Relationships between Carcinogenicity and Theoretical Reactivity
Indices in Polycyclic Aromatic Hydrocarbons

Iden A. Smith, Gregory D. Berger, Paul G. Seybold, and M. P. Servé

Departments of Chemistry [I. A. S., G. D. B., P. G. S., M. P. S.] and Biological Chemistry [P. G. S., M. P. S.], Wright State University, Dayton, Ohio 45435

ABSTRACT

Theoretical reactivity indices have been used to examine the metabolic reactions presumed, on the basis of recent biochemical evidence, to be responsible for the transformation of polycyclic aromatic hydrocarbon precursor carcinogens to ultimate carcinogens. Of a large number of indices examined, several show strong correlations with carcinogenic activity in a set of 25 representative compounds. The results support the belief that specific transformations involving dihydrodiol, "bay-region" epoxide, and carbonium ion intermediates are responsible for the carcinogenic activity of these compounds. Additional implications of the results are discussed, including the suggestion that this type of analysis might provide a rapid and simple means for prescreening compounds for potential carcinogens.

INTRODUCTION

Models suggesting a link between the carcinogenic activities of aromatic compounds and electronic properties date back at least 40 years to the pioneering work of Schmidt (29). In the past quarter century, there have been proposed numerous theories attempting to correlate electronic indices from molecular orbital theory with "carcinogenic activity." A successful theoretical correlation of this nature would be extremely important for several reasons: (a) it could provide important insight into the chemical mechanism(s) by which these compounds cause cancer; (b) when perfected it could provide a rapid and simple screening procedure to supplement the time-consuming and expensive animal and bacterial experiments now required for testing of potential carcinogens; and (c) the information thus provided might eventually be used to guide the design of effective antitumor agents.

Most theoretical models proposed thus far have focused attention on properties of the parent aromatic hydrocarbons, yielding in many cases quite interesting and suggestive results (6, 10, 12, 21-23, 26, 27). However, during the past several years there has been a considerable expansion in our experimental knowledge of the metabolism of PAH's, and the results strongly suggest that the carcinogenic process for these compounds involves a series of metabolic transformations that convert the relatively inert parent hydrocarbons into highly mutagenic and carcinogenic metabolites (4, 5, 7, 8, 13-20, 24, 25, 32, 34-40). Thus, it is possible that the observed differences in carcinogenic activities among these compounds not only reflect characteristics of the parent hydrocarbons but also depend strongly or even crucially on properties of later metabolites and the propensity of these metabolites to undergo specific reactions. Because the nature of selectivity is central to understanding the mechanism of carcinogenesis in this important class of compounds, we have examined the metabolic reactions in question, using a large sample of representative compounds and a variety of reactivity indices from molecular orbital theory. We shall show that strong correlations do exist between the properties of several putative metabolic intermediates and carcinogenic activities, and we shall distinguish those indices that appear most accurately to reflect the carcinogenic process.

Biochemical Evidence. The metabolic transformations of PAH's presumed to be responsible for carcinogenic activity can be illustrated by the example of the "classic" carcinogen BP (Compound I), which has been the subject of extensive experimental examination. The reactions of BP can be regarded as typical for this class of compounds. As a result of a variety of studies, a primary path by which BP is metabolically activated and transformed in vivo from precarcinogen to ultimate carcinogen is believed to consist of the stages shown in Chart 1. These transformations can be regarded as consequences of metabolic attempts to render nonpolar compounds soluble so that they can be excreted.

In Step a of Chart 1, BP is activated via a cellular monooxygenase to form its 7,8-epoxide (Ila). The 4,5-bond in BP forms a classical "K-region" (27), and epoxide formation at this latter bond apparently forms a diversionary pathway from the principal route to carcinogenesis. Recent evidence (38) indicates that cellular epoxide hydrase may act in vivo to detoxify reactants along this pathway. In Step b, epoxide hydrase assists conversion of the 7,8-epoxide to the 7,8-dihydrodiol (Iib). Experimentally, the 7,8-dihydrodiol is found to be more effectively converted to mutagenic and carcinogenic metabolites than is BP itself or dihydrodiols formed at other positions (19, 32, 38). Saturation of the 7,8-bond activates the 9,10-bond, leading to Step c in which Compound Iib is transformed to the 7,8-dihydrodiol-9,10-(bay-region) epoxide (III) (34-40). Strong support for the importance of this step has come from the experimental observation that the product BP diolepoxides have exceptionally high mutagenic activity for both bacterial and mammalian cells (16-18). In Step d the diol-epoxide III converts

---

1 Address after September 15, 1978: Department of Chemistry, University of California at San Diego, La Jolla, Calif. 92037.
2 Present address: Department of Chemistry, Yale University, New Haven, Conn. 06520.
3 To whom requests for reprints should be addressed.
4 The abbreviations used are: PAH, polycyclic aromatic hydrocarbon; BP, benzo(a)pyrene; BA, benz(a)anthracene.

Received February 24, 1978; accepted June 16, 1978.
Theoretical Reactivity Indices and Carcinogenicity

Chart 1. Presumed metabolic steps leading to carcinogenesis by BP. For simplicity only 1 stereoisomeric form is illustrated.

spontaneously to the triol carbonium ion IV. It has been suggested (14) that carbonium ions such as IV act as ultimate carcinogens via electrophilic attack on critical cellular nucleophiles, e.g. DNA. Evidence for this has come with the isolation and identification of specific adducts to guanine (15, 20, 24, 25, 35).

Jerina and Lehr (16), Jerina et al. (17, 18), and Wood et al. (38) have reviewed in much greater detail the experimental evidence supporting the importance of the transformation shown in Chart 1. This evidence has led them to propose the “bay-region theory” of carcinogenic activity for these compounds, which holds that diol-epoxides such as III, formed on saturated angular benzo rings, can be expected to display especially high carcinogenic activity (16-18). Theoretical calculations indicate that bay-region diol-epoxides should form carbonium ions more easily than other diol-epoxides (16-18).

Theoretical Model. In principle any of the steps shown in Chart 1 might be decisive in determining carcinogenicity in BP and other PAH compounds. So too could be the relative efficiencies of competitive, noncarcinogenic metabolic pathways or factors (e.g., size, solubility) largely outside the present analysis. Nonetheless, indices from molecular orbital theory have been exceedingly instructive in our understanding of ordinary chemical reactions (33) and, in so far as the reactions under consideration are subject to the same constraints, such indices should be helpful in elucidating those reactivity differences that might influence carcinogenesis. Of course there are possible hazards in applying this kind of approach to complicated systems (30), but these hazards are not necessarily as great as has sometimes been feared, for reasons that we shall discuss later.

