

## Prediction of hydrocarbon molecular structure using infrared, Raman, and NMR spectroscopy

Collins Obiosa-Maife and J.M. Shaw\*

Department of Chemical and Materials Engineering, University of Alberta, Edmonton, Alberta, T6G 2V4

(\* corresponding author: [jmshaw@ualberta.ca](mailto:jmshaw@ualberta.ca))

Crude oils and other hydrocarbon resources are complex organic mixtures comprising  $10^5$  to  $10^6$  different molecules [1]. Understanding the molecular composition of heavy oil and bitumen is critical for process design calculations and for the selection of production and refining processes, reaction schemes, and conditions that optimize their economic value.

Asphaltenes, for example, are ill-defined on a molecular basis. Although aspects of the molecular building blocks of asphaltenes are known, their nominal molecular structure has been an ongoing subject of debate. Sheremata et al. [2] and Boek et al. [3] generated divergent quantitative molecular representations of asphaltenes using a combination of NMR, elemental and molecular weight analysis obtained from Athabasca bitumen. Sheremata et al. reported archipelago representations (collection of aromatic "islands" connected by aliphatic "bridges"); while Boek et al. who reprocessed Sheremata et al.'s data with a modified molecule assembly algorithm reported pericondensed molecular representations (characterized by a large single aromatic core of carbon atoms from which aliphatic chains extend).

As both pericondensed and archipelago molecular structures can be proposed on the basis of the same analytical data ( $^1\text{H}$  and  $^{13}\text{C}$  NMR, mass spectroscopy, and elemental composition) the hypothesis that  $^1\text{H}$  and  $^{13}\text{C}$  NMR provide ambiguous information concerning large molecules is readily posed. Further, the hypothesis that analytical techniques such as infrared and Raman spectroscopy, which provide information on larger subunits present in molecules, might be a better choice for unambiguous identification of molecules or representative structures present in molecules comprising ill-defined hydrocarbons is readily framed.

The present work assesses the feasibility of using multiple forms of spectroscopy, e.g.: infrared, Raman, and Nuclear Magnetic Resonance (NMR) spectra as a basis for unambiguous or less ambiguous molecule construction of large hydrocarbon molecules. For this proof of concept investigation, all spectra are computed using Density Functional Theory (DFT) and the B3LYP/6-311G basis set. The molecular structure of large molecules is predicted on the basis of a spectral library of small molecules. As a key potential application of the results of this study includes the development of molecular representations for ill-defined hydrocarbons, the large

molecules (Fig.1) comprised compounds with two multiple-ringed polynuclear aromatic hydrocarbons (PAHs) or naphthenic hydrocarbons connected by aliphatic chains of varying lengths. The ringed sub units were typically but not necessarily drawn from the molecules comprising the library. A least squares optimization algorithm was developed to determine the contribution of molecules in the library to the large molecules. More than twenty cases were evaluated.

$^1\text{H}$  and  $^{13}\text{C}$  NMR spectroscopy failed to identify molecular subunits in any of the large molecules unambiguously, even with a heavily constrained library of small molecules. By contrast, IR and Raman spectroscopy identified 40% and 65% respectively and jointly more than 80 % of the aromatic/naphthenic groups and the relative abundance of subunits was semi- quantitative. Further, large compounds not comprising aromatic or naphthenic sub units present in the library were not misidentified. However, the aliphatic chain length was poorly defined.

NMR spectroscopy appears to be over-interpreted in data sets used to estimate molecular structures for ill-defined hydrocarbons, and this may account for the ambiguity of molecular structures proposed for the same samples and data sets. IR and Raman spectroscopy appear to be preferred particularly for the identification of ringed subunits where spectral addition algorithms are employed. Inclusion of additional forms of spectroscopy is expected to further reduce the ambiguity of molecule construction algorithms.

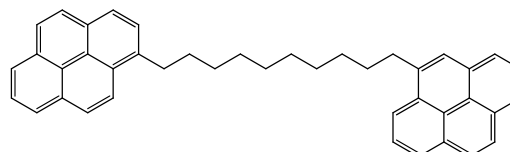


Fig. 1. Model large hydrocarbon compound consisting of two pyrene groups connected by an n-decane bridge.

### References

- [1] Wiehe, I. A.; Liang, K. S., *Fluid Phase Equilibria* 1996, 117, (1-2), 201-210.
- [2] Sheremata, J. M.; Gray, M. R.; Dettman, H. D.; McCaffrey, W. C., *Energy and Fuels* 2004, 18, (5), 1377-1384.
- [3] Boek, E. S.; Yakovlev, D. S.; Headen, T. F., *Energy & Fuels* 2009, 23, (3), 1209-1219.