Anal Bioanal Chem (2010) 397:1893–1901
DOI 10.1007/s00216-010-3739-z

ORIGINAL PAPER

Multiobjective evolutionary optimisation for surface-enhanced Raman scattering

Roger M. Jarvis · William Rowe · Nicola R. Yaffe · Richard O'Connor · Joshua D. Knowles · Ewan W. Blanch · Royston Goodacre

Received: 21 January 2010 / Revised: 8 March 2010 / Accepted: 9 April 2010 / Published online: 4 May 2010

Abstract In most optimisation experiments, a single parameter is first optimised before a second and then third one are subsequently modified to give the best result. By contrast, we believe that simultaneous multiobjective optimisation is more powerful; therefore, an optimisation of the experimental conditions for the colloidal SERS detection of L-cysteine was carried out. Six aggregating agents and three different colloids (citrate, borohydride and hydroxylamine reduced silver) were tested over a wide range of concentrations for the enhancement and the reproducibility of the spectra produced. The optimisation was carried out using two methods, a full factorial design (FF, a standard method from the experimental design literature) and, for the first time, a multiobjective evolutionary algorithm (MOEA), a method more usually applied to optimisation problems in computer science. Simulation results suggest that the evolutionary approach significantly out-performs random sampling. Real experiments applying the evolutionary method to the SERS optimisation problem led to a 32% improvement in enhancement and reproducibility compared with the FF method, using far fewer evaluations.

Keywords SERS · SERRS · PESA-II · Evolutionary · Optimisation · L-cysteine

R. M. Jarvis (✉) · W. Rowe · R. O'Connor · R. Goodacre
School of Chemistry, University of Manchester,
Manchester Interdisciplinary Biocentre,
131 Princess Street,
Manchester M1 7DN, UK
e-mail: roger.jarvis@manchester.ac.uk

N. R. Yaffe · E. W. Blanch
Faculty of Life Sciences, University of Manchester,
Manchester Interdisciplinary Biocentre,
131 Princess Street,
Manchester M1 7DN, UK

J. D. Knowles
School of Computer Science, University of Manchester,
Manchester Interdisciplinary Biocentre,
131 Princess Street,
Manchester M1 7DN, UK

Introduction

Surface-enhanced Raman scattering (SERS) and surface-enhanced resonance Raman scattering (SERRS) are spectroscopic methods that employ roughened metal substrates to enable large increases in Raman scattering intensities. The phenomenon was first reported in 1976, when Fleischmann *et al*. obtained SERS of pyridine from a roughened silver electrode [1]. It is understood that a combination of two physical processes give rise to the enhanced signal. One of these, known as electromagnetic enhancement, occurs due to the interaction between the electric field of the incident light source and the electrons on the metal surface, which induces surface plasmon polariton effects [2]. The other process, called chemical enhancement, is a charge transfer effect that results in an exchange of energy between the metal substrate and adsorbed analytes [3].

Interest in SERS is accelerating [4], resulting in the development of a wide range of new substrates and encouraging the application of the technique across a variety of disciplines. In the biological sciences, there is a great deal of interest in colloidal SERS, which employs multi-faceted nanoscale coinage metal particles as substrates, and can be applied to applications as diverse as intracellular cell labelling [5–8], to DNA [9–11] and protein [12–14] detection and quantification or viral [15] and

bacterial characterisations [16, 17]. A major challenge with this approach is the optimisation of experimental conditions to achieve the largest and most reproducible enhancements for an experiment, which will act to improve detection limits and measurement errors in quantitative analyses. It is difficult to obtain consistently intense and reproducible SERS measurements; the response is influenced by choice of metal used in the substrate preparation, morphology and homogeneity of the nanoparticles, selection of aggregating agent and the relative concentrations of these in a mixture. It should also be noted that pH is of great importance in the optimisation of SERS experiments, as has been shown in [18]. Often, for colloidal SERS applications, the primary interest for improving SERS is in the development of more uniform substrates [19, 20]. Although several studies investigating effects of the experimental setup (colloid, additional reagent, analyte combinations) on SERS have been reported [21, 22], as well as robust multivariate metrics for assessing reproducibility [23] little work has addressed the reproducible optimisation of experimental conditions.

The detection and quantification of small, biologically relevant molecules is of great interest in fields such as metabolomics, where the objective is to obtain a global quantitative profile of the small molecule complement in biological samples [24–26]. This is usually achieved using hyphenated mass spectrometries (MS) or NMR (nuclear magnetic resonance) spectroscopy, although with MS accurate quantification is difficult, and NMR is less sensitive with longer instrument runs required for each sample. SERS has been used as a detection method following the separation of mixtures by LC (liquid chromatography) [27, 28], and therefore has the potential to complement the main analytical methods used in metabolomic studies. However, in the case of very complex mixtures of biological compounds; blood sera, urine or cell extracts for example, it is possible that many hundreds of chemical species will be present in the sample, and the task of making a global quantitative analysis of each of these using SERS would be difficult with a single set of experimental conditions. Therefore, the best approach in developing SERS for metabolomics is to optimise the method on a case-by-case basis.

