Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 330 (2025) 125700

Contents lists available at ScienceDirect

Spectrochimica Acta Part A:
Molecular and Biomolecular Spectroscopy

journal homepage: www.journals.elsevier.com/spectrochimica-acta-part-a-molecular-and-biomolecular-spectroscopy

Trace detection of antibiotics in wastewater using tunable core–shell nanoparticles SERS substrate combined with machine learning algorithms

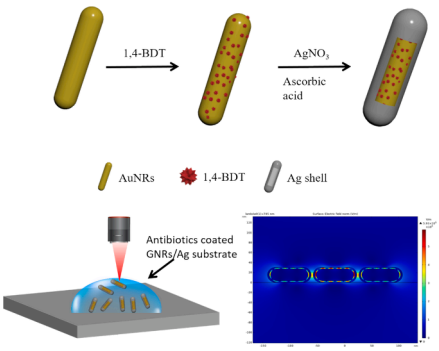
Muhammad Usman [a], Wajid Ali [c], Saleh S. Alarfaji [d], S. Tamulevičius [a,b,*]

[a] Institute of Materials Science, Kaunas University of Technology, K. Baršausko St. 59, LT-51423 Kaunas, Lithuania
[b] Department of Physics, Kaunas University of Technology, Studentų St. 50, LT-51423 Kaunas, Lithuania
[c] Key Laboratory for Micro-Nano Physics and Technology of Hunan Province, College of Materials Science and Engineering, Hunan University, Changsha, Hunan 410082, PR China
[d] Department of Chemistry, Faculty of Science, King Khalid University, P.O. Box 9004, Abha 61413, Saudi Arabia

HIGHLIGHTS

- GNRs@1,4-BDT@Ag sandwiched with 1,4-benzenedithiol as an internal standard for trace detection of antibiotics.
- Clustering analysis of SERS spectra enables the classification of antibiotic wastewater into distinct groups.
- Support Vector Machine (SVM) achieved a significant prediction accuracy of up to 99 %.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Surface-enhanced Raman spectroscopy
Machine learning algorithms
GNRs@Ag core–shell nanoparticles
Antibiotics
Detection sensitivity

ABSTRACT

Surface-enhanced Raman scattering (SERS) show great potential for rapid and highly sensitive detection of trace amounts of contamination from the environment in the surface aquatic ecosystem. The widespread use of antibiotics has resulted in serious degradation of the water environment in the past few years, and their substantial residual contamination of wastewater has a harmful effect on ecosystems, which is associated with the development of antibiotic-resistant bacterial strains. However, in this study, a novel approach of core–shell nanoparticles GNRs@1,4-BDT@Ag was used for the quantitative measurement of the concentration of antibiotics in wastewater solutions using the SERS technique coupled with computational methods. In our experiments, we selected commonly used antibiotics such as ciprofloxacin and levofloxacin in wastewater solutions. We then obtained SERS spectra for each antibiotic and its various combinations at varying concentrations. We combined it with machine learning algorithms to accurately identify and quantify the SERS spectra of the residual antibiotics in the system. Principal Component Analysis (PCA) and Orthogonal Partial Least Squares Discriminant Analysis (OPLS-DA) were subsequently employed for clustering analysis of the SERS spectral datasets. To evaluate the performance of machine learning algorithms five metrics were applied. The classification results demonstrate that while most algorithms achieved over 95 % accuracy in antibiotics status prediction, the Support Vector

* Corresponding author at: Institute of Materials Science, Kaunas University of Technology, K. Baršausko St. 59, LT-51423 Kaunas, Lithuania.
E-mail address: Sigitas.Tamulevicius@ktu.lt (S. Tamulevičius).

https://doi.org/10.1016/j.saa.2025.125700
Received 14 October 2024; Received in revised form 11 December 2024; Accepted 2 January 2025
Available online 5 January 2025

Machine (SVM) model had the best performance, attaining a remarkable prediction accuracy of up to 99 %. This developed approach helps as a simple and expeditious tool for the analysis of antibiotics in wastewater and exhibits potential for broader applications in various domains.

1. Introduction

The widespread use of antibiotics in clinical and agricultural settings has resulted in the release of massive quantities of antibiotics into wastewater systems [1]. Most antibiotic contamination is attributed to hospital and pharmaceutical wastewater, which is often directly discharged into hospital sewage, urban water systems, and wastewater treatment facilities [2]. Wastewater serves as a unique interface between the environment and human exposure, the transfer of pathogens and antimicrobial drug resistance so now its monitoring is more crucial than ever for public health [3]. Antibiotics contribute to the development of bacterial resistance, posing unpredictable risks to human health and safety [4]. These compounds enter the environment and affect ecosystem functions as they can kill or inhibit the growth of microorganisms, potentially disrupting the microbial balance in ecosystems [5], leading to reduced biodiversity and ecosystem degradation [6]. For these reasons, the accurate detection of antibiotics in wastewater and the assessment of their impact on human health and the ecological environment are of paramount importance [7]. A systematic and targeted monitoring of antibiotics in wastewater can better inform proactive prevention, investigate the effectiveness of intervention measures, and even prevent disease spread globally [3]. To combat antibiotic resistance, there is an urgent need for a sensitive and reliable method to detect antibiotics in wastewater. There have been several reported techniques for the detection of antibiotics, such as photometric, gas chromatography, high-performance liquid chromatography, Bratton-Marshall, and electrochemical detection techniques [8–11]. However these methods are time-consuming and require a lot of tedious steps to achieve entire detection.

It is important to highlight that there is potential in the development of efficient and advanced analytical instruments for the rapid detection of these contaminants so novel detection method like surface-enhanced Raman scattering (SERS) has been employed recently. The SERS has been recognized as a promising label-free detection method for multifarious applications due to its advantages of high sensitivity, time-saving experiment process special molecular fingerprint along with non-destructive nature [12,13]. According to the current theoretical research on SERS, the enhancement of SERS is mainly due to electromagnetic effect, which is a creation of a resonant effect of surface plasmon of noble metal nanostructures and incident light, since the frequency of incident light matches collective oscillation of valence electrons in coinage metal nanoparticles. This resonance behavior is called localized surface plasmon resonance [14–16].

