Predicting accurate fluorescent spectra for high molecular weight polycyclic aromatic hydrocarbons using density functional theory

Jacob Powell, Emily C. Heider, Andres Campiglia, James K. Harper *

Department of Chemistry, University of Central Florida, 4111 Libra Drive, Orlando, FL 32816, United States

A R T I C L E   I N F O

A B S T R A C T

The ability of density functional theory (DFT) methods to predict accurate fluorescence spectra for polycyclic aromatic hydrocarbons (PAHs) is explored. Two methods, PBE0 and CAM-B3LYP, are evaluated both in the gas phase and in solution. Spectra for several of the most toxic PAHs are predicted and compared to experiment, including three isomers of $C_{24}H_{14}$ and a PAH containing heteroatoms. Unusually high-resolution experimental spectra are obtained for comparison by analyzing each PAH at 4.2 K in an n-alkane matrix. All theoretical spectra visually conform to the profiles of the experimental data but are systematically offset by a small amount. Specifically, when solvent is included the PBE0 functional overestimates peaks by 16.1 ± 6.6 nm while CAM-B3LYP underestimates the same transitions by 14.5 ± 7.6 nm. These calculated spectra can be empirically corrected to decrease the uncertainties to 6.5 ± 5.1 and 5.7 ± 5.1 nm for the PBE0 and CAM-B3LYP methods, respectively. A comparison of computed spectra in the gas phase indicates that the inclusion of n-octane shifts peaks by +11 nm on average and this change is roughly equivalent for PBE0 and CAM-B3LYP. An automated approach for comparing spectra is also described that minimizes residuals between a given theoretical spectrum and all available experimental spectra. This approach identifies the correct spectrum in all cases and excludes approximately 80% of the incorrect spectra, demonstrating that an automated search of theoretical libraries of spectra may eventually become feasible.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Polycyclic aromatic hydrocarbons (PAHs) are compounds composed of two or more aromatic rings containing only carbon and hydrogen. These products are found in meteors [1], interplanetary dust particles [2], and interstellar grains [3]. In our terrestrial environment, PAHs often occur as products of incomplete combustion. PAHs formed by the birth of stars are ubiquitous in nature, and pose great interest in a variety of fields – from measuring the rate of star formation [4] to possible uses as semiconductors [5]. While the presence of PAHs in space inspires studies into the origin of biological molecules and mechanism for PAH growth [6], in Earth's ecosystem real concerns exist regarding the interaction of PAHs with biomolecules. Epidemiological studies have revealed high toxicity and carcinogenicity for many of these compounds and sixteen are now included on the Environmental Protection Agencies (EPA) list of Priority Pollutants for routine monitoring. Anthropogenic sources of PAHs include wood and coal burning, and their presence in coal and crude oil insures their extraction from the ground and potential for distribution. These factors make PAHs one of the most omnipresent pollutants worldwide. Developing sensitive methods to quantify and identify PAHs in air, water, and soil are therefore of critical importance.

While remote detection of PAHs has revolved around measured and calculated vibrational spectra [7], work in the Campiglia laboratory and by others has [8] exploited the fluorescence [9] and phosphorescence [10] emission of PAHs to quantify and identify PAHs in terrestrial environmental samples including water [11] and soil [12]. The use of luminescence provides several advantages over vibrational spectroscopy including improved sensitivity and specificity with fluorescence detection of PAHs routinely providing part-per-trillion limits of detection. Specificity in fluorescence can be further enhanced with time resolved line-narrowing spectroscopy. This process involves diluting PAH samples in solvents with minimal solute interaction (generally n-alkane solvents) then freezing to cryogenic temperatures (liquid nitrogen or helium) to produce vibrationally resolved fluorescence emission spectra with sufficient resolution to identify individual vibronic transitions. This cryogenic technique is generally referred to as Shpol'skii spectroscopy [13]. Additional specificity can be attained with the measurement of fluorescence lifetimes [14].

In previous work, we have measured Shpol'skii spectra for several of the EPA PAHs and most have been found to exhibit unique spectra [11,12]. These significant differences between spectra mean that complex mixtures of PAHs (e.g. environmental samples) can potentially be characterized. Moreover, since at least some unique emission frequencies are observed for most PAHs, analysis of mixtures is possible without prior chromatographic separation in analyses requiring only a few minutes [15]. Recently, this methodology has been employed to successfully characterize multiple PAHs in complex mixtures available as "reference materials" from the National Institute of Standards and Technology (NIST) [15]. One of the challenges identified in these studies was the discovery that many of the samples contained emission peaks at frequencies not associated with any of the 16 EPA-PAH standards. This is not surprising as the EPA-PAHs are only a small fraction of the total number of PAHs known to exist in environmental samples. Further limiting the monitoring of environmental PAHs is the modest commercial availability of certain pure standards. Many cases also exist where the cost of pure PAH standards is prohibitive. This is particularly true for dibenzopyrene isomers with molecular weight 302. Several dibenzopyrenes are more toxic than benzo[a]pyrene, which is the most toxic PAH in the EPA priority pollutants list [16]. Synthesis is always an alternative to overcome the lack of commercial standards. Unfortunately, the existence of numerous isomers with the same molecular weight and very similar molecular structures often challenges the synthesis and purification of individual standards.