Given that this represents a combinatorial problem for each analyte that will scale exponentially with the number of components investigated, fully mapping the search space to find the best conditions for SERS becomes impossible. A conventional experimental method for addressing such a problem would involve mapping the solution space through a full factorial design, often followed with interpolation by simplex optimisation (iterative interpolation through selection of the most optimal solutions). Simplex optimisation relies on the assumption of additivity between components, such that each component contributes independently to the overall properties of the solution. However, in the physical world, the assumption of additivity does not hold true often, and interactions occur regularly between components. The consequence of this effect for an optimisation process is that the simplex algorithm will repeatedly converge to conditions that are suboptimal.

Therefore, for the optimisation of SERS experiments, alternative methodologies are required to explore this search space rapidly. We have applied the PESA-II (Pareto envelope-based selection) multiobjective evolutionary algorithm (MOEA), to define the experimental parameters for the optimisation of colloidal SERS for L-cysteine (2-amino-3-thiol propane carboxylic acid), a hydrophilic amino acid. PESA-II is one of many MOEAs developed in the late 1990s [29] and has been widely applied in many applications, for recent examples to the optimisation of analytical approaches see [30, 31]. The principle behind this approach is that numerical objective functions are defined which reflect the desired experimental outcome (for SERS we defined objectives that quantified signal enhancement and reproducibility). Through an iterative process of experiment and evaluation of the objectives against the experimental data, the MOEA is capable of evolving new values at which to set the variables defined for a particular problem, and typically this results in substantial improvements in results. The evolutionary aspect of this computational approach uses functions that are analogous to the processes encountered in natural evolution, such as mutation and recombination (crossover); which in conjunction with a Darwinian-inspired survival-of-the-fittest strategy drive, a directed search of the solution space (often referred to in evolutionary optimisation as a landscape) towards global optima. This methodology will be developed in the future, in conjunction with automated robotic systems, to define the optimal conditions for SERS measurements rapidly against a wide range of biologically important compounds, thus improving the capabilities of SERS to be applied for the parallel quantitative metabolite analysis.

Materials and methods

Preparation of silver colloids

Citrate-reduced silver colloid was prepared following the Lee and Meisel [32] procedure, whereby 90 mg of $AgNO_3$ was dissolved in 500 mL of dH_2O and heated until boiling. Following this, 10 mL of a 1% solution of trisodium citrate was added dropwise to the mixture with vigorous stirring, and a change in colour from a clear to grey-green solution was noted. The solution was then held at boiling point for 90 min with continuous stirring. Following preparation, the colloidal solution was stored in the dark at room temperature.

Borohydride-reduced silver colloid was prepared according to Creighton *et al*. [33]. Whereby 14 mg of $NaBH_4$ was dissolved in 300 mL of dH_2O, to which 38 mL of 2.0×10^{-3} mol dm^{-3} $AgNO_3$ was added drop wise with vigorous stirring. Both of these reagents were cooled in an ice bath before mixing. The colloids were stirred continuously for 45 min, to allow for adjustment to room temperature. Borohydride-reduced silver colloid was stored at 4 °C and maintained at a low temperature throughout the analysis.

Hydroxylamine reduced silver colloid was prepared according to [34]. Whereby, 0.017 g of $AgNO_3$ was dissolved in 90 mL of dH_2O and in a separate vessel, 0.021 g of hydroxylamine hydrochloride ($NH_2OH{\cdot}HCl$) was dissolved in 5 mL of H_2O to which 4.5 mL of a 0.10 M NaOH solution was added. The two resultant solutions were combined and a grey-brown solution was observed immediately. The hydrosol was stored in the dark at room temperature. All reagents were purchased from Sigma-Aldrich UK and used without further purification.

Preparation of SERS active mixtures

L-cysteine was chosen as a target analyte for SERS against which to assess the multiobjective optimisation method, since it is of biological relevance and has been studied previously in a SERS optimisation study [22]. Three colloidal silver hydrosols were prepared as detailed above and 0.500 M stock solutions of six aggregating agents; NaCl, KCl, Na_2SO_4, K_2SO_4, $NaNO_3$ and KNO_3 were also prepared. For the initial 216 conditions assessed in the full factorial study, mixtures were made up to a total volume of 3.07 mL in 4-mL glass sampling vials. The full factorial searchspace was defined from a simple linear interpolation between extreme values for the variables being assessed, based upon practical laboratory constraints. Each mixture consisted of 3.26×10^{-6}M of L-cysteine; 14.7, 44.0 or 73.3% *v*/*v* of colloid; and 0.98, 58.63 or 97.72 mM of one of the six ionic solutions. An additional ionic solution concentration at 9.77 mM was added to the full factorial design as a second lower bound when it was found that signal was often not observed at 0.98 mM. For the further experiments, performed in the iterative evolutionary optimisation process, the algorithm determined the values for these variables. All reagents were purchased from Sigma-Aldrich UK and used without further purification.

