CSIRO PUBLISHING
www.publish.csiro.au/journals/ajc

Advances in the Application of Array Detectors for Improved Chemical Analysis, Part I. Comparison of Qualitative Analyses Using Large, Computer-Based Raman Spectral Libraries

M. Bonner Denton,[A,C] *Roger P. Sperline,*[A] *Jeffrey H. Giles,*[A] *Daniel A. Gilmore,*[A]
Carolyn J. S. Pommier[A] *and Robert T. Downs*[B]

[A]Department of Chemistry, University of Arizona, Tucson 85721, USA.
[B]Department of Geosciences, University of Arizona, Tucson 85721, USA.
[C]Author to whom correspondence should be addressed (e-mail: mbdenton@u.arizona.edu).

Although a mature technique, Raman spectroscopy remains not widely used as a routine analytical technique. The advantages of Raman spectroscopy, such as minimal sample preparation and relatively uncomplicated spectra, combined with technical advances, including more stable diode lasers, higher quality filters, and better polychromators and array detectors, have led to a renaissance in sensitive and versatile Raman spectrometers. The lack of appropriate databases for routine analysis, which represents the last barrier to the technique's widespread use, is beginning to be overcome, as described in this review.

Manuscript received: 27 September 2002.
Final version: 18 December 2002.

Introduction

Raman spectroscopy has matured as a specialized research tool but is not widely used as a routine technique in the analytical laboratory. Recent incorporation of stabilized diode lasers for excitation, high quality holographic filters for Rayleigh rejection, a new generation of polychromators for increased resolution, and advanced array detectors for improved detection limits, have created a new class of high sensitivity, versatile Raman spectrometers.[1] The remaining barrier to widespread use of Raman spectroscopy—the availability of appropriate databases—is starting to be overcome.

To quote F. A. Miller '…Raman spectra are still not as generally useful as infrared ones for several reasons: there are not as many good reference spectra, there is not as much information on characteristic group frequencies, and quantitative analyses are much more difficult because there is no equivalent Beer–Lambert law.'[2]

For an overview of Raman spectroscopy, see Skoog, Holler, and Nieman.[3] Raman yields molecular vibrational spectra complementary to those of infrared (IR) absorption.[4,5] Raman spectroscopy has several important advantages over the IR absorption technique:[2]

- The low Raman cross-section for water allows the analysis of aqueous solutions not amenable to IR absorption techniques.[6]
- Nearly any sample, solid, liquid, or gaseous, can be analyzed without modification. Raman spectra can also be taken through the walls of glass containers.[7]
- Overtone and combination bands are far less intense in Raman than in IR. Raman spectra are less complex. Band identification and qualitative analysis of mixtures are simplified.[5] With fewer bands in the Raman spectra, loss of discrimination due to band overlap in the spectra of mixtures is less likely, but when band overlap occurs, there is no advantage for Raman over IR spectra.
- Fibre-optic instruments allow sampling in locations remote from the actual instrumentation.[5,8–10]

Raman spectroscopy also has some distinct disadvantages from IR spectroscopy:[5,6]

- The same relative lack of overtone and combination bands in Raman spectroscopy makes it less sensitive to small structural differences than IR is.
- The relative insensitivity of Raman spectroscopy demands very strong excitation sources (in itself a potential disadvantage from a sample-heating standpoint), which increases the broadband baseline emission intensity due to sample fluorescence. This may be overcome in most cases by changing to longer wavelength excitation, e.g. at 1.06 μm from Nd:YAG (yttrium aluminium garnet)[11] or at 785 nm extra-cavity stabilized semiconductor lasers (see below). There are also new methods for removal of fluorescence by using two excitation wavelengths and subtraction, called shifted excitation Raman difference spectroscopy.[12]
- Because Raman is an emission technique, quantitative analysis can be more difficult. In IR spectroscopy, a background scan can readily be performed and the

10.1071/CH02200 0004-9425/03/020117

spectrum compared against the sample scan to give corrections for instrumental wavelength-dependent response and at least partial correction for sample-cell effects. In Raman spectroscopy, single-wavelength working curves need to be established on the user's instrument, much as is done for ultraviolet- and visible-spectrum fluorescence determinations.

- Again, because Raman is an emission technique, the familiar quantitative transmission techniques used in IR spectroscopy are not applicable. Sampling considerations in Raman have recently been discussed by Hendra,[7] and photometric standards by McCreery.[13]

The Chemical Abstracts Service has marked 675 references as reviews of Raman spectroscopy since the beginning of 2000. Raman and IR spectroscopy have very recently been exhaustively discussed in Chalmers and Griffiths' *Handbook of Vibrational Spectroscopy*.[14] Several comprehensive reviews have appeared on general Raman spectroscopy covering publications during 1991–1993,[15] and 1995–1997.[16] Additional general reviews have emphasized new techniques during 1997–1999,[17] and on process analytical chemistry including Raman Spectroscopy.[18] One review included many references to Raman spectroscopy although it concentrated on IR spectroscopy.[19]

During the last decade, advances in instrumentation have made Raman spectrometers far more accessible and reliable than in the past.[17,20,21] Improvements in laser technology,[22] improvements in high-quality volume-phase holographic filters,[23] improvements in high throughput imaging polychromators,[24] the availability of high quantum efficiency low noise array detectors,[2] and the advancement of Fourier-transform (FT) Raman spectrometers[25] have all contributed to this trend.[16] Field-portable Raman instruments are now available.[9,26]

Spectrum Libraries

Raman spectroscopy, long seen as a powerful analytical technique, has been overshadowed by IR spectroscopy for identification of unknown compounds, largely due to the availability of computer-based libraries of known IR spectra

Bonner Denton's research interests include applying the latest technological advances in electronics, physics, optics, acoustics, engineering, and computer science to developing new methods of chemical analysis and automation. Projects underway include selective single photoionization of complex mixtures, new modes of quadrupole MS, element-selective detection in GC and LC, and new modes of automation, among others. He received BSc and BA degrees from Lamar University (Beaumont, TX) and PhD from University of Arizona (Tucson, AZ); he attained a full professorship at the latter in 1980. He has received awards from many sources, including the ACS Analytical Division and the Society for Applied Spectroscopy. He is also interested in scuba diving and racing automobiles, recently setting a new land speed world record for sports cars.

Roger Sperline received his Ph.D. from Washington State University in inorganic chemistry in 1977 and held a post-doctoral fellowship at California Institute of Technology in electrochemistry. Following six years in analytical chemistry in industry and 13 years in infrared spectroscopy and solvent extraction with Prof. H. Freiser at Univ. of Arizona, he joined Prof. M. B. Denton's group as senior scientist. This work has involved development and characterization of novel semiconductor detectors for mass- and ion mobility spectrometries, and the computer modelling of Raman spectra, infrared spectra, and ion flight trajectories.

Daniel Gilmore received a B.S. in chemistry from Le Moyne College in 1979 and a Masters in Physical Chemistry from Syracuse University in 1983. He held a position of Research Specialist in the chemistry department at the University of Arizona from 1983 to 1999, consulted with Spectral Instruments, Inc. from 1994 until 1999 and has held a position of Application Scientist and Software Engineer with Spectral Instruments, Inc. from 1999 to present.

Carolyn Pommier received her B.S. in chemistry at Pittsburg State University in 1998. She currently is a graduate student in the Denton group at the University of Arizona. Her area of specialization is Raman spectroscopy of materials, including minerals. Her current area of research is the utilization of Raman spectroscopy to study the bonding changes associated with pressure-induced phase transitions of pyroxenes.

Robert Downs obtained his PhD in 1992 at Virginia Tech in mineralogy with an emphasis on the analysis of thermal motion of atoms in crystals. He completed a post-doc at the Geophysical Lab of the Carnegie Institution of Washington, specializing in high-pressure crystallography and mineralogy. In 1996 he was appointed an assistant professor in Geosciences at the University of Arizona. He is the crystallographic editor for the American Mineralogist *and* The Canadian Mineralogist*, a Fellow of the Mineralogical Society of America and winner of the Berry Medal for significant contributions to the Mineralogical Association of Canada.*

and partly due to lack of operator familiarity with Raman as a characterization technique. This topic has been recently reviewed by Lowry.[27] Today, nearly all manufacturers of IR instruments offer spectrum search capabilities and extensive IR spectrum libraries on their computer-controlled spectrometers.[28–30] The Sadtler company, for instance, introduced their Standard Infrared Collection and Commercial Infrared Collections in 1980.[31] Today these libraries contain well over 130 000 spectra.

We will use the nomenclature 'library' to denote any computer searchable collection of spectra, including 'databases' of both spectra and images. At this stage we have not discriminated among possible sampling techniques used to take the spectra, except to label those used for gas analyses. A search of the Chemical Abstracts in late 2002 revealed approximately 100 unique citations involving libraries and Raman spectroscopy. We have cited all the pertinent ones.

Generalized Raman spectrum libraries, along with appropriate search software, have recently become available from several companies, including Sadtler, Thermo–Nicolet, and SpecInfo. The integrated spectral database system for organic compounds (SDSB) library included 28 500 substances.[32–36] In addition, smaller specialized libraries are being developed, covering pharmaceuticals,[37] gems and minerals,[38–44] microbial organisms,[41,45–49] biological compounds,[50–52] biological spectrochemical images,[52,53] steel corrosion products, [43,54] nickel corrosion products,[55] uranium fuels corrosion products,[56] paper coatings,[57] contaminants in the microelectronics industry,[58–60] photographic dyes,[61] pigments and artists' materials,[42,62,63] hazardous materials,[64] polymeric materials,[65–67] lamps and light bulb materials,[68] human calculi, such as kidney stones,[69,70] coherent anti-Stokes Raman (CARS) spectra,[71–73] simulated CARS spectra,[74–76], gases,[77] and materials of interest in tribology.[78] One of these libraries, a large searchable database of FT Raman spectra containing over 14 000 compounds taken at 1.06 μm, has become commercially available from Thermo–Nicolet. Five different comparison algorithms are available for spectroscopic comparisons within this library.[79]

Applications and Algorithm Studies

In addition to the references revealing the development of searchable spectrum libraries, this type of analysis has been used in analyses of pulp and paper,[80] contamination in microelectronics,[81] pigments,[82] document identification (patents),[83] antibiotics,[84] microbial identification,[85] and thin-layer chromatography.[86]

Several references address developments in spectrum comparison algorithms and programming. Citations of spectrum search algorithm development include reviews by Davies and McIntyre,[87] Brown et al.,[88] and Lavine et al.[89,90] Further discussions concern the effects of pre-search spectrum corrections,[91] potentially specialized algorithms for aerosol analysis,[92] the effect of digital precision,[93] principal component analysis of pigment spectra,[94] substructure/subspectrum correlation,[95] wavelet transform algorithms,[96] neural networks,[46,47,49] and general applications.[97–99] The novel Raman spectrometer, based on a multichannel FT technique, differs fundamentally from dispersive spectrometers and FT-Raman systems.[100] Spectra from this instrument have been successfully analyzed by spectrum search against a library of spectra obtained on a dispersive instrument.

