

DESCRIPTION
METHODS AND SYSTEMS FOR RAMAN SPECTRA-BASED IDENTIFICATION
OF CHEMICAL COMPOUNDS

CROSS-REFERENCE TO RELATED APPLICATIONS

5 **[0001]** This application claims the benefit of priority to U.S. Provisional Patent Application Serial No. 63/156,846, filed March 4, 2021, hereby incorporated by reference in its entirety.

FIELD

10 **[0002]** This application relates to the classification and identification of chemical compounds using the Raman spectra of the chemical compounds, and more particularly, by comparing the measured Raman spectra of the chemical compounds to a reference library of *in-silico* simulated Raman spectra.

INTRODUCTION

15 **[0003]** Raman spectroscopy is a spectral measurement technique where light incident on a sample is inelastically scattered, i.e., the frequency of the scattered light is different from the frequency of the incident light. The change in frequency, or equivalently change in energy, is due the interaction of the light with the low energy modes of the chemical bonds of the sample. The Raman spectrum of the sample can be determined from the scattered light, and can be used to identify and classify the sample.

SUMMARY

20 **[0004]** The following summarizes various embodiments of the present disclosure to provide a basic understanding of the discussed technology. This summary is not an extensive overview of all contemplated features of the disclosure and is intended neither to identify key or critical elements of all embodiments of the disclosure nor to delineate the scope of any or all embodiments of the disclosure. Its sole purpose is to present some concepts of one or more
25 embodiments of the disclosure in summary form as a prelude to the more detailed description that is presented later.

30 **[0005]** According to a first aspect of the present invention, there is provided a method, comprising: obtaining a Raman spectrum, of an unknown chemical compound, extracted from a Raman spectroscopic measurement of the unknown chemical compound; generating a Raman barcode corresponding to peaks of the Raman spectrum based at least in part on intensity levels of the peaks of the Raman spectrum; comparing the obtained Raman spectrum to a reference simulated Raman spectrum of a plurality of simulated Raman spectra contained in a simulated Raman spectrum library, each simulated Raman spectra generated via a

quantum-mechanical computation of a known chemical compound, wherein the comparing includes comparing the generated Raman barcode to a reference simulated Raman barcode corresponding to the reference simulated Raman spectrum of the plurality of simulated Raman spectra; and identifying an identity of the unknown chemical compound based on a comparison of the obtained Raman spectrum to the reference simulated Raman spectrum.

[0006] According to a second aspect of the present invention, there is provided a system, comprising: a non-transitory memory storing instructions; and one or more hardware processors coupled to the non-transitory memory and configured to read the instructions from the non-transitory memory to cause the system to perform operations comprising: obtaining a Raman spectrum, of an unknown chemical compound, extracted from a Raman spectroscopic measurement of the unknown chemical compound; generating a Raman barcode corresponding to peaks of the Raman spectrum based at least in part on intensity levels of the peaks of the Raman spectrum; comparing the obtained Raman spectrum to a reference simulated Raman spectrum of a plurality of simulated Raman spectra contained in a simulated Raman spectrum library, each simulated Raman spectra generated via a quantum-mechanical computation of a known chemical compound, wherein the comparing includes comparing the generated Raman barcode to a reference simulated Raman barcode corresponding to the reference simulated Raman spectrum of the plurality of simulated Raman spectra; and identifying an identity of the unknown chemical compound based on a comparison of the obtained Raman spectrum to the reference simulated Raman spectrum.

[0007] There is disclosed a non-transitory computer-readable medium (CRM) having stored thereon computer-readable instructions executable to cause performance of operations. In various embodiments, the operations comprise obtaining a Raman spectrum, of an unknown chemical compound, extracted from a Raman spectroscopic measurement of the unknown chemical compound. In various embodiments, the operations further comprise comparing the obtained Raman spectrum to a reference simulated Raman spectrum of a plurality of simulated Raman spectra contained in a simulated Raman spectrum library. In various embodiments, each simulated Raman spectra is generated via a quantum-mechanical computation of a known chemical compound. The operations further comprise identifying an identity of the unknown chemical compound based on a comparison of the obtained Raman spectrum to the reference simulated Raman spectrum.