Surface-enhanced Raman scattering

SERS spectra were recorded on a NIR 785 nm First-Defender portable Raman spectrometer (Ahura Scientific Inc., Wilmington, MA, USA); using 3 s integrations with an incident laser power of ~10 mW. From each sample, five repeat measurements were acquired and automatically saved on the spectrometer in the Thermo Scientific SPC file format. Data were exported from the device to a PC running the Windows XP operating system, and converted to ASCII format for data processing. For the five repeat SERS spectra recorded under each of the experimental conditions, the average areas under four well-defined L-cysteine SERS peaks (*ca*. 653, 798, 909 and 1,030 cm^{-1}) were calculated and used as a measure of spectral intensity. The average of the Pearson correlation coefficient (Eq. 1) was applied as a measure of reproducibility across the five replicate spectra. Pearson correlation was chosen as a metric as unlike relative standard deviation it does not assume that the level of noise is proportional to the level of enhancement.

$$R = \left(\frac{Cov(A_1, A_2)}{s_{A_1}, s_{A_2}}\right)^2 \tag{1}$$

where A_1 and A_2 represent vectors of SERS spectra and s_{A_1}, s_{A_2} their respective variances. The R value is a real valued number between 0 and +1, for which a value of zero would indicate no correlation between two SERS spectra, with increasingly strong correlations as R tends to +1. All calculations were performed in MATLAB 2007b (The Mathworks, Natick, MA, USA) using either built-in functions or scripts developed in-house.

For the iterative evolutionary optimisation, the non-dominated points from each generation were re-analysed in the subsequent evolutionary generation, and for the final generation of experiments, the non-dominated points from the full factorial landscape were also re-analysed. This allowed for the values of average peak intensities and R to be normalised across all of the experiments, and these values were used to plot the figures displayed in this article. This step was taken to ensure that any factors such as colloid ageing, fluctuation in laser fluence over time, and variability due to colloid heterogeneity did not bias the interpretation of the results.

Implementation of PESA-II for searching the full factorial landscape

To demonstrate the ability of an MOEA to find optimal SERS conditions rapidly, we used a full factorial landscape (as defined above) as an abstraction of the true landscape for testing purposes. PESA-II was selected as the MOEA in this study because it has proven to be effective in other applications within the field of analytical chemistry [31], and the implementation of this algorithm has been reported in detail elsewhere [35, 36]. Points within the solution space were selected at random to form the starting population, and the algorithm affected mutation by either randomly selecting alternatives, in the case of substrate and aggregating agent variables; or selecting proximal values

for substrate and aggregating agent concentrations. In addition, random searches for optimal SERS conditions were run by sampling points within the full factorial landscape, as a means of validating the performance of PESA-II and justifying the use of this method for a SERS optimisation.

For PESA-II, the mutation rate was fixed at either $1/L$ or $2/L$ (where L is the number of variables evaluated); which meant that there was either a 25% or 50% probability that each variable was mutated for each individual in the internal population. Some optimisation of the algorithm parameters was required, since it has been shown that for typical directed evolutionary studies that run over only a few generations, there is a substantial effect on algorithm performance [37]. PESA-II allows individuals to be selected from the external population (the non-dominated points) and placed into the internal population without modification (owing to the nature of bitwise mutation rates). In standard genetic algorithms this is encouraged because it promotes fixation; however, PESA-II utilises elitism through the external population and so this simply wastes evaluations. The source code was modified to prevent any individual in the internal population being identical to an individual in the external population (by repeating the process of chromosomal mutation). Selection from the external population was based on binary tournament selection from boxes with a grid size of 10; and crossover was not introduced in any of these initial evaluations. Comparison between algorithm performance was made using the hypervolume of solutions. Hypervolume (also known as the Lebesgue measure or S-metric) is a direct measure of the volume of the set of solutions (the grey shaded area in Fig. 1), and is considered to be the preferred method of comparison between solutions as it combines the properties of expansion (optimisation of each parameter) and spread across the front in a single scalar [38]. Hypervolume was calculated using code written in-house, which implements the unary hypervolume indicator as proposed in [39].