Other Search Programs

The 'KnowItAll' informatics system has recently been announced.[101] The SCANNET program has been developed for organic compound structure/spectrum correlation, including Raman, NMR (^{13}C, ^{1}H), IR, mass, and UV/vis spectra.[102–106] The earliest reference to a Raman spectrum search program is by Degen in 1976.[107]

Searching of commercial spectrum libraries is also available over the Internet on a pay-per-search basis for both IR and Raman via FTIRsearch.com.[108,109] In October 2000, over 71 000 FT-IR spectra and 16 000 Raman spectra were available for on-line searching, based on the Nicolet databases.

Objectives

In FT techniques, all optical frequencies are observed simultaneously and their wavelengths are calculated from the interferogram via reference to a laser—most commonly a helium-neon laser.[25] In contrast, a dispersive instrument distributes the wavelengths spatially across a detector—the placement of the wavelengths on the detector is referenced to standard neon lines. The laser wavelength is calculated by reference to a standard cyclohexane spectrum. The calculated laser wavelength is then used to calculate the Raman shift of the sample spectra.

These differences in methodology lead to subtly different spectra. The small differences do have an effect on identification of samples when dispersive spectra are compared with FT spectra available in commercial libraries. As demonstrated herein, such comparisons have been successful, indicating that it will not be necessary to construct large, expensive, general libraries of dispersive spectra.

There are three purposes to this paper: (1) Review the utility of the Thermo–Nicolet Raman spectrum library search package as an example of qualitative analysis using Raman spectra of both pure compounds and mixtures. (2) Demonstrate that uncorrected spectra obtained on a dispersive Raman instrument, using 785 nm excitation, can be successfully analyzed using an uncorrected library of spectra obtained on FT Raman instruments using 1.06 μm excitation despite the instrumental differences. (3) Demonstrate that library spectra obtained on a dispersive instrument at 514 nm, in conjunction with library spectra acquired at 785 nm, can be used to identify minerals using spectra acquired at 785 nm.

The studies presented here demonstrate that Raman spectroscopy, combined with library searching, can be a powerful tool for rapid identification of organic compounds and minerals.

Results

The Omnic 5.0 spectral search software, available from Thermo–Nicolet, was used for sample identification. The algorithms available in the software represent the most commonly available search methods. Other chemometric methods, such as partial least-squares and neural network algorithms, are being developed,[90] but were not available in Omnic.

Out of the five comparison algorithms available in the Omnic search software, correlation (CO) and absolute derivative (AD) were found to provide the best identification. These algorithms are described in the Experimental section. Tables 1 and 2 show the results of Raman library searches on the five pure solvent spectra. Tables 3–8 show the results of library searches on the six mixtures to compare the accuracy of Raman search results with the accuracy of comparable searches for IR spectra. Table 9 shows the result of Raman library searches on the mineral sample spectra.

Pure Solvent Raman and IR Spectra

The Raman spectra of five pure solvents—cyclohexane (CyH), benzene, toluene, trichloroethylene (TCE), and [D$_6$]acetone—were subjected to a Raman library search (Tables 1 and 2). The highest quality match was correct, with all the available search algorithms yielding the same result. The five IR spectra of the pure solvents were similarly correctly matched in the IR library. The Raman spectra match values for the five pure solvents were all above 70. Values for obviously bad matches tended to be below 60.

Pure CyH was found to give a match value of 76.1 against the pure CyH spectrum in the HR Aldrich Raman library. This led us to expect that correct compound matches between these experimental dispersive Raman spectra and the higher resolution FT spectra in the HR Aldrich Raman library would rarely give match values greater than approximately 76–78.

Raman Spectra of Mixtures

The AD algorithm was found to yield the best results and was subsequently used for all mixture spectra. The conditions used to obtain the library database spectra, which were uncorrected for instrumental throughput effects (a 1.06 μm source and an FT instrument), were very different conditions from those used to obtain the unknown spectra (an uncorrected dispersive instrument and 785 nm excitation).

Table 1. Results of Raman spectral library searches for spectra of pure CyH, benzene, and toluene
Matches listed in order of quality, highest quality matches first. Correct matches listed in boldface

Rank	CyH Match	Raman, absolute derivative 2305–252 cm^{-1}	Benzene Match	Raman, absolute derivative 2305–252 cm^{-1}	Toluene Match	Raman, absolute derivative 2305–252 cm^{-1}
1	76.02	**Cyclohexane**	72.62	**Benzene**	72.89	**Toluene**
2	53.64	*N*-Cyclohexyl-1,3-propanediamine	62.96	Aluminum sulfate $(H_2O)_{18}$, 98+%	60.29	*N*,*N*′-Dibenzylethylenediamine 97%
3	53.33	2-(Cyclohexyloxy) ethanol	51.39	Ethoxytrimethylsilane, 98%	59.84	Benzyl α–[^{13}C]alcohol, 99 atom% ^{13}C
4	51.86	4-Nitro-1,2-phenylenediamine	50.94	Bromochloromethane	59.78	Benzylamine
5	51.73	3-Cyclohexyl-1-propanol	50.82	Sodium trimethylsilanolate, 97%	59.68	Benzyl ether
6	51.72	Cyclohexanebutyric acid	50.8	Germanium(IV) methoxide, 97%	59.67	Dibenzylamine, 97%
7	51.54	Trimethylene oxide	50.73	*N*,*N*′-Bis(trimethylsilyl)-1,4-butanediamine, 95%	59.39	Benzyl alcohol
8	51.46	Urea phosphate	50.49	Propoxytrimethylsilane, 98%	59.32	*N*-Benzyl-2-phenethylamine
9	51.43	Azetidine	50.46	Cacodylic acid, sodium salt hydrate, 98%	59.04	*N*,*N*′,*N*′-Tribenzyltris(2-aminoethyl)amine
10	51.41	4-Cyclohexyl-1-butanol	50.42	(2-(1,3-Dioxolan-2-yl)ethyl) triphenylphosphonium bromide	58.91	*N*-Benzylethanolamine, 95%

Table 2. Results of Raman spectral library searches for spectra of pure trichloroethylene and [D$_6$]acetone
Matches listed in order of quality, highest quality matches first. Correct matches listed in boldface

Rank	TCE Match	Raman, absolute derivative 2305–252 cm^{-1}	[D$_6$]Acetone Match	Raman, absolute derivative 2305–252 cm^{-1}
1	78.08	**Trichloroethylene**	71.36	**[D$_6$]Acetone**, 99.5 atom% D
2	55.74	Dichlorobis(trimethylsilyl)methane, 97%	57.55	Propargylamine
3	52.92	Methyltrimethoxysilane, 98%	53.10	1,3-Dimethyl-1-butyne, 98%
4	51.85	Tris(trimethylsilyl)phosphine, 95%	52.84	3-Ethynylcyclopentanol, 98%
5	51.33	1,3-Bis(trimethylsilyl)urea, 95%	52.39	Cyclopentane
6	51.18	9-*cis*-Retinal, 98%	51.9	1,1-Dimethylpropargylamine
7	50.9	Bromotrimethylsilane, 97%	51.61	Ethyl sulfide, 98%
8	50.77	(Chloromethyl)trichlorosilane, 97%	51.37	Sodium thiomethoxide, 95%
9	50.7	Dibromodichloromethane, 95%	51.28	Cyanuric acid, compound with melamine
10	50.58	Trimethylsilyltrifluoroacetate	50.71	Chlormethyldichloromethylsilane

Table 3. Results of Raman and infrared spectral library searches for a spectrum of a hexane, heptane, and octane mixture
Matches listed in order of quality, highest quality matches first. Correct matches listed in boldface

Rank	Match	Raman, absolute derivative 2300–200 cm^{-1}	Match	Infrared in NaCl cell, correlation 3995–598 cm^{-1}	Match	Infrared in KBr cell, correlation 3995–455 cm^{-1} [A]
1	67.09	**Heptane**	96.74	**Octane**	76.83	**Octane**
2	65.61	**Octane**	95.52	**Hexanes**	73.22	**Hexanes**
3	64.89	**Hexanes**	91.20	Undecane	72.14	Undecane
4	64.51	**Hexane**	90.71	**Heptane**	70.33	2-Methylnonane
5	64.11	Dodecane	89.93	Ligroin	69.86	Xylan, from larch wood
6	63.87	Undecane	89.73	2-Ethylhexyl sulfide	69.18	Dodecane
7	63.53	Tetradecane	89.39	Amylamine	69.15	Amylamine
8	63.44	Tridecane	89.08	Dodecane	68.71	**Heptane**
9	63.42	Heptylamine	87.82	2-Methylnonane	68.51	Butyltin chloride dihydroxide, 96%
10	63.41	Nonane	86.81	Hexylamine	67.91	Tetradecane

[A] Excessive cell path length.

