A mathematical approach to deal with nanoparticle polydispersity in surface enhanced Raman spectroscopy to quantify antineoplastic agents

Antoine Dowek, Laetitia Minh Mai Lê, Tom Rohmer, François-Xavier Legrand, Hynd Remita, Isabelle Lampre, Ali Tfayli, Marc Lavielle, Eric Caudron

A mathematical approach to deal with nanoparticle polydispersity in surface enhanced Raman spectroscopy to quantify antineoplastic agents

Antoine Dowek[1,2], Laetitia Minh Mai Lê[1,2], Tom Rohmer[3,4], François-Xavier Legrand[5], Hynd Remita[6], Isabelle Lampre[6], Ali Tfayli[2], Marc Lavielle[3,4], Eric Caudron[1,2]

[1]Service de Pharmacie, Hôpital européen Georges Pompidou, APHP.Centre Université-Paris, 20 rue Leblanc, 75015 Paris, France.

[2]Université Paris-Saclay, Lipides, Systèmes Analytiques et Biologiques, 92296, Châtenay-Malabry, France.

[3]Inria, France.

[4] CMAP, Ecole Polytechnique, CNRS, Institut Polytechnique de Paris, France.

[5]Université Paris-Saclay, CNRS, Institut Galien Paris Sud, 92296, Châtenay-Malabry, France.

[6]Université Paris-Saclay, CNRS, Institut de Chimie Physique, UMR 8000, 91405, Orsay, France.

Corresponding author: Antoine Dowek, antoine.dowek@universite-paris-saclay.fr

Abstract:

Antineoplastic agents are, for most of them, highly toxic drugs prepared at hospital following individualized prescription. To protect patients and healthcare workers, it is important to develop analytical tools able to identify and quantify such drugs on a wide concentration range. In this context, surface enhanced Raman spectroscopy (SERS) has been tested as a specific and sensitive technique. Despite the standardization of the nanoparticle synthesis, a polydispersity of nanoparticles in the suspension and a lack of reproducibility persist. This study focuses on the development of a new mathematical approach to deal with this nanoparticle polydispersity and its consequences on SERS signal variability through the feasibility of 5-fluorouracil (5FU) quantification using silver nanoparticles (AgNPs) and a handled Raman spectrophotometer. Variability has been maximized by synthetizing six different batches of AgNPs for an average size of 24.9 nm determined by transmission electron microscopy, with residual standard deviation of 17.0 %. Regarding low performances of the standard multivariate data processing, an alternative approach based on the nearest neighbors were developed to quantify 5FU. By this approach, the predictive performance of the 5FU concentration was significantly improved. The mean absolute relative error (MARE) decreased from 16.8 % with the traditional approach based on PLS regression to 6.30 % with the nearest neighbors approach (p-value < 0.001). This study highlights the importance of developing mathematics adapted to SERS analysis which could be a step to overcome the spectral variability in SERS and thus participate in the development of this technique as an analytical tool in quality control to quantify molecules with good performances, particularly in the pharmaceutical field.

Key words: Surface Enhanced Raman Spectroscopy, quantitative analysis, antineoplastic agents, nanoparticle polydispersity, non-linear regression

Introduction

According to the World Health Organization (WHO), cancer is globally the second leading cause of death in 2018, and about one in six deaths is due to cancer. Different strategies are used to treat cancer. The main treatments involve surgery, chemotherapy, immunotherapy, targeted therapy and/or radiotherapy. Although oral treatments tend to develop, chemotherapy is mainly administered by injectable route in healthcare settings. Those drugs are individually prepared in aseptic conditions in specific units, under pharmaceutical supervision in accordance with the medical prescription. In order to limit toxic effects, antineoplastic drug doses are adjusted for each patient according to the type of cancer, their weight, body surface area and clinical condition. Commercialized formulations in the form of freeze-dried powder or high concentrations of drugs are dissolved and/or diluted with sodium chloride (NaCl 0.9 %) or glucose 5 % to obtain the appropriate concentration prescribed by physicians. A control has to be performed on these preparations, in order to guarantee to the patient the right drug at the right dose and therefore reduce medication errors and their consequences, which can lead to increased risk of morbidity and mortality. Even if analytical control of individualized cytotoxic preparations is not required by pharmaceutical regulations, different strategies of control [1] have been developed to ensure the chemical and physical quality of preparations before administration to patient. As a non-invasive, non-destructive and rapid technique, Raman spectroscopy was recently explored to control antineoplastic drug preparations [2–6]. Despite the optimization of this technique, Raman spectroscopy use remains limited by laser-induced fluorescence and low Raman sensitivity that limit the analysis of drugs at low concentrations. Raman spectroscopy was therefore associated to metallic nanoparticles to enhance the spectral signal of the molecule [7]. This technique known as Surface Enhanced Raman Spectroscopy (SERS) represents a promising tool to analyze drugs at low concentrations. Although the SERS theory is still not perfectly understood, two types of contributions are described in the SERS exaltation [8]: a main contribution independent of the probed molecule due to electromagnetic phenomena induced by the interaction of the laser with metallic nanoparticles, and a low contribution due to chemical mechanisms depending on the nature of the molecule that involve the formation of new molecular states and chemical bonds by direct interaction with the SERS surface. Various metallic nanostructures are described for producing SERS signal [9,10], they are called SERS substrates. In recent years, researches have resulted in extremely low detection limits [11] largely compatible with the analysis of pharmaceutical product. Despite numerous

approaches studied to use SERS in pharmaceutical and medical domain [12], only few studies reported consistent SERS analysis for the quantification of analyte [13–15]. The limited number of quantitative studies available reflects the complexity of SERS analysis. Currently, the most important bottleneck remains the management of signal variability induced mainly by two phenomena: the difficulty of synthesizing SERS substrates in a reproducible manner and controlling the contact between the analyte and the SERS substrate. Although optimization of these steps is possible, it is to be expected that a latent variability still remains. This work is therefore a feasibility study on the various key points to be mastered in SERS analysis, specifically on the data processing, in order to minimize the spectral variability, allowing the quantification of anti-cancer drugs at low concentrations.

