Int J Thermophys (2018) 39:79
https://doi.org/10.1007/s10765-018-2391-2

ATPC 2016

Enhanced Quality Control in Pharmaceutical Applications by Combining Raman Spectroscopy and Machine Learning Techniques

J. C. Martinez[1] · J. R. Guzmán-Sepúlveda[2] · G. R. Bolañoz Evia[1] · T. Córdova[4] · R. Guzmán-Cabrera[3]

Received: 30 November 2016 / Accepted: 18 April 2018

Abstract In this work, we applied machine learning techniques to Raman spectra for the characterization and classification of manufactured pharmaceutical products. Our measurements were taken with commercial equipment, for accurate assessment of variations with respect to one calibrated control sample. Unlike the typical use of Raman spectroscopy in pharmaceutical applications, in our approach the principal components of the Raman spectrum are used concurrently as attributes in machine learning algorithms. This permits an efficient comparison and classification of the spectra measured from the samples under study. This also allows for accurate quality control as all relevant spectral components are considered simultaneously. We demonstrate our approach with respect to the specific case of acetaminophen, which is one of the most widely used analgesics in the market. In the experiments, commercial samples from thirteen different laboratories were analyzed and compared against a control sample. The raw data were analyzed based on an arithmetic difference between the nominal active substance and the measured values in each commercial sample. The principal component analysis was applied to the data for quantitative verification (i.e., without considering the actual concentration of the active substance) of the difference

Selected paper from Asian Thermophysical Properties Conference.

✉ R. Guzmán-Cabrera
guzmanc@ugto.mx

1 Instituto Politécnico Nacional-Unidad Profesional Interdisciplinaria de Ingeniería Campus Guanajuato, 36275 Silao de la Victoria, Guanajuato, México

2 CREOL, The College of Optics and Photonics, University of Central Florida, Orlando, FL 32816, USA

3 Universidad de Guanajuato-División de Ingenierías, 36885 Salamanca, Guanajuato, México

4 Universidad de Guanajuato-División de Ciencias e Ingenierías, 37150 León, Guanajuato, México

Published online: 07 May 2018

in the calibrated sample. Our results show that by following this approach adulterations in pharmaceutical compositions can be clearly identified and accurately quantified.

Keywords Acetaminophen · Machine learning · Polymorph detection · Principal components analysis · Raman spectroscopy

1 Introduction

The normativity for the quality control of medications, despite having some differences from country to country, is in accordance with international legislations that facilitate and ensure trade of safe medication for patients and their medical treatment [1]. In Mexico, for instance, there exist norms such as NOM-001-SSA1-2010 and NOM-059-SSA1-2015 that provide a legal framework for the assessment of the medication, an adequate manufacturing process and methods that are to be followed for quality control [2].

The trade of counterfeit drugs, which takes place predominantly in Asian countries such as China, India, and Russia, has increased due to the availability of cheaper options and the willingness to get treatment without prescriptions from a wide variety of drugs, both branded and generic [3, 4]. According to the World Health Organization, 25 % to 50 % of cases in developing countries have shown that undeclared drugs that are purchased are in many cases counterfeit [5, 6].

In general, the main criteria of evaluating the quality of medications are purity, efficacy, uniformity of pharmaceutical formula, bio-disposability, and stability [3]. Medication with low or deficient quality may, in the best case, fail to achieve the desired effect. In many situations, however, low-quality medications can be severely harmful for the patient and may sometimes result in new medical conditions. Unfortunately, such low-quality medications are commonly encountered in today's pharmaceutical market.

Adequate quality control is complicated due to several factors such as complex manufacturing process, packaging, storage and aging. These are some of the factors among others that might adversely influence the composition of the final product [4]. The problem becomes even more complicated with the availability of polymorphs, i.e., when certain substances exist in different natural structures, and with different properties owing to the difference in their physical origins. For instance, polymorphs can result from differences in manufacturing, processing, or due to improper storage, and aging [7]. This is critical because active ingredients of the formula may transition from a stable, efficacious form to unwanted structural analogues.