Very recently a new alternative to synthesis has become available for identifying PHAs when standards are unavailable. This approach relies on theoretical methods to calculate vibrationally resolved emission spectra for candidate structures [17]. Historically, accurate descriptions of electronic excited states were difficult to obtain because conventional density functional theory (DFT), i.e. Hohenberg-Kohn, was limited to ground states [18]. The development of time-dependent DFT (TDDFT) [19] extends conventional DFT to excited states and allows for studies of emission spectra. A second obstacle arose from the large number of excited vibrational states that must be evaluated in all but the smallest molecules. A key breakthrough in treating these transitions came from the recognition that the majority of the possible vibronic transitions do not, in fact, contribute to the spectrum. A set of rules has recently been developed to identify non-negligible transitions, allowing calculations to be restricted to only these levels [20]. This development has led to a powerful process for computing DFT emission spectra in larger molecules and a more complete description of this approach is given elsewhere [17]. This methodology allows the influence of temperature and a wide range of solvents to be included. Perhaps most significantly, this computational methodology has now been implemented into the widely available computational package, Gaussian, and results in a facile prediction process that non-experts can employ to effectively predict spectra in larger molecules. These techniques have the potential to provide a unique path to structure of higher molecular weight PAHs in environmental samples.

At the present time, these computational methods have been employed to predict spectra for several fluorescent molecules [21], but less has been done to evaluate the accuracy of these methods when computing emission spectra for PAHs. To our knowledge, computational methods have been employed to study a total of 15 PAHs [20a,22] and 3 of these have included solvent effects. Of the predicted spectra, 8 involve PAHs on the EPA list and most have molecular weights of 228 or less. Here, one major aim is to evaluate the ability of these DFT methods to calculate accurate emission spectra for higher molecular weight PAHs in a manner that includes solvent effects. A secondary objective is to develop methods for comparing these predicted spectra to

unusually high-resolution experimental data obtained at 4.2 K with the aim of consistently identifying the correct PAHs based on statistical figures-of-merit. Compounds selected for analysis are shown in Fig. 1 and include benzo[a]pyrene, dibenzo[a,l] pyrene, dibenzo[a,e]pyrene, dibenzo[a,i]pyrene, and (−)-7R,8S,9 R,10S-7,8,9,10-tetrahydroxy-7,8,9,10-tetrahydrobenzo[a]pyrene (referred to hereinafter as "benzo[a]pyrene tetrol"). Benzo[a] pyrene was selected for analysis because it is one of the most toxic compounds on the EPA list [23]. However, dibenzo[a,l]pyrene and dibenzo[a,i]pyrene are estimated to be roughly 10 times more toxic that benzo[a]pyrene while dibenzo[a,e]pyrene is considered equivalent in toxicity to benzo[a]pyrene [24]. Currently, however, none of these dibenzopyrenes are included on the EPA list. Since these compounds are presently recommended for monitoring in the European Union [24] and may be monitored by the EPA in the future, they were also studied here. The inclusion of the three dibenzopyrene isomers is also significant because it allows for a rigorous evaluation of the DFT methods' ability to distinguish structurally similar isomers. Likewise, benzo[a]pyrene tetrol tests the ability of the theoretical methods to predict spectra of polycyclic aromatic compound with heteroatoms (e.g. O or S) in their molecular structure. Benzo[a]pyrene-7,8,9,10-tetrol is a well-known metabolic product of benzo[a]pyrene often used as a biomarker of human exposure to PAHs [25].

In the following discussion, two theoretical methods, PBE0 [26] and CAM-B3LYP [27], are evaluated for their ability to accurately reproduce experimental spectra. Because each of these methods is found to exhibit systematic errors, an empirical correction is introduced. These adjusted theoretical spectra are ultimately compared to experimental data using a process that minimizes residuals. The correct experimental spectrum is found to match the predicted spectra in all cases and the majority of the incorrect spectra are eliminated as possible matches. All spectra are computed in an environment that simulates solvent and in all cases
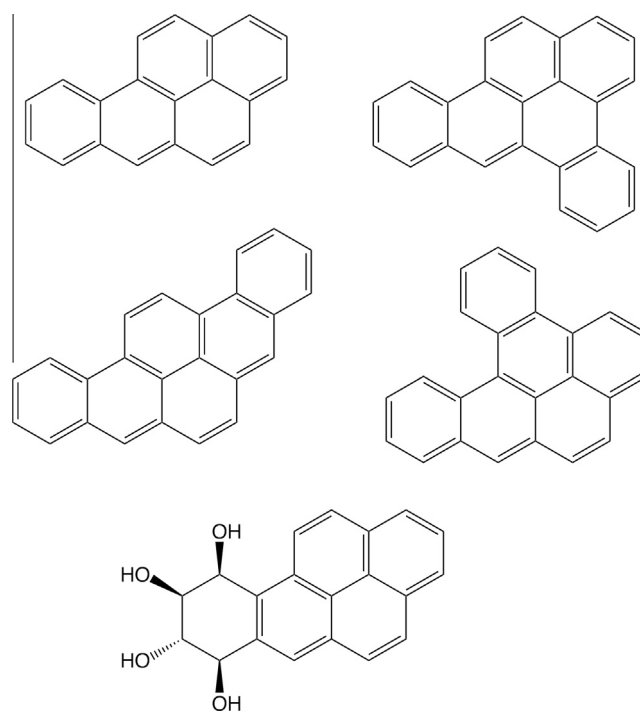

Fig. 1. Structures studied herein include benzo[a]pyrene (top left), dibenzo[a,e] pyrene (top right), dibenzo[a,i]pyrene (middle left), dibenzo[a,l]pyrene (middle right) and Benzo[a]pyrene–7,8,9,10-tetrol (bottom).