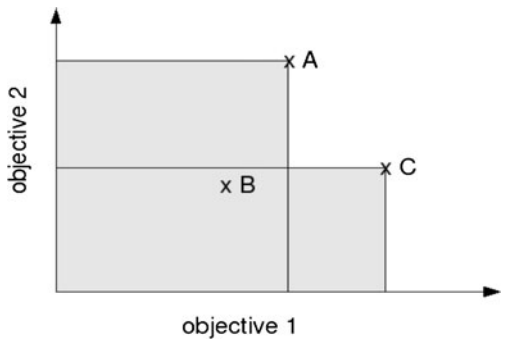


Fig. 1 Demonstrating the concept of Pareto optimality. Point A is superior to point B in terms of both objective 1 and objective 2 and so Point A dominates point B. Point A is superior to point C in terms of objective 2, but not in terms of objective 1, in this case it can be said that neither point dominates

Implementation of PESA-II for searching the real landscape

The parameters applied to PESA-II for optimisation of L-cysteine SERS beyond the constraints of the full factorial landscape were as follows. The selection process used to pick individuals from the external population remained the same. For each generation, an internal population of 20 individuals was used (dictated more by experimental convenience than tuning), with a mutation rate of $1/L$ for both the aggregating agent and colloid. Twenty points from the full factorial experiment were randomly selected to represent the starting population. Subsequently, the aggregating agent and colloid concentrations under this optimisation were treated as continuous values as opposed to pseudo-discrete, and therefore an alternative form of mutation strategy had to be employed. This involved mutating these variables at each stage of evolution with a probability of $2/L$. The proportion that these values were varied was determined by a Gaussian random number generator, with the mean of the distribution of values set to 10% of the maximum from the full factorial landscape. Finally, a low level of uniform crossover operator with a probability of 20% was also employed, in order to recombine features within the non-dominated set. Four generations of this optimisation procedure were performed, and an incremental improvement in the SERS response for L-cysteine was demonstrated at each step.

Results and discussion

In optimising the parameters within a SERS experiment it is not necessarily the case that the conditions optimal to signal enhancement are also optimal to reproducibility, and so there must be a trade-off between these two objectives. If one set of conditions induces greater signal enhancement and has equal or greater reproducibility than another set of conditions, then we can say that it dominates and in real terms is better (see Fig. 1). However, if one set of conditions induces greater signal enhancement but lower reproducibility than another set of conditions (or *vice versa*), we cannot say that either set of conditions is truly superior. The set of solutions within the entire search space (all combinations of components) that are not dominated by other solutions are termed the Pareto optimal front. In MOEAs, this concept can be used as a means of selecting non-dominated points from successive iterations of an optimisation process, that are used as kernels against which to apply evolutionary operators. This selection process is known as elitism, and allows for the exploration of the

Table 1 List of algorithm parameters screened

Parameter Set	Description[a]
A	Internal population size = 4, mutation rate = 1/*L*
B	Internal population size = 4, mutation rate = 2/*L*
C	Internal population size = 8, mutation rate = 1/*L*
D	Internal population size = 8, mutation rate = 2/*L*
E	Internal population size = 20, mutation rate = 1/*L*
F	Internal population size = 20, mutation rate = 2/*L*

[a] *L* is the number of variables evaluated

Pareto front by evolutionary methods using considerably fewer experiments than would be necessary in a full factorial study.

For the experiments reported herein, SERS conditions were evaluated against the two objective values of signal enhancement and reproducibility, since achieving accurate quantification with a high level of sensitivity is crucial to the application of SERS within quantitative biology. In the initial full factorial study 216 SERS conditions for the measurement of L-cysteine were assessed. The purpose of this preliminary study was to define a landscape where the optimal solutions were known; such that the MOEA could be tuned to achieve a rapid convergence to the optima, and its effectiveness compared against a simple random search.

Tuning of MOEA performance

The choice of operating parameters, such as mutation rate, population size and crossover type will greatly influence the performance of the MOEA, and can dramatically affect the number of costly experimental evaluations required. In most optimisations, these parameters are derived from evaluations on *in silico* problems (which in some way may mirror the properties of the true problem), metrics derived from limited sampling of the solution space, or simply an inherent understanding of the problem. In this instance, we can reveal a more accurate representation of the real problem through the full factorial landscape. This limited number of evaluations can be used as a surrogate for the true solution space, and screen MOEAs with various operating parameters to determine those that are most effective.

One of the critical measures in algorithm performance is whether it produces superior results than random search. This is not assured, particularly when the data are noisy and the number of evaluations limited. We have compared the performance of the PESA-II MOEA with different operating parameters (see Table 1), against a random search with increasing numbers of total evaluations (internal population × number of generations). The algorithms were ranked against each other based on *t* test comparisons of 100 independent evaluations. The analysis of algorithm performance on the full factorial landscape (see Table 2) indicates that with increasing numbers of evaluations the MOEAs consistently outperform random search, this is despite the algorithm being constrained to optimising between points on the full factorial landscape. This is very encouraging since it would be expected that when allowed to interpolate through the entire search space, MOEA performance will be enhanced still further. In addition, the results indicate that the algorithm generally performs better with lower mutation rates, and there is also a trend demonstrating that algorithms with smaller populations (evaluated by increasing the number of generations) produce superior results to those with larger populations.