Although identical in chemical composition, polymorphs differ in chemical and physical properties including bioavailability, solubility, dissolution rate, chemical and physical stability, melting point, filterability, density, flow rate [8]. Solubility, for instance, is of importance in pharmaceutics as it can affect drug efficacy, bioavailability, and safety. In this context, there is a need for techniques of advanced analysis and characterization that could provide reliable means of ascertaining the quality of medications, including effective identification and discrimination of polymorphic forms for adequate quality control.

One well-established technique for identification of polymorphs is infrared spectroscopy and X-ray diffraction [9]. Other techniques such as thermo-microscopy [10], a variant of polarized light microscopy, have proved extremely powerful and significantly simpler to implement. Moreover, the mathematical description obtained from combining thermo-microscopy together with differential scanning calorimetry and differential thermal analysis [11] is useful for understanding the thermodynamic nature of polymorphism [12, 13].

Raman spectroscopy (RS) is an analytical technique based on inelastic scattering of monochromatic light [14–16]. This technique has become important in recent years due to its variety of applications in areas such as mineralogy, forensics, biology, and medicine [14]. It is also used, for instance, in the characterization of drugs, bone structures, and organic pesticides [17]. In general, RS and its various applications have proved effective in the detection of polymorphs and the characterization of their physical properties [18–22]. Use of RS has become very widely popular in pharmaceutical applications [23–25]. It plays an important role in quality control and manufacturing, including assessment of pharmacological formula for a product and for comparing it to certified spectroscopic calibration curves provided by certified agencies. Typical spectroscopic applications relate to supervision and control of large-scale manufacturing, especially in order to outline the distribution of active pharmaceutical ingredients and excipients in the different stages of a development cycle [26, 27]. This technique offers unparalleled discrimination of materials, both for analysis of liquid and solid samples. In addition, it is particularly suitable for combination with other analytical techniques since it provides a non-destructive assessment with minor requirements of sample preparation [28, 29]. For instance, RS has been applied to study various crystal forms and solvates of ampicillin and griseofulvin [30]. Some other examples include popular drugs like acetaminophen [31–33].

It has been recognized that, among the number of different techniques of characterization mentioned above, RS allows for fast assessment as it requires only a few seconds for each measurement cycle. Nevertheless, this advantage typically vanishes in practical applications due to the low reliability of data processing and the absence of an accurate system of classification, with the consequent need to perform backup measurements for the same instances by using more than one technique [34].

In this work, we propose an improvement on the typical use of RS, i.e., the identification of biochemical differences among supposedly similar commercially available medications. Machine learning techniques combined with the Raman spectra method may be used to develop an approach for characterization and classification of manufactured pharmaceutical products. Measurements were taken with commercial equipment, for accurate assessment of variations with respect to a calibrated control sample. In our approach, the principal components of the Raman spectrum are used together as attributes within the machine learning algorithms. This permits comparison and classification of the spectra obtained from the studied samples. This permits accurate quality control since all the relevant spectral components are considered simultaneously.

We experimentally demonstrate our approach for the specific case of acetaminophen, which is the most popular analgesic and antipyretic available in the market. It has been popularized in combating diseases such as influenza and common

cold. Unfortunately, this medication is highly unstable and can be easily hydrolyzed in aqueous solution [30]. During manufacture, this drug may experience alterations that, if undetected, may result in severe adulterations.

In our experiments, commercial samples from thirteen different laboratories were analyzed and compared against a control sample. The raw data were analyzed by means of the arithmetic differences between the pure active substance and samples from each different laboratory. The principal component analysis (PCA) was applied to the set of experimental measurements to qualitatively evaluate the differences in each pharmaceutical formulation.