Table 4. Results of Raman and infrared spectral library searches for a spectrum of a methanol, ethanol, and acetone mixture
Matches listed in order of quality, highest quality matches first. Correct matches listed in boldface

Rank	Match	Raman, absolute derivative 2300–200 cm^{-1}	Match	Infrared in NaCl cell, correlation 3995–607 cm^{-1}	Match	Infrared in NaCl cell, correlation 3995–455 cm^{-1}
1	61.12	**Ethyl [1-^{13}C]alcohol**, 98 atom%^{13}C	53.48	**Ethyl alcohol**, denatured	52.96	**Ethyl alcohol**, denatured
2	59.82	Ethyl nitrite, 15 wt.% solution in **ethyl alcohol**	51.07	Ethyl nitrite, 10–20 wt.% solution in **ethyl alcohol**	50.76	Ethyl nitrite, 10–20 wt.% solution in **ethyl alcohol**
3	57.99	**Ethyl alcohol**, reagent, denatured	50.07	*N*-Ethyldiethanolamine, 98%	49.55	*N*-Ethyldiethanolamine, 98%
4	57.52	**Acetone**	48.35	*N*-Methyldiethanolamine, 99%	47.91	4-Hydroxy-3-methyl-2-butanone
5	56.16	**Methyl alcohol**	48.23	4-Hydroxy-3-methyl-2-butanone	47.59	*N*-Methyldiethanolamine
6	54.59	**Methyl [D$_1$]alcohol**, 99.9 atom% D	47.89	4-Dimethylamino-3-methyl-2-butanone	47.40	4-Dimethylamino-3-methyl-2-butanone
7	53.76	Methylcyclopentane, 98%	47.47	**Methyl alcohol**	47.10	**Methyl alcohol**
8	53.71	**Ethyl [D$_1$]alcohol**, 99.5+ atom% D	47.21	Poly(vinylacetate-*co*-butylmaleate-*co*-isobornylacrylate)	46.85	Poly(vinylacetate-*co*-butylmaleate-*co*-isobornylacrylate)
9	53.57	NH$_4$Ge^{IV} oxalate hydrate	46.84	**Ethyl [1-^{13}C]alcohol**, 98 atom% ^{13}C	46.53	**Ethyl [1-^{13}C]alcohol**, 98 atom% ^{13}C
10	53.36	Cyclopentane	46.38	3-Hydroxy-3-methyl-2-butanone	45.81	3-Hydroxy-3-methyl-2-butanone

Table 5. Results of Raman and infrared spectral library searches for a spectrum of an *m*-xylene, *o*-xylene, and *p*-xylene mixture
Matches listed in order of quality, highest quality matches first. Correct matches are listed in boldface

Rank	Match	Raman, absolute derivative 2300–200 cm $^{-1}$	Match	Infrared in NaCl cell, correlation 3995–602 cm^{-1}	Match	Infrared in KBr cell, correlation 3995–455 cm^{-1}
1	59.94	**Xylenes** + ethylbenzene, 98.5+%	83.09	**Xylenes** + ethylbenzene, 98.5+%	87.78	**Xylenes** + ethylbenzene, 98.5+%
2	54.99	***o*-Xylene**, 97%	50.82	4-(4-Biphenylyl)-2-methylthiazole	51.57	4-(4-Biphenylyl)-2-methylthiazole
3	54.82	***p*-Xylene**	46.38	Dibenzothiophene, 98%	47.41	***m*-Xylene**
4	54.79	***m*-Xylene**	45.05	***p*-Xylene**	46.04	***p*-Xylene**
5	54.68	*m*-Xylylenediamine	44.09	2-Phenylindole, Tech., 95%	45.24	***o*-Xylene**, 97%
6	54.42	3-Methylbenzylamine, 98%	43.99	2-Ethylphenyl hydrazine hydrochloride, 98%	44.13	Dibenzothiophene, 98%
7	53.82	3-Methylbenzyl alcohol, 98%	43.34	1,1,1,2,2,3,3-Heptachloropropane	42.27	2-Phenylindole, tech., 95%
8	53.5	4-Ethyltoluene, tech., 90%	43.28	***m*-Xylene**	41.28	2-Ehylphenyl hydrazine hydrochloride, 98%
9	53.35	3-Methoxybenzyl alcohol, 98%	42.45	1,2-Dianilinoethane	40.77	1,2-Dianilinoethane
10	53.19	Polyvinyltoluene, mixed isomers, M_W 80 000	42.36	***o*-Xylene**, 97%	40.41	2-Methylbenzyl cyanide, 98%

Table 6. Results of Raman and infrared spectral library searches for a spectrum of a benzene, toluene, and CyH mixture
Matches listed in order of quality, highest quality matches first. Correct matches listed in boldface

Rank	Match	Raman, absolute derivative 2300–200 cm^{-1}	Rank	Infrared in NaCl cell, correlation 3995–593 cm^{-1}	Rank	Infrared in KBr cell, correlation 3995–455 cm^{-1}
1	59.76	**Benzene**	53.56	Molybdenum(II) acetate dimer	55.34	Molybdenum(II) acetate dimer
2	57.16	Aluminum sulfate $(H_2O)_{18}$	52.48	Sodium fumarate, 98%	51.12	**Benzene**
3	54.67	Pyridine	48.44	Diethyldithiophosphate, ammonium salt	49.01	Diethyldithiophosphate, ammonium salt
4	54.05	**Cyclohexane**	47.69	1,4,7,10-Tetrathiacyclododecane	47.22	Phenylmercuric nitrate, basic
5	52.99	**Toluene**	47.32	3-Chlorophenylhydrazine hydrochloride	46.26	Sodium fumarate, 98%
6	52.7	Lanthanum(III) sulfate hydrate	45.93	Davy reagent methyl, tech., 90%	44.59	**Cyclohexane**, 99+%
7	52.35	Anisole	45.05	5-Amino-2-naphthalenesulfonic acid	43.95	Neodymium(III) acetate hydrate
8	52.29	*N*-phenylbenzylamine	42.57	Hexachloroethane, 99%	43.78	1-Methyl-DL-tryptophan
9	51.72	(*S*)-(–)-α-Methylbenzyl isocyanate	42.00	1-Decyl-2-methylimidazole	43.29	Davy reagent methyl, tech., 90%
10	51.71	(*R*)-(+)-α-Methylbenzyl isocyanate	41.30	Neodymium(III) acetate hydrate	43.01	1-Decyl-2-methylimidazole

Table 7. Results of Raman and infrared spectral library searches for a spectrum of a trichloroethylene, dichloromethane, and chloroform mixture
Matches listed in order of quality, highest quality matches first. Correct matches listed in boldface

Rank	Match	Raman, absolute derivative 2300–200 cm^{-1}	Rank	Infrared in NaCl cell, correlation 3995–602 cm^{-1}	Rank	Infrared in KBr cell, correlation 3995–455 cm^{-1}
1	56.29	**Trichloroethylene**	63.34	**Trichloroethylene**, 98%	60.6	**Trichloroethylene**, 98%
2	54.54	**Chloroform**, 99.8%	36.61	Chloroacetonitrile	38.47	**Chloroform**
3	54.23	**Dichloromethane**, 99.6%	35.77	**Chloroform**	38.11	Rubber, chlorinated
4	52.23	1,3-Bis(trimethylsilyl)urea	34.68	Alpha-angelicalactone, 98%	35.52	Chloroacetonitrile
5	51.82	Dichlorobis(trimethylsilyl) ethane, 97%	34.58	Pinacolone oxime, 98%	34.48	3,5-Dibromo toluene, 97%
6	51.74	Methyltrimethoxysilane, 98%	34.53	1,4-Diisopropenyl benzene, 97%	33.04	1,4-Diisopropenyl benzene, 97%
7	51.73	Methyl sulfoxide	34.28	3,5-Dibromotoluene, 97%	32.65	Polypropylene, chlorinated, M_W 230 000
8	51.51	2-Mercaptoethyl ether, 95%	32.75	Rubber, chlorinated	32.2	Pinacolone oxime, 98%
9	51.12	Trimethylsilyltrifluoroacetate, 96%	31.64	$[(1,5\text{-cod})\text{Rh}^{\text{I}}\ (\text{PPh}_3)_2]\text{PF}_6 \cdot \text{CH}_2\text{Cl}_2$ (1:1)	32	Alpha-angelicalactone, 98%
10	50.9	2-Iodo-1,1,1-trifluoroethane	31.37	**Dichloromethane**, 99.6%	31.35	**Dichloromethane**, 99.6%

Table 8. Results of Raman and infrared spectral library searches for a spectrum of an Excedrin tablet
Matches listed in order of quality, highest quality matches first. Correct matches are listed in boldface

Rank	Match	Raman, absolute derivative 1820.0–272.0 cm^{-1}	Match	Infrared in KBr pellet, correlation 4000–700 cm^{-1}
1	60.94	**Acetylsalicylic acid**	66.23	**4-Acetamidophenol**, 98%
2	55.81	**4-acetamidophenol**, 98%	50.91	**Acetylsalicylic acid**
3	53.17	**Caffeine**	33.02	Chlorobis(4-fluorophenyl)methane
4	52.75	4′-Fluoroacetanilide, 98%	32.01	Poly(bisphenol A-*co*-epichlorohydrin), glycidyl endcapped
5	52.5	Bis(trimethylsilyl)methane, 97%	31.71	2,4,6-Collidine
6	52.39	Perfluoroisopropyliodide	31.24	4-(Octyloxy)benzonitrile, 98%
7	52.29	4-Fluorophenol	31.11	α-Cyclopropyl-4-fluorobenzyl alcohol, 90%
8	52.15	1,4-Dibromobutane	30.01	4-Fluoro-α-methylbenzyl alcohol
9	52.04	4′-(Octyloxy)-4-biphenylcarbonitrile	29.85	1-(3-Hydroxyphenyl)urea, 98%
10	51.99	3-(2-Hydroxyphenyl)propionic acid	29.43	4′-Methoxychalcone, 97%

The natural biases resulting from such search comparisons are effectively reduced by the use of the AD algorithm. The search results for pure benzene also showed that the AD algorithm gives a large weight to narrow bands. The superior performance of the AD algorithm is not surprising, given the differences in relative peak heights between the spectra of mixtures and those of the pure compounds. These differences appear because of the presence of unresolved

Table 9. Results of Raman spectral library searches for spectra of pink sapphire, the inclusion in the pink sapphire, and the Allen Hills meteorite

Matches are listed in order of quality with the highest quality matches first. Wavelength at which library spectrum was acquired is in parentheses. Correct matches are listed in boldface

Rank	Pink sapphire Match	Raman, correlation 1000–200 cm^{-1}	Inclusion sapphire subtracted Match	Raman, correlation 1267–200 cm^{-1}	Meteorite Match	Raman, correlation 1850–200 cm^{-1}
1	81.17	**Sapphire.009 (785)**	57.88	**Hauyne.1m (514)**	43.12	**Diopside.1m (514)**
2	78.47	**Sapphire.007 (785)**	44.92	Alunite.001 (785)	37.06	**Diopside.006 (785)**
3	69.57	**Sapphire.003 (785)**	40.94	Alunite.002 (785)	33.83	**Enstatite.5m (514)**
4	52.52	Alexandrite (treated spinel).001 (785)	39.84	Thenardite.002 (785)	28.80	**Diopside.3u (514)**
5	60.52	Cristobalite.1 (514)	38.58	Barite.002 (785)	28.35	Rhodonite.003 (785)
6	58.90	Cristobalite.3 (514)	38.21	Thenardite.003 (785)	27.89	Kornerupine.001 (785)
7	56.31	Gahnite.1e (514)	37.63	Thenardite.001 (785)	27.46	Kornerupine.002 (785)
8	55.57	Ruby.004 (785)	37.02	Alunite.003 (785)	27.44	Rhodonite.004 (785)
9	53.38	Alexandrite (treated spinel).001 (785)	36.81	Barite.001 (785)	27.35	Kanoite.3p (514)
10	45.20	**Sapphire.004 (785)**	35.63	Thenardite.005 (785)	26.14	Gypsum.007 (785)

peaks and because the spectra from the dispersive and FT-Raman instruments have not been corrected for spectrometer throughput.