Internal standards have been recently considered a promising technique for correcting Raman signal fluctuations from target analytes. The Internal standards is positioned between the core and the shell and can calibrate the signal fluctuation under varied test settings while not affecting the signal from the external environment [17]. Internal standard approaches have been classified into three detection modes such as internal standard addition mode, internal standard tagging detection mode, and ratiometric SERS indicator-based detection mode [18,19]. Internal standard addition and detection mode have the drawback of affecting competitive adsorption on the surface of SERS substrates or dynamic exchange between internal standards and target molecules. The ratiometric SERS indicator-based detection mode can overcome the drawbacks of the other two approaches based on selected internal standard molecules that can specifically bind to target molecules. The primary drawback of this approach is that to locate or synthesis a suitable internal standard to integrate with the target molecule is extremely challenging.

In this study, we synthesized core–shell nanoparticles (GNRs@Ag) sandwiched with 1,4-benzenedithiol as an internal standard (GNRs@1,4-BDT@Ag) for trace detection of antibiotics and generated a strong and stable Raman signal from avoiding the drawbacks of the previously discussed general internal standard approaches. The antibiotic SERS spectra are further used for computational analysis. In particular, five predictive modes were constructed based on these SERS spectra via different ML algorithms, including random forest (RF), decision tree (DT), GBoost, AdaBoost and Support Vector Machine (SVM). The results indicated that all the models achieved a prediction accuracy of over 95 %, among which the SVM model had the best predictive capacity for discriminating antibiotics with an accuracy (ACC) of 99 % and an area under the curve (AUC) close to 1. In summary, the machine learning algorithm that has been suggested effectively addresses the challenges of quantitative detection and accurate prediction of antibiotics in wastewater. These innovations support the development of SERS sensors and their real-life application in these areas.

2. Material and methods

2.1. Materials

Hexadecyltrimethylammonium bromide (CTAB, >98.0 %), 5-bromosalicylic acid (>98.0 %), silver nitrate ($AgNO_3$, >99 %), L-ascorbic acid (BioUltra, ≥99.5 %), sodium borohydride ($NaBH_4$, 99 %), chloroauric acid ($HAuCl_4{\cdot}4H_2O$, 99.9 %), and ethanol (C_2H_6O, 99 %) were purchased from Aladdin. Standards of norfloxacin and ciprofloxacin were purchased from Sigma-Aldrich. All chemicals were used without further purification. We used deionized (DI) water in all experiments.

2.2. Preparation of GNRs

The GNRs were prepared using a seed-mediated growth method following our previously established protocol [20]. In a glass bottle, 2.5 mL of $HAuCl_4$ (0.5 mM) solution and 2.5 mL of CTAB (0.2 mM) solution were combined and gently stirred at 200 rpm for 10 min. After that, 10 mM of fresh $NaBH_4$ solution was then quickly added to the $HAuCl_4$-CTAB solution while vigorously stirring at 1200 rpm for 2 min. The color altered from yellow to brownish-yellow and the solution continued to stir for 30 min. To prepare the growth solution 3.6 g of CTAB with 0.45 g of 5-bromosalicylic acid were dissolved in 100 mL warm deionized water (50 °C–60 °C) in a 250 mL glass bottle with vigorous stirring (1000 rmp). After the solution was cooled to room temperature, 480 μL of 20 mM silver nitrate solution was added. The mixture was then left undisturbed at room temperature for 15 min before 100 mL of $HAuCl_4$ (1 mM) was added. Afterwards, 15 min of stirring (500 rpm), 600 μl of AA solution (0.1 M) was added, followed by vigorous stirring (1200 rpm) for 30 s, until the entire solution was colorless. Finally, 0.1 mL of seed solution was then added into the growth solution with vigorous stirring (1200 rmp) for 1 min and allowed undisturbed at 30 °C for 12 h. The final solution was centrifuged at 10,000 rpm for 20 min to remove the supernatant, and the precipitates were disseminated in 0.05 M CTAB aqueous solution after 12 h of reaction time.

2.3. Preparation of GNRs@Ag core–shell with embedded internal standards

The GNRs@Ag core–shell nanoparticles prepared following the previous research [21–24]. The 10 mL of GNRs solution was centrifuged at 11,000 rmp for 20 min and diluted to 30 mL with deionized water to

remove the additional CTAB molecules in the GNRs solution. The GNRs solution was then dispersed in 30 mL CTAB aqueous solution (0.01 M) for 5 min while vigorous stirring (1000 rmp). In addition, 300 μl of 1,4-BDT aqueous solution(5×10^{-3} M) was added to GNRs solution while another stirring for 20 min. Then, 2 mL of 0.1 M AA was injected to GNRs@1,4BDT solution under vigorous stirring (1000 rmp). After 2 min, 3 mL of 2 mM $AgNO_3$ solution and 1 mL of NaOH solution were added to above solution while vigorous stirring (1000 rmp) respectively. The color of the solution gradually turned to bright sable, demonstrating that the silver layer was successfully coated on the surface of the GNRs. Finally, the GNRs@1,4-BDT@Ag solution was centrifuged at 11,000 rmp for 20 min to remove excess CTAB molecules.

2.4. Raman detection of antibiotics using GNRs@1,4-BDT@Ag NPs

The ciprofloxacin and levofloxacin were selected as target analytes to investigate the SERS activities of GNRs@1,4-BDT@AgNPs. The composite solution containing these antibiotics with different concentrations of 10^{-6} M, 10^{-7} M, 10^{-8} M, 10^{-9} M, and 10^{-10} M was prepared. In the Raman spectroscopy experiments, 10 μl of each antibiotic solution, was combined with the GNRs@1,4-BDT@AgNPs solution and deposited onto a glass slide. The Raman spectra of the samples were captured through a Cora100TM Raman spectrometer using an excitation wavelength of 785 nm. Throughout all experiments, the laser operated at 25mW, offering a spectral resolution of 1 nm, a spectral wavenumber resolution of 10 cm^{-1}, and a detection range spanning 400–2300 cm^{-1}. To calibrate the SERS spectra, the Raman peak at 520 cm^{-1} served as the reference point, with dark current correction achieved through integration time. The resulting SERS spectra per sample were used for subsequent analysis in this study.

3. Machine learning analysis

3.1. Spectral deconvolution and characteristic peaks

All of the information on molecule vibration is contained in the distinctive peaks of the Raman spectrum [25]. We use the Vogit tool in Origin software to extract important peak information so that we can carefully analyse the biological significance of spectral distinctive peaks. Vogit represents the convolution of Lorentzian and Gaussian functions. It is accomplished of efficiently extracting both distinct and identical characteristic peaks in the Raman spectral signals of both pure and mixed antibiotics. In particular, the Lorentzian and Gaussian widths of all characteristic peaks in the Voigt linear function were adjusted to 1 and then met until fitting. The spectral deconvolution bands were produced based on the FitPeakCurve results, and the distinctive peaks corresponding to pure antibiotics in the two-way were identified by different colors. The molecular components that corresponded to the distinctive peaks were found through examining the investigation literature of others.