In this study, 5-fluorouracil (5FU), one of the main antineoplastic drug used to treat solid tumors and colorectal carcinoma [16], has been analyzed. It is an antimetabolic drug with a pyrimidine ring known to respond in Raman spectroscopy and SERS [17]. Previous studies have established a low limit of quantification (LLOQ) for 5FU in Raman Spectroscopy of 1.47 mg/mL [4] while the low limit of detection (LOD) in SERS has been set at 0.15 ng/mL using silver doped sol-gel as SERS substrate [18]. To introduce variability in our spectral data, six consecutive batches of silver nanoparticles in suspension (AgNPs) were prepared and used without size or shape variability criteria, then an experimental protocol was optimized. Finally, two data processing methods were employed: a standard multivariate approach by partial least square (PLS) regression, and a new mathematical approach trying to reduce spectral variability between batches of AgNPs.

Material and Methods

Chemicals

Silver nitrate ($AgNO_3$, crystal, European Pharmacopeia quality) was obtained from Cooper (Melun, France). Trisodium citrate dihydrate ($C_6H_5Na_3O_7$, $2H_2O$, European Pharmacopeia quality, 98 %) was obtained from VWR Chemicals (Leuven, Belgium). Citrate buffer were prepared with citric acid ($C_6H_8O_7$, European Pharmacopeia quality, Merck, Darmstadt, Germany) and sodium hydroxide (NaOH, European Pharmacopeia quality, VWR Chemicals, Fontenay-sous-Bois, France). The pharmaceutical formulation of 5-fluorouracil at 50 mg/mL (5FU, PubChem CID: 3385) was obtained from Teva (La Defense, France). Ultra-pure water was generated from Milli-Q system (Millipore, Bedford, MA, USA). Aqua Regia used to clean up glassware was obtained by mixing 3 volumes of hydrochloric acid 37 % (HCl, European Pharmacopeia quality, VWR Chemicals, Fontenay-sous-Bois, France) and 1 volume of nitric acid 65 % (HNO_3, suprapur, Merck, Darmstadt, Germany).

Silver nanoparticles (AgNPs) synthesis

The AgNPs used for SERS measurements were synthetized according to the modified Lee and Meisel method [19]. Briefly and each day, 45 mg of $AgNO_3$ were dissolved in 250 mL of ultra-pure water and heated to the boiling point. Then, 2.5 mL of a freshly prepared 2 % trisodium citrate (w/v) solution, used as a reducing agent, was added to the $AgNO_3$ solution at a flow rate of 2.5 mL/min with vigorous stirring at 500 rpm. The mix solution was left to boil for one hour while refluxing to obtain AgNPs suspension. Prior to each synthesis, all glass material was cleaned with aqua regia. Six independent batches of AgNPs were prepared according to this protocol.

Characterization of AgNPs

The six batches of AgNPs were characterized by three complementary techniques: transmission electron microscopy (TEM), UV-visible spectroscopy (UV-Vis) and Dynamic Light Scattering (DLS). Drops of nanoparticle suspensions were deposited onto freshly glow-discharged carbon-coated Cu grids. Samples were observed with a JEM 2100-Plus microscope

(JEOL, Croissy, France) operating at 200 kV. The electron microscopy images were recorded with a Rio 16 (Gatan®) camera. The average AgNPs size was calculated with ImageJ software based on the measurement of 200 AgNPs for each synthesis, measured on different fields. UV-visible spectra from 300 nm to 800 nm of the AgNPs suspension were recorded with one cm path length plastic cells using a Varian Cary 50 UV-Vis (Agilent, Les Ulis, France) and ultra-pure water as reference sample. AgNPs suspensions were diluted ten times with ultrapure water before analysis. Based on UV-Vis spectra, the maximum wavelength (λ_{max}) and the full width half maximum (FWHM) were determined to characterize the size of the nanoparticles and their dispersion respectively. The hydrodynamic diameter of particles was determined at 25°C by DLS using a Zetasizer Nano ZS 90 (Malvern Instruments, Malvern, UK) operating at fixed scattering angle at 90°. Measurements were performed in disposable 4 mL cuvettes with 1 cm optical pathway and four optical faces (Sarstedt, Marnay, France) containing an appropriate volume (1 mL) of sample.

The mean results were expressed by the 95 % confidence interval (CI95) and the residual standard deviation (RSD).

Sample preparation before SERS analysis

SERS signal enhancement was first optimized. The pH of the 5FU commercialized solution used is around 9.0. At this pH, the 5FU is negatively charged as well as AgNPs due to citrate at their surface [20]. At pH 9.0, repulsion between 5FU and AgNPs occurs. To promote interactions between 5FU and AgNPs, pH was lowered to make the 5FU neutral and promote AgNPs aggregation. Three different pH values (pH 3.0, 4.0 and 6.0) close to the pKa of citrate were tested for a 5FU solution at 1 mg/mL. Citrate buffer was chosen to avoid adding new chemical specie to the mixture. The maximal SERS signal enhancement was obtained for a buffered pH of 6.0.

The 5FU solutions were prepared by diluting the stock solution in ultra-pure water to obtain solutions from 1 to 50 µg/mL. SERS samples were prepared by adding 100 µL of 5FU solution to 500 µL of freshly prepared AgNPs suspension and vigorously shaking for five seconds. Then, 50 µL of 12.5 mM citrate buffer at pH 6.0 were added to the mixture and stirred for 5 seconds. The same operator prepared a total of six series of SERS samples, one set per day for six days.

RS and SERS measurements

Spectra were acquired using the handled Metrohm Instant Raman Analyzer (MIRA, Metrohm, France) equipped with a 785 nm diode generating a maximum of 75 mW on the sample. Analyses were performed using the vial module at a focal length of 0.85 mm using the orbital raster scan (ORS) system. SERS signals were acquired from 400 cm^{-1} to 2,300 cm^{-1} with a spectral resolution between 12 cm^{-1} to 14 cm^{-1}. Samples were analyzed in glass vials in triplicate. SERS spectra were acquired after 8 minutes of contact time in order to stabilize SERS signal. The acquisition time was 3 sec with 3 accumulations.