2 Methodology

In our experiments, we used a commercial Thermo Scientific DXR Raman spectroscopy equipment. This equipment has an excitation laser of 780 nm and 24 mW, and a built-in microscope objective (50X). Samples from thirteen different pharmaceutical laboratories were analyzed. The samples consisted of commercially available tablets (pills) from each of these manufacturers. A sample of acetaminophen provided by the Mexican National Health Secretary was used as control. As part of the procedure for quality control in the country, this same sample is the one provided to all pharmaceutical manufacturers for comparing their products with a certified reference. For each of the laboratories, twenty-five different Raman spectra were collected from twenty-five different tablets, i.e., twenty-five pills per laboratory, one Raman spectrum per tablet, and then each one measured at a random position on the pill. Each of these spectra was measured in the range from 100 cm^{-1} to 1800 cm^{-1} with resolution of 0.965 cm^{-1}. The dynamic range of the measurement was calibrated by following standard approaches using least-squared methods for the elimination of any background noise, and by normalizing them with respect to the peak of maximum amplitude [35, 36]. The PCA was applied to all the Raman spectra recorded so that a qualitative measure of similarity with respect to the control could be obtained. For PCA, we considered the whole spectral range of the instrument (100–1800 cm^{-1}) and the analysis was carried out using all 1764 data points recorded for each Raman spectrum.

2.1 Raman Spectroscopy Experiments

Acetaminophen shows three polymorphic forms [7]. Forms I and II are known to reveal a packing polymorphism in which molecular conformations are the same, but crystal packing is different. The commercial form of acetaminophen polymorph is form I. Acetaminophen form II is slightly more soluble than form I and suitable for direct compression but is less stable and susceptible to transformation into form I during compression and storage [6]. Form II can be obtained by crystallizing solids in benzyl alcohol at high temperature, by cooling the melt, and by adding carboxylic acid additives to it, or by using the evaporation method. Form III is known to be highly unstable, and it is obtained by cooling the melt. It undergoes solid-state polymorphic transformation to form II within hours [6]. Recently, it was shown that form III can

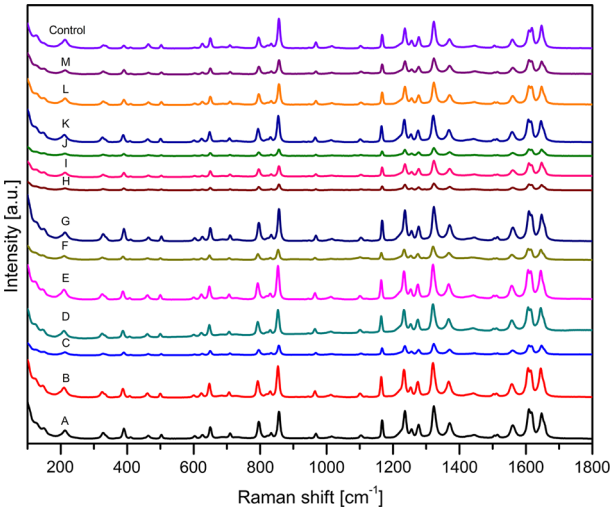

Fig. 1 Averaged Raman spectra measured in the experiments. The curves represent the average of 25 independent measurements as described in the *Methodology*. Different laboratories are indicated with letters from A to M. The control (active substance) is shown as the top line. The curves were shifted vertically for visualization purposes

easily crystallize in nano-confined structures with a pore size ranging from 10 nm to 103 nm [18].

2.2 Preliminary Comparative Analysis

The variations in the Raman spectra both within the subsets of the same laboratory as well as between different manufacturers are directly determined by mechanisms of quality control used during manufacture. As shown in Fig. 1, variations with respect to the control can be significant in some cases.

These variations represent differences in the relative fractions of the real active substances, i.e., differences in the amplitude of the peaks (at the "right" locations) and differences in the actual composition of the formula, i.e., in new peaks in the spectrum (at other frequencies). In any case, it is important to quantify these differences as they may have an impact on patient recovery. Even from the simplest measures of variability, i.e., point-by-point differences, we could observe significant variations with respect to the control (data not shown).

