

DIRECTED REACTION GRAPHS AS POSETS

D. J. Klein & L. Bytautas[†]

Texas A&M University @ Galveston, Galveston, Texas 77553-1675

[†] On leave from the Institute of Theoretical Physics & Astronomy, Vilnius, Lithuania

Abstract

Reaction diagrams are considered especially for the circumstance of progressive substitution (or addition) on a fixed molecular skeleton, and it is noted that these naturally form Hasse diagrams for a partially ordered set (or poset) of the substituted structures. Ultimately several examples of such posetic reaction graphs are presented, but as a first prototypical illustration the case of progressive chlorination of the hexagonal benzene skeleton is described, and some ideas are mentioned for the utilization of the reaction diagram's posetic nature.

1. Introduction

There are many examples of reaction graphs occurring throughout chemistry. Those involving degenerate rearrangements have received much formal attention by chemical graph theorists, in as much as they can be rather large and seemingly involved whilst in fact being much simplified because of the frequent occurrence of high symmetries. Many researchers have made contributions in this area, but one of the first was Balaban, who also has recently reviewed [1] work on a variety of such degenerate rearrangement graphs. But the degeneracy aspect typically leads to undirected reaction graphs. Of course there are many other kinds of reaction graphs which are directed, and perhaps they are even more common in current chemistry. For instance, there typically occur directed reaction graphs in (directed) syntheses, as reviewed in Corey & Cheng's seminal book [2] *The Logic of Chemical Synthesis*. The structures of synthesis graphs are crucial in differentiating between "linear" and "convergent" synthetic approaches [3], though there has been only a little formal graph-theoretic work [4,5] on such graphs, so that there could perhaps be some further investigation incorporating more fully the directed aspect of their edges (and consequent possibility for being interpretable as posets). In many cases [2] it seems these synthesis graphs are simply linear chains or nearly so, whence the posetic structure is especially simple (a total order). The seemingly less common synthesis graphs for so-called

“convergent” syntheses [3,4,5] feature some degree of branching, but still very often these graphs are trees (even when the directions are eliminated from the edges), though this is not a logical necessity. Another commonly occurring and important type of directed reaction graph is found in molecular biological applications, often showing cycles (or “hypercycles”) say in the area of “enzyme kinetics”, again with directions on the edges. These reaction schemes have been much considered, even in a formal sense, e.g., as in refs. [6,7]. But in these cases a poset interpretation is complicated by the general occurrence of cycles. General complex sets of chemical reactions (as considered in [8]) typically similarly exhibit cycles.

Here it is pointed out that a fairly general class of directed reaction diagrams can be neatly viewed as *partially ordered sets*, or *posets*, such as defined in the prolegomenon. That is, in some cases there is an intrinsic natural order, say as for the possible results of substitutional chlorination of benzene, as illustrated in figure 1. There only the hexagon of carbons is shown, and the Cl-substituted carbon vertices are shown as (larger) black dots. An arrow is directed from one structure ξ to a second ζ , if ζ can be obtained from ξ by the replacement of one H-atom by one Cl-atom (without moving around any other Cl-atoms which might already be attached). This arrow then represents a single minimal step of chlorination. And the diagram as a whole is the *Hasse diagram* [9] associated to a poset of different chloro-benzenes. The ordering relation $\xi \geq \zeta$ then means that there is some non-negative number of Cl-atoms which can be (substitutionally) added to structure ξ so as to obtain structure ζ .

The general class of poset reaction diagrams are then those for which there is a progressive degree of reaction (substitution, addition, dissociation, or local rearrangement). The progressive substitution as in figure 1, is just one example, which is discussed further in the next four sections (though somewhat tangentially in section 3). Progressive additions could also be entertained, say involving the hydrogenation of benzene, as indicated in section 6, where also other examples of types of reactions based on the benzene skeleton are mentioned.

Evidently there are many possible examples of progressive reaction graphs, which then may be viewed to impart a poset structure. It may be emphasized that the occurrence of poset reaction graphs is not limited to the identification of chemical species as graphs, but may also incorporate geometric (or stereo-structural) information. Thence for substitutional chlorination of at the corners of Ladenberg benzene, viewed as a trigonal prism of CH groups, one has the result of figure 2, if one pays attention to the 3-dimensional structure, distinguishing enantiomeric (or mirror-image) species (which otherwise have the same graph). Such benzene substituent posets are discussed further in the next section, with special reference to distinctions between the posets for the regular-hexagonal & trigonal prismatic skeletons, such as was of some historical relevance in the valence theory of molecular structures in chemistry.

Several further examples of posets are found in sections 6 & 7, first with reference to the regular hexagonal benzene skeleton, and then more generally, so that one perceives some of the numerous possibilities for poset reaction graphs. Indeed one may find a number of such graphs in the literature, including our benzene substituent example [10,11], but the emphasis (or even mention) of the poset nature seems not to have been previously made. Sections 4 & 5 focus on two possible approaches to utilize such posets in forming correlations for molecular properties, again in the context of the benzene substitution diagram of figure 1. But it is presumed that there are many additional examples and also additional possible uses of such poset reaction diagrams.

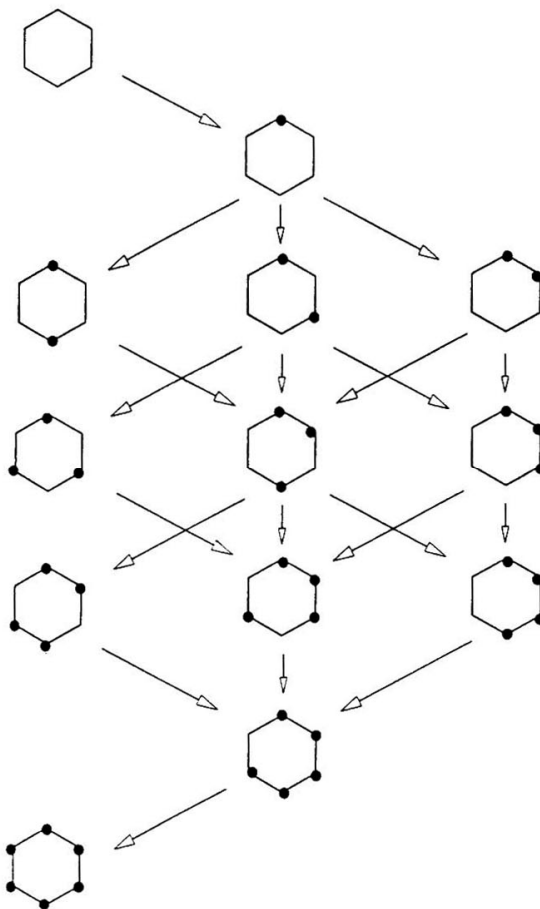


Figure 1 - The posetic reaction diagram for successive chlorination of the benzene skeleton, with black dots identifying Cl-substituted sites.

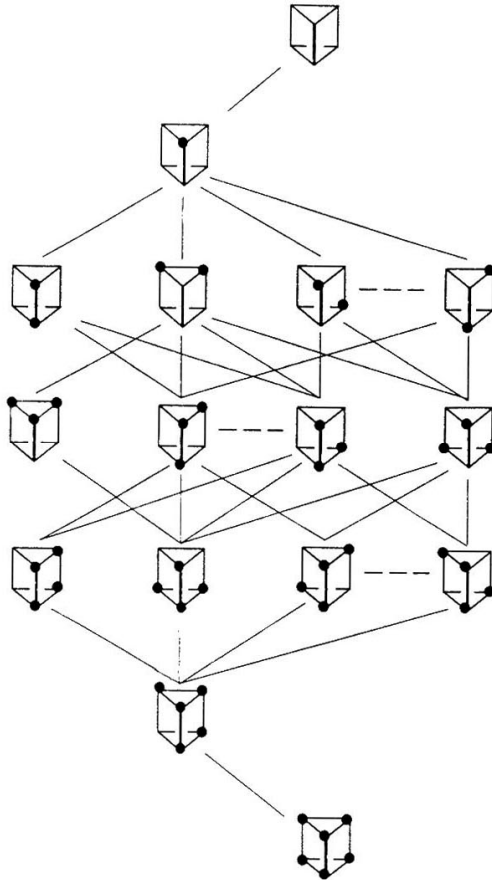


Figure 2 - Reaction diagram for chlorination of a trigonal-prismatic skeleton.

2. Posetic Reaction Diagram & Chemical Structure for Benzene

An interesting question concerns how much the posetic diagram might determine about the molecular structure of the species involved. In particular, what does the substitutional reaction diagram for benzene imply about the skeleton being a regular hexagon? In fact the overall question of the structure of benzene was of great historical relevance, with Kekule arguing [12] for a regular hexagonal skeleton, as based on several points including:

- * first, the atom-count formula;
- * second, the various possible isomeric substituents; and
- * third, substitutional reaction graphs ordered as in our posetic reaction diagrams.