The enhancement of the Raman intensity was estimated by the enhancement factor (EF) [21]. EF is defined by the equation (1) and equates to the ratio of intensities obtained with SERS and Raman conditions corrected by the concentrations. EF was calculated under the same SR and SERS experimental conditions for 5FU at 50 mg/mL for RS acquisition and 5 µg/mL for SERS acquisition.

$$EF = \frac{I_{SERS}/C_{SERS}}{I_{RS}/C_{RS}} \qquad (1)$$

Quantitative analysis

Partial Least Squares regression

In order to develop a quantitative model to predict the 5FU concentration in SERS analysis, a multivariate linear approach based on PLS regression was first explored as the standard method largely used in spectral data processing. PLS consists in constructing the regression model as presented in equation (2), were Y is the response matrix (*i.e.* concentrations), X a matrix of predictors, B a matrix of coefficients and E a residual matrix.

$$Y = XB + E \qquad (2)$$

The objective is to construct new variables, called latent variables (LVs), which are linear combinations of initial variables that contain most of the variability and that best correlate these new variables to the response being studied. This technique is particularly appreciated for the processing of spectral data because it allows to get rid of collinearity phenomena and to

quickly and concisely visualize the most important spectral areas, *i.e.* those which have the most weight in the construction of the model.

According to European Medicines Agency (EMA)[22] all the data were divided into two distinct datasets: a calibration dataset to construct the model and to find the regression coefficients and then a validation dataset for calculation of predictive performance. To illustrate the impact of the AgNPs batches on the predictive performance, different calibration datasets were consecutively constructed with SERS intensities from five batches of AgNPs, the intensities from the last batch serving as validation dataset. In order to construct calibration model with the maximum of robustness and to choose the appropriate number of LVs, a k-fold cross validation were used and consisting to extract randomly k groups of ten spectra from the calibration dataset as a test set to calculated root mean square error of cross validation (RMSECV). A permutation test was then applied to choose the right number of LVs such as none difference between two consecutive RMSECV was calculated.

Non-linear regression

Because variability in response across batches of AgNPs was expected, an alternative approach was explored. The objective was to find the function f which best fitted the SERS intensities for different concentrations of 5FU and different wavenumbers.

Let $y_{ijk\ell}$ be the intensity for the i^{th} series, the j^{th} concentration of 5FU c_j, the k^{th}repetition and the ℓ^{th} wavenumber. Then, the regression model writes

$$y_{ijk\ell} = f(c_j, \theta_{i\ell}) + e_{ijk\ell} \qquad (3)$$

where $\theta_{i\ell}$ a vector of parameters and $e_{ijk\ell}$ the residual error. Regarding the large concentration range investigated, the following nonlinear regression based on a sigmoidal model was selected (equation (4).

$$f(c, \theta_{i\ell}) = S_{i\ell} + \frac{(A_{i\ell} - S_{i\ell})}{1 + \exp(-\gamma_{i\ell}\,(\log(c) - \tau_{i\ell}))} \qquad (4)$$

Here, $\theta_{i\ell} = (S_{i\ell}, A_{i\ell}, \tau_{i\ell}, \gamma_{i\ell})$, for each series i and each wavenumber ℓ. The residual errors $e_{ijk\ell}$ are assumed to be independent and to follow a Gaussian distribution

$$e_{ijk\ell} \sim \mathcal{N}(0, \sigma_{i\ell}^2)$$

where

$$\sigma_{i\ell}^2 = a_{i\ell}^2 + b_{i\ell}^2 f(c, \theta_{i\ell})^2 \quad (5)$$

Maximum Likelihood Estimation (MLE) was used with the training dataset to estimate the parameters of the model for each batch. Once the parameters estimated, an inverse problem has to be solved in order to determine the concentration of a new sample on the basis of experimental acquisition y_ℓ^{new}. For each series *i*, the estimated concentration $\hat{c}_i$ minimizes the weighted quadratic error:

$$e_i(c) = \sum_{\ell=1}^{L} \left(\frac{y_\ell^{new} - f(c, \theta_{i\ell})}{\sigma_{i\ell}} \right)^2 \quad (6)$$

$$\hat{c}_i = Argmin\ e_i(c)$$

Because of the important experimental variability between batches of AgNPs, only the series $\hat{\imath}$ of the training dataset which is the closest from the new acquisition is selected and used for estimating the unknown concentration. In other words, the estimated concentration is $\hat{c} = \hat{c}_{\hat{\imath}}$ where $\hat{\imath} = Argmin\ e_i(\hat{c}_i)$.

A total of 432 spectra were acquired and used to develop quantitative models. For both models, different pre-treatments were evaluated to improve SERS spectral data such as standard normal variate (SNV), normalization by a specific band, first and second derivatives. To evaluate the predictive performance of the proposed method, the relative error (ER) and the mean absolute relative error (MARE) were calculated using *N* test series

$$MARE = \frac{1}{N} \sum_{i=1}^{N} \frac{|\hat{c}_i - c_i|}{c_i} \quad (7)$$

where $\hat{c}_i$ is the estimated concentration and c_i the nominal concentration. The coefficient of determination R^2 was also calculated, where R is the linear correlation between nominal and predictive concentrations. Finally, to compare both approaches, a statistical test based on variance analysis (ANOVA) were applied with alpha equal to 0.05. All analyses were made with R 3.6.1 software (R *Foundation for Statistical Computing*).

Results and discussion

Characterization of AgNPs

A total of six batches of AgNPs suspensions were synthetized and characterized.

UV-Vis spectra presented on the Fig.1 have shown important spectral differences between batches with a decrease of the maximal absorbance intensity and an increase of FWHM. The measure of λ_{max} = 421 nm [CI95: 411-431 nm], FWHM = 127 nm [CI95: 99-156 nm] and of maximum of absorbance with a RSD of 18.9 % confirmed these differences. DLS analyses have shown for the six batches a polydispersity index of 0.56 that is characteristic of inhomogeneous AgNPs size. The average hydrodynamic diameter was equal to 41.20 nm [CI95: 33.80 – 48.60 nm] with RSD = 17.2 %.

Secondly samples were analyzed by TEM. TEM images (Fig.2) showed a majority of spherical AgNPs but also demonstrated the heterogeneous particles size and shape of AgNPs suspensions with nanorods or nanofilaments. The average size of AgNPs determined for spherical nanoparticles was 24.9 nm [CI95: 20.9 – 28.9 nm] (RSD 17.0 %). The repartition of size between the six AgNPs batches is represented in the Fig.2 and confirmed conclusions from UV-Vis and DLS about the variability of AgNPs from one batch to another.