A key element in this approach is a proper choice of theoretical indices, since usually several alternatives are available for analyzing each reaction. We have therefore examined several potentially useful indices at each stage to find which of these, if any, can be used in the present model. Furthermore, although schemes more complete than the simple Huckel molecular orbital theory are available, we see no reason to resort to these more complicated and time-consuming procedures at this early stage. For ease of analysis we shall identify certain key regions in the compounds under consideration. The K- and L-regions are as customarily defined (27), but it is necessary to define two new regions, the A- and B-regions. The definitions of the regions are as follows: K-region, the electron-rich region typified by the 9,10-bond in phenanthrene (27) and containing the highest molecular orbital order; L-region, a region consisting of 2 para-carbon atoms, as the meso-9,10-carbon atoms of anthracene [These carbon atoms display the highest free-valence indices (27).]; A-region, the presumptive initial epoxidation site, on the terminal ring of the bay region, in the metabolic pathway leading to carcinogenesis (This corresponds in most cases to what has sometimes been called the M-region.); B-region, the site of final epoxidation on the terminal ring of the bay region on a bond adjacent to the bay region. These regions are illustrated in Chart 2 for some representative compounds.

To obtain a broadly based analysis, we have examined the properties of 25 representative PAH’s and their presumptive metabolites, a set of compounds previously examined by Jerina et al. (17, 18) in their study of carbonium ions. The compounds are illustrated in Chart 3. Relative values for carcinogenic activity have been estimated by using principally the report of Arcos and Argus (2) and are summarized in Table 1.
RESULTS

Examination of Possible Indices

We now examine the individual stages of metabolic transformation (Chart 1) and associated theoretical indices.

**Compound I, the Parent Compound or Precarcinogen.** Prior theories of carcinogenesis have dealt largely with electronic structures of the precarcinogenic parent compounds. Because many of the parameters to be discussed later bear some relation to molecular size, it seemed desirable first to determine whether some simple feature related to this might not itself correlate with carcinogenic activity. This seemed especially relevant since Arcos and Arcos (1) and Arcos and Argus (2) have shown that most PAH carcinogens (but also many noncarcinogens) have surface areas ("incumbrance areas") within a certain range. One simple measure of molecular size (or mass) is the number of carbon atoms in the compound, and in Table 2 this number is compared with carcinogenic activity. Apparently, although most strong carcinogens are relatively large and the smallest compounds are not carcinogens, the correlation is far from complete. BP is a notable exception.

Another gross molecular feature and one sometimes held responsible for the action of certain drugs and for carcinogenic activity (9, 11) is solubility. Experimentally, this is conveniently described in terms of the partition coefficient $P$ between a polar and a nonpolar phase. As demonstrated by Rogers and Cammarata (28), this feature can be treated within the framework of molecular orbital theory in terms of sums of atomic $\pi$-electron densities $Q_\pi$ (polar phase) and superdelocalizabilities $S_\pi$ (nonpolar phase) for the mole-

---

### Table 1

**Numbering and characteristics of the compounds examined (see Chart 3)**

<table>
<thead>
<tr>
<th>Compound</th>
<th>Compound name</th>
<th>K-region?</th>
<th>L-region?</th>
<th>Bay region</th>
<th>Carcinogenicity index</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Naphthalene</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>Anthracene</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>Tetracene</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>Pentacene</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>Hexacene</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>?</td>
</tr>
<tr>
<td>6</td>
<td>BA</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>Benzo(a)tetracene</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>Phenanthrene</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>Benzo(c)phenanthrene</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>Chrysene</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>Benzo(b)chrysene</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>Picene</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>Triphenylene</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>Benzo(g)chrysene</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>17</td>
</tr>
<tr>
<td>15</td>
<td>Dibenz(a,c)anthracene</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>3</td>
</tr>
<tr>
<td>16</td>
<td>Dibenz(a,j)anthracene</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>4</td>
</tr>
<tr>
<td>17</td>
<td>Dibenz(a,h)anthracene</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>26</td>
</tr>
<tr>
<td>18</td>
<td>Naphtho(2,3-b)pyrene</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>27</td>
</tr>
<tr>
<td>19</td>
<td>BP</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>73</td>
</tr>
<tr>
<td>20</td>
<td>Benzo(e)pyrene</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>2</td>
</tr>
<tr>
<td>21</td>
<td>Dibenz(o,p)pyrene</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>33</td>
</tr>
<tr>
<td>22</td>
<td>Dibenz(a,j)pyrene</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>74</td>
</tr>
<tr>
<td>23</td>
<td>Dibenz(a,e)pyrene</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>50</td>
</tr>
<tr>
<td>24</td>
<td>Dibenz(a,h)pyrene</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>70</td>
</tr>
<tr>
<td>25</td>
<td>Tribenz(a,e,i)pyrene</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>16</td>
</tr>
</tbody>
</table>

*a* This compound does not strictly possess a bay region but does contain a "pseudo" bay region.

*b* Jerina et al. (16, 17) have assigned this as $++++$. 

---
Theoretical Reactivity Indices and Carcinogenicity

Table 2
Comparison of carcinogenic potency and the number of carbon atoms

<table>
<thead>
<tr>
<th>Compound</th>
<th>No. of carbon atoms</th>
<th>Carcinogenicity index</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>28</td>
<td>++</td>
</tr>
<tr>
<td>5</td>
<td>26</td>
<td>?</td>
</tr>
<tr>
<td>24</td>
<td>24</td>
<td>+ + + +</td>
</tr>
<tr>
<td>22</td>
<td>24</td>
<td>+ + + +</td>
</tr>
<tr>
<td>21</td>
<td>24</td>
<td>+ + + +</td>
</tr>
<tr>
<td>23</td>
<td>24</td>
<td>+ + + +</td>
</tr>
<tr>
<td>21</td>
<td>23</td>
<td>+ + + +</td>
</tr>
<tr>
<td>21</td>
<td>22</td>
<td>+ + + +</td>
</tr>
<tr>
<td>16</td>
<td>22</td>
<td>+ + + +</td>
</tr>
<tr>
<td>15</td>
<td>22</td>
<td>+ + + +</td>
</tr>
<tr>
<td>7</td>
<td>22</td>
<td>+ + + +</td>
</tr>
<tr>
<td>4</td>
<td>22</td>
<td>+ + + +</td>
</tr>
<tr>
<td>12</td>
<td>22</td>
<td>+ + + +</td>
</tr>
<tr>
<td>11</td>
<td>22</td>
<td>+ + + +</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>+ + + +</td>
</tr>
<tr>
<td>7</td>
<td>18</td>
<td>+ + + +</td>
</tr>
<tr>
<td>10</td>
<td>18</td>
<td>+ + + +</td>
</tr>
<tr>
<td>9</td>
<td>18</td>
<td>+ + + +</td>
</tr>
<tr>
<td>13</td>
<td>18</td>
<td>+ + + +</td>
</tr>
<tr>
<td>3</td>
<td>18</td>
<td>+ + + +</td>
</tr>
<tr>
<td>8</td>
<td>14</td>
<td>+ + + +</td>
</tr>
<tr>
<td>2</td>
<td>14</td>
<td>+ + + +</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>+ + + +</td>
</tr>
</tbody>
</table>

The K-, L-, and A-regions are shown in Table 4. The K- and L-region values agree with those given previously by Mainster and Memory (21, 22), who suggest that a compound is expected to be a carcinogen if $I_L \geq 2.05$ and $I_L \leq 2.30$. Note that, for compounds with more than a single K-, L-, or A-region, only the highest value is listed. It can be seen that in