The information from the first and second points concerning isomer counts have in fact received much attention even in a general formal manner, in 1930 by Lunn & Senior [13], and Balaban [14] has emphasized that the Polya-theoretic isomer counts ignoring chirality must be the same no matter the number of substituents (because of the identity of the relevant “cycle indices” for benzene and prismane). Even more recently Hässelbarth [15] emphasizes a general theorem that the collection of isomer counts for all different degrees of substitution turns out to characterize the molecular-skeleton permutation group in terms of the numbers of symmetry elements of each different cycle type. That is, just from the isomer counts for all different degrees of substitution, much of the nature of the permutational symmetry associated with the point group of the molecular skeleton is determined. A second question concerns to what extent the numbers of symmetry elements of each different cycle type determine the permutation group, and Hässelbarth [15] suggests that this is fairly unique, for practical cases. Notably even in 1874 Körner [16] investigated permutational questions for the case of benzene using (at least implicitly) substitutional reaction diagrams with an effort to avoid inadvertent geometrical assumptions, though some of the details of the argument to make it “geometry-free” have only more recently been completed by McBride [17]. (Indeed McBride establishes from purely classical chemical arguments much about what the skeletal permutation group must be.) But granted a permutation group one might inquire further as to what features (say as regards point-group symmetry) of the underlying molecular skeleton might be determined. Perhaps the reaction poset yields even more information. In fact, Heilbronner & Dunitz in their charming discourse [10] *Reflection and Symmetry* indicate that such reaction graphs were crucial in making the choice for Kekule’s regular hexagonal skeleton in preference to Ladenberg’s proposal [18] of a trigonal prism. In particular Heilbronner & Dunitz say that Ladenberg’s trigonal-prism structure “gives the correct number of isomers but leads to a different reaction graph from what is actually observed”. But there is some confusion here, because the counts of substitutional isomers only match if mirror images are not distinguished, it being that when first the two proposed structures came to opposition (in 1869), it was a standard view that it was overly presumptive to make geometric interpretations (as implicit in making chiral distinctions), whereas the graph-theoretic structures (i.e. connectivity patterns) were more chemically accessible, and perhaps were sometimes believed to be all that was ultimately relevant. A view of this sort was apparently taken by Körner [16], and even after Vant Hoff [19] & le Bel [20] proposed the tetrahedral geometric nature of carbon atoms (in 1874), there was often a significant degree of reluctance to forego the simple austerity of the (non-geometric) graphical approach, as discussed in section IX.2 on “Opposition to directed Valencies” in Russell’s history [21]. Indeed there was even ridicule (e.g., in the writings of Kolbe, in [22]) of the geometric proposals, and

more reasoned objections by others [23]. Surely the purely graph-theoretic view explained much, and mirror image structures have many properties the same (i.e., all scalar properties are the same), and the non-scalar property of rotation of polarized light presumably was at these early times rather esoteric in the view of many. Even le Bel [24] had reservations for some time about accepting what we now see as straightforward consequences of the tetrahedral model of carbon. Anyway if one accepts the restriction to purely combinatoric ideas without reference to mirror images, then not only do chlorobenzene isomer counts match, but also the reaction diagrams match. To see this from the chirality-attendant figure 2 for the trigonal-prismatic skeleton, condense the enantiomeric pairs (which are joined by horizontal dotted lines) to a single chirally non-distinguished structure, whereupon the reaction diagram becomes isomorphic to that of figure 1 (for the hexagonal skeleton). If on the other hand mirror images are distinguished (which entails not only additional experimental evidence but also going beyond graphs in the interpretative phase), then not only are the reaction diagrams non-isomorphic, but also the isomer counts are different.

One may wonder what happens if the reaction diagram idea is extended in some other manner. For instance, the idea of a single-substituent reaction diagram could be extended to a double-substituent diagram, say for Cl- & F- substituents on the hexagonal benzene skeleton. This reaction graph is much more complicated with 92 nodes (as further considered in section 5), so that this 2-substituent reaction diagram would seemingly contain much more information about benzene. But notably it turns out that without distinguishing mirror images, this 2-substituent reaction diagram for trigonal-prismatic Ladenberg benzene is in one-to-one correspondence (i.e., is isomorphic) with that for hexagonal benzene. To proceed to explicit reaction posets for even more substituents is daunting, there being 430 nodes for the M=3-substituent case, and for successively higher numbers M of substituents there are 1505, 4291, 10528, 23052, 46185, 86185, 151756, etc nodes (such numbers scaling as $\sim(M+1)^6/12$ being, e.g., determinable via standard Polya enumerative techniques [25,26]).

As an alternative to working through the detailed reaction diagrams for such different (and evidently complex) cases with different numbers of substituents, one may make another very informative argument that is fully conclusive. For C_6H_6 there are 6 substituent positions, and one may imagine a labelling as

$$\begin{pmatrix} l_1 & l_2 & l_3 & l_4 & l_5 & l_6 \\ s_1 & s_2 & s_3 & s_4 & s_5 & s_6 \end{pmatrix} \equiv \begin{pmatrix} l(6) \\ s(6) \end{pmatrix} \equiv p[l(6)|s(6)]$$

to indicate that the ligand (or substituent) l_i is to be located at the skeletal site s_i (and $l(6)$ & $s(6)$ are just abbreviations for the respective lists $l_1 l_2 \dots l_N$ & $s_1 s_2 \dots s_N$). But such *permutamer* symbols are "redundant" in that several assignments of skeletal site labelling are equivalent because of skeletal point-group symmetry (and several ligands may be equivalent because of their "identity"). The effect of skeletal point-group symmetry is neatly described in terms of a permutational group of skeletal sites. For the regular hexagonal skeleton H the skeletal sites may be numbered in cyclic order, as in figure 3a, and permutations represented as a list of these site labels enclosed inside a pair of parentheses being cyclically permuted, whence the permutational image (or representation) of the point group is represented as

$$\begin{aligned} P[H; \mathcal{D}_6] &= \{1, (123456), (135)(246), (14)(25)(36), (153)(264), (165432)\} \{1, (26)(35)\} \\ &= P[H; \mathcal{D}_{6h}] \end{aligned}$$

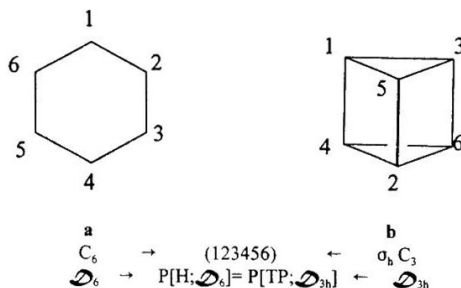


Figure 3 - Site numbering for hexagonal (H) & trigonal-prismatic (TP) skeletons in a & b.

Here it is realized that the larger point group \mathcal{D}_{6h} provides no further images of different permutations of sites than does \mathcal{D}_6 . That is, e.g., the reflection in the plane normal to the molecular plane and containing sites 1 & 4 effects the same interchange of the same sites as does the 2-fold rotation about the axis containing sites 1 & 4 - of course for the electronic wave-function, there are electrons out of the molecular plane and these two point-group operations then have different effects. For the trigonal-prismatic skeleton TP with the site labelling of figure 3b the permutational images of the purely rotational and full point groups turn out to be

$$\begin{aligned} P[TP; \mathcal{D}_3] &= \{I, (135)(246), (153)(264)\} \{I, (14)(23)(56)\} \\ P[TP; \mathcal{D}_{3h}] &= \{I, (123456), (135)(246), (14)(25)(36), (153)(264), (165432)\} \{I, (26)(35)\} \end{aligned}$$

And it may be seen that the permutational images $P[H; \mathcal{D}_{6h}]$ & $P[TP; \mathcal{D}_{3h}]$ are isomorphic, whereas $P[H; \mathcal{D}_6]$ & $P[TP; \mathcal{D}_3]$ are clearly distinguishable, having different numbers (12 & 6) of elements. That is, if mirror images are not to be distinguished, then the skeletal permutation groups of relevance to identifying different permutational isomers are essentially the same (e.g., with the same isomer counts for each degree of substitution). Of course, the geometric symmetries \mathcal{D}_{6h} & \mathcal{D}_{3h} are different (while their permutational representations are the same). But the isomorphism of the permutational groups $P[H; \mathcal{D}_{6h}]$ & $P[TP; \mathcal{D}_{3h}]$ implies that as abstract combinatorial structures these two cases are the same and indistinguishable (within this context). In fact it should perhaps be mentioned that the cage for the historically considered Claus-Armstrong [27] "centric" formula (often depicted with bonds between *para* carbons rather than with double bonds) also gives the same permutational group result - indeed quite clearly because the geometry naturally imagined for this case is also regular hexagonal.

