INTERLABORATORY AND CROSS-INSTRUMENT COMPARISON OF RAMAN SPECTRA OF 96 MINERALS.

M. D. Dyar[1], L. B. Breitenfeld[1], CJ Carey[2], P. Bartholomew[3], T. J. Tague, Jr.[4], P. Wang[4], S. Mertzmann[5], S. A. Byrne[1], M. C. Crowley[1], C. Leight[1], E. Watts[1], J. Caleb Campbell[6], A. Celestian[7], B. McKeeby[8], S. Jaret[8], T. Glotch[8], G. Berlanga[9], and A. K. Misra[9]. [1]Mount Holyoke College, Dept. of Astronomy, South Hadley, MA, 01075, mdyar@mtholyoke.edu, [2]Univ. of Massachusetts Amherst, Amherst MA 01003, [3]Dept. of Biology and Environmental Science, Univ. of New Haven, West Haven, CT 06516, [4]Bruker Optics, Inc., 19 Fortune Dr., Billerica, MA 01821, [5]Franklin and Marshall College, Lancaster, PA, 17603, [6]Western Kentucky Univ., Bowling Green, KY, 42101, [7]Natural History Museum of Los Angeles County, Los Angeles, CA, 90007 [8]Stony Brook Univ., Stony Brook, NY 11794, [8]Univ. of Hawaii, Honolulu, HI 96822.

Introduction: The emerging importance of Raman spectroscopy fo*r in situ* planetary exploration requires an infusion of work into development of databases for mineral identification. Key to that endeavor is an understanding of the characteristics of various commercial and built-from-scratch instruments, their data processing software, and the presumption of consistency of spectra from identical minerals across varying instruments and platforms. This study tests the latter of these assumptions by presenting comparative results on a suite of 96 pure mineral powders from 11 different instruments using an array of geometries and laser energies. Our results assess the usefulness of existing databases built on powder or single crystal data and suggest work needed to align datasets from varying instruments to provide the deepest possible databases for impending Mars exploration use.

Background: Planetary scientists are spoiled by the availability and breadth of reflectance spectral libraries available for commonly used techniques [1-3]. Data from all these sources and time periods can be relied upon to be comparable across the different platforms and instruments used to produce them. *The same cannot be said for Raman spectroscopy.* The example in **Figure 1** shows that differences in peak presence/absence, position, and relative intensity are evident in existing datasets. These differences can be the result of sample factors (e.g. grain size, transparency, crystallographic orientation, grain surface/boundary effects [4]), instrument factors (e.g. laser wavelength, power, and spot size, spectrometer resolution and sensitivity), experimental factors (e.g. angle of incidence and takeoff and the use/absence of polarizers) and data gathering factors (e.g. integration time, averaging, method/frequency of calibration). The questions we seek to answer include: which of these spectral differences are most critical to mineral identification (which is based upon spectral pattern matching), the degree to which spectral differences can be corrected for with data processing techniques, and whether instrument/laboratory corrections can be formulated.

Our work has shown that some of these differences may be mitigated through pre-processing steps such as normalization, smoothing, and squashing [5,6]; the latter uses a transformation function, *f*, that is applied to each wavelength of a spectrum independently to produce a new spectrum with smaller distances between strong and weak spectral features. However, tests of the robustness of these methods depend on the availability of sufficient data for interlaboratory comparisons. This project lays the groundwork by obtaining those data in support of upcoming Mars missions.

Samples: We initially selected 225 pure mineral samples from collections at Mount Holyoke and Smith Colleges, the University of Massachusetts, the authors' collections, and various museums (Harvard Mineralogical Museum and the National Museum of Natural History). Each sample was crushed by hand and carefully handpicked to remove impurities. Subsequent crushing and sieving produced separates with sizes between 38 and 45 μm. Powder x-ray diffraction was undertaken at Franklin and Marshall College; those results indicated that only 96 of the 225 samples were 100% pure single phases. So subsequent measurements proceeded on
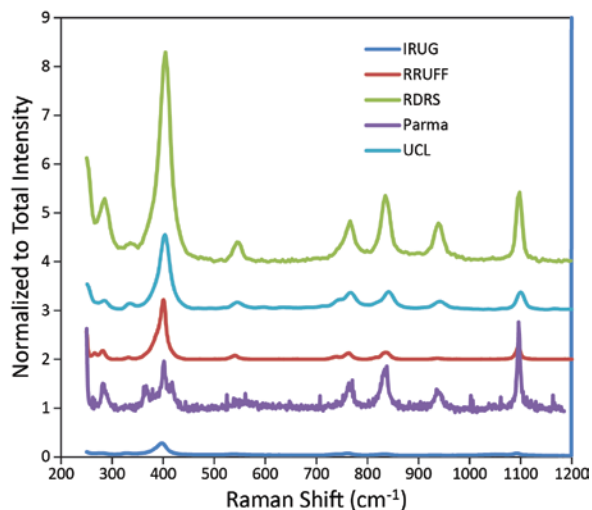