Table 3
Comparison of calculated partition coefficients and carcinogenic potencies

<table>
<thead>
<tr>
<th>Compound</th>
<th>$\sum S_r$</th>
<th>$In \tilde{\rho}$</th>
<th>Carcinogenicity index</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>31.402</td>
<td>14.876</td>
<td>?</td>
</tr>
<tr>
<td>25</td>
<td>25.479</td>
<td>12.321</td>
<td>++</td>
</tr>
<tr>
<td>4</td>
<td>24.758</td>
<td>12.010</td>
<td>–</td>
</tr>
<tr>
<td>18</td>
<td>23.716</td>
<td>11.561</td>
<td>+ +</td>
</tr>
<tr>
<td>24</td>
<td>23.417</td>
<td>11.432</td>
<td>+ + + +</td>
</tr>
<tr>
<td>22</td>
<td>22.822</td>
<td>11.175</td>
<td>+ + + +</td>
</tr>
<tr>
<td>21</td>
<td>22.045</td>
<td>10.840</td>
<td>+ +</td>
</tr>
<tr>
<td>23</td>
<td>21.719</td>
<td>10.699</td>
<td>+ +</td>
</tr>
<tr>
<td>7</td>
<td>21.672</td>
<td>10.679</td>
<td>–</td>
</tr>
<tr>
<td>11</td>
<td>20.504</td>
<td>10.175</td>
<td>–</td>
</tr>
<tr>
<td>17</td>
<td>19.981</td>
<td>9.949</td>
<td>++</td>
</tr>
<tr>
<td>16</td>
<td>19.968</td>
<td>9.944</td>
<td>–</td>
</tr>
<tr>
<td>12</td>
<td>19.658</td>
<td>9.810</td>
<td>–</td>
</tr>
<tr>
<td>15</td>
<td>19.655</td>
<td>9.809</td>
<td>–</td>
</tr>
<tr>
<td>14</td>
<td>19.354</td>
<td>9.679</td>
<td>++</td>
</tr>
<tr>
<td>19</td>
<td>18.904</td>
<td>9.485</td>
<td>+ + + +</td>
</tr>
<tr>
<td>3</td>
<td>18.786</td>
<td>9.433</td>
<td>–</td>
</tr>
<tr>
<td>20</td>
<td>17.789</td>
<td>9.004</td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>16.663</td>
<td>8.518</td>
<td>+</td>
</tr>
<tr>
<td>10</td>
<td>16.050</td>
<td>8.255</td>
<td>+</td>
</tr>
<tr>
<td>9</td>
<td>16.019</td>
<td>8.240</td>
<td>+</td>
</tr>
<tr>
<td>13</td>
<td>15.540</td>
<td>8.033</td>
<td>–</td>
</tr>
<tr>
<td>2</td>
<td>13.486</td>
<td>7.147</td>
<td>–</td>
</tr>
<tr>
<td>8</td>
<td>12.353</td>
<td>6.659</td>
<td>–</td>
</tr>
<tr>
<td>1</td>
<td>8.876</td>
<td>5.159</td>
<td>–</td>
</tr>
</tbody>
</table>

* $a ln \tilde{\rho} = 0.4314 \sum S_r + 1.3297$.

Table 4
Parent compound bond superdelocalizability indices and carcinogenic potencies

<table>
<thead>
<tr>
<th>Compound</th>
<th>$I_K$</th>
<th>$I_L$</th>
<th>$I_A$</th>
<th>$I_{K/A}$</th>
<th>Carcinogenicity index</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>2.206</td>
<td>2.724</td>
<td>1.943</td>
<td>0.8800</td>
<td>+ + + +</td>
</tr>
<tr>
<td>18</td>
<td>2.196</td>
<td>2.724</td>
<td>2.022</td>
<td>0.9208</td>
<td>+ + + +</td>
</tr>
<tr>
<td>24</td>
<td>2.175</td>
<td>1.986</td>
<td>1.945</td>
<td>0.9089</td>
<td>+ + + +</td>
</tr>
<tr>
<td>19</td>
<td>2.140</td>
<td>2.386</td>
<td>1.865</td>
<td>0.8719</td>
<td>+ + + +</td>
</tr>
<tr>
<td>7</td>
<td>2.139</td>
<td>2.863</td>
<td>1.951</td>
<td>0.9229</td>
<td>+ + + +</td>
</tr>
<tr>
<td>21</td>
<td>2.114</td>
<td>2.512</td>
<td>1.871</td>
<td>0.8914</td>
<td>+ + + +</td>
</tr>
<tr>
<td>11</td>
<td>2.099</td>
<td>2.438</td>
<td>1.848</td>
<td>0.8869</td>
<td>+ + + +</td>
</tr>
<tr>
<td>6</td>
<td>2.079</td>
<td>2.438</td>
<td>1.848</td>
<td>0.8967</td>
<td>+ + + +</td>
</tr>
<tr>
<td>17</td>
<td>2.061</td>
<td>2.292</td>
<td>1.875</td>
<td>0.9142</td>
<td>+ + + +</td>
</tr>
<tr>
<td>16</td>
<td>2.051</td>
<td>2.294</td>
<td>1.875</td>
<td>0.9142</td>
<td>+ + + +</td>
</tr>
<tr>
<td>23</td>
<td>2.033</td>
<td>1.908</td>
<td>1.908</td>
<td>0.9385</td>
<td>+ + + +</td>
</tr>
<tr>
<td>20</td>
<td>2.038</td>
<td>1.813</td>
<td>1.813</td>
<td>0.8940</td>
<td>+ + + +</td>
</tr>
<tr>
<td>12</td>
<td>2.008</td>
<td>1.853</td>
<td>1.853</td>
<td>0.9228</td>
<td>+ + + +</td>
</tr>
<tr>
<td>25</td>
<td>2.004</td>
<td>1.904</td>
<td>1.904</td>
<td>0.9501</td>
<td>+ + + +</td>
</tr>
<tr>
<td>10</td>
<td>1.998</td>
<td>1.856</td>
<td>1.856</td>
<td>0.9289</td>
<td>+ + + +</td>
</tr>
<tr>
<td>8</td>
<td>1.994</td>
<td>1.837</td>
<td>1.837</td>
<td>0.9213</td>
<td>+ + + +</td>
</tr>
<tr>
<td>9</td>
<td>1.975</td>
<td>1.859</td>
<td>1.859</td>
<td>0.9413</td>
<td>+ + + +</td>
</tr>
<tr>
<td>14</td>
<td>1.972</td>
<td>1.870</td>
<td>1.870</td>
<td>0.9483</td>
<td>+ + + +</td>
</tr>
<tr>
<td>15</td>
<td>2.306</td>
<td>2.110</td>
<td>2.110</td>
<td>0.8400</td>
<td>+ + + +</td>
</tr>
<tr>
<td>5</td>
<td>4.010</td>
<td>2.192</td>
<td>2.192</td>
<td>0.9123</td>
<td>+ + + +</td>
</tr>
<tr>
<td>4</td>
<td>3.590</td>
<td>2.145</td>
<td>2.145</td>
<td>0.8890</td>
<td>+ + + +</td>
</tr>
<tr>
<td>3</td>
<td>3.010</td>
<td>2.083</td>
<td>2.083</td>
<td>0.9228</td>
<td>+ + + +</td>
</tr>
<tr>
<td>2</td>
<td>2.626</td>
<td>1.995</td>
<td>1.995</td>
<td>0.9501</td>
<td>+ + + +</td>
</tr>
<tr>
<td>13</td>
<td>1.803</td>
<td>1.803</td>
<td>1.803</td>
<td>0.9413</td>
<td>+ + + +</td>
</tr>
<tr>
<td>1</td>
<td>1.867</td>
<td>1.867</td>
<td>1.867</td>
<td>0.9483</td>
<td>+ + + +</td>
</tr>
</tbody>
</table>