This permutational-symmetry result implies that the single-, double-, triple-, & further-substituent reaction posets are isomorphic, and do not then aid in choosing between H & TP skeletons for benzene. To be a little more formal about this point one can explicitly characterize

a general progressive reaction poset in terms of these permutation representations. As a first step toward this we consider the equivalence of the different permutamers $p[l(N)|s(N)]$ for an N -vertex skeleton S with geometric symmetry group \mathcal{G} . First there is equivalence, as already noted, under the action of the permutation group $P[S;\mathcal{G}]$ on the positions of the site labels. That is, permutamer labels obtained from one $p[l(N)|s(N)]$ via permutation of the skeletal positions $s(N)$ about via any $\pi \in P[S;\mathcal{G}]$ are equivalent. But there is generally additional equivalence from permutations of the $l(N)$ if some of the ligands are nondistinguished. Most simply the ligands are each independent each making connection to a single skeletal site (i.e., they are monodentate), whence the equivalence amongst ligands is specified by a second permutation group $P[L]$ which is simply a product of disjoint symmetric groups each such component symmetric group associated to a set of nondistinguished ligands. If the ligands themselves have more complicated structure as connections between pairs, then the group $P[L]$ can have a more complicated structure. Anyway there is equivalence amongst all permutamer labels obtained from $p[l(N)|s(N)]$ by permuting the ligand positions $l(N)$ about via any $\pi \in P[L]$. And finally the permutamer labels $p[l(N)|s(N)]$ are equivalent under simultaneous permutation of ligand & site label positions by the same permutation π for any π in the full N -index permutation group \mathcal{S}_N . Because of this last feature we can re-express the effect of the site-symmetry group $P[S;\mathcal{G}]$, and fix by convention the ordering of the site labels in $p[l(N)|s(N)]$. Then the equivalence classes of permutamers are in one-to-one correspondence [28,29] with so-called $P[S;\mathcal{G}],P[L]$ -double cosets

$$P[S;\mathcal{G}]pP[L] \equiv \{ \pi\rho\sigma \mid \pi \in P[S;\mathcal{G}], \sigma \in P[L] \} \quad , \rho \in \mathcal{S}_N$$

That is, the application of any permutation in $P[S;\mathcal{G}]pP[L]$ to the ligand positions in $p[l(N)|s(N)]$ gives the equivalent permutamers (with the same fixed ordering for the site labels $s(N)$). And the different $P[S;\mathcal{G}],P[L]$ -double cosets (of equivalence transformations) partition \mathcal{S}_N into disjoint parts each in correspondence with a possible isomer (for the given $P[S;\mathcal{G}]$ & $P[L]$). Now for the case of a single-atom substituent (replacing H-atoms, as in benzene), the form of the ligand symmetry group is

$$P[L] = \mathcal{S}_n \times \mathcal{S}_{N-n}$$

if there are n substituted sites (and $N-n$ unsubstituted ones). In dealing with the progressive reaction poset we might identify this $P[L]$ as that $P_{\text{react}}[L]$ for a reactant species. With further substitution one of the unsubstituted sites is singled out to be replaced by the substituent, and at an intermediate stage the effective ligand-symmetry group can be identified as

$$P_{\text{int}}[L] = \mathcal{S}_n \times \mathcal{S}_1 \times \mathcal{S}_{N-n-1} = \mathcal{S}_n \times 1 \times \mathcal{S}_{N-n-1}$$

And a final product species has a ligand-symmetry group

$$P_{\text{prod}}[L] = \mathcal{S}_{n-1} \times \mathcal{S}_{N-n-1}$$

Evidently the criterion that a given reactant permutamer equivalence class (i.e., a given n -fold substituted species corresponding to a particular $P[S;\mathcal{G}],P_{\text{react}}[L]$ -double coset) can give rise to an intermediate (associated with a $P[S;\mathcal{G}],P_{\text{int}}[L]$ -double coset) is that the $P[S;\mathcal{G}],P_{\text{react}}[L]$ -double

coset contain the $P[S; \mathcal{G}]P_{\text{int}}[L]$ -double coset. Likewise, the condition that a product can arise from a given intermediate is that the $P[S; \mathcal{G}]P_{\text{prod}}[L]$ -double coset contain the $P[S; \mathcal{G}]P_{\text{int}}[L]$ -double coset. Thence the overall condition that a reactant with double coset $P[S; \mathcal{G}]P_{\text{react}}[L]$ can give rise to a product with double coset $P[S; \mathcal{G}]P_{\text{prod}}[L]$ is simply that the intersection

$$P[S; \mathcal{G}]P_{\text{react}}[L] \cap P[S; \mathcal{G}]P_{\text{prod}}[L]$$

is nonempty (i.e., that it contain one or more $P[S; \mathcal{G}]P_{\text{int}}[L]$ -double cosets). Indeed we may even view the order of this intersection as a “symmetry number” for the reaction, with steps having higher symmetry numbers being entropically more favored. Evidently the intersection criterion depends solely on the permutation-group-theoretic structure. And the argument directly extends to multi-substituent cases, with for the two-substituent case ligand-symmetry groups of the form $\mathcal{S}_\alpha \times \mathcal{S}_\beta \times \mathcal{S}_{N-\alpha-\beta}$ and intermediates singling out one of the unsubstituted sites to be shifted in the product to one of the other sets of ligands. That is, if for a given N two skeletal groups are isomorphic, then all multi-substituent reaction posets are isomorphic too. In particular so long as chirality is not considered, the multi-substituent reaction posets for hexagonal benzene and for the trigonal prism are the same.

3. More on Benzene Structure

To distinguish between the regular-hexagonal & trigonal-prismatic skeletons something beyond isomer counts and (achiral) reaction diagrams needs to be done. One approach would be to count stereoisomers, with some substitutional patterns being chiral for the TP case, whereas all are achiral for the regular-H case. But so far as we know this recourse to chiral structures (and the relevance of the distinct groups $P[H; \mathcal{D}_6]$ & $P[TP; \mathcal{D}_3]$) had little early impact. The convoluted history of the problem of benzene structure is reviewed by Rocke [30] and by Brush [31]. From an extensive collection of experiments A. von Baeyer [32] argued that some chemical reactions could be better understood if the two substituted sites in the *ortho* isomer were geometrically adjacent (the *ortho* derivatives being distinguishable as that giving two trichlorobenzenes upon further chlorination, as seen in our reaction poset). Brush [31] suggests (on page 33 of his review) that “no one else had a precise understanding of Baeyer’s argument”, though Brush notes that several textbook writers of the late 1800s acknowledged Baeyer’s contribution. In fact these and related arguments seem to have convinced Ladenburg (who was intensely involved), since in his generally very interesting chemical history [33] of 1886 he acknowledges such geometric arguments and accepts benzene as hexagonal - and Rocke [30] says that Ladenburg [34] had reasons even in 1874 to view Kekule’s formulation to be “at least as appropriate, if not more so than the prism”. Balaban [14] who also has looked into the historical aspects agrees that Baeyer’s arguments largely convinced his contemporaries. Another piece of evidence (intimated on page 33 of Brush, and discussed by Balaban) concerned the hydrogenation of benzene to add two more H atoms, whereafter what remains seem to be two ordinary localized double bonds. Brush’s generally substantive historical account [31] continues up to present times, through electron- & x-ray diffraction experiments and quantum-theoretic developments. It seems to us that Kekule’s hexagonal formula was widely accepted even before 1900 (though often there were qualifications).

Perhaps arguments of historical impact for the preference for the regular-H skeleton were via a sort of chemical consistency for a wide range of chemical phenomena, with geometric considerations entering in different subtle manners. Baeyer's arguments [32] that the *ortho* sites be adjacent is one example of this nature - with Baeyer's repeated such examples being overall convincing. Further, for instance, with the regular-H view one might suspect just one naphthalene, whereas with the TP view one might suspect three naphthalene isomers, as in figure 4. Geometry enters here in that rather naturally one might (inadvertently) exclude other "geometrically peculiar" alternative naphthalenic structures, say with an *ortho*-pair of one benzene fused to a *meta*- or *para*-pair of another benzene. But there are many such chemically plausible considerations based on the hexagonal pattern - and though the implicit assumptions contrasting with the TP case seem often not to have been explicitly made, such comparisons perhaps were not published because everything turned out (ultimately) so neatly consistent within the regular-H case. *E.g.*, the oxidation (by acidic chromate) of naphthalene yields *o*-phthallic acid, with *o*-phthallic acid having a position in the reaction diagram with just two ways to add a chloro-substituent to the ring, and *o*-phthallic acid being the one dicarboxylic benzo acid to form an acid anhydride (as expected on geometric grounds if the two carboxylic acid groups are adjacent). Another geometric argument notes that no enantiomeric substituted benzenes are observed as should be possible were it to have the prismatic structure - but here the first such argumentation which Brush [31] notes is in a text-book of Markwald in 1898. It seems to us that various (sometimes implicit) geometric reasons played an important role in the preference for the regular hexagonal structure. Of course there were complications: the absence of two *ortho* disubstituted derivatives (across single or double bonds), often imagined to be accounted for by oscillations between the two possibilities; the fact that the properties of benzene were not like that of (a mixture of two) conformations with localized double bonds; and the difficulties with the centric formula in that reliance solely on the connectivity pattern would indicate that *ortho* & *para* disubstituted derivatives should be equivalent. The fully unambiguous verification of the hexagonal skeleton would await the development of x-ray diffraction, and a full understanding of the behavior would await quantum mechanical insight. All this is discussed by Brush [31].

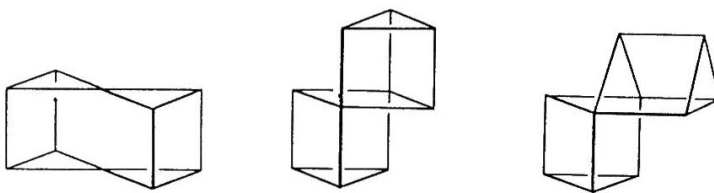


Figure 4 - Three naphthalenic analogues for the trigonal-prismatic skeleton.