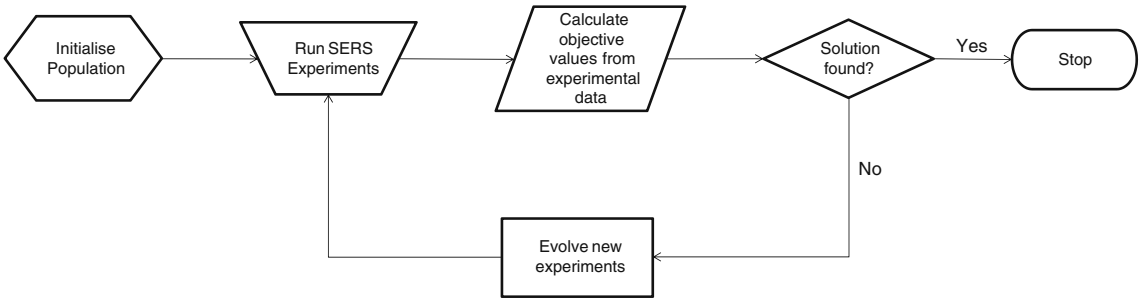
Optimising SERS for L-cysteine

By performing the full factorial study we have shown that using a MOEA can greatly improve the speed with which SERS can be optimised for an analyte. In reality, this technique would be applied without partially mapping the search space; and an iterative process of experiment, evaluation and evolution (as depicted in Scheme 1) would be performed. Therefore, we used this approach in a further optimisation of L-cysteine SERS, to demonstrate that the MOEA could be effective across a real, unconstrained search space. A total of four generations, consisting of 20 experiments in each were performed, and a considerable

Table 2 Performance of MOEAs with different parameter settings (listed in Table 1) and random search on the full factorial landscape

40 evaluations		80 evaluations		120 evaluations		160 evaluations		200 evaluations	
Para.	Rank	Para.	Rank	Para.	Rank	Para.	Rank	Para.	Rank
A	1(-)	A	1(-)	A	1(-)	A	1(-)	C	1
B	1(-)	C	1(-)	C	1(-)	C	1(-)	A	2(-)
C	1(-)	D	1(-)	B	3(-)	E	1(-)	B	2(-)
Rand.	1(-)	E	1(-)	E	3(-)	B	4(-)	D	2(-)
D	5	B	5	D	5(-)	D	4(-)	E	2(-)
E	6(-)	F	6	F	5(-)	F	4(-)	F	2(-)
F	6(-)	Rand.	7	Rand.	5(-)	Rand.	7	Rand.	7

Rand. indicates a random experiment



Scheme 1 Flow diagram depicting the multiobjective evolutionary optimisation process

improvement in spectral intensity and reproducibility was noted. In Fig. 2, the search landscape is shown as a scatterplot of the objective function values observed for each set of SERS conditions tested. As a comparison for performance, the non-dominated points from the full factorial search space are shown as black coloured circles. It can be seen clearly that by the second generation (red circles) little improvement is observed in the SERS response for L-cysteine. With an MOEA approach, it might be anticipated that the optimisation would progress more swiftly; and it is possible that performance could have been hindered by constraints in the starting population of experiments, since we used randomly selected points from the full factorial experiments as seeds for the evolution of the first generation. However, by the third generation (yellow symbols), there is a clear improvement in both SERS enhancement and reproducibility, and with the fourth generation of experiments (blue symbols) even greater progress is seen. Overall, enhancement is improved by a factor of 1.6; however, there appears to be a trade-off in achieving greater enhancement at the expense of reduced spectral reproducibility.

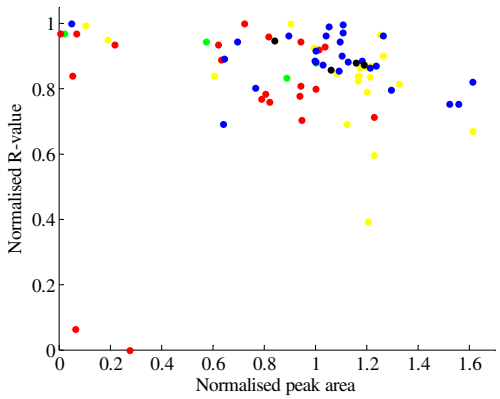


Fig. 2 A biplot of the SERS enhancement *vs.* reproducibility. *Black circles* represent non-dominated points from the FF study; *green circles* represent non-dominated points from the first generation; *red circles* represent all points from the second generation; *yellow circles* represent all points from the third generation and blue circles represent all points from the fourth generation

