects and promptly frozen in dry ice. Urine samples collected by the Policlinico San Giorgio were transported to CRO-Biobank in dry ice. All samples were then stored at −80 °C (within 4–5 h after collection) until analysis. Urine samples were thawed at 37 °C in a thermostatically controlled water bath and immediately filtered (Amicon Ultra 3K centrifugal filter devices, cutoff 3 kDa) by centrifugation at 14,000*g*, at 4 °C for 15 min, in order to remove traces of proteins (e.g., hemoglobin) which might interfere with SERS analysis. In this preliminary study, we focused on the detection of urine metabolites, ruling out any interfering signal due to proteins. All samples were then stored at −80 °C until SERS analysis.

SERS substrate preparation and characterization

The SERS spectra were acquired using Au nanoparticles (AuNP), prepared using the Turkevich method [6]. Briefly, 10.6 mg of $HAuCl_4$ was dissolved in 25 mL of Milli-Q water and heated until boiling. 750 μL of a 1 % sodium citrate solution was then quickly added under vigorous stirring. The solution was kept boiling under stirring for 20 min. The AuNP colloids obtained were stored in the dark at RT and were stable for several months. All chemicals were purchased from Sigma-Aldrich and used as received. All glassware used for AuNP preparation was carefully cleaned with aqua regia (HNO_3+HCl 1:3) and thoroughly rinsed with Milli-Q water. For all cleaning procedures and preparation of solutions, Milli-Q water was used. AuNP colloids were characterized by UV-visible absorption spectroscopy after each preparation (data not shown), using a Lambda 20bio UV-vis spectrometer (PerkinElmer, Monza, Italy). The extinction band maximum at 535 nm indicates an average particle diameter of 50 nm [7].

SERS instrumentation

SERS spectra were collected using an inVia Raman microscope (Renishaw plc, Wotton-under-Edge, UK) equipped with NIR diode laser emitting at 785 nm (Toptica Photonics AG, Germany) delivering 120 mW of power at the sample, and a spectrograph with a 1800-l/mm grating. For data acquisition, the laser was focused on the sample via a ×10 microscope objective (N.A. 0.25). The spectrograph was calibrated using the lines of a Ne lamp, and the calibration was checked prior to each measurement using the 520 cm^{-1} band of a silicon reference sample.

Sample preparation for SERS measurements

Filtered urine aliquots were mixed with AuNP using a micropipette in a 1.5-mL polypropylene tube, in a biofluid-substrate ratio of 1:9 for a total volume of 50 μL (i.e., 5 μL+45 μL); the AuNP were those obtained as previously described and they acted as SERS substrates. The mixtures were rapidly transferred on a UV-vacuum quality CaF_2 microscope slide (Crystal GmbH, Berlin, Germany) using a micropipette, and instantly placed under the Raman microscope for spectral acquisition. The laser was then focused on the top of the drop, and SERS spectra were immediately acquired with an exposure time of 10 s.

Data preprocessing

All data preprocessing and analysis was performed within the R software environment for statistical computing and graphics [8]. In particular, data import and export, preprocessing, and visualization were performed with the *hyperSpec* package [9] for R.

The preprocessing consisted of four steps: (1) baseline correction, (2) smoothing interpolation to project the spectra on an evenly spaced wavenumber axis (using the function *spc.loess* from package *hyperSpec*), and (3) intensity vector normalization. For the baseline correction, a polynomial