All mixtures, except that of the Excedrin tablet, consisted of approximately equal proportions (by volume) of the three reagents. The ten highest quality matches for each mixture are shown in Tables 3–8; the number one match is the best, that is, it has the highest mathematical score according to the comparison algorithm. Matches that correspond to known components of the mixture are shown in bold. For the Excedrin tablet, the known components were taken to be those main ingredients listed on the package. For all mixtures analyzed, the known components were within the top five matches. Figure 1 shows the IR and Raman spectra of the Excedrin tablet.

IR Spectra of Mixtures

These also show the results of a library search performed on the IR spectra of the same six mixtures. The CO algorithm was used in all cases. It is the most commonly used algorithm, and there was insufficient difference between the resolutions of the experimental and library spectra to justify using the AD algorithm. The tables illustrate the best ten matches; correct matches are in bold. Raman analysis identified all mixture components within the five unique best matches. IR analysis was equally successful on two of the mixtures, identified all three within the first ten hits on another mixture, and missed one component in each of the remaining mixtures.

The use of a NaCl transmission cell and all data presented by the spectrometer (4000–400 cm^{-1}) produced uniformly poorer matches, except for the hexane/heptane/octane mixture (Table 3). NaCl transmission falls off below 600 cm^{-1}, and the spectrometer software completes the spectra from approximately 550 cm^{-1} to 400 cm^{-1} with meaningless data. The hexane/heptane/octane mixture had no significant features between 600 and 400 cm^{-1}, so the NaCl cell spectrum was adequate. On the other hand, inclusion or exclusion of meaningless data between 600 and 455 cm^{-1} made no significant difference to the identification of the methanol/ethanol/acetone mixture (Table 4).

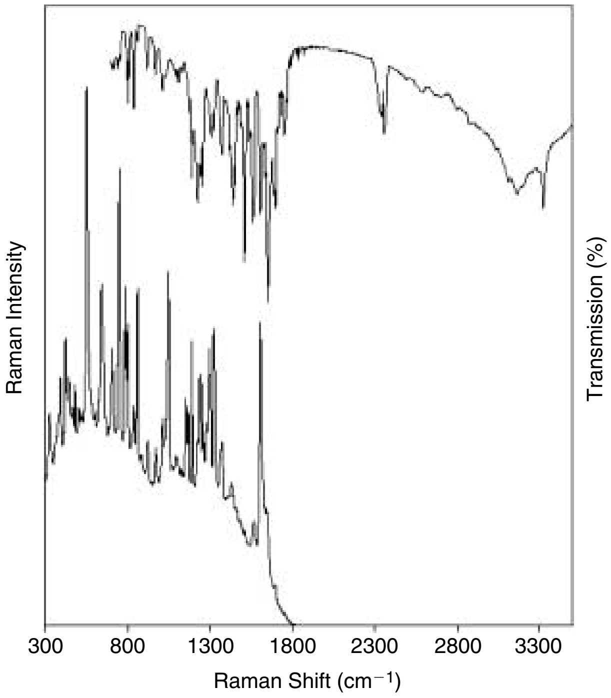

Fig. 1. IR (top) and Raman (bottom) spectra of the Excedrin tablet. The Raman spectrum is of the unmodified tablet, the IR spectrum is of a KBr pellet.

The xylenes mixture was perfectly identified by both Raman and IR, only when IR analysis was performed with a KBr cell (Table 5). The xylenes spectra taken with KBr windows contain real features down to 400 cm^{-1}. Proper identification of the xylenes mixture thus required the bands between 600 and 455 cm^{-1} (the library spectra extended to 450 cm^{-1}). Figure 2 shows the IR spectra of the 1 : 1 : 1 v/v/v mixture of *o*-, *m*-, and *p*-xylenes taken through both NaCl and KBr windows. The spectrum in Figure 2, taken with NaCl windows, presents a false broad peak near 525 cm^{-1}.

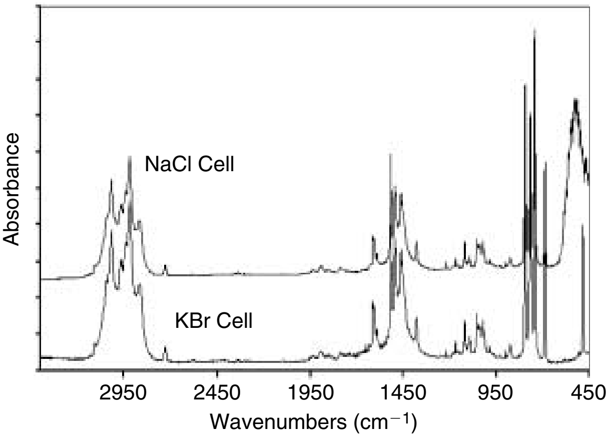

Fig. 2. IR spectra of 1 : 1 : 1 v/v/v xylenes mixture in NaCl and KBr capillary transmission cells.

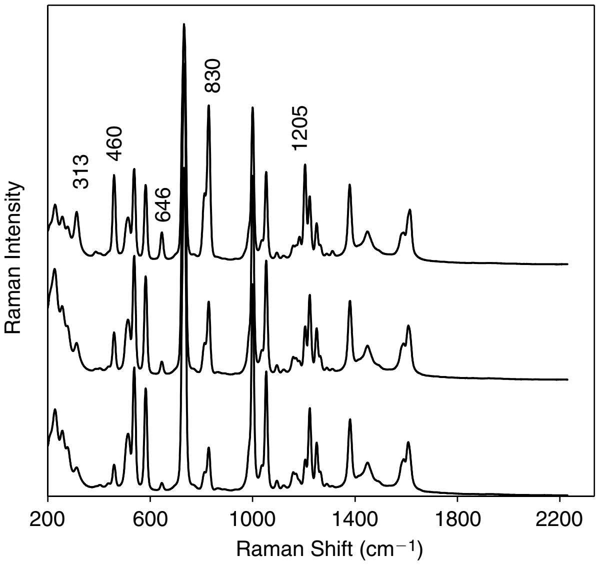

Fig. 3. Raman spectra of three mixtures of *p*-, *m*-, and *o*-xylene. The top spectrum is of an equal mixture by volume, the middle of 15% *p*-xylene, and the bottom of 9% *p*-xylene. The wavenumber position of the most prominent distinguishing *p*-xylene peaks are labelled.

p-Xylene has a strong sharp peak near 485 cm^{-1}, *o*-xylene has two weak sharp peaks (580 and 505 cm^{-1}), whereas *m*-xylene has no peaks in this region. The additional information present in the spectra taken in KBr was sufficient to give higher match values to the xylene isomers. The xylenes library spectrum, however, had excellent correspondence in all the rest of the spectrum, as it should have, being approximately the same mixture as the experimental sample.

To further demonstrate a pitfall of IR library search identification of mixtures, the hexane/heptane/octane spectrum was purposely performed in a KBr cell with an excessively long path length (Table 3). With truncated absorbances for the CH stretching bands, identification of the mixture component was poorer than in the NaCl cell, which had a shorter path length, and the match values fell from approximately 90 to 75.

For the hexane/heptane/octane mixture using the CO algorithm (Table 3), and for several other mixtures when using other search algorithms, several unexpected but highly ranked search hits appeared. During an IR search on the xylenes mixture spectrum taken in the NaCl cell case and using the AD algorithm, K_2TaF_7, which has a very strong broad band mimicked by the broad 'feature' near 531 cm^{-1} in the xylenes spectrum, was ranked third. In the KBr cell case, using the AD algorithm, K_2SiF_6, which has a strong sharp band similar to that of *p*-xylene near 482 cm^{-1}, was ranked fifth. In addition, library searches on the hexane/heptane/octane mixture using the AD algorithm gave matches for wholly inorganic solids with bands in the CH_2 scissor region. The spectra in the library for these and many other solids appeared to have been taken as Nujol-like mulls, although this was not indicated in the library, nor could such spectra be excluded from the lists of possible matches.

Curiously, for the benzene/toluene/CyH mixture, toluene was not in the hit list for IR spectra taken with either window material, and no correct matches were made in the NaCl cell despite limiting the data range (Table 6). Toluene is responsible for a medium-intensity sharp band near 464 cm^{-1}. We also noted the complete absence of acetone from both IR match lists for the methanol/ethanol/acetone mixture (Table 7). The search did, however, offer butanones as possible matches for the carbonyl function. Similarly interesting was the absence of caffeine from the IR match list for the Excedrin tablet (Table 8). Pure caffeine has a strong Raman band near 555 cm^{-1}, which is apparently identifiable in the Raman library, but no sufficiently unique bands to be identifiable in the IR libraries.

Raman Spectra of Xylene Mixtures

As the relative concentration of mixture component decreases, it should become more difficult for a library search to correctly identify that component. In order to assess the limits of such compound identification, mixture samples of *m*-, *o*-, and *p*-xylene were prepared in which the relative concentration of *p*-xylene decreased by gradations from an amount equal to that of the other two components. *p*-Xylene was chosen because it has several intense distinguishing peaks that are completely resolved with respect to *m*- and *o*-xylene. These mixtures should have represented a favourable case for identification of the components. Furthermore, the structural similarities of the three compounds should have insured that the Raman cross-sections and densities were also approximately similar. The Raman spectra of three of these mixtures are shown in Figure 3. The most significant spectroscopic features corresponding to *p*-xylene are labelled. The AD algorithm was used for all spectra searches.