To elaborate more detailed comparisons between different laboratories, a regression analysis to find the best subsets was run with each of the 25 Raman spectra (of the active substance) as response variables and two randomly selected spectra (per laboratory) as prediction variables. In each run, five of the best spectra were selected among the total of 26 included. The coefficient of determination, R^2, was then calculated to quantify the variability on the dependent variable as it would be explained by the variations on the independent variables. Table 1 shows the value of R^2 obtained for 125 (out of the total 325) randomly selected spectra. In order to track the particular spectrum involved

in the comparison, the following code was used: In the column "Variable," the first letter indicates the laboratory (labeled from A to M), while the number following this letter indicates the index of the measurement for that particular sample (from 1 to 25). Results are shown in Table 1.

Based on this R^2 analysis, the presence of a laboratory spectrum on the best subset indicates that the spectra of this laboratory sample have little differences when compared to the control spectra. Thus, high percentages indicate no significant difference between the laboratory samples and control for several comparisons and therefore the presence of this laboratory sample within the best subsets is not random.

The correlation analysis shown in Table 2 confirms that laboratory L is closest to the control as in 100 % of times at least one sample of laboratory spectra L was found on the best subsets. In each run of a regression model, two samples of each laboratory were included. In the case of laboratory L, 88 % of the samples were placed on the best subsets. Subsequent places are occupied by laboratories J, G, C, and A, respectively.

2.3 Principal Component Analysis

A more accurate quantification of the spectral variations between different laboratories and the control was performed by means of a PCA. The principal components refer to the peaks of greater intensity in the Raman spectra such as contain information on the absence or presence of spectral components. The first three principal components of the recorded Raman spectra were plotted in order to visualize the characteristic dispersion of the samples from different laboratories, a phenomenon inherently associated with compositional homogeneity.

Figure 2 shows the features of the dispersion group corresponding to each laboratory, including control. Figure 2a–c shows the comparisons between different pairs of principal components, respectively, while Fig. 2d shows the three principal components simultaneously. For clarity, the control is plotted in color (blue), while all the laboratories are plotted in black markers.

It can be observed that laboratories A, C, G, I, L, and M are closer to the control. It can also be noticed that samples from these laboratories have little dispersion, i.e., high composition homogeneity. Conversely, laboratories J, H, F, and K are both more dispersed and more different from the control. Interestingly, laboratories D, E, and B significantly differ from the control but show good compositional homogeneity.

2.4 Data Classification

The goal of the present work is not just to present differences in quantification but also provide a classification for them and to measure their potential impact. In doing so, we followed the standard classification of defects in quality as indicated in the Spanish Health Society for drug management in which the substance under study is assigned a category from 1 to 3, depending upon how it differs from the control. In this classification, category 1 represents the largest difference with respect to the control and, therefore, contains samples that are potentially spurious.

Table 1 Comparisons between active substance (control) and the best subsets Raman spectra

Active substance									
CO1		CO2		CO3		CO4		CO5	
R^2	Variable	R^2	Variable	R^2	Variable	R^2	Variable	R^2	Variable
95.7	L19	95.6	L14	98.1	L23	97.2	G1	99.1	L18
94.8	G19	95.1	H25	96.5	T12	97.1	L13	97.2	A6
94.6	C19	94.7	L25	95.6	A12	96.7	J1	96.7	J18
94.2	L7	94.1	J25	95.1	C23	95.6	J13	95.7	G6
93.9	J7	93.7	G14	94.5	L12	95.5	L11	95.5	G18
CO6		CO7		CO8		CO9		CO10	
94.6	C11	94.9	J4	98	L23	96.5	L20	96.5	L9
94	J3	93.6	C4	95.4	C23	95.4	A7	96.1	L8
93	L3	93.3	L4	94.9	L21	95.1	J20	95.9	G9
92.5	C3	93.1	A4	93.7	A21	94.7	L7	95.6	J9
91.9	L11	93	H17	92.8	M23	94.5	J7	95.3	J8
CO11		CO12		CO13		CO14		CO15	
96.5	L15	98	L13	96.3	L14	98	L10	96.9	G1
94.4	K2	94.4	L21	93.9	M22	93.6	C16	95.2	L24
93.7	A2	93.9	J13	93.7	L22	93.2	H10	95.0	J1
93.6	G2	93.3	A13	92.7	C22	91.7	L16	94.5	L1
92.9	L2	92.8	G13	92.6	A22	91.4	I16	94.5	G24