3.2. Clustering analysis of SERS spectra

In this study, clustering analysis comprises the combination of identical SERS spectra into different groups to demonstrate the primary structures among the spectral data [26,27]. The two clustering techniques used to analyse the SERS data mixed with two types of antibiotics (CPX and LVX) are PCA and OPLS-DA. The PCA approach was used with the function from the Python sci-kit learn module. The fit_transform technique from the PCA function was used to fit the different SERS signals and reduce the dimension of the entire collection of SERS spectral data.

The two main components with the highest contribution values, PCA1 and PCA2, were chosen to represent the general features of the SERS data. Although PCA is an unsupervised learning technique, it merely reveals underlying patterns in the distribution of the data that can be affected by factors like Raman signal intensities and feature shifts. To overcome this drawback, the supervised learning algorithm OPLS-DA was used. OPLS-DA uses previous information collected from data labels to learn certain data patterns and separate different SERS signals. In this work, a commercial multivariate analytic programme called SIMCA software was used to perform OPLS-DA in order to reduce the impact of confusing factors on the classification results. The descriptive and predictive properties of the OPLS-DA model were evaluated using three evaluation indices: R2X, R2Y, and Q2.

3.3. Model performance evaluation and interpretive analysis

This study evaluates the classification model's performance using five key metrics: accuracy, precision, recall, F1-score, and 5-fold cross-validation. To validate the model, a different test data set is used, specifically focusing on True Positive and False Positive outcomes, apart from the training set. In the context of feature performance evaluation, accuracy is computed using the *accuracy_score* function, which represents the proportion of properly predicted samples relative to the total variety of positive and negative samples. The *precision_score* (average = 'micro') function is used to determine Precision, which signifies the proportion of correctly predicted positive instances against all predicted positive instances. The *recall_score* (average = 'macro') function is utilized to measure Recall, which denotes the proportion of correctly predicted positive instances against all actual positives. While Precision and Recall serve as a balanced pair of evaluation metrics, the F1-score serves as their harmonious average, providing a comprehensive reflection of the model's overall performance in terms of classification quality. The harmonious performance of these two indicators is taken into account using the *f1_score* (average = 'weighted'). To gauge the model's generalization capacity, the *cross_val_score* function is implemented. The model efficiency is noteworthy in practical applications, so the time function is employed to calculate the model's fitting time, with the aim of maximizing efficiency while minimizing computing resources. Additional evaluation techniques employed include the utilization of the *roc_curve* function to compute the ROC curve for each category and visually inspect its AUC value. The construction of a confusion matrix through the *confusion_matrix* function allows the examination of the sample distribution per category in the model's predictive outcomes.

4. Results and discussion

4.1. Characterizations and simulation studies of GNRs@1,4-BDT@AgNPs

GNRs@1,4-BDT@AgNPs were synthesized by placing 1,4-BDT as an internal standard between GNRs and Ag shell. The overall synthesis mechanism is depicted in Fig. 1(a). The synthetic strategy for GNRs@1,4-BDT@Ag NPs of different sizes is focused on the gradual reduction of Ag^+ ions with AA as a reducing agent. The silver nitrate solution was then eventually applied to the GNRs which had been centrifuged three times. At the same time, the reduction ability of AA was insufficient to reduce Ag^+ ions at low PH, therefore, it was critical to adjust the PH to 8.5 through the addition of NaOH solution. Actually, Ag nanoparticle was easily produced during the process of forming silver layer on the surface of GNRs, therefore, it is very important to use the mild reducing agent AA as well as the gradual addition of silver nitrate solution [21].

Fig. 1(b), and (c) show the TEM images of GNRs and GNRs@1,4-BDT@Ag NPs respectively. The morphology of GNRs was uniform and the average dimensions were about $(45 \pm 3\ nm) \times (12 \pm 3\ nm)$. Fig. 1(c) shows that the silver shell was successfully coated on the surface of GNRs after the addition of $AgNO_3$ solution to the gold nanorods solution in the AA environment. The size of GNRs and the thickness of the silver layer could be used to control the final size of GNRs. The size of GNRs could be determined by the amount of silver nitrate solution in the

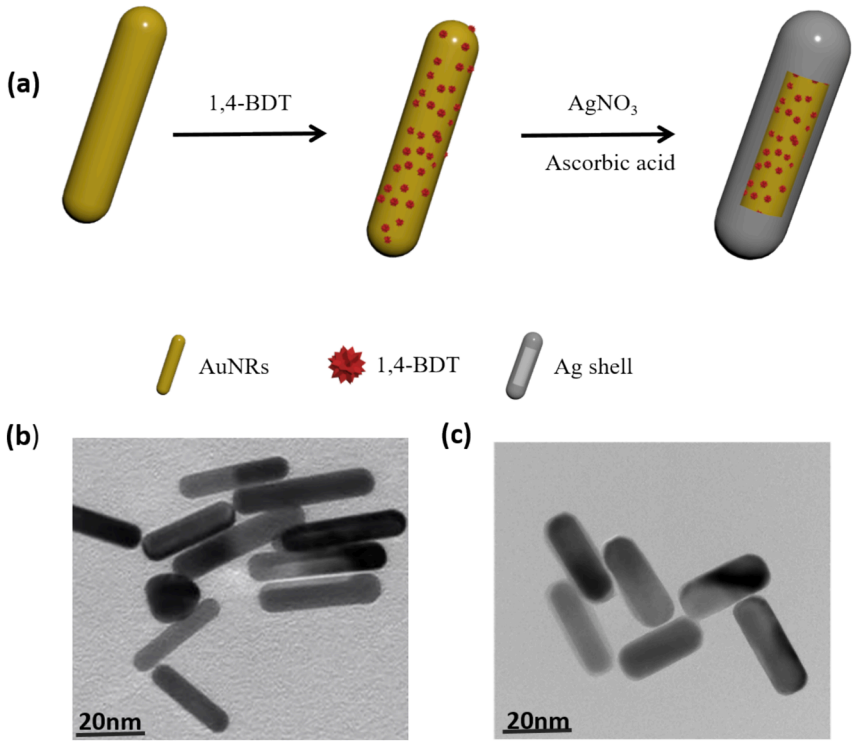


Fig. 1. (a) Schematic fabrication procedure for the core–shell GNRs@1,4-BDT@Ag NPs. (b) TEM image of GNRs. (c) TEM image of GNRs@1,4-BDT@Ag NPs.

growth solution or adding different additives such as 5-bromosalicylic acid, sodium salicylate and sodium 3-methylsalicylate to the growth solution. The goal is to accomplish maximum SERS efficiency for the internal standards by optimizing the synthesis of GNRs@1,4-BDT@Ag NPs through varying the amount of internal standard molecules on GNRs cores and the thickness of the Ag shell [28–30].