4. Molecular Properties & the Substitution-Reaction Poset

The posetic reaction diagram for progressive reaction, as for the substitution of chlorine on benzene, may be correlated with various properties, in different ways. For instance, the presentation of figure 1 may be viewed to be correlated with (mean) polarizability & (molecular) dipolarity in the vertical & horizontal directions, respectively. First, *mean polarizability* is generally accepted to correlate roughly with molecular volume, as in the present case increases with the number of Cl atoms. (The full polarizability is a tensor so that the mean considered here is just 1/3 the trace of this tensor.) Second, the dipole moment may be viewed to be (at least approximately) a vectorial sum of bond dipoles, which are most significant for the C–Cl bonds, and the molecular dipolarity would be the magnitude of the net molecular dipole. (Evidently the dipolarity is more than a graph invariant since it depends on some aspects of the geometry.) In as much as the strength of intermolecular interactions should increase with increasing polarizability and with increasing dipolarity, one then expects that the boiling points of chlorobenzenes should be increasing as one moves either downward or to the right. Such a correlation is indicated in figure 5, where in fact we have entered the boiling points, from the *78th CRC Handbook of Chemistry & Physics* [35]. It seems from this figure that the polarizability aspect is more increases in the rightward direction, whence we have added the left-downward slanting lines to indicate that the boiling point increases in the downward direction typically outweigh those in the rightward direction, thereby suggesting a fuller possible poset. The boiling point for the 1,2,3,5-tetrachlorobenzene is enclosed in parentheses because it is out of order with the 1,2,4,5-derivative, and because the boiling point of the 1,2,3,5-derivative is not in the recent Handbook but instead was taken from an older edition [36]. These two “anomalies” together suggest that this value for the 1,2,3,5-derivative may be treated with some suspicion. In figure 6 the boiling points (of [35]) for methyl-substituted benzenes are displayed, and it is seen that here there is full compliance with our partial ordering, even as regards the extension with lines slanted downward to the left. In figure 7 critical temperature values for the methyl-benzenes are inserted into the poset, with the upper value at each position being from one reference [35] while the second value is from another reference [36]. Here it is seen that though there is rather frequent disagreement in values between the two references, all is in agreement with our partial ordering. We have not found a very complete set of critical temperatures for any other substituents, and so do not presently display any further such diagrams with critical temperatures.

For boiling points we have found in [35] one other substituent such that most of the positions of the poset are filled in, this being for the fluorobenzenes, excepting the 1,2,3-derivative for which we did not find a boiling point. However, these fluorobenzenes evidently are exceptional in as much as there is frequent disagreement with the posetic ordering pattern, so that we do not presently illustrate it. Though the poset of figures 5, 6, & 7 is not the same as that of figure 1, it is clear that the two are intimately related, and evidently the boiling points & critical temperatures often are ordered in accordance with the present reaction-graph-derived posets.

Of course the general idea of correlating molecular properties with different aspects of molecular structure (or geometry) goes far back in history, with some notable works in the first half of the 1800s, even pre-dating the now conventional structural formulas. The book *Chemical*

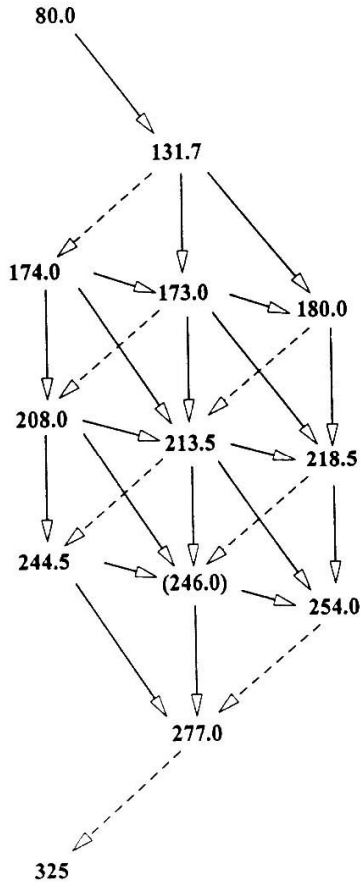


Figure 5 - Boiling points (in °C) for the various chlorobenzenes (situated as in figure 1).

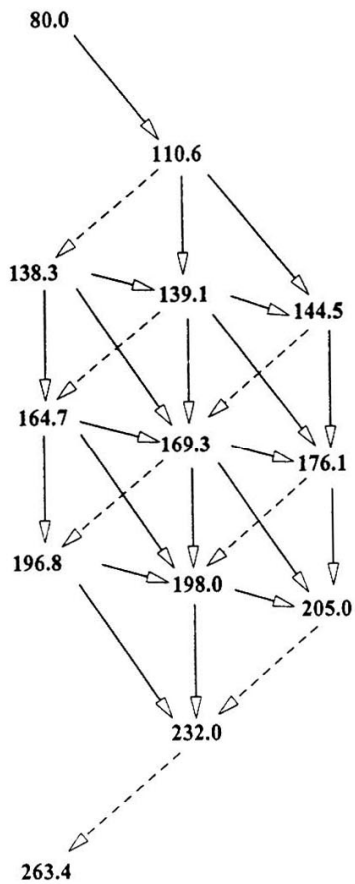


Figure 6 - Boiling points (in °C) for the various methylbenzenes (situated as in figure 1).

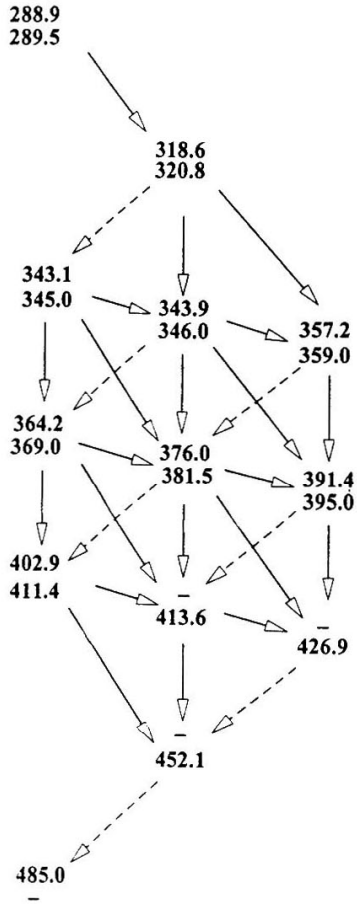


Figure 7 - Critical temperatures (in °C) for the various methylbenzenes.

Constitution and some Physical Properties by Smiles [38] gives a wide-ranging overview as of 1910, and the handbook [39] gives a similar overview as of 1982. And currently the areas of "quantitative structure-property relationships" and of "quantitative structure-activity relationships" remain of great interest. The present reaction-graph-related considerations indicate that at least in certain cases that different properties correlate with the position in a (structure-sensitive) poset, especially if some attention is paid to correlating the horizontal positions with some simple graph-theoretic or structural invariant. As such this idea is akin to that of Randić & Wilkins [40] in developing a "periodic table of alkanes", where for a given class of (alkane) isomers, one locates structures horizontally & vertically in consonance with two simple graph invariants, and looks for property correlations. The posetic work of Bartel [41] is also closely related, especially since in Bartel's approach in application to hetero-substituted benzenes the number of heteroatoms would presumably be chosen to specify vertical position in a Hasse diagram (while the type of heteroatom associates to horizontal position). Another general approach to deal with properties in a posetic manner is via "cluster expansions", as discussed in the next section.

5. Properties via Posetic Cluster Expansion

Another sort of property related characteristic of posetic reaction diagrams is by way of "cluster expansions" [42-46]. In a so-called "additive" framework, a scalar property X with value $X(\xi)$ for structure ξ may be decomposed in terms of contributions $x(\gamma)$ from (connected) sub-structures γ as

$$X(\xi) = \sum_{\gamma}^{\xi} f(\xi, \gamma) x(\gamma)$$

where the sum is over a prescribed type of sub-structures, the $x(\gamma)$ are cluster coefficients (dependent on the particular property) and $f(\xi, \gamma)$ is a *cluster function*, which satisfies

$$f(\xi, \gamma) \begin{cases} = 1 & , \gamma = \xi \\ \neq 0 & , \gamma \subset \xi \\ = 0 & , \gamma \not\subset \xi \end{cases}$$

Here the substructure relation corresponds to that of the directed reaction poset, as of figure 1. For X a heat of formation the $x(\gamma)$ for a one-substituent substructure γ may be viewed as a difference between a C-Cl & C-H bond energy, the three $x(\gamma)$ for two-substituent substructures γ make corrections for *ortho*-, *meta*-, or *para*-placements, and any higher terms make further (perhaps minor) adjustments. Ordinarily $f(\xi, \gamma)$ is a simple (property independent) invariant of the pair (ξ, γ) such as given by the zeta function

$$\zeta(\xi, \gamma) = \begin{cases} 1 & , \gamma \subset \xi \\ 0 & , \gamma \not\subset \xi \end{cases}$$