The top spectrum in Figure 3 was taken from an equal mixture of the three components. Xylenes was the best match for this mixture, whereas *p*-, *m*-, and *o*-xylene were the second-, third-, and fourth-best matches. In the middle spectrum, the relative concentration of *p*-xylene had been reduced to 15%, and the library search of this mixture showed that *p*-xylene had dropped to the fifth-best match. For a 12% solution (not shown), p-xylene dropped to sixth position. For the bottom spectrum, of a 9% solution, p-xylene was no longer included as one of the ten best matches.

Difference Spectroscopy for Identification of Trace Components

The identification of mixture components at less than 15% concentration was clearly not feasible by a simple search of the raw spectrum. Once a component is known to be present in a mixture, however, the pure spectrum of that component could be subtracted from the raw spectrum (‘spectrum stripping’), enabling a search that should yield better results for the remaining components.

To demonstrate this procedure, a mixture consisting of 90% o-xylene, 9% m-xylene, and 1% p-xylene was prepared. The spectra of pure o- and m-xylene used for subtraction could be obtained from the library but could not be used, because the FT-Raman spectrometer used to acquire the library spectra had a different instrument response function from the dispersive instrument used in this study. The Raman spectra of pure m- and o-xylene were therefore collected on the same instrument as the spectra for the xylene isomer mixture.

Figure 3 shows the spectrum of the initial xylenes mixture. The Raman spectrum library search results for this mixture were as follows:

1. o-xylene
2. 2-methylbenzylamine
3. thiourea
4. 2-ethyltoluene
5. 1,2-diodoethane

Note that o-xylene was the only correct match within the first five hits.

Figure 4 shows the difference spectrum produced by subtraction of the o-xylene spectrum. Spectrum subtraction was accomplished manually using the Omnic software. The subtraction factor was changed to minimize the intensities of peaks at 1057 and 584 cm^{-1} due to o-xylene. Minimization was decided visually by operator, rather than by any calculation. The search results for this spectrum were as follows:

1. m-xylene
2. xylenes, isomers plus ethylbenzene
3. 3-methylbenzylamine
4. m-xylenediamine
5. 3-ethylaniline

Figure 5 shows the results after spectrum subtraction of m-xylene from the previous difference spectrum. The subtraction was changed to minimize the intensity of the peaks at 1000 and 538 cm^{-1} due to m-xylene. Minimization was decided visually by operator, rather than by any calculation. It can be seen from this spectrum that the previous subtractions were not perfect. A few over-subtraction features were visible, as well as residual spectroscopic features associated with the other two isomers. The spectrum exhibited peaks distinctive of p-xylene, however. The search results were as follows:

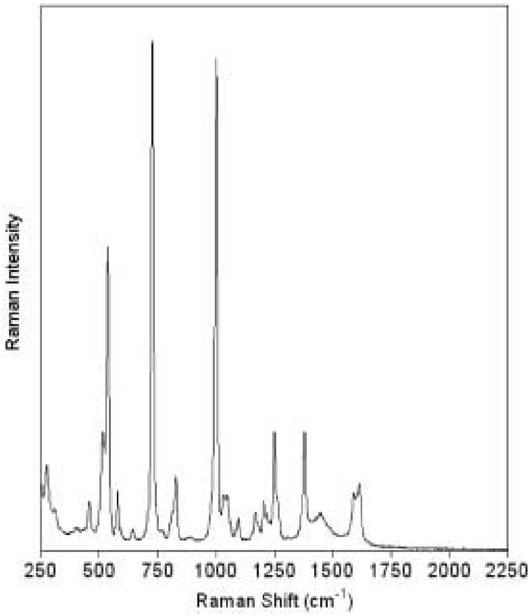

Fig. 4. Mixture of xylenes with o-xylene subtracted.

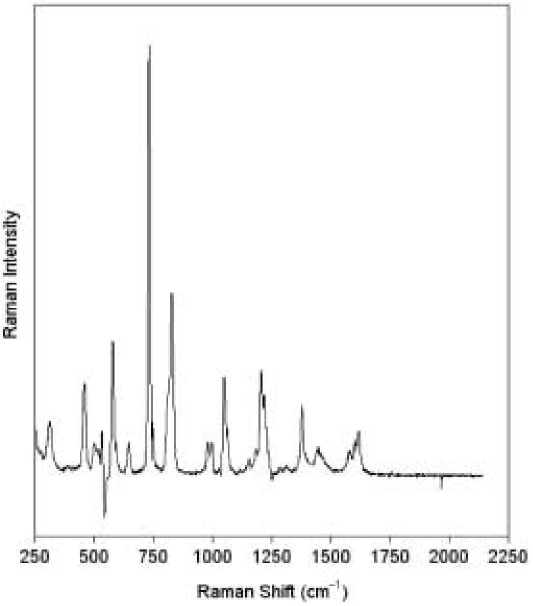

Fig. 5. Mixture of xylenes with o-xylene and m-xylene subtracted.

1. p-xylene
2. 4-ethyltoluene
3. 4-methylbenzylamine
4. dibromomethane
5. thiourea

Note that p-xylene is now identified even at the 1% level.

In a more rigorous treatment, an aggregate residual could be calculated for the wavelength regions corresponding to all bands of the spectrum to be removed—minimization of the residual would provide an objective subtraction factor. For the purposes of demonstration of a user-friendly method, however, subjective subtraction was adequate.

The procedure of spectral subtraction of a scaled pure component spectrum is analogous to background subtraction in a transmission experiment. Non-level baselines will be partially levelled during subtraction, depending on the scale factor used. Wavelength-dependent sensitivity will, however, not be addressed by subtraction of spectra. The identification of *p*-xylene at the 1% level shows that further response correction is not necessary for successful qualitative and semi-quantitative analyses.

Identification of Minerals

Table 9 shows the results of a library search on the spectra of three minerals: a pink sapphire, a small blue inclusion completely contained within the pink sapphire, and a sample of a meteorite. The inclusion was identified after subtraction of contributions to the spectrum from the pink sapphire that surrounded the inclusion. The CO algorithm was used in all cases. Correct matches are in bold.

Raman analysis correctly identified various sapphire spectra as the top three matches. The inclusion spectrum distinctly matched the sole library spectrum of the mineral hauyne ($(\mathrm{Na,Ca})_{8-4}[(\mathrm{SO_4})_{2-1}|(\mathrm{AlSiO_4})_6]$), despite the fact that the library spectrum was acquired at 514 nm and the sample spectrum at 785 nm. The meteorite was identified as either diopside ($\mathrm{CaMgSi_2O_6}$) or enstatite ($\mathrm{Mg_2Si_2O_6}$). These library spectra represent an assortment of reference, mainly single crystal samples, but the top matches for both diopside and enstatite were spectra of massive samples, like the meteorite. The meteorite spectrum acquired at 785 nm matched spectra acquired at both 785 nm and at 514 nm.

Discussion

The additional features made available in the 600–450 $\mathrm{cm^{-1}}$ region through the use of KBr cells, rather than NaCl cells, was crucial to the IR analysis of the xylenes mixture. Similarly, the elimination of a false 'peak' near 530 $\mathrm{cm^{-1}}$ in the NaCl cell spectra, caused by the cut-off of NaCl transmittance near 600 $\mathrm{cm^{-1}}$, eliminated the appearance of wildly incorrect compounds in the match lists. These observations reveal a trap awaiting neophyte analysts; the use of IR cells with limited wavelength ranges of transparency can introduce spectroscopic artifacts that can lead to inappropriate library matches. On the other hand, Raman spectroscopy/spectrum searching does not present this difficulty.

To improve the match quality between library spectra taken on one instrument and sample spectra taken on another, several attempts were made to identify the prime causes for poor match values and to create simple correction procedures to be applied to sample spectra prior to library searching. First, a sample spectrum was divided into several wavelength ranges and each range individually scaled to try to match peak heights with corresponding bands in the library spectrum. This produced no better results than using the amplitude-insensitive AD algorithm. Second, a sample spectrum was convolved with a continuously varying Gaussian function to try to remove the wavelength dependence of the band width across spectra collected on the dispersive instrument. Band widths in the Raman library spectra were constant because the FT process does not introduce wavelength-dependent bandwidths, so spectra from the dispersive instrument were widened progressively across the wavelength range. This process slightly improved the match values, but not dramatically.

Further improvement in match values and library search accuracy would no doubt be obtained by correction of both the library spectra and the sample spectra for the respective instruments' response functions by the methods discussed by McCreery.[13] The results of the experiments described here, however, indicate that quite good results may be had without such tedious corrections.

The poor discrimination between potential components in sample mixtures is not unexpected. There are no single compounds in the library combining all the spectroscopic features of the mixtures. Hence there is not a good correspondence between the spectra, and many spectra give similar match values.

The algorithms utilized in this study are rather simple compared with chemometric routines. The authors chose to not perform a comparison of the search algorithms with chemometric routines for several reasons. First, the routines and, except for the mineral database, the libraries utilized are commercially available. Second, the commercial libraries were not in a format that is convenient to import into a chemometric software package. Third, use of the software is simple and requires no specific knowledge of the math behind the user interface, and therefore is more accessible to an average user.

In reality there is rarely a complete lack of information about an unknown sample. The available information can be used to limit the possibilities presented by Raman or IR library searching. Solid samples (usually) do not contain liquid compounds and library matches with liquid compounds are incorrect. These results must be examined carefully, nonetheless. With a single-component sample, there is a large difference between the match value for the correct match compound to the second best match—here the difference was over 10 points for Raman library searches. For the chlorocarbons experimental mixture, the difference between the first and second match values in the IR library search was over 25 points. If such a mixture were examined in a blind test, one could incorrectly conclude that it contained only TCE. In contrast, in the Raman analysis, the chlorocarbons experimental mixture gave values from 56 to 54 for the correct matches (the first three), but the fourth place match was only 1.5 points behind third place. If such a mixture were examined in a blind test, one could not establish from the library search alone that the mixture contained only three components. Small noise features or minor mixture components could shift the possible match values enough to reorder the possible compounds suggested by the search procedure.

Additional information, either non-spectroscopic or from other types of spectroscopy, can be utilized to substantiate identification of minerals. For example, the inclusion, shown in Figure 6, was identified as hauyne. This is a reasonable match, considering that the shape of the inclusion was cubic, the colour was blue, and that hauyne is commonly found in the same locality where the sapphire was collected.