Table 1 continued

CO16		CO17		CO18		CO19		CO20	
94.9	H25	98.4	L18	94.7	J3	96.5	L23	95.4	K2
94.0	C11	95.1	L21	93.9	A7	95.9	C23	94.3	G22
93.6	L25	94.9	H18	93.5	L7	95.7	L5	94.2	A2
93.3	J25	93.7	G18	93.3	J7	94.1	J5	93.8	M22
93.2	L11	93.3	J21	93.3	L3	92.9	M5	93.4	L2
CO21		CO22		CO23		CO24		CO25	
98.0	L19	94.7	L8	98.1	L23	97.5	L5	96.2	G1
95.7	G19	94.3	J8	97.2	C23	95.3	K2	95.4	G9

Table 2 Significance on best subsets regression models

Laboratory	A	B	C	D	F	I	G	H	I	J	K	L	M
Percentage of presence on best subsets	44	0	44	0	0	0	48	32	4	72	20	100	16
Percentage of samples on best subsets	22	0	26	0	0	0	32	16	2	44	10	88	8

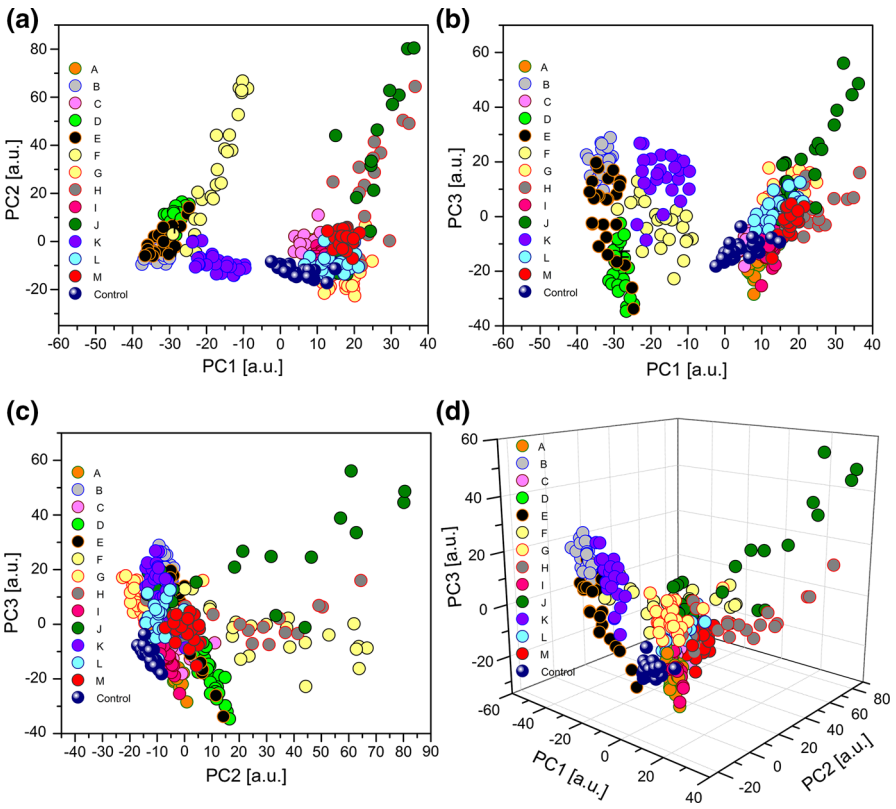

Fig. 2 (a) PCA of the acetaminophen spectra. Scatter plots of the scores from the first two principal components (PC 1 vs PC 2) for the different laboratories (thirteen groups: A-M) versus the control group (active substance: blue dots). (b) Plots of the first principal component (PC 1) versus the three principal component (PC 3) for the different laboratories (thirteen groups: A-M) versus the control group (active substance: blue dots). (c) Plots of the first principal component (PC 3) versus the three principal components (PC 1) for the different laboratories (thirteen groups: A-M) versus the control group (active substance: blue dots). (d) Scatter plots of the scores from the first three principal components (PC 1 vs PC 2 vs PC 3) for the different laboratories (thirteen groups: A-M) versus the control group (active substance: blue dots) (Color figure online)

The data classification was carried out automatically by means of a Naïve Bayes algorithm in which the attributes of the classification were considered as the first principal components of the Raman spectra. As mentioned above, the so-called principal components refer to the peaks of greater intensity in the Raman spectra which contain information on the absence or presence of spectral components. The control spectra (25 spectra from the control sample) were used as the training set.

As part of this classification approach, a two-factor experiment was designed with the two factors being the laboratories and the Raman shift correspondingly. The first factor has fourteen levels, each level corresponding to one of the laboratory samples used in the experiment. Raman shifts have four levels. Levels 1, 2, and 3 correspond to Raman shifts in the ranges of 300–550 cm^{-1}, 75–900 cm^{-1}, 1100–1500 cm^{-1}, and

Table 3 Performance of the classification approach proposed

Class	Variation (%) (with respect to control)	Precision (%)	Recall (%)
3	0–5	98	97
2	5–15	96	98
1	>15	96	97

level 4 to 1500–1700 cm^{-1}, respectively. The design was made with four runs. For each run, a Raman shift was randomly selected. The response variable is the median intensity of the samples of each laboratory on the randomly selected Raman shift.