TEM can be considered as the gold standard method of characterization because it is possible to directly observe the size and the shape of nanoparticles. It should be considered as the most accurate method to estimate the size of AgNPs. However, TEM could only visualize a tiny fraction of the suspension and cannot be used easily because it requires a specific platform with sophisticated materials and specialized workers. UV-Vis and DLS are faster and easier to use but don't give a direct measurement of size and shape. Even so, they can be used to reveal the quality of a synthesis and to compare different batches of AgNPs. UV-Vis allowed to see graphically the quality of synthesis through the thickness of maximum absorbance peak and the value of λ_{max}. It can also be used to confirm Nps aggregation by following the shift of maximum absorption wavelength. DLS can approximate the size by the calculation of hydrodynamic diameter for spherical nanoparticles, which is always an overestimation of the real size. The higher AgNPs diameter found in DLS can be explained by the heterogeneity in shape of nanoparticles for which the apparent diameter can be much higher than the real diameter; furthermore, the hydrodynamic diameter is always an overestimation of the real one.

Nevertheless, the RSD measured with DLS is close from this measured with the UV-Vis maximum of absorbance. By monitoring these indicators, this study confirms the difficulty of producing AgNPs of controlled sizes and shapes by this synthesis method despite same conditions of preparation and respect of a rigorous protocol.

Description of the SERS signal

In accordance with Farquharson [23] the SERS spectrum of 5FU presented in Fig. 3 is dominated by few peaks characteristic of the 5FU molecule assigned to the pyrimidine ring breathing mode at 786 cm^{-1}, the trigonal ring plus C-F stretching mode at 828 cm^{-1}, the ring stretching mode at 1270 cm^{-1}, the trigonal ring plus C-H stretching mode at 1339 cm^{-1}, the ring plus N-H wagging mode at 1400 cm^{-1} and the C=O stretching mode at 1667 cm^{-1}.

The comparison of SERS and Raman spectra for the 5FU at 5 µg/mL and 50 mg/mL respectively (Fig.3) has shown significant difference in intensities with a maximal enhancement factor observed around 10^4 for the strongest band at 786 cm^{-1}, corresponding to the stretching vibration of the pyrimidine ring breathing that confirmed the interaction between 5FU and AgNPs. In addition to signal enhancement, a displacement of the Raman shift has been observed suggesting conformational changes of the molecule during interaction. This statement associated with shift of the band position (shift of 16 cm^{-1} in comparison to RS spectrum described by Farquharson) suggested a strong interaction of the pyrimidine group with AgNPs.

The optimization of experimental conditions, in particular thanks to the choice of a pH allowing the analyte and the nanoparticles to be brought closer together [24], has enabled us to reach an enhancement factor of more than 10^4. This EF should be considered as an indicator of the amplification capacity of the Raman intensity of 5FU by colloidal suspension of AgNPs at pH of 6, keeping in mind that it can change from one wavenumber to another or from one concentration to another as it had been demonstrated by Ricci and *al.* [11].

Quantitative model

The best models were obtained for SERS data corrected by Asymmetric Least Square correction (ALS) [25] and normalized by the intensity of the citrate peak at 950 cm^{-1} which is present in all samples at constant concentration. Consequently, all SERS spectra were first

corrected by these pretreatments for both approaches. An example of SERS spectra obtained with the first batch of AgNPs is represented on Fig.4.

PLS model

The PLS regression was optimized by a permutation algorithm based on the evolution of the RMSECV calculated with the calibration dataset. As an example, for batches B2 to B6 used as calibration dataset, the best model was obtained for 7 latent variables, which corresponded to a RMSECV of 3.640 µg/mL. The first latent variable explained 96.9 % of the total variability and seemed to be linked to the 5FU concentration. As shown in Fig.5 this first latent variable is well representative of the 5FU SERS signal comparing with mean SERS spectra of 5FU at 20 µg/mL for each batch. Based on this model, a MARE of 16.8 % was obtained for a R^2 of 0.9377 between theorical and predicted concentration. The correlation curve is represented on the Fig.6.

Non-linear regression

Based on the method previously proposed (Material and methods/Quantitative analysis/Non-linear regression), the mean absolute relative error has significantly decreased to 6.30 % (p-value < 0.001) for a R^2 of 0.9790 between theorical and predicted concentration. The correlation curve is displayed Fig.6.

The comparison of prediction quality through calculation of relative error between mathematical models (Fig.7) highlights that the new model is more accurate specifically on the extreme concentrations. The comparison of the prediction capacities by synthesis has shown that, for both models, syntheses S4 and S6 had the most important error.

Results show an important variability of ER regarding to AgNPs batches, probably induced by the lack of reproducibility of synthesis. This finding seems to link the AgNPs polydispersity to the quality of concentration prediction from one synthesis to another due to a lack of reproducibility of enhancement signal. Polydispersity and experimental conditions can lead to changing the excitation plasmon wavelength or the condition of aggregation [26] to reach such excitation wavelength, compatible with the wavelength of the laser used to acquire SERS spectra *i.e.* 785 nm here. This impact on the variability is directly highlighted in the PLS model where all batches were used to construct the calibration model. Despite the development of a new mathematical approach, the error of prediction still remained higher for the batches with the most polydisperse AgNPs *i.e.* S4 and S6. It would be interesting to set standard

specifications from indicators previously described (specially UV-Vis and DLS) to assert the quality of a synthesis before being used in SERS analysis, but also to accumulate more data, especially on the repetition of SERS experiments for a synthesis at a given concentration to highlight the predictive quality of a synthesis.

Conclusion

According to this study, we investigated the surface enhanced Raman scattering using a handheld Raman spectrometer at 785 nm and silver colloidal nanoparticles for the quantification of 5FU in aqueous solution at concentrations a thousand time lower than the low limit of quantification of 5FU with RS. The spectral variability involved by variability of AgNPs from one batch to another and its consequence on prediction have been highlighted. If an improvement of the synthesis reproducibility is possible, the investigation of new mathematical approach taking into account the spectral variability induced by different synthesis of AgNPs seems very convenient for the development of quantitative model in SERS. Further investigations should be conducted to improve prediction performances, specifically by acquiring more experimental data from new AgNPs batches and by applying selection criteria on the quality of batches in terms of size and shape of AgNPs Other synthesis methods as radiolysis will be also investigated in order to control more precisely the size and the polydispersity of nanoparticles.