With MOEA methods, the search landscape is partially mapped as a consequence of the optimisation process. Therefore, an opportunity arises to depict the landscape in terms of the variables that have been optimised (see Fig. 3), and to try to observe patterns within these results. In this figure, aggregating agents are encoded using six different symbols, the three colloids employed in the study are indicated by coloured dots overlaid onto these, the proportion of colloid and concentration of aggregating agent in a mixture are given on the abscissa and ordinate axes respectively, the same colour coding as for Fig. 2 is used to represent the generation at which each mixture was assessed; and finally, the size of each symbol is scaled to the product of SERS enhancement and reproducibility metrics, with larger symbols representing better SERS conditions. Immediately, it can be seen that a trend is observed with respect to the balance between the proportion of colloidal

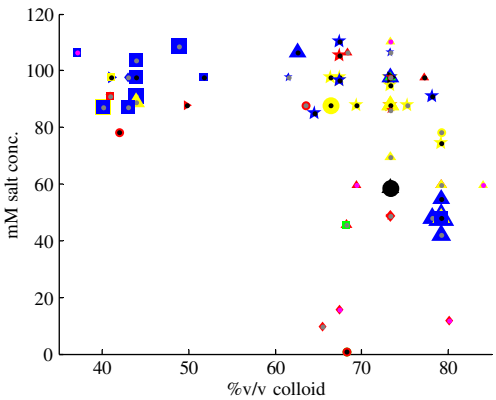


Fig. 3 A biplot showing the distribution of solutions w.r.t. the % *v/v* composition of colloid *vs.* concentration of aggregating agent, where the size of each symbol is scaled to the ratio of enhancement to reproducibility. *Black symbols* represent non-dominated points from the FF study; green symbols represent non-dominated points from the first generation; *red symbols* represent all points from the second generation; *yellow symbols* represent all points from the third generation and blue symbols represent all points from the fourth generation. Aggregating agents are indicated as follows, *filled circle* K_2SO_4, *filled upright triangle* $NaNO_3$, *filled square* KCl, *filled star* KNO_3, *filled diamond* Na_2SO_4, *filled right-pointing triangle* NaCl. *Superposition of coloured dots* indicate the colloid used; *black* hydroxylamine, *grey* citrate; *magenta* borohydride

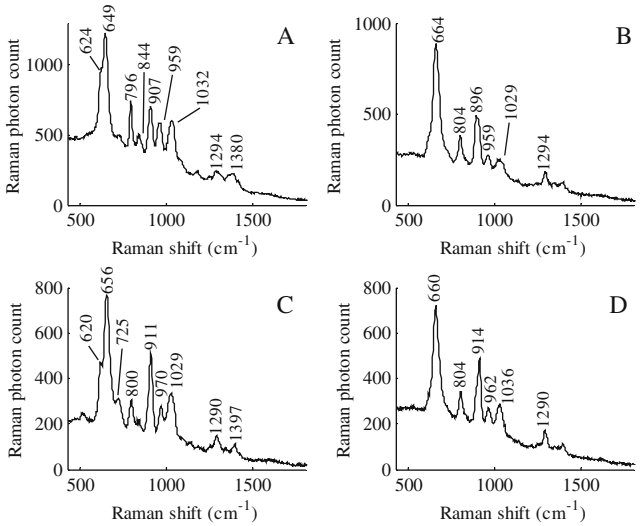


Fig. 4 Example SERS spectra of L-cysteine. **a** Citrate-reduced silver colloid with $NaNO_3$ (79.2% *v/v* and 47.9 mM), representing the strongest enhancement obtained from the EO; **b** citrate-reduced silver colloid with KCl (44.0% *v/v* and 90.9 mM), representing the most reproducible SERS spectra obtained from the EO; **c** hydroxylamine reduced silver colloid with $NaNO_3$ (73.3% *v/v* and 58.6 mM), representing the strongest enhancement obtained from the FF experiment; **d** citrate-reduced silver colloid with NaCl (73.3% *v/v* and 97.7 mM), representing the most reproducible SERS spectra obtained from the FF experiment

substrate in a mixture, and the concentration of aggregating agent. In general, where $NaNO_3$ is selected, a larger proportion of colloid with a lower concentration of the aggregating agent is favoured. Conversely, where KCl is used for aggregation, this is usually selected in higher concentration combined with a lower concentration of colloid. In both cases, either citrate or hydroxylamine reduced colloidal silver are favoured; in general, the borohydride-reduced silver sol did not yield a SERS signal, and in this case could have been due to the particular method we chose to prepare the hydrosol which did not leverage the stabilising effects of trisodium citrate. Where both a large relative amount of colloid and a high concentration of aggregating agent are used in the mixture, it is with KNO_3 that the best SERS response is typically achieved.

The conditions that achieved the greatest enhancement over all of the experiments used were 79.2% *v/v* of citrate-reduced silver colloid, with 47.9 mM of $NaNO_3$ aggregating agent. The most reproducible spectra were achieved when citrate-reduced silver colloid was mixed with KCl (44.0% *v/v* and 90.9 mM). Finally, conditions that achieved a combination of good reproducibility and enhancement included, citrate-reduced silver with KCl (48.9% *v/v* and 108.5 mM) and hydroxylamine reduced silver with $NaNO_3$ (73.3% *v/v* and 97.7 mM).