Classifiers can be developed by using several techniques such as neural networks, logistic regression, heuristics, decision trees, and Bayesian methods. Because of the various implementations and their sensitivity to the particular set on which they are applied, a general quantification of the performance is typically pursued in terms of the following parameters: true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN). By counting each of these occurrences, final measures of performance such as precision, recall, or the so-called f1 score (combination of precision and recall) can be calculated. In our case, the metrics used were that of precision and recall defined as the number of correct classifications based on the number of attempts made, i.e., $P = \mathrm{TP}/(\mathrm{TP}+\mathrm{FP})$, and the number of correct classifications based on the number of total possible attempts $R = \mathrm{TP}/(\mathrm{TP}+\mathrm{FN})$, respectively.

All spectra from laboratories (325 in total) were used as test sets in the automatic classification systems; the results of the performance metrics obtained are shown in Table 3.

The final results of the proposed characterization and classification system show that in all cases the samples can be categorized with precision of at least 96 %. Thus, the confidence in the outcome of the analysis is high.

3 Conclusions

Unlike the typical use of Raman spectroscopy in pharmaceutical applications, in our approach the principal components of the Raman spectrum are used concurrently as attributes in the machine learning algorithms. This is used for comparison and classification of the spectra measured from the studied samples. This allows performance of an accurate quality control as all relevant spectral components are considered simultaneously. The main goal of our work is to show that PCA applied on Raman spectra can be a suitable tool for clearly detecting quantitative differences between different pharmaceutical products. We experimentally demonstrate our approach for the specific case of acetaminophen, one of the most broadly used analgesics. Our results show that differences with respect to a control sample can be accurately identified and quantified, and that their potential impact can be classified with an accuracy of at least 96 %. In principle, this type of analysis can be extended to other medications or pharmaceutical products, and quality settings can be defined according to particular rules. Future

work may relate to the inclusion of this classificatory procedure in mobile spectroscopic devices, in order to perform *in situ* quality control. In Mexico, this inspection is necessary as significant variations are revealed in the analyzed commercial samples and in their difference from prescribed controls.