Figure 1. Raman spectra of azurite taken from five different on-line databases: RDRS: http://rdrs.uaic.ro/, UCL: http://www.chem.ucl.ac.uk/resources, RRUFF: rruff.info, IRUG: http://www.irug.org/search-spectral-database, and Parma: http://www.fis.unipr.it/phevix/ramandb.html. All data were normalized to total intensity over this wavelength range. There are clear differences in instrument sensitivity at the various institutions, but the diagnostic peaks associated with azurite are present in all cases.
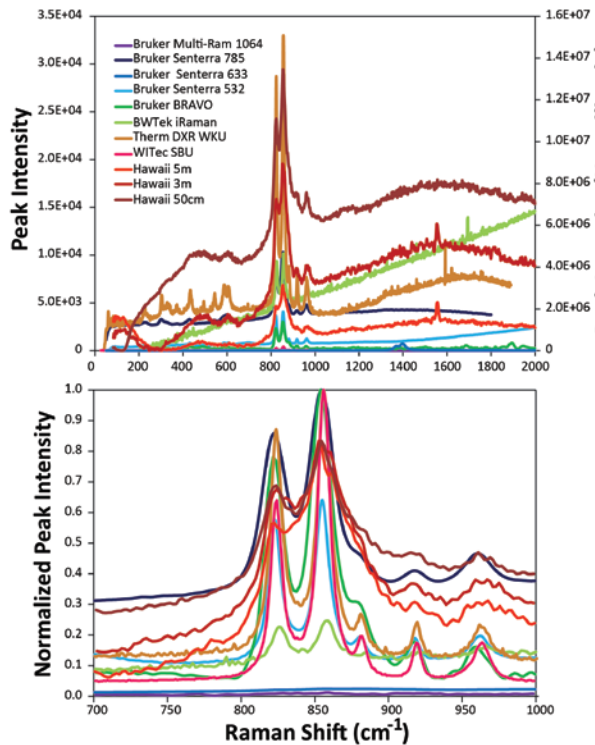
Figure 2. Raw and normalized Raman data collected on forsterite sample 194.

only those 96 samples, which were packaged in 1 dram glass vials for distribution. Sample quantities ranged from ~100-2000 mg.

Methods: To date, we have collected 11 complete datasets on the 96-sample suite; additional measurements are in process at other labs. Instruments used include B&W Tek i-Raman (532 nm laser) and Nano-Ram (785 nm); Bruker Senterra (532, 633, and 785 nm lasers), MultiRAM 1064, and BRAVO (758 and 852 nm lasers simultaneously); an in-house system using a HoloPlex grating, a custom gated thermo-electrically cooled mini-ICCD detector, and 2.5" Meade ETX-125 telescope at the University of Hawaii (532 nm laser)
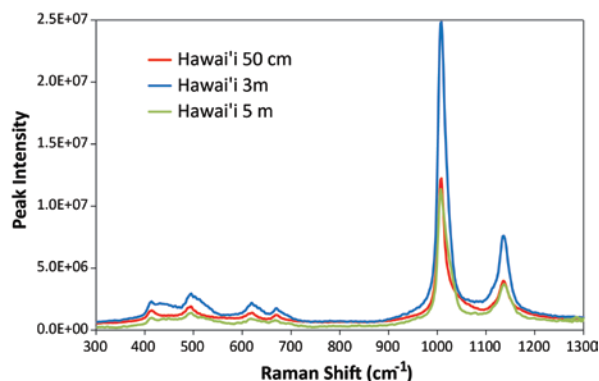

Figure 3. Raw Raman data collected on sample collected at varying stand-off distances on the Hawaii system.

where the samples were studied at 0.5, 3, and 5 cm standoff distances; the WITec alpha 300R Micro-Imaging spectrometer (532 nm laser) at Stony Brook University; and the Thermo DXR dispersive microscope at Western Kentucky University (780 nm laser). Data were processed using standard protocols at each institution. These typically include noise, white light, CCD dark-field, substrate and fluorescence corrections and subtraction of cosmic ray events. Spectra were not baseline corrected so as to allow us to manipulate the algorithm used for data from each instrument.

Results: This data set allows many useful comparisons to be made. It was immediately clear that raw data (without baseline removal or normalization) from different instruments and laser energies are dramatically variable (**Figure 2**); normalization to total intensity does not mitigate those differences, though mineral identification is feasible in data from all spectrometers (ranging in price from $21K to > $0.5 M).

Our data set also allows us to inspect the magnitude of various instrumental parameters, such as sample geometry, laser wavelength, intensity, integration time, and sensitivity. For example, **Figure 3** shows the difference in peak intensity in the Hawaii data as a function of varying stand-off distance. Further study of this data set should enhance our understanding of instrument differences and inform the application of multiple terrestrial databases to interpretation of remote Raman data acquired on other planets such as Mars.

Our group is developing methods that overcome these differences through a generalized approach to using preprocessing steps to make disparate spectra of the same mineral more similar [5,6]. It uses different types of intensity normalization, smoothing using parameterizations of the Savitzky–Golay filter, and squashing, a transformation function that is applied to each wavelength of a spectrum independently to produce a new spectrum with smaller distances between strong and weak spectral features. This will enable successful integration of multiple Raman mineral data sets an improve the accuracy of mineral identification in diverse applications.

References: [1] Abrams M. and Hook S. J. (2000) *IEEE Trans. Geosci. Remote Sens.*, 33, 692-699. [2] Clark R. N. et al. (1990) *J. Geophys. Res.*, 95, B8, 12653-12680. [3] Pieters C. M. et al. (1983) *J. Geophys. Res.*, 88, 9534-9544. [4] Haskin L. M. et al. (1997) *LPS XXVII*, Abstract #1454. [5] Carey CJ et al. (2015) *J. Raman Spectrosc.*, 46, 894-903. [6] Carey CJ et al. (2015) *AI in Space, IJCAI 2015*.