In the study of plasmonic nanostructures, particularly GNRs and their composite systems with 1,4-BDT@Ag NPs, two-dimensional wave optics simulations provide insight into the distribution and enhancement of electric fields at the nanoscale. Here we used a finite element

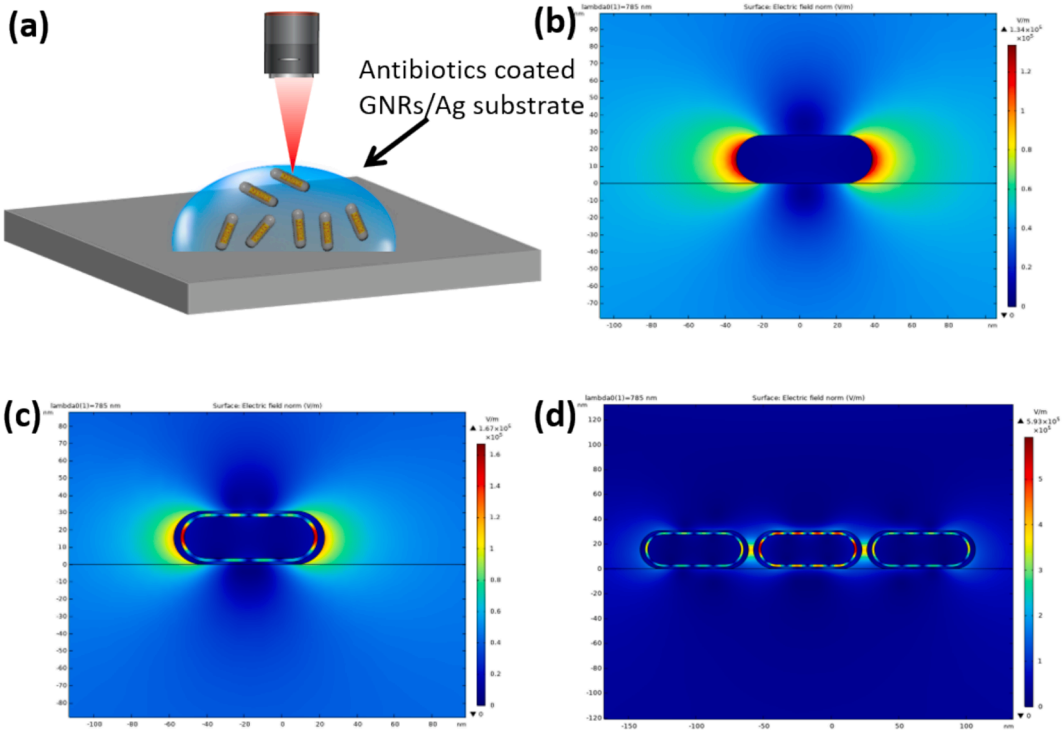


Fig. 2. COMSOL Multiphysics modeling of the GNRs@1,4-BDT@AgNPs. (a) Schematic diagram of gold nanorod coupling with the 1,4-BDT@Ag NPs as a strategy to increase electromagnetic field. (b – d) FEM simulated electric field distribution around the GNR, GNRs@1,4-BDT@Ag and assembly of the nanostructure on glass substrate to the incident wave at the wavelengths of 785 nm, respectively.

method (FEM) based software with a dedicated 2D wave optics module, the modeling commences with the construction of an appropriate computational domain representing the glass substrate, air and overlying GNRs shown in Fig. S1. The dimensions and arrangement of the GNRs within the array are meticulously defined to replicate the experimental conditions precisely.

To map the electric field distribution near individual gold nanorods (GNRs), specifically the GNRs@1,4-BDT@Ag NPs, and their array on a glass substrate, we employed the Finite Element Method (FEM) for numerical simulation, further details can be found supporting section S. Under the excitation wavelength of a 785 nm laser in Fig. 2b illustrates the plasmonic response of GNRs on the glass substrate, while Fig. 2c and d display the electric field intensity across the single and assembled GNRs@1,4-BDT@Ag NPs, respectively. To quantify the electromagnetic field enhancement, we observed an increased plasmon intensity indicates enhanced excitation and emission, leading to improved electromagnetic field. Notably, the electric field strength of GNRs@1,4-BDT@Ag NPs on the glass substrate surpasses that of isolated GNRs. The cavity effects, particularly the Fabry–Pérot effect, between the nanorods in the assembly shown in Fig. 1d, significantly enhance the electric field confinement. This simulation affirms the significant local field enhancement by the GNRs@1,4-BDT@Ag NPs array on the glass substrate, which is pivotal for plasmonically enhanced lasing. Based on our model, fluorophores are drawn to regions with strong electric fields, especially where GNRs are attached to the 1,4-BDT@Ag NPs. This amplified electric field directly boosts the excitation of fluorophore molecules and, consequently, enhances the SERS efficiency of the 1,4-BDT molecule. Therefore, the overall increase in SERS enhancement for the GNRs@1,4-BDT@Ag NPs array is attributed to the combined effect of increased surface area and intensified electric fields where fluorophores are bound.