References

1. Organización Mundial de la Salud, La falsificación de medicamentos: una amenaza creciente. OMS **88**, 241–320 (2010)
2. http://www.cofepris.gob.mx/MJ/Páginas/Normas-Oficiales-Mexicanas.aspx
3. C.S. Gautam, A. Utreja, G.L. Singal, Spurious and counterfeit drugs: a growing industry in the developing world. Postgrad. Med. J. **85**, 251–256 (2009)
4. S. Nitin, S. Tanushree, Generic drug industry in India: the counterfeit spin. J. Intellect. Prop. Rights **14**, 236–240 (2009)
5. K. Theodore, K. Iosif, I.R. Petros, E.F. Matthew, Counterfeit or substandard antimicrobial drugs: a review of scientific evidence. J. Antimicrob. Chemother. **60**, 214–236 (2007)
6. A. Tariq, C. Imti, S. Helen, Substandard and counterfeit medicines: a systematic review of the literature. BMJ Open **3**, 1–7 (2013)
7. P. Yashwant, Emerging techniques for polymorph detection. Int. J. Chem. Pharm. Anal. **3**(1), 2–3
8. K. Wichmann, A. Klamt, Drug solubility and reaction thermodynamics. Chem. Eng. Pharm. Ind. R&D Manuf. **1**, 457–476 (2010)
9. J. Luyapaert, D.L. Massart, Y. Vander, Near-infrared spectroscopy applications in pharmaceutical analysis. Talanta **72**, 865–883 (2007)
10. S.M. Raoul, S.D. Yung, D. Volker, Z. Renato, Nanoscale chemical analysis by tip-enhanced Raman spectroscopy. Chem. Phys. **318**, 131–136 (2000)
11. S.R. Goodyear, R.M. Aspden, Raman microscopy of bone, in *Bone Research Protocols*, ed. by M.H. Helfrich (Humana Press, Totowa, NJ, 2012), pp. 15–30
12. R.M. El-Abassy, P. Donfack, A. Materny, Visible Raman spectroscopy for the discrimination of olive oils from different vegetable oils and the detection of adulteration. J. Raman Spectrosc. **40**, 1284–1289 (2009)
13. L. Olga, S.C. Nilam, Y.R. Chanda, W.T. Joseph, R.G. Matthew, V.D. Richard, Real-time glucose sensing by surface-enhanced Raman spectroscopy in bovine plasma facilitated by a mixed decanethiol/mercaptohexanol partition layer. Anal. Chem. **19**, 6134–6139 (2005)
14. S. Ruchita, Y. Agrawal, Raman spectroscopy: recent advancements, techniques and applications. Vib. Spectrosc. **57**, 163–176 (2011)
15. L.M. Richard, Raman scattering, Chap. 2, in *The Raman spectroscopy for chemical analysis*, ed. by L.M. Richard (Wiley, Columbus, Ohio, 2000), pp. 15–30
16. S. Ewen, D. Geoffrey, Raman spectroscopy, Chap. 3, in *The modern Raman spectroscopy: a practical approach*, ed. by E. Smith, G. Dent (Wiley, Chichester, West Sussex, 2005), pp. 71–91
17. C.L. Haynes, A.D. McFarland, R.P. Van Duyne, Surface-enhanced Raman spectroscopy. Am. Chem. Soc. **77**, 338 (2005)
18. S.Y. Lin, An overview of famotidine polymorphs: solid-state characteristics, thermodynamics, polymorphic transformation and quality control. Pharm. Res. **31**, 1619–1631 (2014)
19. S.Y. Lin, W.T. Cheng, The use of hot-stage microscopy and thermal micro-Raman spectroscopy in the study of phase transformation of metoclopramide HCl monohydrate. J. Raman Spectrosc. **43**, 1166–1170 (2012)
20. T. Kojima, S. Onoue, N. Murase, F. Katoh, T. Mano, Y. Matsuda, Crystalline form information from multiwell plate salt screening by use of Raman microscopy. Pharm. Res. **23**, 806–812 (2006)
21. A. Heinz, M. Savolainen, T. Rades, C.J. Strachan, Quantifying ternary mixtures of different solid-state forms of indomethacin by Raman and near-infrared spectroscopy. Eur. J. Pharm. Sci. **32**(3), 182–192 (2007)
22. A.P. Ayala, Polymorphism in drugs investigated by low wavenumber Raman scattering. Vib. Spectrosc. **45**, 112–116 (2007)
23. C.M. McGoverin, T. Rades, K.C. Gordon, Recent pharmaceutical applications of Raman and terahertz spectroscopies. J. Pharm. Sci. **97**, 4598–4621 (2008)