Indeed often [42,43,45,46] this is the only considered choice for f . In any event because of the presumed form of the general cluster function $f(\xi, \gamma)$, it has a functional inverse f^{-1} ,

$$\sum_{\gamma} f(\xi, \gamma) f^{-1}(\gamma, \xi') = \delta(\xi, \xi') = \sum_{\gamma} f^{-1}(\xi, \gamma) f(\gamma, \xi')$$

such that f^{-1} is also a cluster function. Basically this all works because f looks a lot like a matrix \hat{f} with (ξ, γ) th element $f(\xi, \gamma)$: first since the diagonal of \hat{f} is all non-zero, f must have be non-singular (and have an inverse), and second since \hat{f} is triangular, the inverse \hat{f}^{-1} is triangular too. Then the decomposing relation can be inverted to give the cluster coefficients in terms of the principle properties,

$$x(\xi) = \sum_{\gamma} f^{-1}(\xi, \gamma) X(\gamma)$$

The inverse ζ^{-1} of the zeta function is called [47] the Möbius function μ . For our case of the chlorobenzene poset, one possible cluster function f identifies $f(\xi, \gamma)$ as the number of ways that the sub-structure γ can be embedded in the structure ξ whence one obtains

$$\hat{f} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 3 & 0 & 3 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 3 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 3 & 0 & 1 & 2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 4 & 2 & 2 & 2 & 0 & 4 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 4 & 1 & 3 & 2 & 1 & 2 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 4 & 1 & 2 & 3 & 0 & 2 & 2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 5 & 2 & 3 & 4 & 1 & 6 & 3 & 1 & 2 & 2 & 1 & 0 & 0 & 0 \\ 1 & 6 & 3 & 6 & 6 & 2 & 12 & 6 & 3 & 6 & 6 & 6 & 1 & 0 & 0 \end{pmatrix}$$

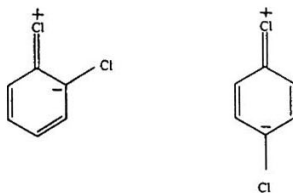
(where the structures for the rows & columns here are ordered as one reads the poset of figure 1, from left to right first and then from top to bottom). The inverse then is

$$\hat{f}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 3 & 0 & -3 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 3 & -1 & -1 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 3 & 0 & -1 & -2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -4 & 2 & 2 & 2 & 0 & -4 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -4 & 1 & 3 & 2 & -1 & -2 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -4 & 1 & 2 & 3 & 0 & -2 & -2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 5 & -2 & -3 & -4 & 1 & 6 & 3 & -1 & -2 & -2 & 1 & 0 & 0 & 0 \\ 1 & -6 & 3 & 6 & 6 & -2 & -12 & -6 & 3 & 6 & 6 & -6 & 1 & 0 & 0 \end{pmatrix}$$

The cluster expansion may be used in developing approximants for the property X, by truncating the expansion at substructures of a certain "size". Here *size* is conveniently measured in terms of the row number of the Hasse diagram, starting from 0 at the top. Then size *s* is just the number of chlorines substituted, and the *s*th order approximant is

$$X_s(\xi) = \sum_{\gamma}^{\xi, s} f(\xi, \gamma) x(\gamma)$$

where the sum is over all substructures of ξ of size no more than *s*. For the approximation scheme to be of the most value it should converge rapidly. For our chloro-benzene poset and X corresponding to heats of formation one might expect that truncation at *s*=1 could be reasonable for substituents which do not participate significantly in π -interactions. Chlorine does participate significantly, whence the *s*=2 approximation should probably be needed - because, e.g., for the dichlorobenzenes the ionic resonance structures with a + charge at one of the Cl-atoms tends to be unfavorable if the (counterbalancing) - charge ends up on the C-atom to which the other Cl-atom is attached. That is, for the *ortho*- & *para*-derivatives the resonance structures



end up being unfavorable (while there is no such problem with the *meta*-structure).

Such cluster expansion schemes though common are often presented in less formal fashion. They are for instance implicit in the wide-spread use of bond energies for estimating heats of formation or in the use of atomic Pascal's constants for estimating magnetic susceptibilities. Formal developments of cluster expansions have tended [42-46] to focus on an open-ended (potentially infinite) set of graph-theoretic structures, with properties to be expressed in terms of subgraph structures (perhaps with some decoration to account for different hetero-atoms). Hässelbarth [46] focuses on substituent structures of a fixed skeleton, much the same as here, but he implements a cluster expansion in fitting globally all terms to data for all the structures, without so much reference to the reaction poset considered here. The (ideal) fitting procedure which Hässelbarth proposes just recasts the full set of property values for all substituent structures into an equinumerous set of expansion coefficients. More commonly [42-45] truncated approaches are the point of focus. The current (posetic) approach is perhaps closest to that of ref. [44] and yields (hopefully worthwhile) approximants given just the property values for the smaller size members of the poset.

6. Further Benzene-Based Progressive Reaction Posets

There are several possible further sorts of progressive reaction diagrams even with restriction to the context of benzene. Indeed one could consider the successive stages of hydrogenation of benzene (without breaking any C–C σ -bonds). Reaction products can (much as in figure 1) still be indicated by a hexagon decorated with black dots at the sites of hydrogenation (thereby indicating saturated $-\text{CH}_2-$ groups), whence some of the structures of figure 1, might naturally be deleted (as unstable radical species) if there is no allowed pattern for a classical distribution of double bonds to accommodate the tetravalence of carbon. In particular there then would be no hexagon structures with an odd number of black dots (identifying $-\text{CH}_2-$ groups). Anyway the consequent posetic diagram appears as in figure 8, where each elementary step is now viewed to involve addition of two black dots.

Another possible substitution reaction is to consider ligands which are bidentate. To represent such pictorially one may mark the substituted sites with black dots and then connect

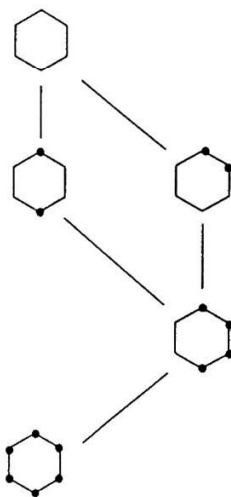
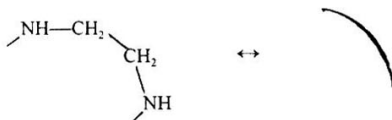


Figure 8 - The reaction diagram for successive hydrogenation of a benzene skeleton, with the black dots identifying $-\text{CH}_2-$ groups.

such dots by a curve if they correspond to the two ends of a symmetric bidentate ligand, say perhaps



which is to be connected to two carbons of the benzene ring. The result is illustrated in figure 9, where attention is paid only to structural isomerism. Notably the double-coset labelling of equivalence classes of isomers still applies [48], though now the form of the ligand-symmetry group is to reflect the more complicated structural connections between pairs. E.g., if there are four sites and two nondistinguished bidentate ligand structures l_1, l_2 & l_3, l_4 with l_1, l_2 connected together while also l_3 & l_4 are connected together, then the portion of $P[L]$ concerned with these

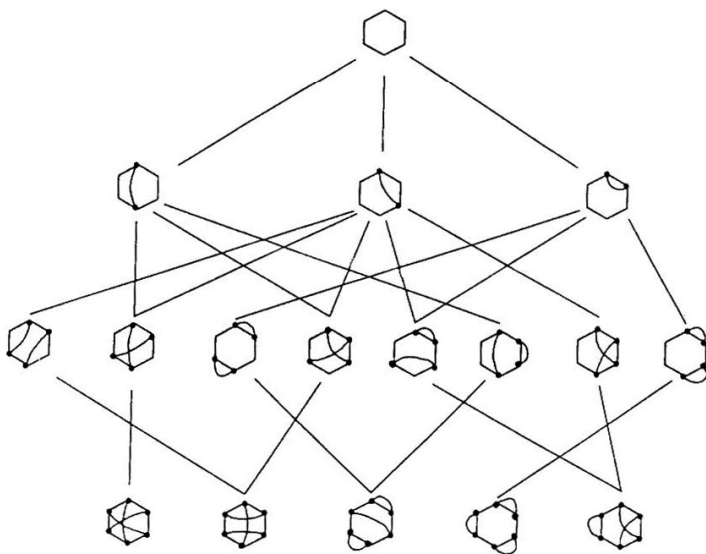


Figure 9 - Reaction poset for substitution of symmetric bidentate "chelates".

ligands takes the form

$$P[L] = \{I, (13)(24)\}$$

if the ends of these two bidentate ligands are distinguished. And if for this same 4-site circumstance the ends are not distinguished (i.e., l_1 is equivalent to l_2 & l_3 is equivalent to l_4), then the portion of $P[L]$ concerned with these ligands takes the form

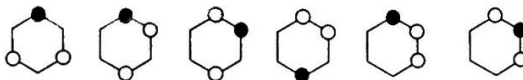
$$P[L] = \{I, (13)(24)\} \{I, (12)\} \{I, (34)\}$$

For the hexa-substituted benzene case

$$P[L] = \{I, (13)(24), (15)(26), (35)(46), (135)(246), (153)(264)\} \{I, (12)\} \{I, (34)\} \{I, (56)\}$$

Here in this 4-fold product of subgroups, the first subgroup permutes the three bidentate substituents about as a whole, and the three final 2-element groups permute within the three different bidentate substituents. Of course in chemical reality some ways of adding some bidentate substituents may be forbidden, say if the bidentate substituent is such as not of sufficient length to stretch between *para*-positions (when skeletal geometry and bidentate bond lengths & bond angles are taken into account).