4.2. Evaluation of SERS sensitivity and reproducibility

The SERS spectra of the antibiotics exhibit significantly enhanced Raman signals at different concentrations in wastewater as shown in Fig. 3(a) and (b). Importantly, the limit of detection (LOD) of antibiotics was determined to be 10^{-10} M. Furthermore, calibration curves have been established for the characteristic Raman peaks of CPX at 1394 cm^{-1} and LVX at 1396 cm^{-1}, as illustrated in Fig. 3(c) and (d) plotting the intensity-concentration relationship. The calibration curve was constructed using characteristic peaks at various concentrations. Notably, a strong linear correlation was observed between the SERS signal and concentration ratio, with correlation coefficients of $R2_{CPX} = 0.9914$ and $R2_{LVX} = 0.9931$. To evaluate the reproducibility of the SERS signals in this investigation, a comprehensive analysis was conducted to assess the impact of variations in signal acquisition points. To evaluate the reproducibility of the SERS signal investigation, a detailed analysis

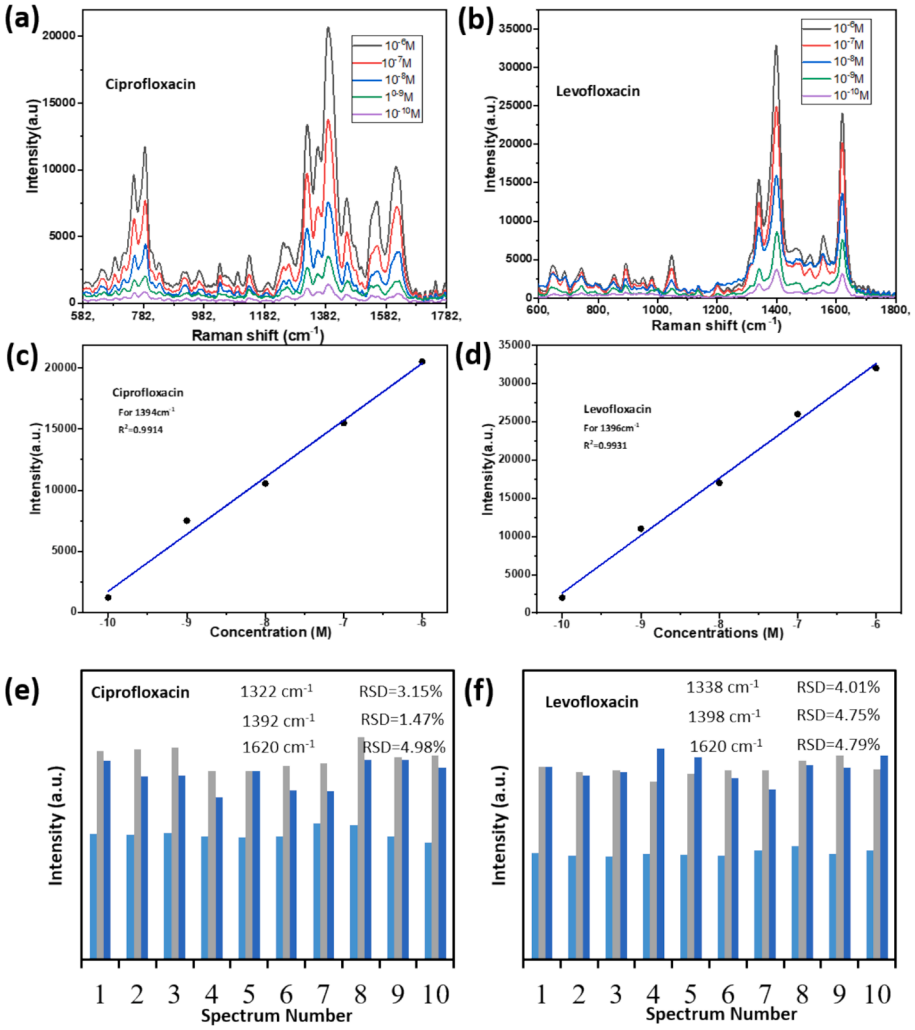


Fig. 3. (a–b) SERS spectra of CPX and LVX in wastewater at concentrations of 10^{-6} M to 10^{-10} M. (c–d) Linear correlation between CPX and LVX concentrations in wastewater. (e–f) The relative standard deviation (RSD) of the SERS intensity was determined for the three different peaks of CPX and LVX based on 10 randomly selected spots.

was performed to assess the effect of variations in signal acquisition points. The RSD of the antibiotics in the primary characteristic peaks confirms that the SERS spectra acquired in this study were consistently represented. The relative standard deviation (RSD) of the SERS intensity for the 1322 cm^{-1}, 1392 cm^{-1}, and 1620 cm^{-1} peaks of CPX was evaluated from 100 randomly selected spots, as illustrated in Fig. S2. The RSD values for these peaks were found to be 3.15 %, 1.47 %, and 4.98 %, respectively. In addition, for LVX, the intensity of the 1338 cm^{-1}, 1398 cm^{-1}, and 1620 cm^{-1} peaks was measured at 100 distinct random places to assess the SERS substrate performance. This resulted in RSD values of 4.01 %, 4.75 %, and 4.79 %, as illustrated in Figure (e) and (f) [31]. These findings demonstrate the reproducibility of the suggested SERS technique in both clinical and biological settings. Therefore, these results determine the feasibility of the method, which reveals excellent reproducibility for the detection of trace antibiotics in real-life applications.

4.3. Deconvoluted SERS spectra and characteristic peaks

However, the presence of coexisting components in antibiotic mixtures, along with noise, baselines, and system variations among Raman spectrometers, renders the identification of antibiotic components in Raman spectra of mixtures challenging. To address this issue, we initially employed a spectral deconvolution approach that established on molecular vibrations, which has been successfully utilized for fine-tuning and distinguishing differences between spectra [32]. Spectral deconvolution is the process of deconstructing of the spectral signal of the combination. The deconvoluted bands reveal that the distinct peaks of each antibiotic are clearly separated (Fig. 4). Furthermore, the pure antibiotics in various mixture spectra are shown in different colors. These antibiotics exhibit the same bands at 1598 cm^{-1} and 1628 cm^{-1} that are associated with the C═O stretching vibration [33] and the C═O stretching [34]. The potential to distinguish between the SERS spectra of various antibiotics is based on the distinct peaks that correspond to each pure antibiotic, as well as common characteristic peaks. The CPX contains four unique peaks, 754 cm^{-1} corresponds to C—S stretching [35], 784 cm^{-1} as the leading vibrational frequency of the CPX corresponds to the symmetric vibration of the C—F bond [36], the peak at 1260 cm^{-1} is attributed to the stretching vibration of the C—O bond [37], and the symmetric O—C—O of the carboxylic acid group appears at a wavelength of 1390 cm^{-1} [34]. In addition, the LVX also exhibited two distinctive peaks, 1344 cm^{-1} corresponds to CH2 wagging [38], and 1400 cm^{-1} attributed to the stretching of the pyridine ring [33]. The structural changes in these signature peaks across different spectra can serve as biomarkers for sample identification [39].