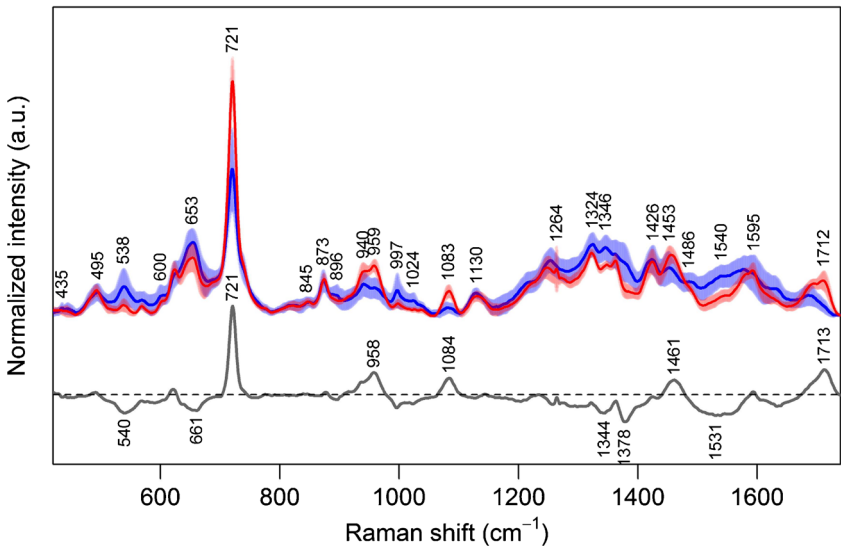

Fig. 1 Average normalized SERS spectra of urine samples of the prostate cancer group (in *red*, $n=9$) and of the control group (in *blue*, $n=9$), together with their difference spectrum (average prostate cancer−average control, in *gray*). Together with the average spectra, intensity standard deviations are reported as *shaded areas*

baseline (4th order) was fit automatically to the whole spectral range and was subtracted from each spectrum of the dataset using function *modpolyfit* from package *baseline* [10], and peak picking was performed using *detectPeaks* from package *MALDIquant* [11].

PCA-LDA classification of SERS spectra

LDA was performed using the *lda* function of package *MASS* [12]. The PCA-LDA model was then validated using the "leave-one-out" (LOO) cross-validation technique [4, 5], i.e., dividing the n spectra of the dataset in a "training set" of $n-1$ spectra (used to build the PCA-LDA model), and a "test set" of 1 spectrum (used to test the model), and repeating this operation n times, until all spectra have been left out as test set once. Since there was one spectrum per patient, the validation corresponds to a "leave-one-patient-out." For each cross-validation step, training and test sets were kept independent of each other with respect to PCA, which was carried out on the training set only; conversely, the PCs for the test samples were calculated using the *preProcess* function of package *caret* [13].

Results and discussion

Average SERS spectra of urine samples of the prostate cancer and control groups are shown in Fig. 1, along with their spectral difference. Literature on SERS spectra of urine is scarce [14], and thus, most bands in the spectra of Fig. 1 are still unassigned. Based on available data, the bands at 1426 and 1453 cm^{-1} are attributed to creatinine [15], and bands at 495,