Another interesting case concerns the possibility of allowing more than one type of substituent. Indeed in section 2 the possibility was mentioned of using such multi-substituent reaction diagrams to distinguish between the regular-hexagonal & trigonal-prismatic skeletons for benzene. Such progressive reactions for several substituents may be considered to lead to posets also. For simultaneous fluorination or chlorination of (hexagonal-skeleton) benzene the consequent reaction poset has 92 vertices (each corresponding to a structural isomer $C_6H_mCl_nF_p$, for m, n, p being non-negative integers which sum to 6), so that this poset is of somewhat greater complexity than our earlier examples. To present it graphically is a challenge, and here in figure 10 we present it only in abbreviated form. We suppress the display of the molecular structures, and list only overall formula, with the benzene nuclei (including whatever H-atoms are attached) abbreviated to ϕ . Thence whole collections of structures are compressed to a single entry. For instance, for the entry $6 \phi \text{-ClF}_2$, one has $n=1$ & $p=2$ with the six structures



where black & white dots respectively indicate Cl & F atoms. The different rows of figure 10 identify different total degrees ($n+p$) of substitution, proceeding from the unsubstituted top (with $n+p=0$) to the fully substituted bottom (with $n+p=6$). Further the different n & p values are arranged so that chlorination reactions slope downward & to the left while fluorination reactions slope downward to the right. The numbers of structural isomers for each overall formula are indicated, and further the numbers on the reaction arrows indicate the numbers of arrows in the parent 92-node diagram now compressed into a single arrow of our abbreviated diagram.

For a greater number M of types of substituents the parent uncompressed diagrams become even more formidable (with 430, 1505, 4291, 10528, 23052, 46185, 86185, 151756, etc nodes at $M=3, 4, 5, 6, 7, 8, 9, 10$, etc). Evidently algebraic/combinatoric schemes to deal with such large (and complex) diagrams may be of use to develop.

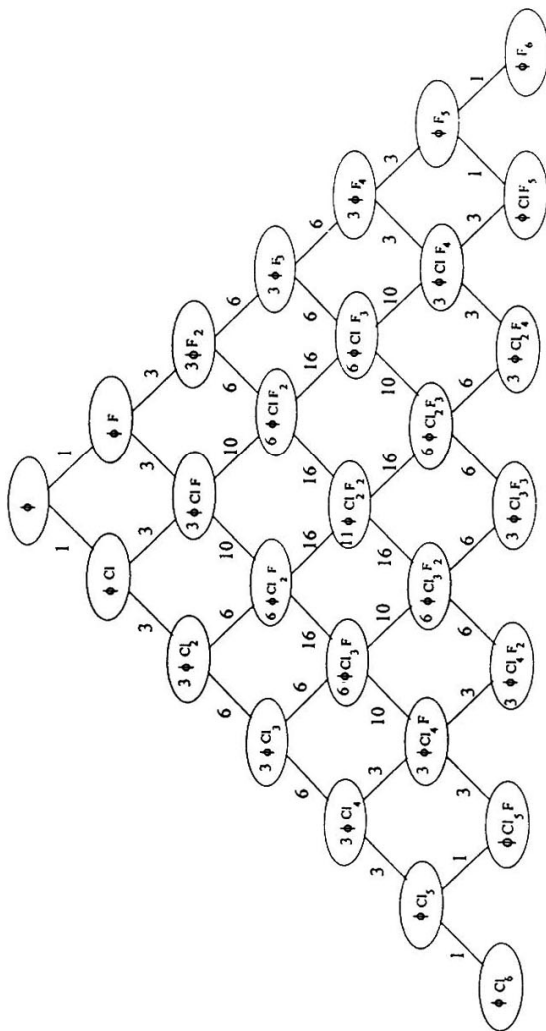


Figure 10 - The compressed double-substituent reaction poset for benzene.

7. Overview of Progressive Reaction Diagrams

Evidently there are numerous examples of progressive reaction graphs which may be viewed to define posets, and there are some possibilities for the utilization of the posetic structure. The benzene-chlorination poset presented here is just an illustrative example, and there may be many more posetic manipulations of chemical interest beyond those indicated here. An example displaying condensation “paths” for aromatic hydrocarbons appearing as pyrolytic depositions is given in ref. [49], and is redrawn in figure 11, though presumably there could be more nodes in the full reaction graph.

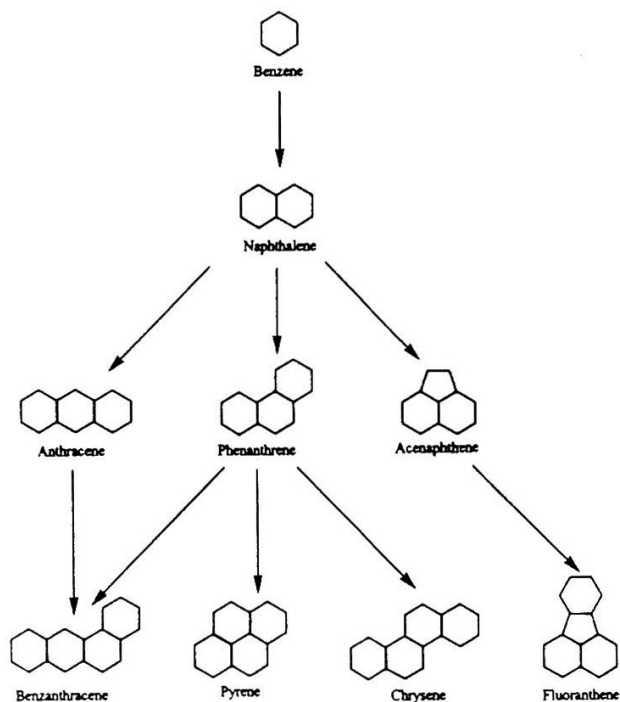


Figure 11 - A progressive reaction diagram for the pyrolytic condensation of aromatic hydrocarbons.

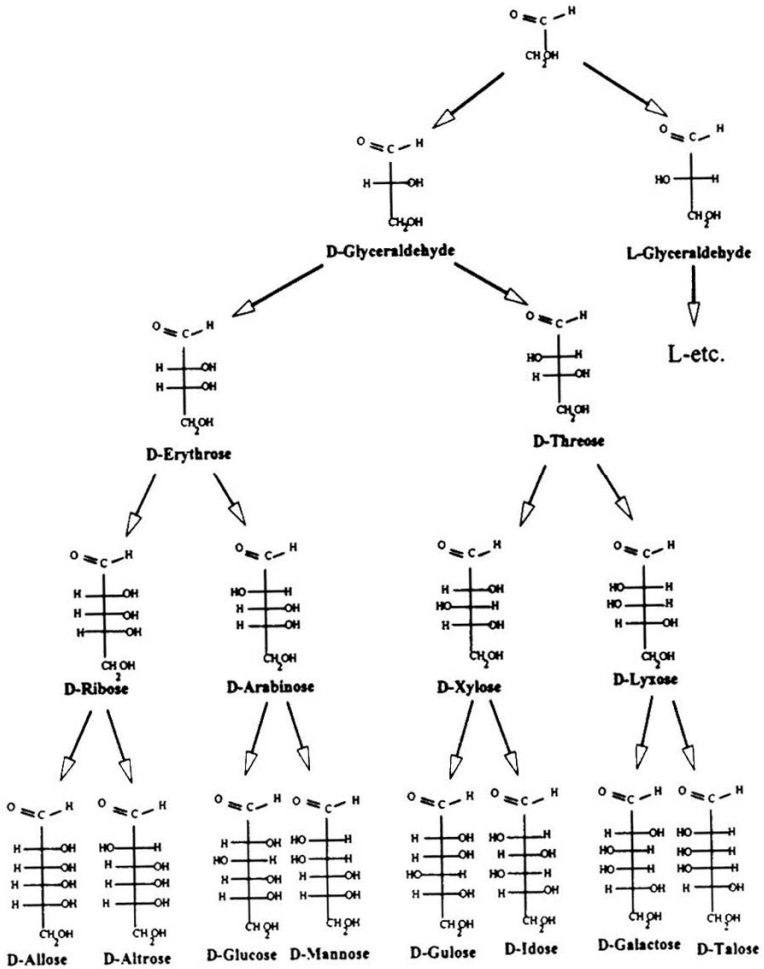


Figure 12 - Posetic organization of carbohydrates.

An example of a posetic diagram is that of saturated acyclic hydrocarbons (i.e., of alkanes) ordered according to a subgraph relation. Indeed the restriction to acyclics need not be made, though with this restriction the poset is found in figure 1 of Bertz [50]. This subgraph case can be viewed as an ordering according to another chemical characteristic, and cluster expansions as discussed previously [42-46] often have been viewed to apply primarily for such cases. Such a poset can be viewed as ordering according to a special type of chemical reaction, involving the addition of an additional saturated C-atom into the network (via a single C-C bond, for the acyclic case).

Another posetic example which can be viewed to involve a rather unusual type of reaction involves a classification of carbohydrates, e.g., as shown in figure 26.3 of [51] or the figure of [52], and as also shown in the present figure 12 here, where the chirality of the different centers is indicated by the standard Fischer-projection convention. This does not involve a simple subgraph relation, but may still be viewed to involve a rather unusual reaction where a -HCOH- group is inserted (in either a D- or L-chiral form) directly adjacent to the aldehyde end of a carbohydrate.