4.4. Cluster analysis of SERS spectra of antibiotics

Clustering analysis is an effective approach for detecting similarities among data [34]. Therefore, two clustering techniques, PCA and OPLS-DA, were used for conducting clustering analysis on SERS signals in this work. The findings of the PCA technique showed a significant overlap between the two antibiotic SERS spectra, demonstrating that clusters were comparatively not distinct. This highlights the inadequate performance of the PCA algorithm for this specific dataset (Fig. 5a). On the other hand, OPLS-DA, a supervised statistical method for data clustering, provides information into the division of data groups according to multi-dimensional spectral measurements. The antibiotic spectra were clearly clustered into different groups according to the OPLS-DA findings, which can be identified through reduced intra-genus variance and clear group boundaries. (Fig. 5b). Moreover, the assessment indices, R2X = 0.921, R2Y = 0.904, and Q2 = 0.891, established the ability of OPLS-DA in efficiently differentiating between various SERS signals.

4.5. Machine learning hyperparameter optimization

In this study, we examined the performance of five supervised machine learning algorithms and investigated their efficacy by analyzing their performance on a specific dataset. Five evaluation metrics were employed to examine the performance of different models. The discrimination accuracy of each model reached over 0.95 according to the calculations in Table 1, after parameter tuning and iteration. Among them, SVM performed the best of all the other algorithms in this group and achieved nearly perfect scores on all evaluation metrics, with an accuracy of 99.85 %, precision of 99.85 %, recall of 98.24 %, F1 score of 99.85 %, and a high score of 99.14 % in 5-fold cross-validation. This indicates that the SVM model is capable of processing the dataset with high accuracy and efficiency. DT, RF and GBoost also showed good performance, but their scores in 5-fold cross-validation were slightly lower, at 95.00 % and 95.95 %, respectively, suggesting possible mild overfitting. Relatively speaking, AdaBoost performed poorer, which only scored 96.33 % in 5-fold cross-validation, possibly indicating severe overfitting of the decision tree model on this dataset.

The specificity and sensitivity among the five models were compared through ROC curves, as shown in Fig. 6a. The model performance is

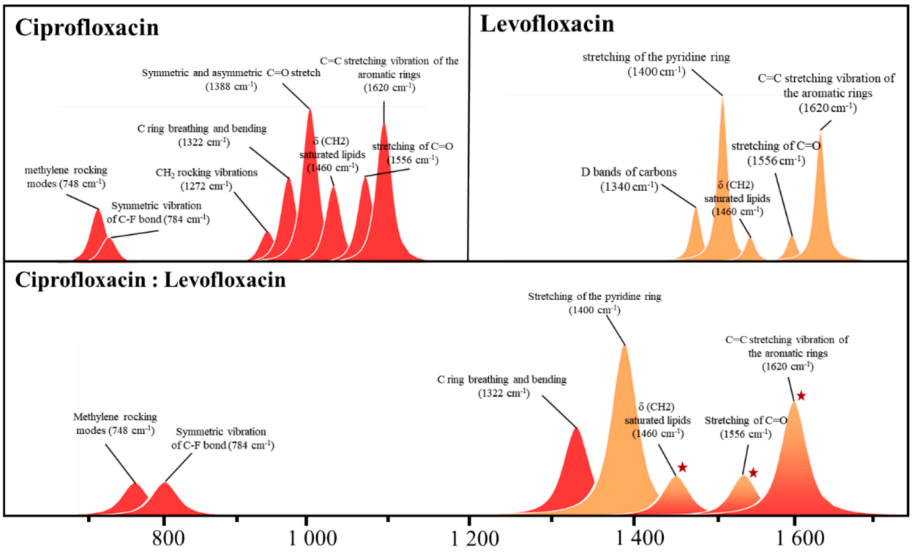


Fig. 4. DE convoluted Raman Gaussian-Lorentzian bands. The Raman shift and biological significance of each characteristic peak are shown.

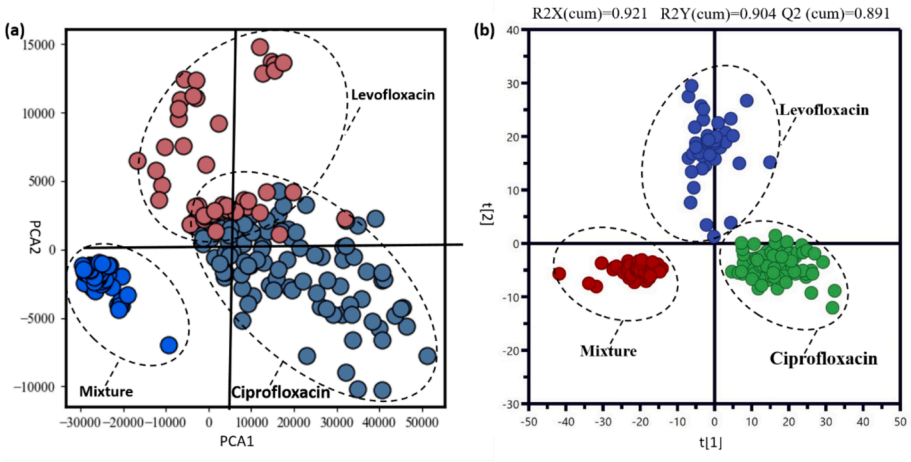


Fig. 5. PCA and OPLS-DA approach using clustering analysis of the SERS spectra of single and mixed antibiotics. (a) The scatter plot of sample points demonstrates a relatively scattered distribution of the PCA analysis. (b) In OPLS-DA analysis, various sample points are separated into distinct clusters.

Table 1
Evaluation of five supervised machine learning methods in predicting SERS spectra of two antibiotic.

Algorithm	Accuracty	Precision	Recall	F1-score	5-Fold
DT	97.77 %	97.77 %	98.41 %	97.80 %	97.72 %
RF	97.77 %	97.77 %	96.66 %	97.64 %	97.53 %
GBoost	97.46 %	97.46 %	97.51 %	97.46 %	97.95 %
AdaBoost	96.55 %	96.55 %	97.72 %	96.62 %	96.55 %
SVM	99.85 %	98.85 %	98.24 %	99.85 %	99.14 %

better when the ROC curve is closer to the upper left corner. Upon examination of the ROC curve and AUC score, there are certain differences in the performance of different machine learning algorithms on this task. The AUC score of the SVM reached a perfect score of 1.00, which shows that the SVM has excellent classification performance on this task. The confusion matrix of SVM depicted in Fig. 6b demonstrates that the model possesses exceptionally high classification performance. The antibiotics samples are correctly predicted completely.

