

# Combinatorics of Reaction-Network Posets

Douglas J. Klein<sup>\*a</sup>, Teodora Ivanciuc<sup>a,b</sup>, Anton Ryzhov<sup>a</sup> and Ovidiu Ivanciuc<sup>b</sup>

<sup>a</sup>Texas A&M University at Galveston, Galveston, Texas 77553-1675, USA

<sup>b</sup>UTMB, Galveston, Texas 77553, USA

**Abstract:** Reaction networks are viewed as derived from ordinary molecular structures related in reactant-product pairs so as to manifest a chemical *super-structure*. Such super-structures then are candidates for applications in a general combinatoric chemistry. Notable additional characterization of a reaction super-structure occurs when such reaction graphs are directed, as for example when there is progressive substitution (or addition) on a fixed molecular skeleton. Such a set of *partially ordered* entities is in mathematics termed a *poset*, which further manifests a number of special properties, as then might be utilized in different applications.

Focus on the overall "super-structural" poset goes beyond ordinary molecular structure in attending to how a structure fits into a (reaction) network, and thereby brings an extra "dimension" to conventional stereochemical theory. The possibility that different molecular properties vary smoothly along chains of interconnections in such a super-structure is a natural assumption for a novel approach to molecular property and bioactivity correlations. Different manners to interpolate/extrapolate on a poset network yield quantitative super-structure/activity relationships (*QSSARs*), with some numerical fits, e.g., for properties of polychlorinated biphenyls (PCBs) seemingly being quite reasonable. There seems to be promise for combinatoric posetic ideas.

**Keywords:** Posets, partial orderings, reaction networks, substitution reactions, QSAR, QSSAR.

## 1. INTRODUCTION

During the last two decades, the area of "combinatorial chemistry" has extensively developed to involve a "guided" set of reactions so as to generate whole libraries of chemical compounds to be investigated – by way of a related and also newly developed field of "chemo-informatics". But there are considerably wider involvements of combinatorics in chemistry than appears with the now standard combinatorial chemistry, or even chemo-informatics. In fact, combinatorics is generally deeply involved in "permutational isomerism", "enzyme kinetics", "QSAR", "chemical graph theory", and many other areas, even within the more physics-related subjects of "statistical mechanics" and "quantum chemistry".

Here attention is directed to reaction networks and their combinatoric structure, such as then may be utilized in the characterization of the network, and of the substances appearing in the network. Combinatorics on reaction networks is relevant for many areas of chemistry, biochemistry and molecular biology. Many reactions involve cycles, but also many manifest an overall "direction" which then provides mathematical (combinatoric) features of potential use. A nice generic example of such a directed network involves substitutions, additions, or eliminations on a fixed skeleton with several locations to make such change. These directed networks turn out to manifest the mathematical feature of a "partially ordered set", or "poset", where the "following from" relation, though ordering many pairs of species in the network, does not so generally relate an arbitrary pair. That is, though every chlorinated benzene is ordered with respect to benzene in that it "follows from" benzene through some succession of chlorinations, there are trichlorobenzenes which

which do not so "follow from" some dichlorobenzenes – e.g., 1,3,5-trichlorobenzene does not result from chlorination of 1,4-dichlorobenzene. Granted such a posetic reaction network, it frequently seems plausible that many molecular properties generally vary gradually over the reaction network or poset. Thence, it is desirable to have some powerful combinatoric procedures for fittings to interpolate/extrapolate properties amongst the different substances appearing in the network. Simple illustrations can be made on the reaction network of chlorinated benzenes, with 13 members (including benzene). A less trivial example involves the 210-member reaction poset resulting from (substitiutional) chlorination of biphenyl – the polychlorinated-biphenyls (PCBs) being of keen interest since they frequently occur as carcinogenic contaminants of waters and soils.