In the atomic superdelocalizability at position $r$ is defined as (10, 23)

$$ S_r = 2 \sum_{j} c_{rj}^2 m_j $$

where the sum is over occupied molecular orbitals, the $c_{rj}$ are carbon atom molecular orbital coefficients for molecular orbital $j$, and the molecular orbital energies are $\epsilon_j = \alpha + m_j \beta$. Since all of the compounds examined here are alternate aromatic hydrocarbons, their $\pi$-electron charge densities are uniformly zero at all carbon positions. For this case Rogers and Cammarata (28) obtained the equation

$$ ln \tilde{\rho} = 0.4314 \sum S_r + 1.3297 $$

that holds reasonably well for a large number of compounds. A comparison between carcinogenic activity and $ln \tilde{\rho}$ is shown in Table 3. No special relationship is evident from the results.

Of the many molecular orbital theories of carcinogenesis, the K- and L-region approach of the Pullmans (26, 27) has received the widest attention. In this approach carcinogenic activity is associated with a reactive K-region and a relatively unreactive L-region, both measured by indices derived from combinations of localization energies in the two regions. The results are found in many although not all cases to yield a suggestive relationship to carcinogenic activity (6, 26, 27). The Pullman indices are somewhat complicated to calculate, and Mainster and Memory (21, 22) have shown that a simpler index $f$, the sum of the two atomic superdelocalizable (10, 23) involved in a particular regional bond, can be used to give comparable results. This 2-center index would appear to be well suited for estimating reactivity with regard to epoxidation. Results for

SEPTEMBER 1978
all cases $I_k > I_L$. As a rule the most potent carcinogens lack L-regions and have $I_k$ values >1.90. An exception is dibenz(a,h)anthracene (Compound 17).

$\pi$-Electron bond orders $p_\pi$ might also be a fair measure of 2-center attack and are shown in Table 5 for the K-, L-, and A-regions. Again, no strong correlation is evident. Most carcinogens have values for the highest occupied molecular orbital (in $\beta$ units) between 0.300 and 0.500 (Table 6).

**Compound Iib, the A-Region Dihydrodiol.** From the viewpoint of elementary $\pi$-electron theory, Compounds Iia and Iib are equivalent since both are saturated at bond A. The efficacy of Step c depends on activation of the B-region bond, for which there are several possible measures. As before, a superdelocalizability index should be appropriate. We call this index $I_k$, the prime indicating the dihydrodiol form. Elsewhere we have already demonstrated that a strong correlation exists between this index for the B-region and carcinogenic activity in these compounds (3). Values of $I_k$ and $I_k'$ for the A-region dihydrodiol forms are summarized in Table 7. In all cases, formation of the A-region dihydrodiol significantly activates the B-region bond, as shown by the high value of index $I_k'$. In many cases, e.g., BP (Chart 2), the K-region bond is deactivated, accenting further the tendency toward reaction at the B-region.

To see whether some other index might describe the bond activation of this stage as well as does $I_k'$, we examined the B-region $\pi$-electron bond orders, as shown in Table 8. It can be seen that this index also correlates rather well with carcinogenic activity.

**Compound IV, the Carbonium Ion.** According to the present picture, trihydrotriol carbonium ions act as the final

<table>
<thead>
<tr>
<th>Table 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Comparison of parent compound mobile bond orders and carcinogenic potencies</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Compound</th>
<th>$p_\pi^a$</th>
<th>$p_K^a$</th>
<th>Carcinogenicity index</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>0.6831</td>
<td>0.7314</td>
<td>++</td>
</tr>
<tr>
<td>23</td>
<td>0.6854</td>
<td>0.7537</td>
<td>+++</td>
</tr>
<tr>
<td>15</td>
<td>0.6864</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.6898</td>
<td>0.7754</td>
<td>+</td>
</tr>
<tr>
<td>13</td>
<td>0.6904</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>0.6918</td>
<td>0.7833</td>
<td>++</td>
</tr>
<tr>
<td>14</td>
<td>0.6931</td>
<td>0.7457</td>
<td>+</td>
</tr>
<tr>
<td>7</td>
<td>0.6981</td>
<td>0.7851</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.7005</td>
<td>0.7833</td>
<td>+</td>
</tr>
<tr>
<td>16</td>
<td>0.7027</td>
<td>0.7797</td>
<td>+</td>
</tr>
<tr>
<td>17</td>
<td>0.7028</td>
<td>0.7790</td>
<td>++</td>
</tr>
<tr>
<td>8</td>
<td>0.7068</td>
<td>0.7747</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>0.7103</td>
<td>0.7322</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>0.7117</td>
<td>0.7613</td>
<td>+</td>
</tr>
<tr>
<td>10</td>
<td>0.7122</td>
<td>0.7540</td>
<td>+</td>
</tr>
<tr>
<td>11</td>
<td>0.7137</td>
<td>0.7480</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>0.7165</td>
<td>0.7209</td>
<td>++++</td>
</tr>
<tr>
<td>19</td>
<td>0.7207</td>
<td>0.7538</td>
<td>++++</td>
</tr>
<tr>
<td>24</td>
<td>0.7238</td>
<td>0.7617</td>
<td>++++</td>
</tr>
<tr>
<td>1</td>
<td>0.7246</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0.7354</td>
<td>0.7466</td>
<td>+</td>
</tr>
<tr>
<td>2</td>
<td>0.7374</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.7409</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.7420</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.7424</td>
<td>?</td>
<td></td>
</tr>
</tbody>
</table>

*If there is more than 1 possible K-region or A-region, the value given represents the K-region or A-region with the lowest bond order.*