The identity of the meteorite can also be suggested as a reasonable match due to additional information. This meteorite has also been previously reported to consist of over 90% pyroxenes.[110] Both diopside and enstatite are members of the pyroxene group. The meteorite spectrum displays characteristics of both diopside and enstatite spectra in the mineral library. Neither of the two minerals came up with a significantly higher match value than the other, so it is difficult to determine which is the correct match solely on basis of the search. Two of the more obvious reasons that diopside matched better than enstatite are the missing peaks in the meteorite spectrum at approximately 400 and 1000 cm^{-1}. On the other hand, in the meteorite spectrum there is a shoulder on the peak at 675 cm^{-1} that is coincident with a peak in the enstatite spectrum, whereas only a single peak appears in this location in the diopside spectra. The reader is presented with the spectra of the meteorite and the top four matches in Figure 7.

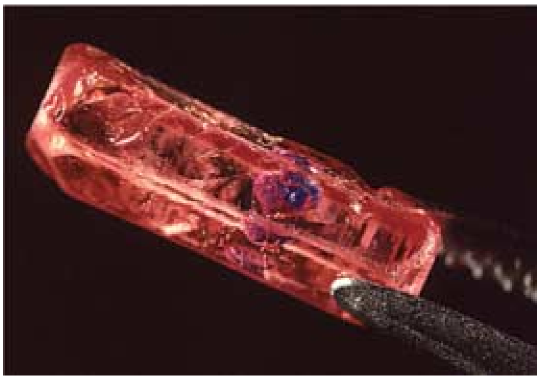

Fig. 6. Pink sapphire (corundum) with blue inclusion (hauyne). Photograph courtesy: American Gemological Laboratory.

Conclusions

These dispersive Raman spectra of organic compounds, mixtures, and minerals were more than adequate for successful analyses using spectrum searching, although the organic compound Raman library spectra were taken using an FT-Raman spectrometer and were uncorrected for instrument response. Raman spectra of minerals were also successfully analyzed using 785 nm excitation, even though some of the library spectra were acquired at 514 nm. Comparisons with libraries obtained using more similar instruments or databases corrected for instrumental throughput would probably have been even more remarkable.

A successful search has been defined here as one yielding correct matches of the experimental mixture components within the first five matches. Successful analysis of the six organic compound mixtures of interest was achieved using Raman spectroscopy. Spectrum searches of IR spectra were less successful, possibly because IR spectra are generally more complicated than Raman spectra—they often contain combination and overtone bands rare in Raman spectra. A given IR peak or spectroscopic feature is less likely to be unique for a specific compound, reducing the likelihood in searching a large spectrum library that the search algorithm will select a correct match. For this reason, Raman spectroscopy may be inherently superior for the qualitative analysis of mixtures.

Raman analysis may also be superior in some cases because the selection rules for Raman are different than for IR. The *n*-alkane mixture used in this study provides an excellent demonstration of this superiority. The IR spectra of hexane, heptane, and octane are so similar that it is difficult to tell them apart. The Raman spectra, on the other hand, exhibit unique features for each of these components. Figure 8 shows the Raman spectrum of this mixture below 600 cm^{-1}. The labelled features correspond to C—C—C bending and are distinctive for the three similar hydrocarbons. These modes are known to be weak or nonexistent in the IR.[111]

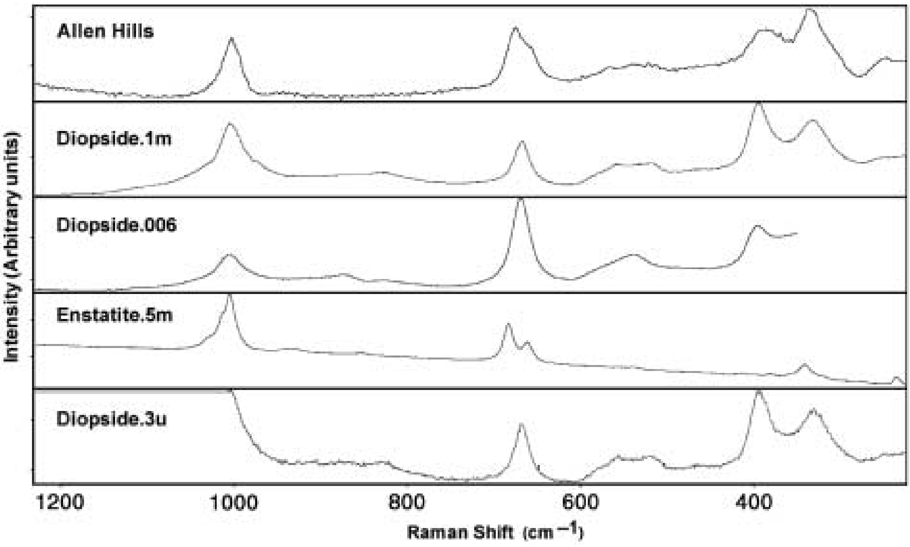

Fig. 7. Spectra of the Allen Hills meteorite and its top four library matches.

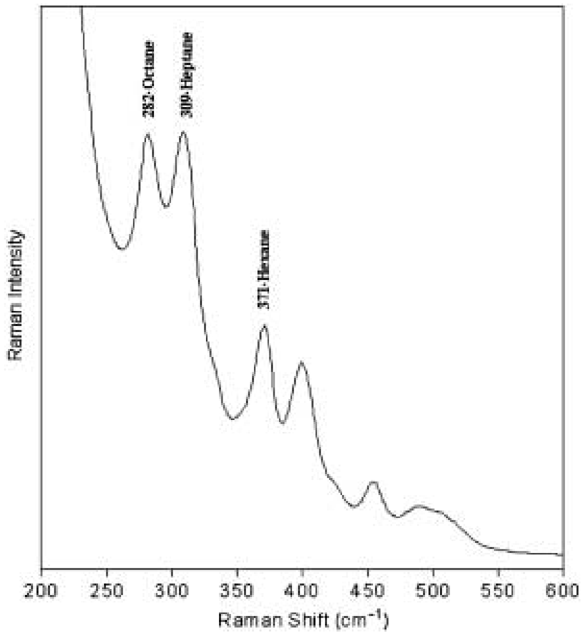

Fig. 8. Low wave-number C–C–C bending modes in the Raman spectrum of a mixture of hexane, heptane, and octane.

Analyses of the xylene mixtures suggest that if a mixture component has approximately the same Raman cross-section as other major constituents, and some resolved spectroscopic features, it should be detectable down to about 15% of total volume. In the xylenes case, the spectrum stripping analysis was able to detect the same analyte at 1% of total volume.

A fair comparison of IR and Raman must also take into consideration the sampling requirements of the techniques. For liquid samples, IR requires that the container be opened so that a few drops can be placed between salt plates or in a cell. Raman spectra can be taken right through the glass bottle, and no contact with the liquid is required. Such sampling could be a great benefit in some applications. For example, the major components in a glass bottle of unknown lab waste could be identified without opening the bottle and exposing technicians to possibly toxic contents. In addition, IR would have been difficult to use in the identification of the mineral inclusion because the sample was buried inside the sapphire and the sapphire would have dominated the resultant spectrum. The Raman microscope, because of its confocal focussing ability, allowed maximization of signal from the inclusion and minimization of signal from the bulk mineral. This permitted identification of the inclusion without having to destroy the sample to perform X-ray powder diffraction analysis.

The probability of success in qualitative analysis of mixtures by spectrum searching can be maximized. We are now in a position to make the following suggestions:

- A marked decrease in match values caused by a low-quality experimental spectrum (here an excessive optical path length resulted in match values falling from 90 to 70) was accompanied by increasingly inappropriate matching compounds. This suggests that path lengths and sample preparation for spectrum searching, particular in IR spectroscopy, should be stringently controlled to minimize misidentification of components.
- Solid sampling is more convenient using Raman instrumentation. The sample can usually be analyzed without modification, as was the Excedrin tablet used in this study. By comparison, preparation of a KBr pellet is significantly more time consuming and difficult, as well as destructive to the sample. Certainly the presence of spectra of solid compounds prepared as hydrocarbon mulls in the IR libraries makes IR library searching somewhat confused. Until this point is resolved, we will state dogmatically that one should exclusively use Raman library searching for qualitative identification of solid samples.
- The AD spectrum matching algorithm was found to give correct identification of mixture components even though the experimental spectrometer had neither the same resolution, nor the same throughput function, as the spectrometer used to produce the library spectra. Other matching algorithms did not identify the components as accurately. For this reason, we suggest that the AD algorithm be used to compare experimental spectra taken on a dispersive Raman spectrometer (or at a different resolution than the library spectra) with a spectrum library database developed on an FT-Raman instrument.

Raman spectroscopy should become a routine analytical technique in the near future. Qualitative identification by Raman can often be accomplished in less than a minute, often with no sample preparation. Successful identification of unknown samples by this means, although not always possible, will avoid the necessity for more time-consuming or destructive analytical techniques. In laboratories where large numbers of samples must be analyzed, a significant increase in throughput could result.

Experimental

All reagents used for analysis were purchased from Aldrich (Milwaukee, WI) and were of 98% or greater purity. The Excedrin tablet was purchased locally. The mixtures were chosen to represent what we anticipated to be spectroscopically distinctive sets of compounds (alcohols plus ketone, arene plus alkarene plus alkane, chlorocarbons), sets of compounds similar in the IR but dissimilar in the Raman (*m*-, *o*-, *p*- xylenes), sets of very similar compounds (hexane, heptane, octane), and at least one solid mixture (Excedrin). The pink sapphire, courtesy of American Gemological Laboratories, was approximately $1.5 \times 0.5 \times 0.5$ cm with a blue inclusion approximately 1 mm beneath the surface of the sapphire. The meteorite sample was a 91 mg portion of the Allan Hills meteorite (ALHA 84 001, 172).