Yet another type of interesting example where the idea of progressive reaction graphs seems to arise involves a reaction of polar species at the surface of a cell or micelle. The polar species may be conceivably adsorbed with either their + or their - end directed to the micelle, and one may change from one adsorption arrangement to another by an elementary process where the direction (with respect to the micelle) of exactly one of the adsorbants is reversed. But of course the micelle prefers one orientation or the other for the adsorbants. Thence one can make a reaction graph showing these elementary steps with a direction appearing because the reactant & product have different numbers of adsorbants in the favored orientation to the micelle. And because of the favoritism of the micelle these elementary steps may also be taken to identify "unit" changes in energy. For four polar adsorbants on a micellar surface one then has the reaction graph of figure 13, where the four arrows indicate the orientations of each of the adsorbants on the surface. Such a picture appears of relevance in ref. [53] where the micelle is just a cavity in a polar liquid, and the arrangement of the dipolar molecules provides a partial description of the structure of the cavity (walls). Such an example parallels closely the progressive substitution on a skeleton - it involves the progressive reversal of arrows (say from in to out) on a fixed skeleton, with no symmetry equivalences presumed. Of course many experimentally realized micelles would have many more than 4 dipolar species at the surface, so that much more complicated reaction diagrams would generally result.

A related example involves step-wise ionization of large species (ionomers), which may extend up to colloidal sizes, or beyond, say to cell surfaces. For instance, the acid dissociation of humic and fulvic acids occurs [54] in a variety of (reversible) stages. And again this should lead to a posetic progressive deprotonation reaction diagram, though especially for humic & fulvic acids there are different structures, so that there is even an ensemble of different reaction diagrams, with different equilibrium constants for each edge of the diagram. It is to be emphasized that even when reactions proceed in both directions, the directed posetic structure frequently is of interest, with, e.g., in the present case the relevant extent of deprotonation determining the direction (or in our benzene-substituent the extent of chlorination determining the direction).

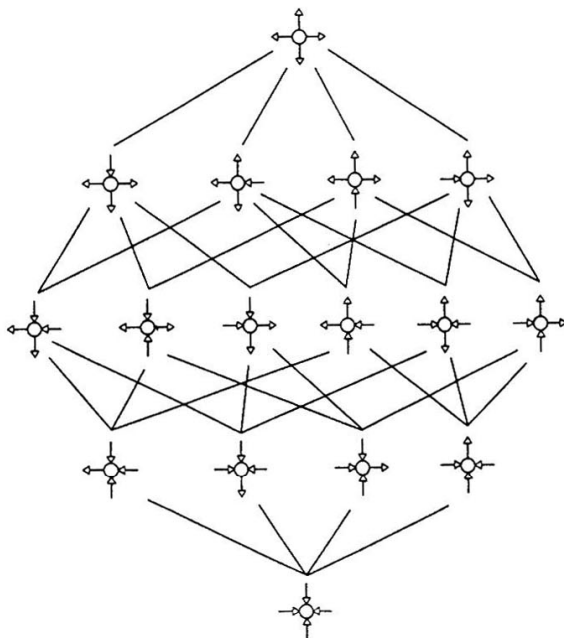


Figure 13 - Posetic organization of orientational possibilities for 4 adsorbants on a micellar surface.

One example involving ionization the particular structures of interest concerns "protein charge ladders". Here [55] a (globular) protein is reacted at different amino groups on the surface to be replaced by acyl groups which then tend to form negative ions. The different patterns of substitution arrangements, occurring on different levels of a progressive reaction poset, have rather different properties and are then readily separated. The "ladder" description of ref. [55] comes in that the levels in the poset are viewed as the rungs of a ladder.

A final example of current interest and more closely similar to our benzene example is that of the reaction poset for progressive hydrogenation of buckminsterfullerene C_{60} . This hydrogenation (via the Birch reaction) is certainly selective in terms of what species are formed, since on arriving at $C_{60}H_{36}$ it seems [56] that most of what occurs is but a single isomer, but presumably along the way several paths may be taken (within the reaction poset), there being

many conceivable structural isomers. Progressive halogenation of C_{60} is also of interest [57], and comparable questions can arise in this context. In any event very little is known about such a reaction poset other than the fact that it is immense. Even the enumeration of stable chemical isomers at each level of the poset is not generally known, the difficult point being the exclusion of unstable radicaloid species from the diagram. That is, though the different (symmetry inequivalent) patterns of placing H-atoms on subsets of sites of C_{60} may be neatly done [58] via Polya enumeration theory [24,25] (or equivalently via double-coset enumeration [27,28]), many of these structures do not permit a Kekule structure on the remaining sites, and thence are presumably radicaloid, and thence so chemically unfavorable as to be desired to be excluded from the enumeration. Certainly the lack of a Kekule structure (and the presumed chemical unfavorableness) applies for $C_{60}H_n$, when n is odd, but perhaps even the bulk of the even- n species are radicaloid. For $n=0$ there are numerous (i.e., 12500) Kekule structures for the single "empty" substitution pattern, for $n=2$ we observe [59] that all of the 23 structural isomers of $C_{60}H_2$ exhibit many Kekule structures, for $n=4$ at least 20 of the 4190 structural isomers of $C_{60}H_4$ exhibit no Kekule structure, and the percentage without a Kekule structure we speculate to increase with (even) n , at least up through $n=30$ where there are $\sim 9.855 \times 10^{14}$ positional isomers. The Polya isomer counts are symmetric to interchange of H-substituted & non-H-substituted sites, but the non-radicaloid counts are not. *E.g.*, at $n=58$ there are 21 non-Kekule structures out of 23 $C_{60}H_{58}$ structures. Thus systematic mathematical tools to effectively deal with the relevant poset (and even the associated more basic relevant isomer enumerations) seem so far to be largely wanting.

In general the combinatoric generation of properties for different structures in such reaction posets should be of chemical interest, if the combinatoric properties are chosen to correlate with chemical properties, perhaps developed in terms indicated in sections 3 & 4 here. Notably standard recursive combinatoric techniques have been so applied for some isomer problems, in recent work of ours [60], however with little attention so far to posetic relations. Presumably such work might be usefully extended.

Overall it appears that there is a wide-spread occurrence of interesting progressive reaction posets, though little has been previously done in any formal sense. Hopefully more will be understood of what to do with these posetic structures, and mathematical questions of how to deal efficiently with even complex cases of these posets will be elucidated. The present discussion hopefully indicates some possibilities, for a range of future works.

The authors acknowledge support from the Welch Foundation of Houston, Texas.

References

- [1] A. T. Balaban, pages 137-158 in *Graph-Theoretical Approach to Chemical Reactivity*, ed. D. Bonchev & O. Mekenyan (Kluwer, Dordrecht, 1994).
- [2] E. J. Corey & X-M. Cheung. *The Logic of Chemical Synthesis* (Wiley, NY, 1989).
- [3] J. B. Hendrickson. *J. Am. Chem. Soc.* **99** (1977) 5439.
- [4] S. Bertz, *J. Chem. Soc., Chem. Comm.* (1984) 218.

- S. Bertz, J. Chem. Soc., Chem. Comm. (1986) 1627.
- [5] S. Bertz & T. J. Sommer, *Organic Synthesis: Theory & Applications* **2** (1993) 67.
- [6] M. Eigen, *Naturwissenschaften* **58** (1971) 465.
- [7] See, e.g.: T. Hill, *Free Energy Transduction in Biology* (Academic Press, NY, 1977), or J. Schnakenberg, *Thermodynamic Network Analysis of Biological Systems* (Springer-Verlag, Berlin, 1977).
- [8] O. N. Temkin, A. B. Zeigarnik, & D. Bonchev, *Chemical Reaction Networks* (CRC Press, Boca Raton, Florida, 1996).
- [9] This mathematical idea is widespread, but the prolegomenon in this special issue is an abbreviated introduction, which also gives many primary references. The article by Bruggemann *et al* as well as others in the present special issue of MatCh also provide many examples of chemically relevant posets.
- [10] E. Heilbronner & J. D. Dunitz, *Reflections on Symmetry* (VCH, Weinheim, 1993).
- [11] J. D. Dunitz, *Proc. Natl. Acad. Sci.* **93** (1998) 14260.
- [12] A. Kekule, *Bull. soc. chim. France* **3** (1865) 98.
A. Kekule, *Annalen* **137** (1866) 129.
A. Kekule, *Annalen* **162** (1872) 77.
- [13] A. C. Lunn & J. K. Senior, *J. Phys. Chem.* **33** (1929) 1027.
- [14] A. T. Balaban, *Rev. Roum. Chem.* **19** (1974) 1323.
- [15] W. Hässelbarth, *J. Comp. Chem.* **8** (1987) 700.
- [16] There evidently are 4 papers with the first [W. Körner, *Sci. Nat. Econ. Palermo* **5** (1874) 212] referenced by McBride in [17] especially difficult to find. The fourth 142-page article [W. Körner, *Gazz. Chim. Ital.* **4** (1874) 305] describes syntheses of nearly 200 substituted benzenes while also making some theoretical arguments, and it is more readily available as a 38-page summary with comments (evidently) by H. E. Armstrong in *J. Chem. Soc.* **29** (1876) 204.
- [17] J. M. McBride, *J. Am. Chem. Soc.* **102** (1980) 4134.
- [18] A. Ladenberg, *Chem. Ber.* **2** (1869) 140 & 272.
A. Ladenberg, *Chem. Ber.* **5** (1872) 322.
- [19] J. H. Van't Hoff, *Voorstel tot Uitbreiding der ... Structuur Formules in de Ruimte* (Utrecht, 1874) published in translation to French the following year with an abbreviated title *La Chimie dans l'Espace* (Paris, 1875), and appearing in English as *The Arrangement of Atoms in Space* (2nd edn, London, 1898).
- [20] J. E. le Bel, *Bull. Soc. Chim. (Paris)* **22** (1874) 337.
- [21] C. A. Russell, *The History of Valency* (Leicester University Press, Leicester, 1971).
- [22] H. Kolbe, *J. prakt. Chem.* **15** (1877) 474.