Acknowledgments:

The authors would like to thank the Center for Interdisciplinary Research on Medicines from Liège University, specifically Ph.D. Eric Ziemons and Charlotte De Bleye. They would also like to thank Ph.D. Jean-Luc Puteaux from the Plant Macromolecule Research Center of Grenoble for the TEM images. Finally, they are grateful to the public hospitals of Paris for the financial support to Ph.D. Laetitia Lê.

[1] C. Bazin, B. Cassard, E. Caudron, P. Prognon, L. Havard, Comparative analysis of methods for real-time analytical control of chemotherapies preparations, Int. J. Pharm. 494 (2015) 329–336.

[2] L. Lê, M. Berge, A. Tfayli, A. Baillet Guffroy, P. Prognon, A. Dowek, E. Caudron, Quantification of gemcitabine intravenous drugs by direct measurement in chemotherapy plastic bags using a handheld Raman spectrometer, Talanta. (2018). https://doi.org/10.1016/j.talanta.2018.11.062.

[3] L.M.M. Lê, A. Tfayli, J. Zhou, P. Prognon, A. Baillet-Guffroy, E. Caudron, Discrimination and quantification of two isomeric antineoplastic drugs by rapid and non-invasive analytical control using a handheld Raman spectrometer, Talanta. 161 (2016) 320–324.

[4] L.M.M. Lê, M. Berge, A. Tfayli, J. Zhou, P. Prognon, A. Baillet-Guffroy, E. Caudron, Rapid discrimination and quantification analysis of five antineoplastic drugs in aqueous solutions using Raman spectroscopy, Eur. J. Pharm. Sci. 111 (2018) 158–166. https://doi.org/10.1016/j.ejps.2017.09.046.

[5] L.M.M. Le, B. Kégl, A. Gramfort, C. Marini, D. Nguyen, M. Cherti, S. Tfaili, A. Tfayli, A. Baillet-Guffroy, P. Prognon, P. Chaminade, E. Caudron, Optimization of classification and regression analysis of four monoclonal antibodies from Raman spectra using collaborative machine learning approach, Talanta. 184 (2018) 260–265. https://doi.org/10.1016/j.talanta.2018.02.109.

[6] P. Bourget, A. Amin, A. Moriceau, B. Cassard, F. Vidal, R. Clement, The Raman Spectroscopy (RS): A new tool for the analytical quality control of injectable in health settings. Comparison of RS technique versus HPLC and UV/Vis-FTIR, applied to anthracyclines as anticancer drugs, Pathol. Biol. (Paris). 60 (2012) 369–379.

[7] M. Fleischmann, P.J. Hendra, A.J. McQuillan, Raman spectra of pyridine adsorbed at a silver electrode, Chem. Phys. Lett. 26 (1974) 163–166.

[8] E. Le Ru, P. Etchegoin, Principles of Surface-Enhanced Raman Spectroscopy: and related plasmonic effects, Elsevier, 2008.

[9] M. Fan, G.F. Andrade, A.G. Brolo, A review on the fabrication of substrates for surface enhanced Raman spectroscopy and their applications in analytical chemistry, Anal. Chim. Acta. 693 (2011) 7–25.

[10] W. Li, X. Zhao, Z. Yi, A.M. Glushenkov, L. Kong, Plasmonic substrates for surface enhanced Raman scattering, Anal. Chim. Acta. 984 (2017) 19–41.

[11] M. Ricci, E. Trombetta, E. Castellucci, M. Becucci, On the SERS quantitative determination of organic dyes, J. Raman Spectrosc. 49 (2018) 997–1005. https://doi.org/10.1002/jrs.5335.

[12] J. Cailletaud, C. De Bleye, E. Dumont, P.-Y. Sacré, L. Netchacovitch, Y. Gut, M. Boiret,

Y.-M. Ginot, Ph. Hubert, E. Ziemons, Critical review of surface-enhanced Raman spectroscopy applications in the pharmaceutical field, J. Pharm. Biomed. Anal. 147 (2018) 458–472. https://doi.org/10.1016/j.jpba.2017.06.056.

[13] C. De Bleye, P.-Y. Sacré, E. Dumont, L. Netchacovitch, P.-F. Chavez, G. Piel, P. Lebrun, Ph. Hubert, E. Ziemons, Development of a quantitative approach using surface-enhanced Raman chemical imaging: First step for the determination of an impurity in a pharmaceutical model, J. Pharm. Biomed. Anal. 90 (2014) 111–118. https://doi.org/10.1016/j.jpba.2013.11.026.

[14] S. Fornasaro, A. Bonifacio, E. Marangon, M. Buzzo, G. Toffoli, T. Rindzevicius, M.S. Schmidt, V. Sergo, Label-Free Quantification of Anticancer Drug Imatinib in Human Plasma with Surface Enhanced Raman Spectroscopy, Anal. Chem. 90 (2018) 12670–12677. https://doi.org/10.1021/acs.analchem.8b02901.

[15] S.L. Kleinman, M.C. Peterman, M. Benhabib, M.T. Cheng, J.D. Hudson, R.E. Mohler, Rapid Quantification of 4,4′-Methylenedianiline by Surface-Enhanced Raman Spectroscopy, Anal. Chem. 89 (2017) 13190–13194. https://doi.org/10.1021/acs.analchem.7b02936.

[16] PubChem, 5-Fluorouracil, (n.d.). https://pubchem.ncbi.nlm.nih.gov/compound/3385 (accessed August 22, 2019).

[17] S. Farquharson, A.D. Gift, C. Shende, P. Maksymiuk, F.E. Inscore, J. Murran, Detection of 5-fluorouracil in saliva using surface-enhanced Raman spectroscopy, Vib. Spectrosc. 38 (2005) 79–84. https://doi.org/10.1016/j.vibspec.2005.02.021.

[18] S. Farquharson, C. Shende, F.E. Inscore, P. Maksymiuk, A. Gift, Analysis of 5-fluorouracil in saliva using surface-enhanced Raman spectroscopy, J. Raman Spectrosc. 36 (2005) 208–212.