SERS spectra of L-cysteine

A wide range of SERS spectral profile was observed for L-cysteine under the different experimental conditions applied throughout the course of this study. Figure 4 gives several examples of these, where the conditions used gave the most favourable results in terms of the enhancement and reproducibility objective functions, in both the full factorial experiment and evolutionary optimisation. There are close similarities between Fig. 4a, b and d, which result from the use of citrate-reduced silver colloid with $NaNO_3$, KCl and NaCl aggregating agents, respectively. The spectrum displayed in Fig. 4c is observed when $NaNO_3$ and the hydroxylamine reduced silver sol are used. Where $NaNO_3$ is used as the aggregating agent, an additional shoulder at ca. 620 cm^{-1} is apparent. For the citrate-reduced silver examples, where the larger anion (NO_3^-) is used for aggregation as opposed to the smaller anion (Cl^-), subtle peak shifts with very obvious changes in relative intensities of SERS bands can be seen across the spectrum are also apparent.

In Table 3, tentative band assignments are provided for each example spectrum shown in Fig. 4. Commonly

Table 3 Tentative SERS band assignments observed for L-cysteine, taken from the optimal experiments observed for the full factorial study and MOEA optimisation [40, 41]

Tentative assignments	Ag citrate and $NaNO_3$	Ag citrate and KCl	Ag hydroxylamine and $NaNO_3$	Ag citrate and NaCl
$\delta(CO_2^-)$	624	–	620	–
ν(CS)	649	664	656	660
$\rho_g(CH_2)$	–	–	725	–
$\delta(CO_2)$	796	804	800	804
$w(CO_2)$	844	–	–	–
ν(CC)	907	896	911	914
δ(SH)	959	959	970	962
$\rho(NH_3)$	1,032	1,029	1,029	1,036
$w(CH_2)$	1,294	1,294	1,290	1,290
$\nu_s(CO_2)$	1,380	1,397	–	–

ν stretch, δ deformation, *w* wag, ρ rock, *s* symmetric, *g* gauche conformation

observed bands include the strong CS stretching vibration at *ca*. 649–664 cm^{-1}, the CO_2 deformation at 796–804 cm^{-1}, and a range of SH, NH_3 and further CO_2 vibrations. The additional shoulder at ca. 620 cm^{-1} in the spectra obtained using $NaNO_3$ arises from the CO_2^- deformation [40], which could indicate a predominance of zwitterionic L-cysteine where $NaNO_3$ is used for aggregation. The CH_2 rocking mode is observed where hydroxylamine reduced silver is mixed with $NaNO_3$, and this correlates with bands calculated for the gauche rotational conformation of L-cysteine [41]. Finally, additional CO_2 vibrations (the wag at *ca*. 844 cm^{-1} and rocking mode at *ca*. 1,380 cm^{-1}) are observed where the SERS response is more intense for citrate-reduced silver with either NaCl or $NaNO_3$ aggregating agents. At this stage, we are unable to offer suggestions as to why these differences are observed, since aggregating agent mediated surface–analyte interactions are not well understood, and these would undoubtedly impact upon the L-cysteine conformation and therefore SERS profile under different conditions. However, we intend to work towards generating larger datasets using MOEA methods on a range of analytes that may aid to a better theoretical understanding of the conditions which give rise to these variations.

Conclusions

We have shown that it is possible to optimise the conditions for SERS of L-cysteine rapidly and robustly using PESA-II, an algorithm that performs better than a simple random search. Within our limited experimental landscape we have shown that SERS of L-cysteine can be optimised for both reproducibility and enhancement with a number of preparations including citrate-reduced silver with KCl (48.9% *v*/*v* and 108.5 mM) and hydroxylamine reduced silver with $NaNO_3$ (73.3% *v*/*v* and 97.7 mM). This directed method requires only a small number of experiments to be performed in order to achieve improvements in spectral reproducibility and signal intensity, and provides opportunities to explore large combinatorial search landscapes. As a computational approach, it will be possible to use this procedure, in conjunction with robot scientist platforms, to optimise SERS for a large library of analytes, and introduce this generic technology as a viable platform for metabolomic and other SERS studies.

Acknowledgment RMJ, WR & RG would like to thank the BBSRC for funding. EB and RG are indebted to EPSRC and the RSC Analytical Trust Fund for funding, and NY thanks these bodies for her studentship.