[0008] Other features, and embodiments of the present disclosure will become apparent to those of ordinary skill in the art, upon reviewing the following description of specific, exemplary embodiments of the present disclosure in conjunction with the accompanying figures. While features of the present disclosure may be discussed relative to certain
5 embodiments and figures below, all embodiments of the present disclosure can include one or more of the advantageous features discussed herein. In other words, while one or more embodiments may be discussed as having certain advantageous features, one or more of such features may also be used in accordance with the various embodiments of the disclosure

CLAIMS

1. A method, comprising:
 - obtaining a Raman spectrum, of an unknown chemical compound, extracted from a Raman spectroscopic measurement of the unknown chemical compound;
 - generating a Raman barcode corresponding to peaks of the Raman spectrum based at least in part on intensity levels of the peaks of the Raman spectrum;
 - comparing the obtained Raman spectrum to a reference simulated Raman spectrum of a plurality of simulated Raman spectra contained in a simulated Raman spectrum library, each simulated Raman spectra generated via a quantum-mechanical computation of a known chemical compound, wherein the comparing includes comparing the generated Raman barcode to a reference simulated Raman barcode corresponding to the reference simulated Raman spectrum of the plurality of simulated Raman spectra; and
 - identifying an identity of the unknown chemical compound based on a comparison of the obtained Raman spectrum to the reference simulated Raman spectrum.
2. The method of claim 1, wherein the simulated Raman spectrum library excludes experimental Raman spectra obtained by performing Raman spectroscopic measurements on chemical compounds.
3. The method of claim 1, wherein the comparing includes applying a scaling algorithm to the obtained Raman spectrum and/or the reference simulated Raman spectrum to correct a wavenumber offset between the obtained Raman spectrum and/or the reference simulated Raman spectrum.
4. The method of claim 1, wherein the identifying includes determining that the unknown chemical compound matches a chemical compound from which the reference simulated Raman spectrum is computed when a barcode overlap level between the generated Raman barcode and the reference simulated Raman barcode exceeds a barcode overlap threshold.
5. The method of claim 4, wherein the barcode overlap level indicates a number of bars located at a first wavenumber location of the generated Raman barcode and at a second wavenumber location of the reference simulated Raman barcode, the first wavenumber location and the second wavenumber location being at least substantially equal to each other.

6. The method of claim 1, wherein the quantum-mechanical computation is based on density-functional theory (DFT).
7. The method of claim 1, wherein the generating includes including a bar in the Raman barcode, corresponding to a peak of the peaks, when an intensity level of that peak exceeds a peak intensity threshold.
8. The method of claim 7, wherein the peak intensity threshold is an average of the intensity levels of the peaks.
9. The method of claim 1, wherein the generating includes generating the Raman barcode corresponding to the peaks of the Raman spectrum based at least in part on wavenumber locations of the peaks of the Raman spectrum.
10. A system, comprising:
 - a non-transitory memory storing instructions; and
 - one or more hardware processors coupled to the non-transitory memory and configured to read the instructions from the non-transitory memory to cause the system to perform operations comprising:
 - obtaining a Raman spectrum, of an unknown chemical compound, extracted from a Raman spectroscopic measurement of the unknown chemical compound;
 - generating a Raman barcode corresponding to peaks of the Raman spectrum based at least in part on intensity levels of the peaks of the Raman spectrum;
 - comparing the obtained Raman spectrum to a reference simulated Raman spectrum of a plurality of simulated Raman spectra contained in a simulated Raman spectrum library, each simulated Raman spectra generated via a quantum-mechanical computation of a known chemical compound, wherein the comparing includes comparing the generated Raman barcode to a reference simulated Raman barcode corresponding to the reference simulated Raman spectrum of the plurality of simulated Raman spectra; and
 - identifying an identity of the unknown chemical compound based on a comparison of the obtained Raman spectrum to the reference simulated Raman spectrum.

11. The system of claim 10, wherein the simulated Raman spectrum library excludes experimental Raman spectra obtained by performing Raman spectroscopic measurements on chemical compounds.
12. The system of claim 10, wherein the comparing includes applying a scaling algorithm to the obtained Raman spectrum and/or the reference simulated Raman spectrum to correct a wavenumber offset between the obtained Raman spectrum and/or the reference simulated Raman spectrum.
13. The system of claim 10, wherein the identifying includes determining that the unknown chemical compound matches a chemical compound from which the reference simulated Raman spectrum is computed when a barcode overlap level between the generated Raman barcode and the reference simulated Raman barcode exceeds a barcode overlap threshold.
14. The system of claim 13, wherein the barcode overlap level indicates a number of bars located at a first wavenumber location of the generated Raman barcode and at a second wavenumber location of the reference simulated Raman barcode, the first wavenumber location and the second wavenumber location being at least substantially equal to each other.
15. The system of claim 10, wherein the quantum-mechanical computation is based on density-functional theory (DFT).
16. The system of claim 10, wherein the generating includes including a bar in the barcode corresponding to a peak of the peaks when an intensity level of that peak exceeds a peak intensity threshold.
17. The system of claim 16, wherein the peak intensity threshold is an average of the intensity levels of the peaks.
18. The system of claim 10, wherein the generating includes generating the Raman barcode corresponding to the peaks of the Raman spectrum based at least in part on wavenumber locations of the peaks of the Raman spectrum.