<table>
<thead>
<tr>
<th>Table 6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Comparison of parent compound highest occupied molecular orbital energies ($\beta$ units) and carcinogenic potencies</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Compound</th>
<th>Highest occupied molecular orbital energy</th>
<th>Carcinogenic activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>0.6840</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>0.6190</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>0.6052</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>0.5676</td>
<td>+</td>
</tr>
<tr>
<td>14</td>
<td>0.5319</td>
<td>++</td>
</tr>
<tr>
<td>10</td>
<td>0.5201</td>
<td>+</td>
</tr>
<tr>
<td>12</td>
<td>0.5019</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.4991</td>
<td>+</td>
</tr>
<tr>
<td>20</td>
<td>0.4970</td>
<td>+</td>
</tr>
<tr>
<td>16</td>
<td>0.4917</td>
<td>+</td>
</tr>
<tr>
<td>17</td>
<td>0.4735</td>
<td>++</td>
</tr>
<tr>
<td>6</td>
<td>0.4523</td>
<td>+++</td>
</tr>
<tr>
<td>23</td>
<td>0.4216</td>
<td>+++</td>
</tr>
<tr>
<td>2</td>
<td>0.4142</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>0.4048</td>
<td>-</td>
</tr>
<tr>
<td>21</td>
<td>0.3983</td>
<td>++</td>
</tr>
<tr>
<td>25</td>
<td>0.3957</td>
<td>++</td>
</tr>
<tr>
<td>19</td>
<td>0.3711</td>
<td>++++</td>
</tr>
<tr>
<td>22</td>
<td>0.3420</td>
<td>++++</td>
</tr>
<tr>
<td>7</td>
<td>0.3270</td>
<td>-</td>
</tr>
<tr>
<td>24</td>
<td>0.3027</td>
<td>++++</td>
</tr>
<tr>
<td>18</td>
<td>0.3026</td>
<td>++</td>
</tr>
<tr>
<td>3</td>
<td>0.2950</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>0.2197</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>0.1694</td>
<td>?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 7</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bond superdelocalizability indices $I_{k'}$ and $I_{k}$ for A-region dihydrodiol forms</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Compound</th>
<th>$I_{k'}$</th>
<th>$I_{k}$</th>
<th>Carcinogenic index</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>2.407</td>
<td>2.180</td>
<td>++++</td>
</tr>
<tr>
<td>24</td>
<td>2.390</td>
<td>2.174</td>
<td>++++</td>
</tr>
<tr>
<td>7</td>
<td>2.375</td>
<td>2.196</td>
<td>-</td>
</tr>
<tr>
<td>25</td>
<td>2.361</td>
<td>2.025</td>
<td>+</td>
</tr>
<tr>
<td>19</td>
<td>2.358</td>
<td>2.051</td>
<td>+++</td>
</tr>
<tr>
<td>21</td>
<td>2.353</td>
<td>2.141</td>
<td>+</td>
</tr>
<tr>
<td>23</td>
<td>2.335</td>
<td>2.056</td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>2.333</td>
<td>2.106</td>
<td>+</td>
</tr>
<tr>
<td>17</td>
<td>2.318</td>
<td>2.088</td>
<td>+</td>
</tr>
<tr>
<td>14</td>
<td>2.309</td>
<td>2.042</td>
<td>+</td>
</tr>
<tr>
<td>15</td>
<td>2.308</td>
<td>2.091</td>
<td>++</td>
</tr>
<tr>
<td>16</td>
<td>2.307</td>
<td>2.079</td>
<td>+</td>
</tr>
<tr>
<td>5</td>
<td>2.303</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>2.301</td>
<td>2.054</td>
<td>+</td>
</tr>
<tr>
<td>13</td>
<td>2.275</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2.275</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>2.274</td>
<td>2.153</td>
<td>+</td>
</tr>
<tr>
<td>12</td>
<td>2.273</td>
<td>2.015</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>2.272</td>
<td>1.970</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>2.264</td>
<td>1.923</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>2.261</td>
<td>2.018</td>
<td>+</td>
</tr>
<tr>
<td>3</td>
<td>2.242</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>2.236</td>
<td>2.000</td>
<td>+</td>
</tr>
<tr>
<td>2</td>
<td>2.205</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2.177</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

*Edel/3, Jerina et al. (16–18) have pointed out the importance of bay-region carbonium ions and examined the ease of carbonium ion formation, using the change in delocalization energy $\Delta E_{deloc}/\beta$, as obtained from a perturbation calculation. The success of this index in correlating metabolic species that attack critical cellular nucleophiles (presumably DNA).*
Table 8
Comparison of π-electron bond orders for the A-region dihydrodiols with carcinogenic potencies

<table>
<thead>
<tr>
<th>Compound (from Chart 3)</th>
<th>$p_a^{*}$</th>
<th>$p_a^{**}$</th>
<th>Carcinogenic index</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>0.8914</td>
<td>0.7450</td>
<td>++++</td>
</tr>
<tr>
<td>24</td>
<td>0.8923</td>
<td>0.7483</td>
<td>++++</td>
</tr>
<tr>
<td>19</td>
<td>0.8934</td>
<td>0.7675</td>
<td>++++</td>
</tr>
<tr>
<td>25</td>
<td>0.8942</td>
<td>0.7453</td>
<td>++</td>
</tr>
<tr>
<td>7</td>
<td>0.8951</td>
<td>0.7265</td>
<td>++</td>
</tr>
<tr>
<td>23</td>
<td>0.8955</td>
<td>0.7870</td>
<td>++</td>
</tr>
<tr>
<td>21</td>
<td>0.8957</td>
<td>0.7833</td>
<td>++</td>
</tr>
<tr>
<td>6</td>
<td>0.8961</td>
<td>0.7235</td>
<td>+</td>
</tr>
<tr>
<td>17</td>
<td>0.8968</td>
<td>0.7170</td>
<td>+</td>
</tr>
<tr>
<td>14</td>
<td>0.8970</td>
<td>0.7498</td>
<td>++</td>
</tr>
<tr>
<td>15</td>
<td>0.8974</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0.8975</td>
<td>0.7185</td>
<td>+</td>
</tr>
<tr>
<td>8</td>
<td>0.8993</td>
<td>0.7123</td>
<td>–</td>
</tr>
<tr>
<td>13</td>
<td>0.8993</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.8997</td>
<td>0.6949</td>
<td>–</td>
</tr>
<tr>
<td>11</td>
<td>0.9003</td>
<td>0.6830</td>
<td>–</td>
</tr>
<tr>
<td>10</td>
<td>0.9004</td>
<td>0.6898</td>
<td>+</td>
</tr>
<tr>
<td>9</td>
<td>0.9030</td>
<td>0.6960</td>
<td>+</td>
</tr>
<tr>
<td>18</td>
<td>0.9047</td>
<td>0.7524</td>
<td>++</td>
</tr>
<tr>
<td>5</td>
<td>0.9050</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.9055</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.9066</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.9088</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.9113</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a When more than 1 B-region (dihydrodiol) exists, the value listed is the lowest bond order.

b By using the dihydrodiol with the lowest B-region bond order, the value given is the lowest K-region bond order.

data for the present set of compounds is impressive. A shortcoming is that this index cannot be used in certain cases (e.g., it does not distinguish methyl derivatives) (17), which has led us to examine alternative indices.

One such index is the net π-electron charge (i.e., $Q_0 = l - q_b$, where $q_b$ is the π-electron density) at the benzylic carbon position $b$ of the trihydrodiol carbonium ion. In Table 9 this index, which we call $Q_0$, is compared with $\Delta E_\text{dielec}/\beta$ and carcinogenic activity. The index $Q_0$ is equally successful with $\Delta E_\text{dielec}/\beta$ in its correlation to carcinogenicity, the two giving almost identical results. An advantage of the index $Q_0$ is that it is more generally applicable; i.e., it can be used for distinctions among methyl derivatives.