Raman spectra used for comparison with the database were taken using a laboratory-built Raman spectrometer. The laser was a semiconductor diode laser (SDL model 8530, San Jose, CA) and was wavelength stabilized at 785 nm. Laser power was set at 200 mW. The beam was directed through a holographic bandpass filter (Kaiser Optical Systems, Ann Arbor, MI) and then focussed onto the sample. A collection mirror was oriented at 180° to the sample and the collimated scattered light was directed through a holographic notch filter (Kaiser Optical Systems) to eliminate Rayleigh scattered light. The resultant Raman scatter was dispersed using an Oriel polychromator (model MS 257, Stratford, CT) with a 100 μm entrance slit and a 600 groove mm^{-1} grating. Detection was accomplished with a Tektronix 1024×1024 backside-illuminated CCD array interfaced with a CC200 controller (Photometrics, Tucson, AZ). The energy range observed was 140–2230 cm^{-1} (793–951 nm). The dispersion of the polychromator and the size of the CCD array limited this energy window. Instrument control and data collection were implemented by a custom LabVIEW program (National Instruments, Austin, TX). The Raman spectra of all liquid samples were taken through the sides of 16 mL bottles (Kimble, Vineland, NJ) made of KG–33 borosilicate glass. The Excedrin spectrum was taken from an unmodified tablet. The spectra of the sapphire and its inclusion were acquired through a microscope objective, allowing minimization of sapphire contribution to the inclusion's spectrum. The meteorite spectrum was also acquired through an Olympus UPlanFl 10 × infinity-corrected microscope objective for ease of handling such a small sample. The spot size of the laser through the objective was approximately one fifth of the size of the inclusion.

Raman spectra acquired at 514.5 nm were excited with a argon ion laser (543 series, Melles-Griot, Carlsbad, CA) and dispersed through a spectrometer (Spex HR 460, Jobin Yvon, Edison, NJ). A 1152 × 256 pixel CCD (Princeton Instruments, Monmouth Junction, NJ), run by WinSpec software, was used for detection.

IR spectra were taken using an FT-IR absorption spectrometer (model 510P, Nicolet, Madison, WI). Liquid samples were held as capillary films, either between two flat KBr discs or between two flat NaCl discs. The Excedrin tablet was ground to powder, mixed (2% by weight) with dry KBr, and then pressed into a pellet. A background was subtracted from the raw spectrum.

All organic compound library searches were performed using Omnic 5.0 software (Nicolet). The 'HR Aldrich Raman library' containing FT spectra of 14 033 compounds or mixtures with a data point spacing of $1.928\,\mathrm{cm}^{-1}$, was used for Raman searches. The 'HR Aldrich FT-IR Collection II' contained spectra of 18 454 compounds or mixtures present at a data point spacing of $1.928\,\mathrm{cm}^{-1}$ and was used for IR searches. The spectra in the Raman libraries were uncorrected for instrument response. The spectra in the IR libraries were corrected by background subtraction. Experimental spectra were subjected to library searching without modification. The Omnic software automatically matches the spectrum comparison ranges to the limits of the sample and library spectra.

The libraries used in the mineral experiments were created at the University of Arizona using either 514 or 785 nm excitation. Several spectra of the same mineral from different localities or from different crystal orientations were incorporated into the library, resulting in a library containing more than 2000 spectra of over 300 different minerals.[38]

The search algorithms employ Equations (*1*) (correlation) and (*2*) (absolute derivative), to yield 'match' values reported by the Omnic software.[112] An exact overlap of two spectra would give a match value of 100.0. To speed the matching process, square roots of the summation terms were not taken in the usual 'Euclidian distance' fashion. This does not change the order of the match values. Points x_i indicate sample spectra; points r_i indicate library reference spectra.

$$M_{\mathrm{CO}} = \frac{N}{D} \times 100; \qquad (1)$$

where $N = \left(\sum_i y_i s_i\right)^2$,

$$D = \sum_i (y_i)^2 \sum_i (s_i)^2,$$

and $y_i = x_i - x_{i+1}$ and $s_i = r_i - r_{i+1}$

$$M_{\mathrm{AD}} = \left(1 - \frac{N}{D}\right) \times 100; \qquad (2)$$

where $N = \sum_i |y_i - s_i|$,

$$D = \sum_i |y_i|,$$

and $y_i = x_i - x_{i+1}$ and $s_i = r_i - r_{i+1}$.

The CO algorithm removes any baseline offset in the unknown spectrum and uses large peaks and peak shape to find good matches. The AD algorithm, most useful for searching by peak position, removes any baseline offset and emphasizes small peaks and peak shifts. The AD algorithm requires highly ranked hits to possess the major bands in the sample spectrum, but does not seem to reduce the rankings of library spectra containing superfluous bands.[79]

Acknowledgements

The authors thank C.R. Beesley and the American Gemological Laboratory for providing the unusual sample shown in Figure 6 and several other samples included in the mineral database.

References

[1] S. Farquharson, W. W. Smith, R. M. Carangelo, C. R. Brouillette, *Proc. SPIE Int. Soc. Opt. Eng.* **1999**, *3859*, 14.

[2] F. A. Miller, in *Handbook of Vibrational Spectroscopy* (Eds J. M. Chalmers, P. R. Griffiths) **2002**, p. v–vi (John Wiley & Sons: Chichester).

[3] D. A. Skoog, F. J. Holler, T. A. Nieman, *Principles of Instrumental Analysis* **1998**, Ch. 18 (Saunders: Philadelphia, PA).

[4] P. Crews, J. Rodriguez, M. Jaspars (Eds), *Organic Structural Analysis* **1998** (Oxford University Press: New York, NY).

[5] J. D. Ingle, S. R. Crouch, *Spectrochemical Analysis* **1988** (Prentice Hall: Englewood Cliffs, NJ).

[6] J. R. Ferraro, K. Nakamoto, *Introductory Raman Spectroscopy* **1994** (Academic Press: Boston, MA).

[7] P. J. Hendra, in *Handbook of Vibrational Spectroscopy* (Eds J. M. Chalmers, P. R. Griffiths) **2002**, p. 1263 (John Wiley & Sons: Chichester).

[8] S. E. Nave, R. R. Livingston, W. S. Prather, *Fiber Optic Probe for Light Scattering Measurements* **1995**, p. 11 (United States Department of Energy: Washington, DC).

[9] N. T. Kawai, J. A. Janni, *Spectroscopy (Eugene, Oreg.)* **2000**, *15*, 32.

[10] D. L. Stokes, T. Vo-Dinh, *Sens. Actuators B* **2000**, *69*, 28.

[11] H. M. M. Wilson, M. V. Pellow-Jarman, B. Bennett, P. J. Hendra, *Vib. Spectrosc.* **1996**, *10*, 89.

[12] J. Zhao, M. M. Carrabba, F. S. Allen, *Appl. Spectrosc.* **2002**, *56*, 834.

[13] R. L. McCreery, in *Handbook of Vibrational Spectroscopy* (Eds J. M. Chalmers, P. R. Griffiths) **2002**, p. 920 (John Wiley & Sons: Chichester).

[14] J. M. Chalmers, P. R. Griffiths (Eds), *Handbook of Vibrational Spectroscopy* **2002** (John Wiley & Sons: Chichester).

[15] D. L. Gerrard, *Anal. Chem.* **1994**, *66*, 547R.

[16] L. A. Lyon, C. D. Keating, A. P. Fox, B. E. Baker, L. He, S. R. Nicewarner, S. P. Mulvaney, M. J. Natan, *Anal. Chem.* **1998**, *70*, 341R.

[17] S. P. Mulvaney, C. D. Keating, *Anal. Chem.* **2000**, *72*, 145.

[18] J. Workman, Jr, K. E. Creasy, S. Doherty, L. Bond, M. Koch, A. Ullman, D. J. Veltkamp, *Anal. Chem.* **2001**, *73*, 2705.

[19] M. L. McKelvy, T. R. Britt, B. L. Davis, J. K. Gillie, F. B. Graves, L. A. Lentz, *Anal. Chem.* **1998**, *70*, 119R.

[20] C. Henry, *Anal. Chem.* **1997**, *69*, 309A.

[21] D. A. Gilmore, M. B. Denton, *Special Publication—R. Soc. Chem.* **1996**, *194*, 145.

[22] M.-W. Pan, R. E. Benner, L. M. Smith, in *Handbook of Vibrational Spectroscopy* (Eds J. M. Chalmers, P. R. Griffiths) **2002**, p. 490 (John Wiley & Sons: Chichester).

[23] H. Owen, in *Handbook of Vibrational Spectroscopy* (Eds J. M. Chalmers, P. R. Griffiths) **2002**, p. 482. (John Wiley & Sons: Chichester).

[24] M. J. Pelletier, in *Handbook of Vibrational Spectroscopy* (Eds J. M. Chalmers, P. R. Griffiths) **2002**, p. 467 (John Wiley & Sons: Chichester).

