Towards a Complete Electronic Database of PAHs and the Identification of Resolved DIBs

Xiaofeng Tan
Space Science Division, NASA Ames Research Center, Moffett Field, CA 94040, E-mail: x.tan@jhu.edu
The Diffuse Interstellar Bands (DIBs)

- Interstellar absorption bands first discovered by Heger in 1921. More than 300 diffuse bands in the visible and near infrared.
- No carriers have ever been identified!
- Providing crucial information on the chemical compositions of galaxies and interstellar medium, on the energy balance, chemical evolution of the universe and origin of life...
The PAH–DIB Hypothesis

- PAHs (neutral or ionic) may be responsible for some of the DIBs.
- Abundance of elements in the universe: H, He, O, C, N, Ne...
- PAHs are remarkably photo stable.
- PAHs have been identified in meteorites.
- PAHs and the infrared emission bands.
We Are Lottery Players

• Scientists are lottery players who are hoping to hit the Jack Pot by chance.
• Unfortunately, up to date only very limited PAHs have been studied either experimentally or theoretically: ~ 0.1% PAHs containing up to 10 fused benzene rings.
• A big question to ask therefore is: “exactly how many PAHs are out there and what PAHs should we study?”
Computer Enumeration of PAHs

- The ideal carbon skeleton of a PAH is a "polyhex" that consists of $h$ fused benzene rings.
- By "ideal", we mean all benzene rings in the polyhex are identical regular hexagons.
- How many polyhexes exist for a given number of $h$ hexagons? Harary offered $100 for the solution to this difficult problem in 1968.
A Working Algorithm

- Cell growth: polyhexes with $h$ hexagons are generated from polyhexes with $h - 1$ hexagons.
- Each time a new polyhex is generated, we transform it using each of the 12 symmetry operations.
- Each transformed polyhex is then converted into its SIR and compared to existing SIRs – the most time consuming step.
- A balanced search tree is used to store the SIRs.
The Honeycomb Grid

- All polyhexes with $h$ hexagons can be contained in a $h \times h$ block honeycomb grid.
- Each hexagon, or “cell”, is represented by one grid point located in its center of mass.

A $4 \times 4$ honeycomb grid.
The Index Representation

- A polyhex can be represented by the cell numbers of its constitution cells, denoted as $C(i, j, ...)$, where $i, j, ...$ are the cell numbers.
- Index Representation: the cells of the T-shape are represented by their indices.
- Standard Index Representation (SIR).

$$C(1,5,2,6) \equiv \begin{pmatrix} 1 & 2 & 1 & 2 \\ 1 & 1 & 2 & 2 \end{pmatrix} \equiv \begin{pmatrix} 1 & 1 & 2 & 2 \\ 1 & 2 & 1 & 2 \end{pmatrix}$$
Cell Connectivity

- Each cell (grid point) is connected with six cells (grid points):
Symmetry Operations

- One translation.
- Six rotations: $E$, $C_6$, $C_6$, $C_3$, $C_3$, $C_2$.
- Six reflections:

The six mirrors of the honeycomb grid: $1 - \sigma_x$, $2 - \sigma_x'$, $3 - \sigma_y$, $4 - \sigma_y$, $5 - \sigma_x$, $6 - \sigma_y$. 
Transformation Matrices

\[
M(E) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad M(C_6) = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}, \quad M(C_{-6}) = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix},
\]

\[
M(C_3) = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}, \quad M(C_{-3}) = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}, \quad M(C_2) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},
\]

\[
M(\sigma_x) = \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix}, \quad M(\sigma'_x) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \quad M(\sigma''_x) = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix},
\]

\[
M(\sigma_y) = \begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix}, \quad M(\sigma'_y) = \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix}, \quad M(\sigma''_y) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.
\]
## Number of Polyhexes & Cost