Two other indices, the free-valence index $F_v$ and the atomic superdelocalizability $S_A$ at the bay-region carbonium ion position also showed correlations with carcinogenic potency. (The free-valence index is

$$F_v = 1.732 - \sum p_{rs}$$

where $s$ refers to atoms bonded to the $r$ atom.) Results for these indices appear in Table 10.

**Composite Energy Indices**

Transformation energies can sometimes be instructive in kinetic analyses (witness the Jerina-Lehr energy index). We examined the following energies: $\Delta E(x^*)^{\text{a}}$, the π-energy loss in going from the parent compound to the A-region epoxide or dihydrodiol (Step a); $\Delta E(x^*)^{\text{b}}$, the π-energy loss in forming the dihydrodiol-epoxide from the A-region dihydrodiol (Step

Table 9
Carbonium ion charge density $Q_0$ at the benzylic carbon position and $\Delta E_\text{dielec}/\beta$ compared with carcinogenic potencies

<table>
<thead>
<tr>
<th>Compound</th>
<th>$Q_0$</th>
<th>$\Delta E_\text{dielec}/\beta$</th>
<th>Carcinogenic index</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>0.3215</td>
<td>0.870</td>
<td>++++</td>
</tr>
<tr>
<td>7</td>
<td>0.3333</td>
<td>0.846</td>
<td>–</td>
</tr>
<tr>
<td>24</td>
<td>0.3334</td>
<td>0.846</td>
<td>++</td>
</tr>
<tr>
<td>25</td>
<td>0.3494</td>
<td>0.818</td>
<td>++</td>
</tr>
<tr>
<td>21</td>
<td>0.3600</td>
<td>0.808</td>
<td>++</td>
</tr>
<tr>
<td>19</td>
<td>0.3637</td>
<td>0.794</td>
<td>++</td>
</tr>
<tr>
<td>5</td>
<td>0.3712</td>
<td>0.782</td>
<td>–</td>
</tr>
<tr>
<td>23</td>
<td>0.3750</td>
<td>0.778</td>
<td>++</td>
</tr>
<tr>
<td>6</td>
<td>0.3810</td>
<td>0.766</td>
<td>+</td>
</tr>
<tr>
<td>17</td>
<td>0.3984</td>
<td>0.738</td>
<td>++</td>
</tr>
<tr>
<td>15</td>
<td>0.4083</td>
<td>0.722</td>
<td>+</td>
</tr>
<tr>
<td>16</td>
<td>0.4084</td>
<td>0.722</td>
<td>+</td>
</tr>
<tr>
<td>14</td>
<td>0.4103</td>
<td>0.728</td>
<td>++</td>
</tr>
<tr>
<td>20</td>
<td>0.4138</td>
<td>0.714</td>
<td>+</td>
</tr>
<tr>
<td>4</td>
<td>0.4167</td>
<td>0.710</td>
<td>–</td>
</tr>
<tr>
<td>18</td>
<td>0.4286</td>
<td>0.690</td>
<td>++</td>
</tr>
<tr>
<td>13</td>
<td>0.4464</td>
<td>0.664</td>
<td>–</td>
</tr>
<tr>
<td>12</td>
<td>0.4476</td>
<td>0.682</td>
<td>–</td>
</tr>
<tr>
<td>8</td>
<td>0.4500</td>
<td>0.658</td>
<td>–</td>
</tr>
<tr>
<td>11</td>
<td>0.4580</td>
<td>0.647</td>
<td>–</td>
</tr>
<tr>
<td>10</td>
<td>0.4630</td>
<td>0.640</td>
<td>+</td>
</tr>
<tr>
<td>3</td>
<td>0.4706</td>
<td>0.628</td>
<td>–</td>
</tr>
<tr>
<td>9</td>
<td>0.4902</td>
<td>0.600</td>
<td>+</td>
</tr>
<tr>
<td>2</td>
<td>0.5294</td>
<td>0.544</td>
<td>–</td>
</tr>
<tr>
<td>1</td>
<td>0.5714</td>
<td>0.488</td>
<td>–</td>
</tr>
</tbody>
</table>

a When more than 1 bay-region carbonium ion is possible, the value given represents the carbonium ion with the lowest $Q_0$.

b From Jerina et al. (16, 17).

Table 10
Comparison of carbonium ion free valence and atomic superdelocalizability indices with carcinogenic potencies

<table>
<thead>
<tr>
<th>Compound</th>
<th>$F_v$</th>
<th>$S_A$</th>
<th>Carcinogenicity index</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>1.0085</td>
<td>0.844</td>
<td>++++</td>
</tr>
<tr>
<td>24</td>
<td>1.0120</td>
<td>0.819</td>
<td>++++</td>
</tr>
<tr>
<td>19</td>
<td>1.0179</td>
<td>0.713</td>
<td>++++</td>
</tr>
<tr>
<td>7</td>
<td>1.0201</td>
<td>0.593</td>
<td>–</td>
</tr>
<tr>
<td>25</td>
<td>1.0207</td>
<td>0.536</td>
<td>++</td>
</tr>
<tr>
<td>21</td>
<td>1.0242</td>
<td>0.784</td>
<td>++</td>
</tr>
<tr>
<td>6</td>
<td>1.0274</td>
<td>0.707</td>
<td>+</td>
</tr>
<tr>
<td>23</td>
<td>1.0276</td>
<td>0.754</td>
<td>+++</td>
</tr>
<tr>
<td>17</td>
<td>1.0310</td>
<td>0.651</td>
<td>++</td>
</tr>
<tr>
<td>14</td>
<td>1.0328</td>
<td>0.504</td>
<td>++</td>
</tr>
<tr>
<td>16</td>
<td>1.0340</td>
<td>0.629</td>
<td>+</td>
</tr>
<tr>
<td>15</td>
<td>1.0341</td>
<td>0.628</td>
<td>+</td>
</tr>
<tr>
<td>20</td>
<td>1.0377</td>
<td>0.628</td>
<td>+</td>
</tr>
<tr>
<td>13</td>
<td>1.0435</td>
<td>0.519</td>
<td>–</td>
</tr>
<tr>
<td>8</td>
<td>1.0440</td>
<td>0.509</td>
<td>–</td>
</tr>
<tr>
<td>12</td>
<td>1.0449</td>
<td>0.529</td>
<td>–</td>
</tr>
<tr>
<td>11</td>
<td>1.0479</td>
<td>0.506</td>
<td>–</td>
</tr>
<tr>
<td>10</td>
<td>1.0485</td>
<td>0.485</td>
<td>+</td>
</tr>
<tr>
<td>5</td>
<td>1.0552</td>
<td>1.056</td>
<td>?</td>
</tr>
<tr>
<td>18</td>
<td>1.0589</td>
<td>0.689</td>
<td>++</td>
</tr>
<tr>
<td>9</td>
<td>1.0597</td>
<td>0.437</td>
<td>–</td>
</tr>
<tr>
<td>4</td>
<td>1.0606</td>
<td>0.793</td>
<td>–</td>
</tr>
<tr>
<td>3</td>
<td>1.0696</td>
<td>0.568</td>
<td>–</td>
</tr>
<tr>
<td>2</td>
<td>1.0835</td>
<td>0.392</td>
<td>–</td>
</tr>
<tr>
<td>1</td>
<td>1.0971</td>
<td>0.304</td>
<td>–</td>
</tr>
</tbody>
</table>

a If more than 1 bay-region carbonium ion is possible, the value given represents the carbonium ion with the highest value for the index.

b For these compounds the bay-region carbonium ion with the highest $S_A$ does not give the highest $F_v$. 