[19] P.C. Lee, D. Meisel, Adsorption and surface-enhanced Raman of dyes on silver and gold sols, J. Phys. Chem. 86 (1982) 3391–3395.

[20] C.H. Munro, W.E. Smith, M. Garner, J. Clarkson, P.C. White, Characterization of the surface of a citrate-reduced colloid optimized for use as a substrate for surface-enhanced resonance Raman scattering, Langmuir. 11 (1995) 3712–3720.

[21] E.C. Le Ru, E. Blackie, M. Meyer, P.G. Etchegoin, Surface enhanced Raman scattering enhancement factors: a comprehensive study, J. Phys. Chem. C. 111 (2007) 13794–13803.

[22] Guideline on the use of Near Infrared Spectroscopy (NIRS) by the pharmaceutical industry and the data requirements for new submissions and variations, (n.d.) 28.

[23] S. Farquharson, A. Gift, C. Shende, F. Inscore, B. Ordway, C. Farquharson, J. Murren, Surface-enhanced Raman spectral measurements of 5-fluorouracil in saliva, Molecules. 13 (2008) 2608–2627.

[24] R.A. Alvarez-Puebla, E. Arceo, P.J. Goulet, J.J. Garrido, R.F. Aroca, Role of

nanoparticle surface charge in surface-enhanced Raman scattering, J. Phys. Chem. B. 109 (2005) 3787–3792.

[25] P.H.C. Eilers, H.F.M. Boelens, Baseline Correction with Asymmetric Least Squares Smoothing, (n.d.) 24.

[26] F. Tian, F. Bonnier, A. Casey, A.E. Shanahan, H.J. Byrne, Surface enhanced Raman scattering with gold nanoparticles: effect of particle shape, Anal Methods. 6 (2014) 9116–9123. https://doi.org/10.1039/C4AY02112F.

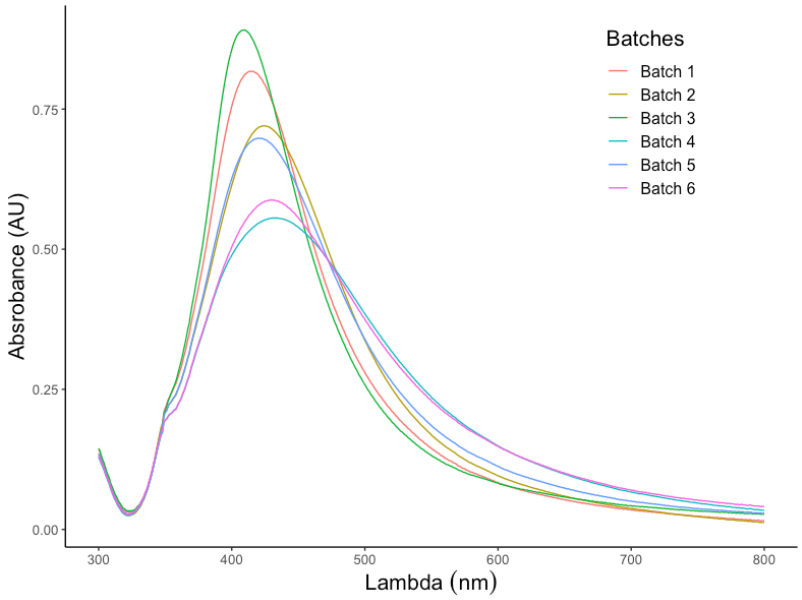

Fig. 1: UV-Vis spectra of the six silver nanoparticle batches.