<table>
<thead>
<tr>
<th>$h$</th>
<th># PHs</th>
<th>Time (s)</th>
<th>$h$</th>
<th># PHs</th>
<th>Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>&lt; 1</td>
<td>7</td>
<td>333</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>&lt; 1</td>
<td>8</td>
<td>1448</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>&lt; 1</td>
<td>9</td>
<td>6572</td>
<td>29</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>&lt; 1</td>
<td>10</td>
<td>30490</td>
<td>167</td>
</tr>
<tr>
<td>5</td>
<td>22</td>
<td>&lt; 1</td>
<td>11</td>
<td>143552</td>
<td>746</td>
</tr>
<tr>
<td>6</td>
<td>82</td>
<td>&lt; 1</td>
<td>12</td>
<td>683101</td>
<td>6527</td>
</tr>
</tbody>
</table>
Polyhexes with $h = 4$

$C(1,2,3,4)$  $C(3,4,5,6)$  $C(1,2,3,7)$

$C(2,3,5,7)$  $C(1,2,3,6)$  $C(2,3,5,6)$  $C(2,6,7,9)$
“Model geometry” – H-C-C angles: $2\pi/3$; C-C bond: 1.397 Å; C-H bond: 1.084 Å.

For geometry optimization, the Austin Model 1 (AM1) Hamiltonian is used.

For electronic excitation energies, the Zerner’s Intermediate Neglect of Diatomic Differential Overlap (ZINDO) method is used.

The transition energies calculated at the model geometries are found to be better than those calculated at the geometries optimized at the AM1 level of theory.
Calibration of the Database

- The current database contains all PAHs up to $\hbar = 10$ (~40,000 PAH molecules).
- Sample space: 10 closed-shell neutral PAHs (18 rotational constants and 15 transition energies).
- Rotational constants: maximum relative error is 0.76% and the standard deviation is 0.39%.
- Electronic transitions energies: the mean is $-1059 \text{ cm}^{-1} (-0.13 \text{ eV})$; the largest deviation is $-2757 \text{ cm}^{-1} (-0.34 \text{ eV})$; standard deviation is $1322 \text{ cm}^{-1} (0.16 \text{ eV})$. 
Where the Database Was Built
Electronic Properties of PAHs

- The rings are in the same plane in the electronic ground state (Cs).
- Conjugated $\pi$ (A'') systems.
- Dominated by the $\pi \rightarrow \pi^*$ and $\sigma \rightarrow \sigma^*$ transitions.
- Close shell: electronic transitions are in the UV and move to longer wavelength as the size grows.
- Open shell: electronic transitions are in the visible and near IR.

![Benzoperylene](image)
Molecular Symmetry

- In general, PAHs are asymmetric tops. $I_c$ is perpendicular to the molecular plane and $I_a$ and $I_b$ lie in the molecular plane.
- Two independent rotational constants since $I_c = I_a + I_b$.
- If there exists a $C_3$ axis, the molecule is a oblate.
- Type-A and type-B band profiles.
Simulated vibronic band profiles using the rotational constants of the $S_0$ and $S_1$ electronic states of perylene [Tan & Salama, J.C.P. 122, 084318, (2005)]. The spectra are convoluted with a Lorentizan function with $\gamma = 0.05$ cm$^{-1}$. $T_{\text{rot}} = 20$ K.
Deciphering The λ6614 DIB

(a) Comparison of the λ6614 DIB (HD149757, Galazutdinov et al 2002) and simulated type-A profiles of planar PAHs, Trot = 54 K, A'' = 0.01 cm⁻¹, γ = 0.4 cm⁻¹. In the three simulations, A' = A'' = A, B' = B'' = B. B/A ratios are shown on the figure. (b). Comparison of the λ6614 DIB and simulated (TDDFT) profiles of 7RT113 and 8RT702.
Conclusions

- It is possible to build a “complete” electronic database of PAHs.
- Survey of all possible PAHs in this database is a very promising method to solve the PAH-DIB problem.
- For the first time, two closed-shell PAH cations are found to meet all constraints put on the λ6614DIB (wavelength, intensity, band profile, and ionization potential).