References

1. Fleischmann M, Hendra PJ, McQuillan AJ (1974) Chem Phys Lett 26:163–166
2. Bao X, Wilcock IC, Muhler M, Ertl G (1994) 72
3. Otto A, Mrozek I, Grabhorn H, Akemann W (1992) Journal of Physics-Condensed Matter 4:1143–1212
4. Graham D, Goodacre R (2008) Chem Soc Rev 37:883–884
5. Kneipp J, Kneipp H, Rajaduraj A, Redmond RW, Kneipp K (2009) J Raman Spectrosc 40:1–5
6. Scaffidi JP, Gregas MK, Seewaldt V, Vo-Dinh T (2009) Anal Bioanal Chem 393:1135–1141
7. Stokes RJ, McKenzie F, McFarlane E, Ricketts A, Tetley L, Faulds K, Alexander J, Graham D (2009) Analyst 134:170–175
8. Tang HW, Yang XBB, Kirkham J, Smith DA (2008) Appl Spectrosc 62:1060–1069
9. Faulds K, Jarvis RM, Smith WE, Graham D, Goodacre R (2008) Analyst 133:1505–1512
10. Graham D, Faulds K (2008) Chem Soc Rev 37:1042–1051
11. Mahajan S, Richardson J, Brown T, Bartlett PN (2008) J Am Chem Soc 130:15589–15601
12. Douglas P, McCarney KM, Graham D, Smith WE (2007) Analyst 132:865–867
13. Fabris L, Dante M, Nguyen TQ, Tok JBH, Bazan GC (2008) Adv Funct Mater 18:2518–2525
14. Lin CC, Yang YM, Chen YF, Yang TS, Chang HC (2008) Biosens Bioelectron 24:178–183
15. Reyes-Goddard JM, Barr H, Stone N (2008) Photodiagnosis and Photodynamic Therapy 5:42–49
16. Jarvis RM, Goodacre R (2004) Anal Chem 76:40–47
17. Jarvis RM, Goodacre R (2008) Chem Soc Rev 37:931–936
18. Shadi I, Chowdhry B, Snowden M, Withnall R (2004) J Raman Spectrosc 35:800–807
19. Sabur A, Gogotsi M (2008) J Raman Spectrosc 39:61–67
20. Jackson JB, Halas NJ (2004) Proc Natl Acad Sci USA 101:17930–17935
21. Abdali S (2007) J Phys, Condens Matter 19:5205
22. Yaffe NR, Blanch EW, Corfu Greece (2007) Elsevier 196–201
23. Jarvis RM, Johnson HE, Olembe E, Panneerselvam A, Malik MA, Azad M, O'Brien P, Goodacre R (2008) Analyst 133:1449–1452
24. Raamsdonk LM, Teusink B, Broadhurst D, Zhang NS, Hayes A, Walsh MC, Berden JA, Brindle KM, Kell DB, Rowland JJ, Westerhoff HV, van Dam K, Oliver SG (2001) Nat Biotechnol 19:45–50
25. Ellis DI, Dunn WB, Griffin JL, Allwood JW, Goodacre R (2007) Pharmacogenomics 8:1243–1266
26. Goodacre R, Vaidyanathan S, Dunn WB, Harrigan GG, Kell DB (2004) Trends Biotechnol 22:245–252
27. Cabalin LM, Ruperez A, Laserna JJ (1996) Anal Chim Acta 318:203–210
28. Kennedy BJ, Milofsky R, Carron KT (1997) Anal Chem 69:4708–4715
29. Coello CAC, Veldhuizen DAV, Lamont GB (2002) Evolutionary algorithms for solving multi-objective problems. Kluwer, Boston
30. O'Hagan S, Dunn WB, Broadhurst D, Williams R, Ashworth JA, Cameron M, Knowles J, Kell DB (2007) Anal Chem 79:464–476
31. O'Hagan S, Dunn WB, Brown M, Knowles JD, Kell DB (2005) Anal Chem 77:290–303
32. Lee PC, Meisel D (1982) J Phys Chem 86:3391–3395
33. Creighton JA, Blatchford CG, Albrecht MG (1979) J Chem Soc, Faraday Trans 2(75):790–798

34. Leopold N, Lendl B (2003) J Phys Chem B 107:5723–5727
35. Come DW, Jerram NR, Knowles JD (2001), Proc. Genetic and Evolutionary Computation Conf. (GECCO-2001), Morgan Kaufmann Publishers 283–290
36. Corne DW, Knowles JD, Oates MJ (2000) Springer LNCS 1917/2000 839–848.
37. Wedge D, Rowe W, Kell D, Knowles J (2009) J Theor Biol 257:131–141
38. Knowles J, Corne D (2002) Cec'02: Proceedings of the 2002 Congress on Evolutionary Computation. 1(2) 711–716
39. Zitzler E, Thiele L (1998) Amsterdam, the Netherlands, September 27th-30th. Springer 292–301
40. Gargaro A, Barron L, Hecht L (1993) J Raman Spectrosc 24:91–96
41. Pawlukojc A, Leciejewicz J, Ramirez-Cuesta AJ, Nowicka-Scheibe J (2005) Spectrochim Acta A Mol Biomol Spectrosc 61:2474–2481