5. Summary

In summary, we have successfully prepared size-tunable GNRs@1,4-BDT@Ag NPs as surface-enhanced Raman scattering substrates for quantitative SERS analysis. The 1,4-BDT molecules as an internal standard were embedded between the Au-Ag nanogap, protected from the external environment by the sliver shell, so that provided a reference Raman signal to correct the SERS signal. The result showed that this GNRs@1,4-BDT@Ag NPs SERS substrate helps to trace the detection of antibiotics and that there is a strong linear relationship between the concentrations of antibiotic residues and the intensity of SERS signals. Furthermore, clustering analysis techniques, PCA and OPLS-DA, are employed to classify and distinguish between the two antibiotics. The results indicate that OPLS-DA outperforms PCA in this task. In addition, the study investigates the use of machine learning techniques to distinguish and predicted SERS spectra obtained from wastewater using various antibiotics. The performance and robustness of these algorithms are evaluated through the plotted grid search gradient. Among all hyperparameter combinations, SVM exhibits the highest prediction accuracy (ACC = 99.85 %), highlighting its excellent analytical capacity for high-dimensional data in small sample sizes. We have presented an efficient and sensitive solution for the quantitative analysis of antibiotics in wastewater. This is crucial for maintaining the ecological balance of public health.

Funding

This project has received funding from the Research Council of Lithuania (LMTLT), agreement No. [S-PD-24-24], and the Deanship of Scientific Research at King Khalid University, Saudi Arabia, under grant number RGP-2/459/45.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

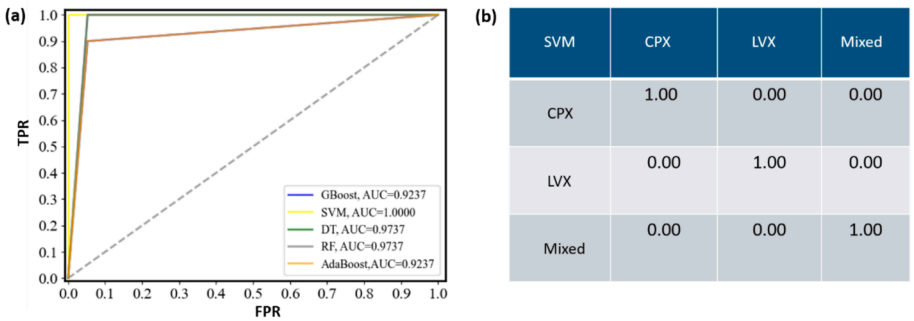


Fig. 6. Performance evaluation of machine learning classification models. (a) ROC curves of seven algorithms. ROC curves of different colors have been quantified with AUC values. (b) Confusion matrix of SVM.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.saa.2025.125700.

Data availability

Data will be made available on request.

References

[1] M. Pazda, J. Kumirska, P. Stepnowski, E. Mulkiewicz, Antibiotic resistance genes identified in wastewater treatment plant systems–a review, Sci. Total Environ. 697 (2019) 134023.

[2] P. Lorenzo, A. Adriana, S. Jessica, B. Carles, F. Marinella, L. Marta, B.J. Luis, S. Pierre, Antibiotic resistance in urban and hospital wastewaters and their impact on a receiving freshwater ecosystem, Chemosphere 206 (2018) 70–82.

[3] L. Cui, H.-Z. Li, K. Yang, L.-J. Zhu, F. Xu, Y.-G. Zhu, Raman biosensor and molecular tools for integrated monitoring of pathogens and antimicrobial resistance in wastewater, TrAC Trends Anal. Chem. 143 (2021) 116415.

[4] E.M. Aghdam, M.S. Hejazi, A. Barzegar, Riboswitches: from living biosensors to novel targets of antibiotics, Gene 592 (2016) 244–259.

[5] P.A. Segura, M. François, C. Gagnon, S. Sauvé, Review of the occurrence of anti-infectives in contaminated wastewaters and natural and drinking waters, Environ. Health Perspect. 117 (2009) 675–684.

[6] M. Qiao, G.-G. Ying, A.C. Singer, Y.-G. Zhu, Review of antibiotic resistance in China and its environment, Environ. Int. 110 (2018) 160–172.

[7] J. Lyu, L. Yang, L. Zhang, B. Ye, L. Wang, Antibiotics in soil and water in China–a systematic review and source analysis, Environ. Pollut. 266 (2020) 115147.

[8] T. You, X. Yang, E. Wang, Determination of sulfadiazine and sulfamethoxazole by capillary electrophoresis with end-column electrochemical detection, Analyst 123 (1998) 2357–2360.

[9] S. Joshi, HPLC separation of antibiotics present in formulated and unformulated samples, J. Pharm. Biomed. Anal. 28 (2002) 795–809.

[10] X. Meng, C.H. Nightingale, K.R. Sweeney, Quantification of in-vitro post-antibiotic effect based on the mean recovery-time. II: a comparison of colony counting versus photometric methods for the determination of bacterial growth, J. Antimicrob. Chemother. 28 (1991) 515–521.

[11] M. Seifrtová, L. Nováková, C. Lino, A. Pena, P. Solich, An overview of analytical methodologies for the determination of antibiotics in environmental waters, Anal. Chim. Acta 649 (2009) 158–179.

[12] J.-F. Li, Y.-J. Zhang, S.-Y. Ding, R. Panneerselvam, Z.-Q. Tian, Core–shell nanoparticle-enhanced Raman spectroscopy, Chem. Rev. 117 (2017) 5002–5069.

[13] N. Khinevich, M. Juodėnas, A. Tamulevičienė, T. Tamulevičius, M. Talaikis, G. Niaura, S. Tamulevičius, Wavelength-tailored enhancement of Raman scattering on a resonant plasmonic lattice, Sens. Actuators B 394 (2023) 134418.

[14] C. Muehlethaler, M. Leona, J.R. Lombardi, Review of surface enhanced Raman scattering applications in forensic science, Anal. Chem. 88 (2016) 152–169.

[15] H. Kang, J.T. Buchman, R.S. Rodriguez, H.L. Ring, J. He, K.C. Bantz, C.L. Haynes, Stabilization of silver and gold nanoparticles: preservation and improvement of plasmonic functionalities, Chem. Rev. 119 (2018) 664–699.

[16] R.L. Garrell, Surface-enhanced Raman spectroscopy, Anal. Chem. 61 (1989) 401A–A411.