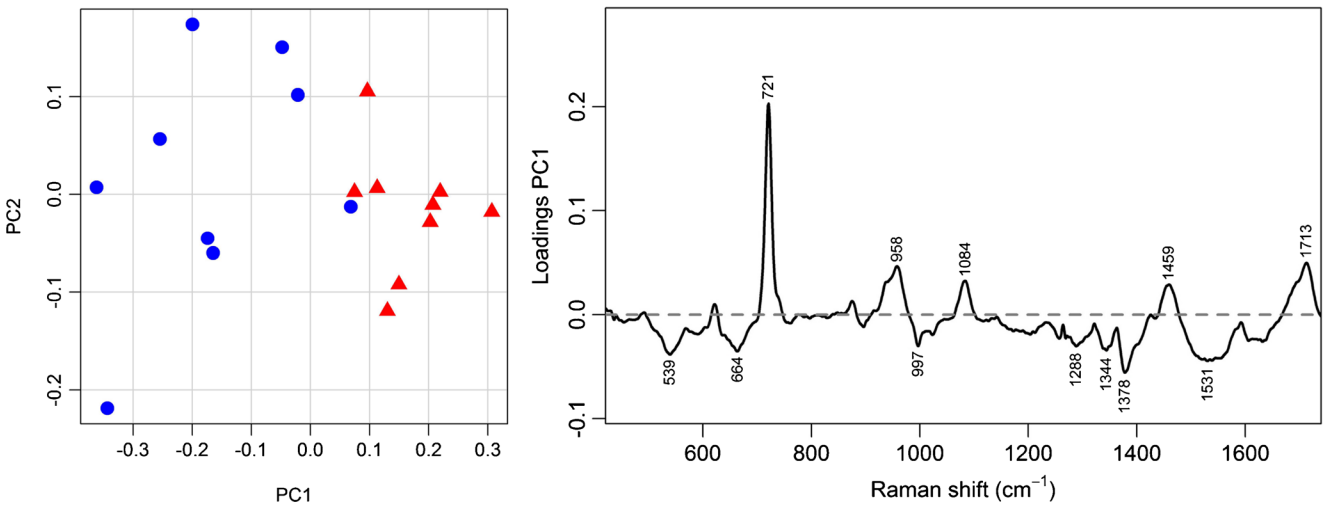

Fig. 2 (*Left*) Scores plot for the first two PCs of the urine SERS spectra (*red* prostate cancer samples, *blue* controls) and (*right*) loadings for the first component (PC1)

Table 1 Confusion table obtained from the LOO cross validation of the PCA-LDA model for the classification of urine SERS spectra

		Reference	
		Control	Cancer
Predicted	Control	8	0
	Cancer	1	9

653, and 1130 cm^{-1} are tentatively assigned to uric acid [16], both metabolites being abundant in urine [17]. The most intense band at 721 cm^{-1} could be due to a purine derivative, such as hypoxanthine [16], which, in spite of its relatively low (μM) concentration, has been observed also in other biofluids, such as serum and tears [18]. The presence of hypoxanthine in urine SERS spectra is further corroborated by the positive bands observed in the difference spectrum in Fig. 1: positive peaks at 721, 958, 1084, and 1461 cm^{-1} can be all assigned to hypoxanthine [16, 18], indicating that this metabolite is more abundant in the prostate cancer samples than in controls.

Spectral differences between SERS spectra of prostate cancer and controls are also found using PCA, a well-known "exploratory analysis" technique [4, 5]. PCA extracts the meaningful information out of a set of spectra, conveying it in a limited set of principal components (PCs). Already, the first principal component (PC1) appears to give a reasonable separation of the prostate cancer group from the controls (Fig. 2, left). As expected, the loadings of the PC1 (Fig. 2, right) are very similar to the difference spectrum of Fig. 1, suggesting that the differences observed between the spectra of the two groups might be sufficient to tell them apart.

LDA is a classification method which, given a dataset is divided into two classes (i.e., in our case, prostate cancer and controls), finds one linear function of the dataset variables that can discriminate between the classes [4, 5]. Since the number of variables should not exceed the number of samples, dimension is usually reduced by summarizing the spectral information into a smaller set of latent variables, such as the PCs. In our case, the first nine PCs, explaining the 95 % of spectral variance, were used as variables in the LDA for the construction of a predictive model, from which the sensitivity, specificity, and overall accuracy of the method can be estimated. The results of the LOO-cross-validated PCA-LDA classification are reported, in the form of a confusion matrix, in Table 1. As can be observed, only one sample was misclassified by the model, leading to a sensitivity of 100 %, a specificity of 89 %, and an overall diagnostic accuracy of 95 %.

In spite of the limited number of subjects involved in this preliminary study, the results are extremely encouraging and comparable to those obtained from blood serum [3], suggesting the need for further studies, extending this approach to a larger number of samples.

Acknowledgments We would like to thank Dr. Sergio Lenarduzzi for coordinating sample collection at Policlinico San Giorgio (Pordenone, Italy). AB would like to thank C. Beleites for the useful discussions on LDA.