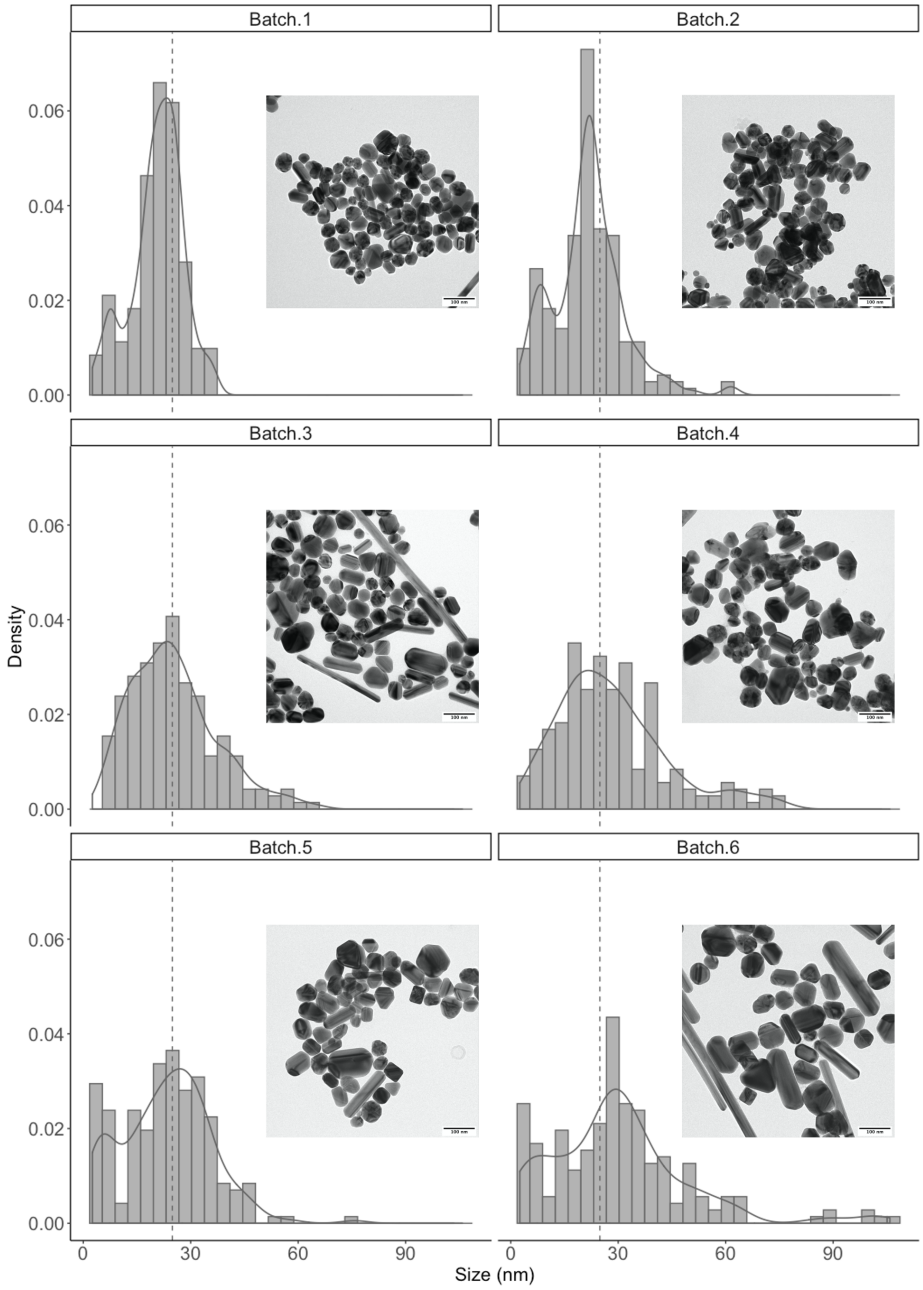

Fig. 2: Repartition of the size associated with transmission electronical microscopy images of the six silver nanoparticle suspensions. The dash line represents the mean size.

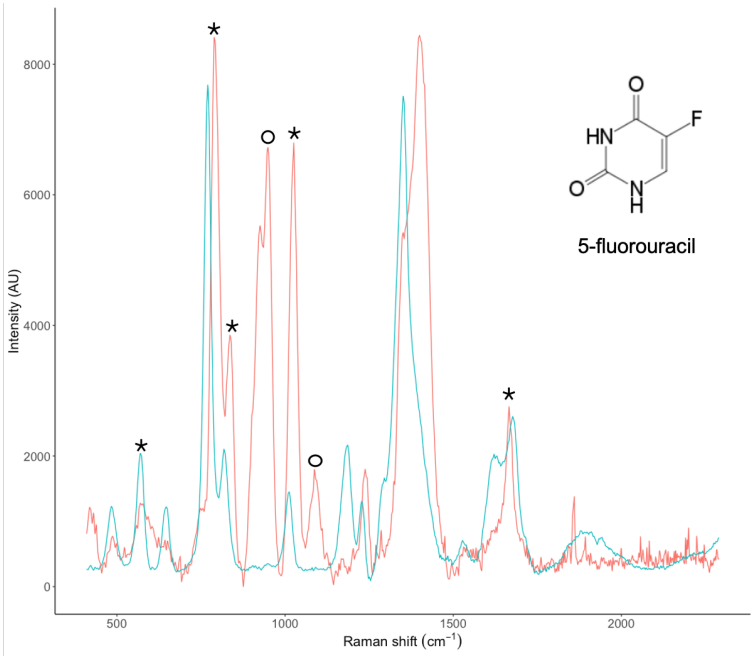

Fig. 3: The (a) Raman spectrum of 5-fluorouracil at 50 mg/mL (blue line) and (b) SERS spectrum (orange line) of 5-fluorouracil at 5 µg/mL after interaction with AgNPs and chemical structure of 5-fluorouracil (AgNPs : silver nanoparticles, SERS : surface enhanced Raman scattering). A pretreatment by asymmetric least square had been done for both spectra. Stars representing Raman shift from 5-fluorouracil, rings representing Raman shift from citrate.

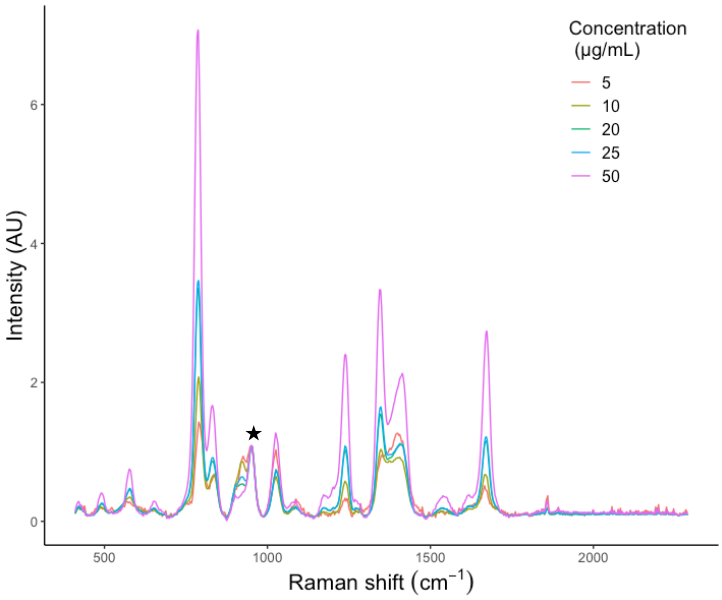

Fig. 4: The SERS spectra of 5-fluorouracil from 5 to 50 µg/mL obtained by interaction with AgNPs of the first colloidal suspension after ALS baseline correction and normalization by the 950 cm^{-1} band (★).