[17] X.-Q. Zhang, S.-X. Li, Z.-P. Chen, Y. Chen, R.-Q. Yu, Quantitative SERS analysis based on multiple-internal-standard embedded core-shell nanoparticles and spectral shape deformation quantitative theory, Chemom. Intel. Lab. Syst. 177 (2018) 47–54.

[18] X.-H. Pham, E. Hahm, K.-H. Huynh, B.S. Son, H.-M. Kim, D.H. Jeong, B.-H. Jun, 4-Mercaptobenzoic acid labeled gold-silver-alloy-embedded silica nanoparticles as an internal standard containing nanostructures for sensitive quantitative thiram detection, Int. J. Mol. Sci. 20 (2019) 4841.

[19] X.-H. Pham, E. Hahm, E. Kang, Y.N. Ha, S.H. Lee, W.-Y. Rho, Y.-S. Lee, D.H. Jeong, B.-H. Jun, Gold-silver bimetallic nanoparticles with a Raman labeling chemical assembled on silica nanoparticles as an internal-standard-containing nanoprobe, J. Alloys Compd. 779 (2019) 360–366.

[20] M. Usman, X. Guo, Q. Wu, J. Barman, S. Su, B. Huang, T. Biao, Z. Zhang, Q. Zhan, Facile silicone oil-coated hydrophobic surface for surface enhanced Raman spectroscopy of antibiotics, RSC Adv. 9 (2019) 14109–14115.

[21] A.K. Samal, L. Polavarapu, S. Rodal-Cedeira, L.M. Liz-Marzán, J. Pérez-Juste, I. Pastoriza-Santos, Size tunable Au@ Ag core–shell nanoparticles: synthesis and surface-enhanced raman scattering properties, Langmuir 29 (2013) 15076–15082.

[22] P.P. Patra, G.P. Kumar, Single-molecule surface-enhanced Raman scattering sensitivity of Ag-core Au-shell nanoparticles: revealed by bi-analyte method, J. Phys. Chem. Lett. 4 (2013) 1167–1171.

[23] L. Wan, R. Zheng, J. Xiang, Au@ 1, 4-benzenedithiol@ Au core-shell SERS immunosensor for ultra-sensitive and high specific biomarker detection, Vib. Spectrosc. 90 (2017) 56–62.

[24] Z. Huang, G. Meng, X. Hu, Q. Pan, D. Huo, H. Zhou, Y. Ke, N. Wu, Plasmon-tunable Au@ Ag core-shell spiky nanoparticles for surface-enhanced Raman scattering, Nano Res. 12 (2019) 449–455.

[25] W. Lu, X. Chen, L. Wang, H. Li, Y.V. Fu, Combination of an artificial intelligence approach and laser tweezers Raman spectroscopy for microbial identification, Anal. Chem. 92 (2020) 6288–6296.

[26] C.C. Aggarwal, An introduction to cluster analysis, in: Data Clustering, Chapman and Hall/CRC, 2018, pp. 1–28.

[27] W. Liu, J.-W. Tang, J.-Y. Mou, J.-W. Lyu, Y.-W. Di, Y.-L. Liao, Y.-F. Luo, Z.-K. Li, X. Wu, L. Wang, Rapid discrimination of Shigella spp. and Escherichia coli via label-free surface enhanced Raman spectroscopy coupled with machine learning algorithms, Front. Microbiol. 14 (2023).

[28] L. Lin, Z. Liu, X. Li, H. Gu, J. Ye, Quantifying the reflective index of nanometer-thick thiolated molecular layers on nanoparticles, Nanoscale 9 (2017) 2213–2218.

[29] N. Blow, Proteins and proteomics: life on the surface, Nat. Methods 6 (2009) 389–393.

[30] X. Jiang, Z. Tan, L. Lin, J. He, C. He, B.D. Thackray, Y. Zhang, J. Ye, Surface-enhanced Raman nanoprobes with embedded standards for quantitative cholesterol detection, Small Methods 2 (2018) 1800182.

[31] J.V. Perales-Rondon, A. Colina, M.C. Gonzalez, A. Escarpa, Roughened silver microtubes for reproducible and quantitative SERS using a template-assisted electrosynthesis approach, Appl. Mater. Today 20 (2020) 100710.

[32] G. Pezzotti, F. Boschetto, E. Ohgitani, Y. Fujita, M. Shin-Ya, T. Adachi, T. Yamamoto, N. Kanamura, E. Marin, W. Zhu, Raman molecular fingerprints of SARS-CoV-2 British variant and the concept of Raman barcode, Adv. Sci. 9 (2022) 2103287.

[33] H. Deng, D. McShan, Y. Zhang, S.S. Sinha, Z. Arslan, P.C. Ray, H. Yu, Mechanistic study of the synergistic antibacterial activity of combined silver nanoparticles and common antibiotics, Environ. Sci. Technol. 50 (2016) 8840–8848.

[34] I. Nugrahani, B. Tjengal, T. Gusdinar, A. Horikawa, H. Uekusa, A comprehensive study of a new 1.75 hydrate of ciprofloxacin salicylate: SCXRD structure determination, solid characterization, water stability, solubility, and dissolution study, Crystals 10 (2020) 349.

[35] L. Ruggiero, A. Sodo, M. Cestelli-Guidi, M. Romani, A. Sarra, P. Postorino, M. Ricci, Raman and ATR FT-IR investigations of innovative silica nanocontainers loaded with a biocide for stone conservation treatments, Microchem. J. 155 (2020) 104766.

[36] Z. Ansari, T.S. Bhattacharya, A. Saha, K. Sen, Block copolymer mediated generation of bimetallic Ni-Pd nanoparticles: Raman sensors of ethyl paraben and ciprofloxacin, React. Funct. Polym. 124 (2018) 1–11.

[37] P. Carey, D. Phelps, Vibrations of the scissile C—O bond in an acyl-chymotrypsin observed by resonance Raman spectroscopy, Can. J. Chem. 61 (1983) 2590–2595.

[38] K. Rajalakshmi, S. Gunasekaran, S. Kumaresan, Vibrational spectra, electronic and quantum mechanical investigations on ciprofloxacin, Indian J. Phys. 88 (2014) 733–744.

[39] G. Pezzotti, M. Kobara, T. Nakaya, H. Imamura, T. Fujii, N. Miyamoto, T. Adachi, T. Yamamoto, N. Kanamura, E. Ohgitani, Raman metabolomics of Candida auris clades: profiling and barcode identification, Int. J. Mol. Sci. 23 (2022) 11736.