24. S. Stewart, R.J. Priore, M.P. Nelson, P.J. Treado, Raman imaging. Annu. Rev. Anal. Chem. **5**, 337–360 (2012)
25. A. Paudel, D. Raijada, J. Rantanen, Raman spectroscopy in pharmaceutical product design. Adv. Drug Deliv. Rev. **89**, 3–20 (2015)
26. Y. Shen, T. Paul, E. Claire, P. Andrew, P. Nikin, Detection of low levels of amorphous lactose using H/D exchange and FT-Raman spectroscopy. Pharm. Res. **25**, 2650–2656 (2008)
27. Z. Huang, X. Chen, Y. Chen, J. Chen, M. Dou, S. Feng, H. Zeng, R. Chen, Raman spectroscopic characterization and differentiation of seminal plasma. J. Biomed. Opt. **16**, 110501 (2011)
28. M. Hoehse, A. Paul, I. Gornushkin, U. Panne, Multivariate classification of pigments and inks using combined Raman spectroscopy and LIBS. Anal. Bioanal. Chem. **402**, 1443–1450 (2012)
29. R. Kast, R. Rabah, H. Wills, J. Poulik, G.W. Auner, M.D. Klein, Differentiation of small round blue cell tumors using Raman spectroscopy. J. Pediatr. Surg. **45**, 1110–1114 (2010)
30. F. Nishioka, I. Nakanishi, T. Fujiwara, K. Tomita, The crystal and molecular structure of the β-cyclodextrin inclusion complex with aspirin and salicylic acid. J. Incl. Phenom. **2**, 701–714 (1984)
31. J. Peng, S. Peng, A. Jiang, J. Wei, C. Li, J. Tan, Asymmetric least squares for multiple spectra baseline correction. Anal. Chim. Acta **683**, 63–68 (2010)
32. T. Vankeirsbilck, A. Vercauteren, W. Baeyens, F. Verport, Applications of Raman spectroscopy in pharmaceutical analysis. Trends Anal. Chem. **12**, 869–877 (2002)
33. W. Siegfried, H.H. Reinhard, Pharmaceutical applications of mid-IR and Raman spectroscopy. Adv. Drug Deliv. Rev. **57**, 1144–1170 (2005)
34. P.D. Schmitt, E.L. DeWalt, X.Y. Dow, G.J. Simpson, Rapid discrimination of polymorphic crystal forms by nonlinear optical Stokes ellipsometric microscopy. Anal. Chem. **88**, 5760–5768 (2016)
35. J.F. Kauffman, L.M. Batykefer, D.D. Tuschel, Raman detected differential scanning calorimetry of polymorphic transformations in acetaminophen. J. Pharm. Biomed. Anal. **48**, 1310–1315 (2008)
36. ChemImage™, Discrimination of acetaminophen polymorphs using Raman chemical imaging. Application Note (2010)