[25] D. B. Chase, in *Handbook of Vibrational Spectroscopy* (Eds J. M. Chalmers, P. R. Griffiths) **2002**, p. 522 (John Wiley & Sons: Chichester).
[26] B. M. Cullum, J. Mobley, A. L. Wintenberg, R. A. Maples, D. L. Stokes, T. Vo-Dinh, *Proc. SPIE Int. Soc. Opt. Eng.* **2002**, *4577*, 65.
[27] S. Lowry, in *Handbook of Vibrational Spectroscopy* (Eds J. M. Chalmers, P. R. Griffiths) **2002**, p. 1948 (John Wiley & Sons: Chichester).
[28] W. A. Warr, *Anal. Chem.* **1993**, *65*, 1087A.
[29] W. O. George, H. A. Willis (Eds), *Computer Methods in UV, Visible, and IR Spectroscopy* **1990** (Royal Society of Chemistry: Cambridge).
[30] G. W. Small, *Anal. Chem.* **1987**, *59*, 535A.
[31] R. H. Shaps, J. F. Sprouse, *Ind. Res. Dev.* **1981**, *23*, 168.
[32] K. Hayamizu, *Mizu Kankyo Gakkaishi* **1992**, *15*, 160.
[33] K. Hayamizu, *Kagaku to Sofutowea* **1992**, *14*, 191.
[34] K. Tanabe, K. Hayamizu, S. Ono, *Anal. Sci.* **1991**, *7*, 711.
[35] J. Hiraishi, *JETI* **1984**, *32*, 54.
[36] T. Nishikawa, *Kagaku Kojo* **1983**, *27*, 24.
[37] R. L. McCreery, A. J. Horn, J. Spencer, E. Jefferson, *J. Pharm. Sci.* **1998**, *87*, 1.
[38] R. T. Downs, M. B. Denton, C. J. S. Pommier, S. J. Bailey, R. E. Garoutte, *Raman Gem and Mineral Database* (University of Arizona: Tucson, AZ). A copy of the database, in Omnic format, is available by writing to R. T. Downs or M. B. Denton.
[39] A. Wang, J. Han, L. Guo, J. Yu, P. Zeng, *Appl. Spectrosc.* **1994**, *48*, 959.
[40] D. Laurent, N. Q. Dao, *Proc. Natl. Cent. Sci. Technol. Vietnam* **1998**, *10*, 47.
[41] D. D. Wynn-Williams, H. G. M. Edwards, *Icarus* **2000**, *144*, 486.
[42] L. Burgio, R. J. H. Clark, *Spectrochim. Acta A* **2001**, *57*, 1491.
[43] K. N. Jallad, D. Ben-Amotz, *Mater. Sci. Technol.* **2001**, *17*, 1479.
[44] J. J. Wylde, G. C. Allen, I. R. Collins, *Appl. Spectrosc.* **2001**, *55*, 1155.
[45] D. D. Wynn-Williams, H. G. M. Edwards, E. M. Newton, J. M. Holder, *Int. J. Astrobiol.* **2002**, *1*, 39.
[46] J. Schmitt, T. Udelhoven, D. Naumann, H. C. Flemming, *Proc. SPIE Int. Soc. Opt. Eng.* **1998**, *3257*, 236.
[47] R. Hafenrichter, R. Froyman, H. Labischinski, G. Schiffer, J. Schmitt, T. Udelhoven, in *PCT Int. Appl.* **2002**, 35; WO 0233400.
[48] D. Wynn-Williams, E. M. Newton, H. G. M. Edwards, *Spec. Publ. ESA* **2001**, *496*, 225.
[49] J. Schmitt, T. Udelhoven, T. Lochte, H. C. Flemming, D. Naumann, *AIP Conf. Proc.* **1998**, *430*, 260.
[50] H. G. M. Edwards, M. J. Falk, M. G. Sibley, J. Alvarez-Benedi, F. Rull, *Spectrochim. Acta A* **1998**, *54*, 903.
[51] H. G. M. Edwards, D. E. Hunt, M. G. Sibley, *Spectrochim. Acta A* **1998**, *54*, 745.
[52] B. M. Allen, A. Plant, E. N. Lewis, P. Colarusso, *Proc. NOBCChE* **1999**, *26*, 11.
[53] C. P. Schultz, *Inst. Phys. Conf. Ser.* **2000**, *165*, 53.
[54] C. Johnston, *Vib. Spectrosc.* **1990**, *1*, 87.
[55] C. Johnston, P. R. Graves, *Appl. Spectrosc.* **1990**, *44*, 105.
[56] M. Amme, B. Renker, B. Schmid, M. P. Feth, H. Bertagnolli, W. Dobelin, *J. Nucl. Mater.* **2002**, *306*, 202.
[57] G. Joseph, D. F. Wood, in *PCT Int. Appl.* **2001**, 34; WO 0150112.
[58] L. K. Ballast, T. Z. Hossain, A. Campion, *Proc. SPIE Int. Soc. Opt. Eng.* **2000**, *4182*, 221.
[59] P. Lang, J. E. Katon, *Microbeam Anal.* **1986**, *21*, 47.
[60] M. Simard-Normandin, *Proc. 4th Symp. Part. Surf.* **1995**, 61.
[61] A. W. T. Bristow, A. W. Strawn, Y. Courbariaux, C. Sewell, *Anal. Commun.* **1998**, *35*, 297.
[62] P. Vandenabeele, L. Moens, H. G. M. Edwards, R. Dams, *J. Raman Spectrosc.* **2000**, *31*, 509.
[63] I. M. Bell, R. J. H. Clark, P. J. Gibbs, *Spectrochim. Acta A* **1997**, *53*, 2159.
[64] S. D. Harvey, M. E. Vucelick, R. N. Lee, B. W. Wright, *Forensic Sci. Int.* **2002**, *125*, 12.
[65] M. Boruta, J. Chalmers, *Polym. Prepr. (Am. Chem. Soc., Div. Polym. Chem.)* **2002**, *43*, 1025.
[66] S. Kumar, E. R. Grant, C. M. Duranceau, *Spec. Publ. Automot. Eng.* **2000**, *1542*, 143.
[67] A. H. Kuptsov, *Proc. SPIE Int. Soc. Opt. Eng.* **2000**, *4069*, 224.
[68] R. J. Forrest, R. Devonshire, C. V. Varanasi, T. R. Brumleve, *Appl. Spectrosc.* **2002**, *56*, 1013.
[69] H. Nguyen Thi Diem, D. Phat, P. Plaza, M. Daudon, D. Nguyen Quy, *Proc. SPIE Int. Soc. Opt. Eng.* **1991**, *1525*, 132.
[70] T. D. N. Hong, D. Phat, P. Plaza, M. Daudon, D. Nguyen Quy, *Clin. Chem.* **1992**, *38*, 292.
[71] D. Ball, H. S. T. Driver, R. J. Hutcheon, R. D. Lockett, G. N. Robertson, *S. Afr. J. Sci.* **1994**, *90*, 343.
[72] T. H. Vu, R. Field, *AIP Conf. Proc.* **1987**, *160*, 693.
[73] D. Ball, H. S. T. Driver, R. J. Hutcheon, R. D. Lockett, G. N. Robertson, *Opt. Eng.* **1994**, *33*, 2870.
[74] J. P. Singh, F. Y. Yueh, W. Kao, R. L. Cook, *Appl. Opt.* **1993**, *32*, 894.
[75] O. Jarrett, Jr, R. R. Antcliff, M. W. Smith, A. D. Cutler, G. S. Diskin, G. B. Northam, *Temp. Its Meas. Control Sci. Ind.* **1992**, *6*, 667.
[76] J. Bood, P. E. Bengtsson, M. Alden, *Appl. Phys. B Lasers Opt.* **2000**, *70*, 607.
[77] F. Lipp, E. P. Hassel, J. Janicka, *3rd Int. Symp. Chemical Propulsion*, **1994**, 179.
[78] J. S. Zabinski, N. T. McDevitt, *Raman Spectra of Inorganic Compounds Related to Solid State Tribochemical Studies* **1995** (Wright–Patterson AFB, OH, USA).
[79] *OMNIC 5.0 Program Instruction Manual* (Thermo–Nicolet: Madison, WI).
[80] J. J. Workman, Jr, *Appl. Spectrosc. Rev.* **2001**, *36*, 139.
[81] J. N. Ramsey, *J. Phys. Colloq.* **1984**, 881.
[82] D. C. Smith, A. Barbet, *J. Raman Spectrosc.* **1999**, *30*, 319.
[83] F. A. Nottke, J. A. Kirchoff, J. N. Lucke, *Laser Verification and Authentication Raman Spectrometer (LVARS)* **1999**, p. 13 (Wizard of Ink: Tucson, AZ).
[84] S. A. Broadbent, G. C. Jones, Jr, *Abstr. Pap. Am. Chem. Soc.* **2001**, *221*, CHED-824.
[85] L. P. Choo-Smith, K. Maquelin, T. Van Vreeswijk, H. A. Bruining, G. J. Puppels, N. A. N. Thi, C. Kirschner, D. Naumann, D. Ami, A. M. Villa, F. Orsini, S. M. Doglia, H. Lamfarraj, G. D. Sockalingum, M. Manfait, P. Allouch, H. P. Endtz, *Appl. Environ. Microbiol.* **2001**, *67*, 1461.
[86] C. Petty, N. Cahoon, *Spectrochim. Acta A* **1993**, *49*, 645.
[87] A. N. Davies, P. S. McIntyre, *Comput. Appl. Mol. Spectrosc.* **1995**, 41.
[88] S. D. Brown, S. T. Sum, F. Despagne, B. K. Lavine, *Anal. Chem.* **1996**, *68*, 21.
[89] B. K. Lavine, *Anal. Chem.* **1998**, *70*, 209R.
[90] B. K. Lavine, C. E. Davidson, A. J. Moores, *Vib. Spectrosc.* **2002**, *28*, 83.
[91] K. G. Ray, R. L. McCreery, *Appl. Spectrosc.* **1997**, *51*, 108.
[92] M. Garland, S. Y. Sin, E. Widjaja, L. Yu, *223rd ACS Natl. Meet.* **2002** (Orlando, FL), ANYL–086.
[93] M. Fuller, R. Rosenthal, *Proc. SPIE Int. Soc. Opt. Eng.* **1993**, *2089*, 440.
[94] P. Vandenabeele, A. Hardy, H. G. M. Edwards, L. Moens, *Appl. Spectrosc.* **2001**, *55*, 525.
[95] J. Yao, B. Fan, J.-P. Doucet, A. Panaye, J. Li, C. Sun, S. Yuan, *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 199.
[96] K. Beckenkamp, O. Mandal, K. Molt, in *Ger. Offen.* **2001**, p. 8; D 10026195.
[97] M. Scandone, G. M. Banik, *222nd ACS Natl. Meet.* **2001** (Chicago, IL), CINF.
[98] W. Windig, Chemom. *Intell. Lab. Syst.* **1997**, *36*, 3.

[99] D. R. Lide, Jr, *J. Mol. Struct*. **1986**, *146*, 1.

[100] J. Zhao, R. L. McCreery, *Appl. Spectrosc*. **1997**, *51*, 1687.

[101] G. M. Banik, T. Abshear, *224th ACS Natl. Meet.* **2002** (Boston, MA), CINF.

[102] B. Debska, *Comput. Chem. (Oxford)* **1995**, *19*, 269.

[103] B. Debska, *J. Mol. Struct.* **1993**, *294*, 17.

[104] B. Guzowska-Swider, Z. S. Hippe, *J. Mol. Struct*. **1992**, *275*, 225.

[105] B. Debska, Z. S. Hippe, *J. Mol. Struct*. **1992**, *267*, 261.

[106] B. Debska, *Anal. Chim. Acta* **1992**, *265*, 201.

[107] I. A. Degen, L. Birmingham, G. A. Newman, *Analyst (Cambridge, U.K.)* **1976**, *101*, 212.

[108] J. W. Rydzak, P. N. Bourassa, *Spectroscopy (Eugene, Oreg.)* **2001**, *16*, 18.

[109] J. W. Rydzak, P. N. Bourassa, *Spectroscopy (Eugene, Oreg.)* **2000**, *15*, 16.

[110] Y. N. Miura, N. Sugiura, *Geochim. Cosmochim. Acta* **2000**, *64*, 559.

[111] D. Lin-Vien, N. B. Colthup, W. G. Fateley, J. G. Grassetti, *The Handbook of Infrared and Raman Characteristics Frequencies of Organic Molecules* **1991** (Academic Press: San Diego, CA).

[112] S. Lowry, personal communication.