- [23] J. Thomsen, *Berichte* **13** (1880) 1806 & 2166.
A. Claus, *Chem. Ber.* **14** (1881) 432.
R. Demuth & V. Meyer, *Chem. Ber.* **21** (1888) 265.
- [24] J. E. le Bel, *Bull. Soc. chim. France* **37** (1882) 300.
J. E. le Bel, *Bull. Soc. chim. France* **47** (1892) 164.
- [25] G. Polya, *Acta Math.* **68** (1937) 145.
G. Polya & R. C. Read, *Combinatorial Enumeration of Groups, Graphs, and Chemical Compounds* (Springer-Verlag, Berlin, 1987).
- [26] See also, e.g.,
F. Harary & E. Palmer, *Graphical Enumeration* (Academic Press, NY, 1973).
K. Balasubramanian, *Chem. Rev.* **85** (1985) 599.
S. Fujita, *Symmetry & Combinatorial Enumeration in Chemistry* (Springer-Verlag, Berlin, 1992).
Or see many combinatorics texts, a nice one being:
C. Berge, *Principles of Combinatorics* (Academic Press, NY, 1971).
- [27] A. Claus, *Berichte Verhand. Naturforsch. Gesellschaft Freiburg* **4** (1867) 116.
H. E. Armstrong in [16] above and also *J. Chem. Soc.* (1887) 264.
- [28] E. Ruch, W. Hässelbarth, & B. Richter, *Theor. Chim. Acta* **19** (1970) 288.
- [29] E. Ruch & D. J. Klein, *Theor. Chim. Acta* **63** (1983) 447.
W. Hässelbarth, *Theor. Chim. Acta* **67** (1985) 427.
- [30] A. J. Rocke, *Ann. Sci.* **42** (1985) 355.
- [31] S. G. Brush, *Stud. Hist. Phil. Sci.* **30** (1999) 21 & 263.
- [32] A. v. Baeyer, *Berichte* **19** (1886) 1797; **23** (1890) 1272.
A. v. Baeyer, *Ann.* **245** (1888) 103; **251** (1889) 257; **258** (1890) 145; & **269** (1892) 145.
- [33] A. Ladenburg, *Lectures on the History of the Development of Chemistry since the Time of Lavoisier* (Simpkin, Marshall, Hamilton, Kent & Co., London, 1905) being a "faithful translation" of the German 2nd edition of 1886 with an added chapter (not concerning benzene).
- [34] A. Ladenburg, *Berichte* **7** (1874) 1133..
- [35] 7^{8th} Edition *CRC Handbook of Chemistry & Physics*, ed. D. R. Lide (CRC Press, Boca Raton, Florida, 1997).
- [36] 3^{8th} Edition *CRC Handbook of Chemistry & Physics*, ed. C. D. Hodgeman (CRC Press, 1957).
- [37] *Physical Properties of Chemical Compounds*, Advances in chemistry series #15 (Am. Chem. Soc., Washington, D. C., 1955).
- [38] S. Smiles, *Chemical Constitution and some Physical Properties* (Longman, Green, & Co., London, 1910).

- [39] *Handbook of Chemical Property Estimation Methods*, ed. W. J. Lyman, W. F. Reehl, & D. H. Rosenblatt (McGraw-Hill, NY, 1982)
- [40] M. Randić & C. L. Wilkins, *J. Phys. Chem.* **83** (1979) 1525.
M. Randić & C. L. Wilkins, *Chem. Phys. Lett.* **63** (1979) 332.
M. Randić & C. L. Wilkins, *Intl. J. Quantum Chem.* **18** (1981) 1005.
- [41] H.-G. Bartel, *MatCh* **30** (1994) 9; **32** (1995) 27; **34** (1996) 79; & **36** (1997) 185.
- [42] E. A. Smolenski, *Russ. J. Phys. Chem. [Engl. trans.]* **38** (1964) 700.
M. Gordon & J. W. Kennedy, *J. Chem. Soc., Faraday Trans. II* **69** (1973) 484.
- [43] J. W. Essam, J. W. Kennedy, M. Gordon, & P. Whittle, *J. Chem. Soc. Faraday II* **73** (1977) 1289.
J. W. Kennedy & M. Gordon, *Ann. N. Y. Acad. Sci.* **319** (1979) 331.
- [44] D. J. Klein, *Intl. J. Quantum Chem. S* **20** (1986) 153.
D. J. Klein, T. G. Schmalz, & L. Bytautas, *SAR & QSAR Environ. Research* **10** (1999) 131.
- [45] I. I. Baskin, M. I. Skvortsova, I. V. Stankevich, & N. S. Zefirov, *Doklady Chem. [English trans.]* **359** (1994) 346.
M. I. Skvortsova, I. I. Baskin, O. L. Slovokhotova, & N. S. Zefirov, *Doklady Chem. [English trans.]* **336** (1994) 496.
- [46] W. Hässelbarth, *MatCh* **31** (1994) 7.
- [47] G.-C. Rota, *Zeit. Wahr. & Verw. Geb.* **2** (1964) 340.
- [48] D. J. Klein & A. H. Cowley, *J. Am. Chem. Soc.* **100** (1978) 2593.
- [49] P. Lucas & A. Marchand, *Carbon* **28** (1990) 207.
- [50] S. H. Bertz, *Disc. Appl. Math.* **19** (1988) 65.
- [51] This example was pointed out by Dr. M. Lesko as occurring on p. 1017 of J. McMurry. *Organic Chemistry*, 4th edn. (Brooks/Cole Pub., Pacific Grove, California, 1996).
- [52] A. D. McNaught, *Chem. Intl.* **17** (1995) 98.
- [53] V. B. P. Leite & J. N. Onuchic, *J. Phys. Chem.* **100** (1996) 7680.
- [54] J. Buffle, *Metal Ions Biol. Systems* **18** (1984) 165.
R. L. Wershaw, *J. Contam. Hydrol.* **1** (1986) 29.
R. R. Engebretson & R. von Wandruszka, *Environ. Sci. Technol.* **28** (1994) 1934.
- [55] J. D. Carbeck, I. J. Colton, J. Gao, & G. M. Whitesides, *Acc. Chem. Res.* **31** (1998) 343.
- [56] R. E. Haufler, J. Conceicao, L. P. F. Chibante, Y. Chai, N. E. Byrne, S. Flanagan, M. M. Haley, S. C. O'Brien, C. Pan, Z. Xiao, W. E. Billups, M. A. Ciufolini, R. M. Hauge, J. L. Margrave, L. J. Wilson, R. F. Curl, & R. E. Smalley, *J. Phys. Chem.* **94** (1990) 8634.
C. Rüchardt, M. Gerst, J. Ebenhoch, H.-D. Beckhaus, E. E. B. Campbell, R. Tellmann, H. Schwartz, T. Weiske, & S. Pitter. *Angew. Chem., Intl. Ed. Engl.* **32** (1993) 584.
M. I. Attalla, A. M. Vassallo, B. N. Tattam, & J. V. Hanna, *J. Phys. Chem.* **97** (1993) 6329.

- [57] J. H. Holloway, E. G. Hope, R. Taylor, G. J. Langley, A. G. Avent, T. J. Dennis, J. P. Hare, H. W. Kroto, & D. R. M. Walton, *J. Chem. soc., Chem. Commun.* (1991) 966.
G. A. Olah, I. Bucci, C. Lambert, R. Aniszfeld, N. J. Trivedi, D. K. Sensharma, & G. K. Surya Prakash, *J. Am. Chem. Soc.* **113** (1991) 9385.
P. R. Birkett, P. B. Hitchcock, H. W. Kroto, R. Taylor, & D. R. M. Walton, *Nature* **357** (1992) 479.
- [58] H. Hosoya, *Gendai-Kagaku* **201** (1987) 38, (as mentioned in Fujita's 1991-article below).
S. Fujita, *Bull. Chem. Soc. Japan* **64** (1991) 3215.
K. Balasubramanian, *Chem. Phys. Lett.* **183** (1991) 292.
K. Balasubramanian, *J. Phys. Chem.* **97** (1993) 6990.
- [59] H.-Y. Zhu, D. J. Klein, & T. G. Schmalz, unpublished work.
- [60] L. Bytautas & D. J. Klein, *J. Chem. Inf. & Comp. Sci.* **38** (1998) 1063.
L. Bytautas & D. J. Klein, *Theor. Chem. Acc.* **101** (1999) 371.
L. Bytautas & D. J. Klein, *J. Chem. Inf. & Comp. Sci.* **39** (1999) 803.
L. Bytautas & D. J. Klein, *Croatica Chemica Acta* (accepted 1999).
L. Bytautas & D. J. Klein, *Phys. Chem. - Chem. Phys.* **1** (1999) 5565.
L. Bytautas, D. J. Klein, & T. G. Schmalz, *New J. Chem.* (submitted 1999).
L. Bytautas & D. J. Klein, *J. Chem. Inf. & Comp. Sci.* (submitted 1999).