References

1. Lumen N, Fonteyne V, De Meerleer G, De Visschere P, Ost P, Oosterlinck W, Villeirs G (2012) Screening and early diagnosis of prostate cancer: an update. Acta Clin Belg 67:270–275. doi:10.1179/ACB.67.4.2062671
2. Kast RE, Tucker SC, Killian K, Trexler M, Honn KV, Auner GW (2014) Emerging technology: applications of Raman spectroscopy for prostate cancer. Cancer Metastasis Rev 33:673–693. doi:10.1007/s10555-013-9489-6
3. Li S, Zhang Y, Xu J, Li L, Zeng Q, Lin L, Guo Z, Liu Z, Xiong H, Liu S (2014) Noninvasive prostate cancer screening based on serum surface-enhanced Raman spectroscopy and support vector machine. Appl Phys Lett 105:091104. doi:10.1063/1.4892667
4. Brereton RG (2003) Chemometrics: data analysis for the laboratory and chemical plant. John Wiley & Sons, Chichester
5. Varmuza K, Filzmoser P (2009) Introduction to multivariate statistical analysis in chemometrics. CRC Press, Boca Raton
6. Kimling J, Maier M, Okenve B, Kotaidis V, Ballot H, Plech A (2006) Turkevich method for gold nanoparticle synthesis revisited. J Phys Chem B 110:15700–15707. doi:10.1021/jp061667w
7. Haiss W, Thanh NTK, Aveyard J, Fernig DG (2007) Determination of size and concentration of gold nanoparticles from UV–vis spectra. Anal Chem 79:4215–4221. doi:10.1021/ac0702084
8. R Core Team (2014) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.
9. Beleites C, Sergo V (2014) 'hyperSpec: a package to handle hyperspectral data sets in R', R package version 0.98, URL http://hyperspec.r-forge.r-project.org
10. Liland KH, Mevik B-H (2015) Baseline: baseline correction of spectra. R package version 1.1-2. URL http://CRAN.R-project.org/package=baseline
11. Gibb S, Strimmer K (2012) MALDIquant: a versatile R package for the analysis of mass spectrometry data. Bioinformatics 28:2270–2271. doi:10.1093/bioinformatics/bts447
12. Venables WN, Ripley BD (2002) Modern applied statistics with S (4th ed.), Springer, New York
13. Kuhn M, Wing J, Weston S, Williams A, Keefer C, Engelhardt A, Cooper T, Mayer Z, Kenkel B, the R Core Team and Benesty M. (2014). Caret: classification and regression training. R package version 6.0-37. URL http://CRAN.R-project.org/package=caret
14. Wang T-L, Chiang HK, Lu H-H, Peng F-Y (2005) Semi-quantitative surface enhanced Raman scattering spectroscopic creatinine measurement in human urine samples. Opt Quant Electron 37:1415–1422. doi:10.1007/s11082-005-4221-6
15. Premasiri WR, Clarke RH, Womble ME (2001) Urine analysis by laser Raman spectroscopy. Lasers Surg Med 28:330–334. doi:10.1002/lsm.1058
16. Bonifacio A, Dalla Marta S, Spizzo R, Cervo S, Steffan A, Colombatti A, Sergo V (2014) Surface-enhanced Raman spectroscopy of blood plasma and serum using Ag and Au nanoparticles: a

systematic study. Anal Bioanal Chem 406:2355–2365. doi:10.1007/s00216-014-7622-1

17. Bouatra S, Aziat F, Mandal R, Guo AC, Wilson MR, Knox C, Bjorndahl TC, Krishnamurthy R, Saleem F, Liu P, Dame ZT, Poelzer J, Huynh J, Yallou FS, Psychogios N, Dong E, Bogumil R, Roehring C, Wishart DS (2013) The human urine metabolome. PLoS ONE 8:e73076. doi:10.1371/journal.pone.0073076
18. Hu P, Zheng X-S, Zong C, Li M-H, Zhang L-Y, Li W, Ren B (2014) Drop-coating deposition and surface-enhanced Raman spectroscopies (DCDRS and SERS) provide complementary information of whole human tears. J Raman Spectrosc 45:565–573. doi:10.1002/jrs.4499