SEPTEMBER 1978  2973
c); and $\Delta E_p^{(3)}$, the $\pi$-energy change in forming the trihydrotriole carbonium ion from the dihydrodiol-epoxide (Step d). The last of these is formally identical to the Jerina-Lehr index $\Delta E_{\text{del}}/\beta$, but it is calculated directly and not by using a perturbation technique. Table 11 shows that $\Delta E_p^{(1)}$ does not yield a sensible correlation. However, both $\Delta E_p^{(2)}$ and $\Delta E_p^{(3)}$ show a strong correlation with carcinogenic activity (Table 12).

Many of these indices show strong correlations among themselves, as can sometimes be expected on theoretical grounds. Some of these relationships are illustrated in Charts 4 to 11. As seen in Chart 4, there is some tendency for $I_B$ and $I_K$ to increase together, but this does always hold true. The Jerina-Lehr index $E_{\text{del}}/\beta$ correlates strongly with $I_B$ and almost exactly with the charge index $Q_B$.

**DISCUSSION**

The results reported here would appear to lend strong theoretical support to the idea that reactions such as those in Chart 1 are related to the carcinogenic activity of these compounds, i.e., to the bay-region hypothesis (16–18). Jerina and Lehr (16) and Jerina et al. (17, 18) in their study of carbonium ion delocalization energies were the first to demonstrate a convincing correlation between carcinogenicity and properties of a metabolite. We have shown elsewhere that the index $I_B$ for the dihydrodiol intermediate also correlates strongly with carcinogenic potency (3). In the present much more extensive examination, it is apparent that several reactivity indices representing later metabolic stages correlate strongly with carcinogenic activity. Because these indices can be sensibly associated with the reactions in Chart 1, the most reasonable interpretation is that such reactions are indeed involved in the carcinogenic process.

It is interesting that properties of the later metabolic intermediates give stronger correlations to carcinogenicity than do those of the parent compounds. In his valence bond study, Herndon (12) concluded tentatively that the parent aromatic hydrocarbons themselves act as ultimate carcinogens. Most previous theories, similarly predating the current metabolic evidence, have also operated at least implicitly from this assumption or at least the assumption that the properties of the parent compounds are determining for the carcinogenic phenomenon. The results obtained in this study lead us to conclude that properties of the later stages, although of course indirectly preordained by the natures of

---

**Table 11**

<table>
<thead>
<tr>
<th>Compound</th>
<th>$\Delta E_{E}^{(1)} (\beta)$</th>
<th>Carcinogenicity index</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>0.8911</td>
<td>-2.4577</td>
</tr>
<tr>
<td>24</td>
<td>0.8824</td>
<td>-2.4565</td>
</tr>
<tr>
<td>19</td>
<td>0.8675</td>
<td>-2.4550</td>
</tr>
<tr>
<td>7</td>
<td>0.8666</td>
<td>-2.4516</td>
</tr>
<tr>
<td>25</td>
<td>0.8614</td>
<td>-2.4536</td>
</tr>
<tr>
<td>21</td>
<td>0.8563</td>
<td>-2.4507</td>
</tr>
<tr>
<td>6</td>
<td>0.8480</td>
<td>-2.4504</td>
</tr>
<tr>
<td>23</td>
<td>0.8457</td>
<td>-2.4519</td>
</tr>
<tr>
<td>17</td>
<td>0.8395</td>
<td>-2.4495</td>
</tr>
<tr>
<td>14</td>
<td>0.8354</td>
<td>-2.4494</td>
</tr>
<tr>
<td>15</td>
<td>0.8333</td>
<td>-2.4484</td>
</tr>
<tr>
<td>16</td>
<td>0.8332</td>
<td>-2.4486</td>
</tr>
<tr>
<td>20</td>
<td>0.8263</td>
<td>-2.4466</td>
</tr>
<tr>
<td>8</td>
<td>0.8121</td>
<td>-2.4462</td>
</tr>
<tr>
<td>22</td>
<td>0.8113</td>
<td>-2.4460</td>
</tr>
<tr>
<td>12</td>
<td>0.8105</td>
<td>-2.4451</td>
</tr>
<tr>
<td>11</td>
<td>0.8048</td>
<td>-2.4445</td>
</tr>
<tr>
<td>10</td>
<td>0.8032</td>
<td>-2.4444</td>
</tr>
<tr>
<td>5</td>
<td>0.8009</td>
<td>-2.4326</td>
</tr>
<tr>
<td>18</td>
<td>0.7905</td>
<td>-2.4337</td>
</tr>
<tr>
<td>4</td>
<td>0.7882</td>
<td>-2.4322</td>
</tr>
<tr>
<td>9</td>
<td>0.7828</td>
<td>-2.4399</td>
</tr>
<tr>
<td>3</td>
<td>0.7687</td>
<td>-2.4308</td>
</tr>
<tr>
<td>2</td>
<td>0.7437</td>
<td>-2.4279</td>
</tr>
<tr>
<td>1</td>
<td>0.7205</td>
<td>-2.4242</td>
</tr>
</tbody>
</table>