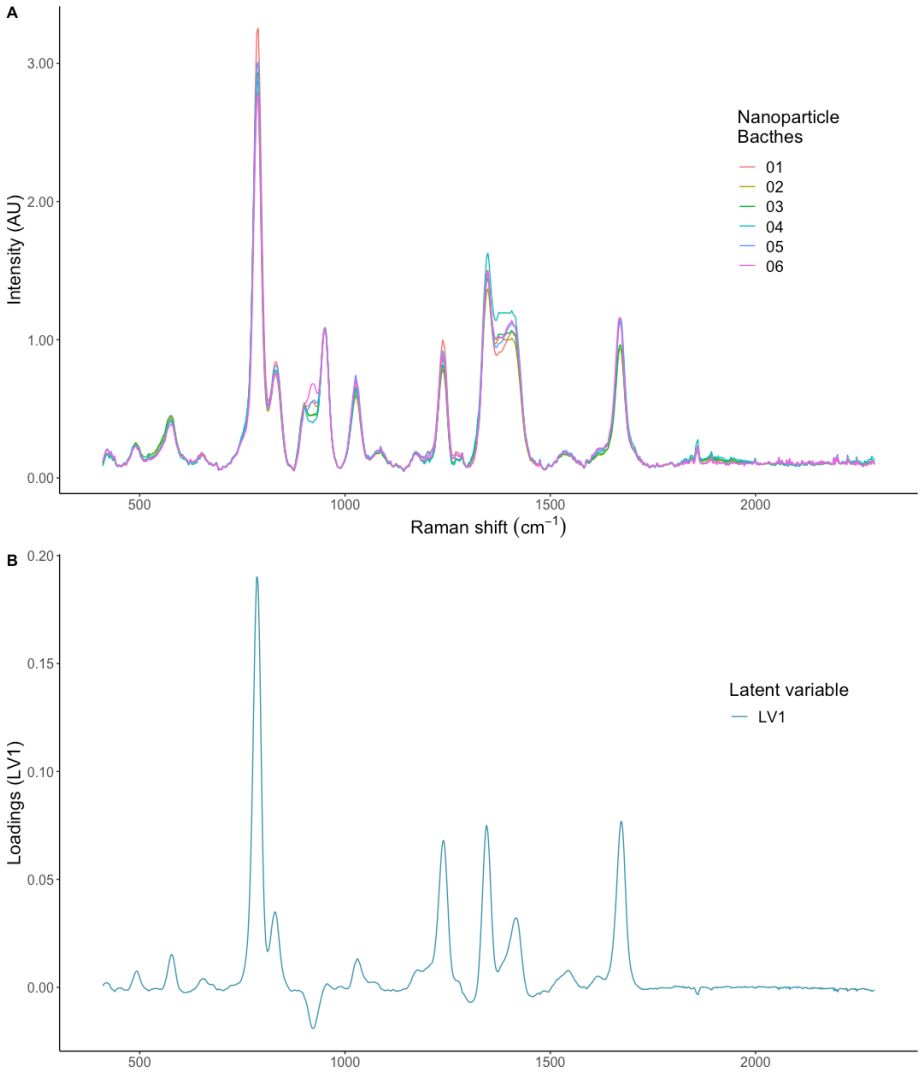

Figure 5: Similarities between the first latent variable (A) and the mean SERS spectra of 5-fluorouracil at a concentration of 20 µg/mL for the six nanoparticle batches (B)

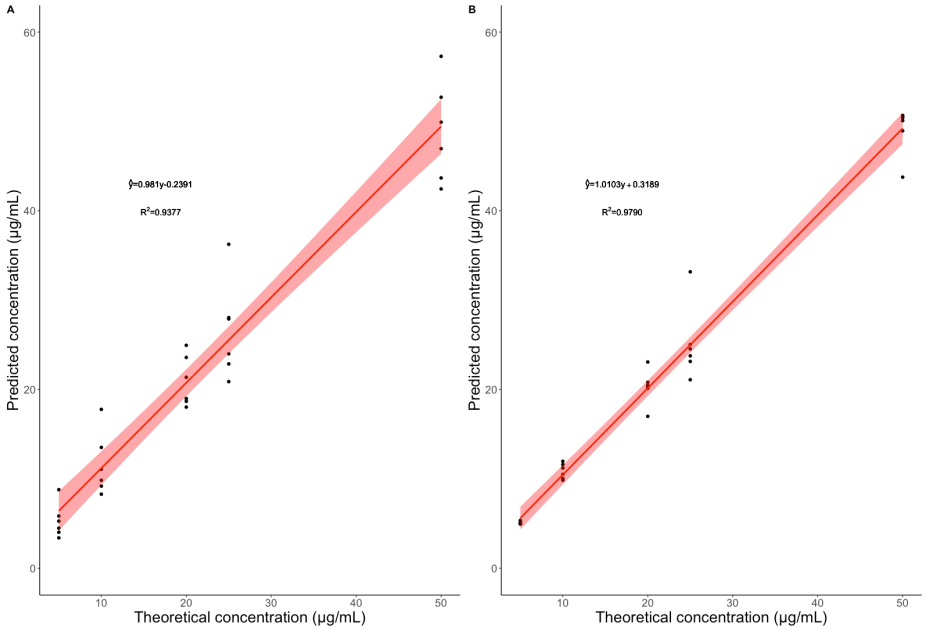

Fig. 6: Correlation curve for A: PLS, B: new approach, $\hat{y}$ represents predicted concentrations, y theorical concentrations and R^2 correlation coefficient. The red filled space represented the interval of confidence of the slope at 95 %.

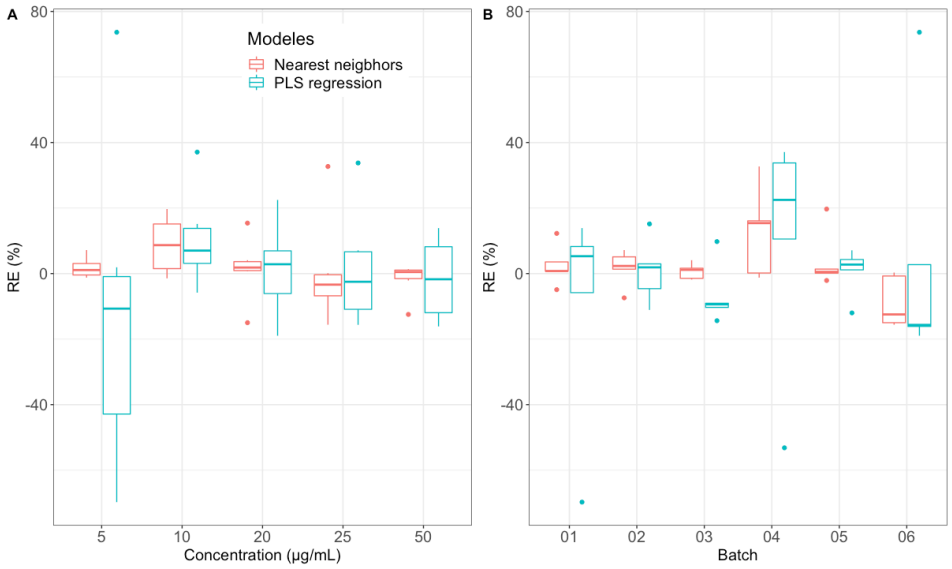

Fig. 7: Relative error (RE) of 5FU predicted concentration as a function of A) the concentrations and B) the silver nanoparticle batches for both the partial least square (PLS) regression and the nearest neighbors approach.