---

**Table 12**

<table>
<thead>
<tr>
<th>Compound</th>
<th>$\Delta E_{E}^{(1)} (\beta)$</th>
<th>$\Delta E_{E}^{(1)b} (\beta)$</th>
<th>Carcinogenicity index</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>0.8911</td>
<td>-2.4577</td>
<td>++</td>
</tr>
<tr>
<td>24</td>
<td>0.8824</td>
<td>-2.4565</td>
<td>++</td>
</tr>
<tr>
<td>19</td>
<td>0.8675</td>
<td>-2.4550</td>
<td>++</td>
</tr>
<tr>
<td>7</td>
<td>0.8666</td>
<td>-2.4516</td>
<td>-</td>
</tr>
<tr>
<td>25</td>
<td>0.8614</td>
<td>-2.4536</td>
<td>++</td>
</tr>
<tr>
<td>21</td>
<td>0.8563</td>
<td>-2.4507</td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>0.8480</td>
<td>-2.4504</td>
<td>+</td>
</tr>
<tr>
<td>23</td>
<td>0.8457</td>
<td>-2.4519</td>
<td>+</td>
</tr>
<tr>
<td>17</td>
<td>0.8395</td>
<td>-2.4495</td>
<td>+</td>
</tr>
<tr>
<td>14</td>
<td>0.8354</td>
<td>-2.4494</td>
<td>+</td>
</tr>
<tr>
<td>15</td>
<td>0.8333</td>
<td>-2.4484</td>
<td>+</td>
</tr>
<tr>
<td>16</td>
<td>0.8332</td>
<td>-2.4486</td>
<td>+</td>
</tr>
<tr>
<td>20</td>
<td>0.8263</td>
<td>-2.4466</td>
<td>+</td>
</tr>
<tr>
<td>8</td>
<td>0.8121</td>
<td>-2.4462</td>
<td>+</td>
</tr>
<tr>
<td>22</td>
<td>0.8113</td>
<td>-2.4460</td>
<td>+</td>
</tr>
<tr>
<td>12</td>
<td>0.8105</td>
<td>-2.4451</td>
<td>+</td>
</tr>
<tr>
<td>11</td>
<td>0.8048</td>
<td>-2.4445</td>
<td>+</td>
</tr>
<tr>
<td>10</td>
<td>0.8032</td>
<td>-2.4444</td>
<td>+</td>
</tr>
<tr>
<td>5</td>
<td>0.8009</td>
<td>-2.4326</td>
<td>+</td>
</tr>
<tr>
<td>18</td>
<td>0.7905</td>
<td>-2.4337</td>
<td>+</td>
</tr>
<tr>
<td>4</td>
<td>0.7882</td>
<td>-2.4322</td>
<td>+</td>
</tr>
<tr>
<td>9</td>
<td>0.7828</td>
<td>-2.4399</td>
<td>+</td>
</tr>
<tr>
<td>3</td>
<td>0.7687</td>
<td>-2.4308</td>
<td>+</td>
</tr>
<tr>
<td>2</td>
<td>0.7437</td>
<td>-2.4279</td>
<td>+</td>
</tr>
<tr>
<td>1</td>
<td>0.7205</td>
<td>-2.4242</td>
<td>+</td>
</tr>
</tbody>
</table>

*If there is more than 1 A-region, the value listed is for the A-region with the most negative $\Delta E_{E}^{(1)}$."

*"Comparison of the energy indices $\Delta E_{E}^{(2)}$ and $\Delta E_{E}^{(3)}$ with carcinogenic potencies (see "Results")*
lower than the corresponding \( I_A \) values. We interpret this to mean merely that the reactions (Step a) of Chart 1 represent a minor chemical pathway compared to reaction at the K-region, consistent with the experimental evidence (2). [This does not necessarily imply the same for the in vivo situation in which enzymatic specificities can influence the relative yields (13).] It is ironic that historically the well-known (both theoretically and experimentally) high reactivities of K-regions may actually have inhibited progress in understanding chemical carcinogenesis by PAH's, because it was natural to focus attention on the predominant reactions rather than on a minor pathway.

An exception to the model also provides insight. Benzo(a)tetracene (Compound 7) has a high \( I_A \) value and a low \( Q_A \) value, marking it as a potential carcinogen, but it is not found to be carcinogenic. An explanation may be that this compound has 2 active L-regions; the idea that the L-region is a deactivating region appears to retain its usefulness in the present model, consistent with earlier theories (27).

Naphtho(2,3-b)pyrene (Compound 18) is apparently more active than predicted by the model, and we have no good explanation for this case. However, the very qualitative nature of the experimental data on carcinogenicity must be recognized and may provide reason not to be overly worried if complete agreement with the experimental results is not obtained at this stage. Development of a coherent, accurate, and general numerical measure of carcinogenic potency (if this ideal is even possible) would greatly facilitate
the development of a more quantitative theory.

Which indices are "best" to use in future studies remains a somewhat subjective decision, but we offer the following comments based on accuracy in representing carcinogenic potency, ease of calculation, general applicability, and physical interpretability. Clearly, for the dihydrodiol either \( I_9' \) or \( p_9' \) appears adequate for describing activation of the B-region bond. For the carbonium ion we favor the index \( Q_6 \) because it correlates strongly with carcinogenicity, is easily calculated, can be applied to substituent effects, and can be satisfactorily interpreted as a measure of carbonium ion stability (lower positive charges indicate greater delocalization of electron density to this position and hence greater stability). In studies to be published elsewhere, we have found that \( I_6' \) and \( Q_6 \) are effective in explaining variations in carcinogenic activities among the methyl derivatives of chrysene and benz(a)anthracene. Thus, we are gaining confidence in the correctness of the present analysis and the effectiveness of its indices. In fact we are now sufficiently confident to suggest that the present type of metabolic analysis appears suitable for use in selected cases as a rapid prescreening procedure, supplementing the more fundamental biological carcinogen tests and possibly alerting investigators to compounds that might be expected to be especially hazardous.

If the present metabolic picture is correct (and the experimental evidence appears to be increasing that it is), it is natural to inquire as to why the original K-region theory was as successful as it was. That is, if the K-region is not on the main path to carcinogenesis and is in fact a diversionary path, it is hard to understand why reactivity of the K-region bond should correlate with carcinogenic potency. Jerina et al. (18) have already pointed out that the angular ring arrangement necessary for a K-region is also a prerequisite for a bay region and that compounds that have one almost always have the other. Also, as shown in Chart 4, there is a slight tendency for \( I_8 \) and \( I_6 \) to increase together. Both tend to increase with molecular size, which itself correlates somewhat with carcinogenic activity. Still, we remain puzzled and suspect that there are more fundamental connections that we have not yet fully perceived.

Finally, there is the interesting question of why the present analysis works as well as it does. In a provocative article Scribner (30) has correctly pointed out some of the dangers inherent in the application of molecular orbital techniques to complicated biological systems. Foremost among these dangers are the unknown influences of solubility factors, transport properties, and catalytic enzymes on the rates of the transformations under consideration. We admit to some degree of surprise ourselves at the extent of agreement achieved here between the theoretical reactivity indices and carcinogenic activity. Taken at face value the results imply that solubility and transport considerations are not crucial to the carcinogenic potency of the compounds under study and that the enzymes operating in these systems are indeed largely nonspecific, as must be assumed for this sort of analysis to be valid.

The analysis presented here must be considered only as a crude first step toward a more complete theoretical examination of the metabolic transformations of these compounds. That all of the subtleties and many of the specific details of the metabolic transformations are completely ignored should be recognized. The concept of carcinogenic activity itself for a given compound depends on a multitude of factors, including the nature of the organism tested, the mode of administration, the target organs examined, etc. The metabolic process occurs in the very complicated environment of a living cell. Almost certainly the metabolic pathway examined here is not the exclusive route to carcinogenicity for these compounds. The calculations ignore many features, including any influence of specific geometrical configurations. The correlations observed apparently result because electronic factors are of greater importance than any of these neglected features.

ACKNOWLEDGMENTS

We thank Dr. Donald Jerina for his comments on this work and for sending us copies of several papers prior to publication. We also thank Professors Roland Lehr, J. D. Memory, and R. Daudel for sending us copies of their papers prior to publication.

REFERENCES

Theoretical Reactivity Indices and Carcinogenicity