In the next section, the ubiquity of some reaction-network ideas is emphasized with focus on the partially ordered case. Several examples of substitution (or addition or elimination) reaction networks are noted to realize posets – such as those that form a recognized area of mathematics. In the section after that, there is some indication of a step toward a formal characterization of substitution-reaction posets (on a fixed skeleton). After this, we briefly address the problem of systematic methods by which to interpolate/extrapolate on such a network. And in the concluding section, it is sought to indicate how reaction-network posets might be viewed in an even broader framework, to encompass all kinds of chemical "periodicities" – such as manifested in Mendeleev's periodic chart.

## 2. CHEMICAL COMBINATORICS AND REACTION POSETS

Traditional *combinatorial chemistry* formally entails: first a sequence  $M = (M_1, M_2, \dots, M_p)$  of molecules  $M_i$ , and second a sequence  $L = (L_1, L_2, \dots, L_q)$  of ligands  $L_j$  to be

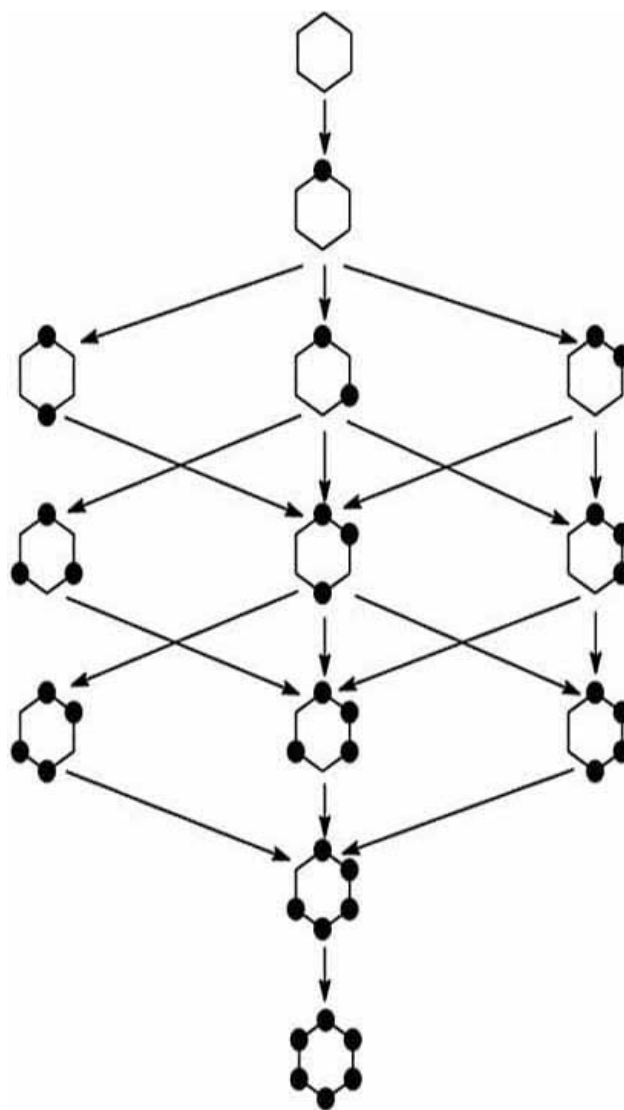
\*Address correspondence to this author at the Texas A&M University at Galveston, Galveston, Texas 77553-1675, USA;  
E-mail: kleind@tamug.edu

substituted (or added) at a given site of each molecule. That is, a  $P \times Q$  grid of reactions for each  $M_i, L_j$ -pair is considered to yield a new substituted molecule  $M_iL_j$ , the collection of which then constitutes a *library* of structures whose properties are to be characterized. Typically this is regarded in an experimental framework, but it may also be treated theoretically, e.g., with there being some such development [1-4]. A further theoretical development follows when some graph-theoretic QSAR techniques (in terms of "topological indices") are suitably adapted [5] to the combinatorial-chemical circumstance, where experimental efficiencies are also reflected as theoretical efficiencies. But the reactions are utilized in such an approach in a special manner, and there might be further interesting ways to utilize reactions in a combinatorially useful manner.

Indeed, there are many more combinatorial structures resulting from reactions, especially if reactions are "piled up" one after another. One ubiquitous example (in synthetic organic chemistry) is provided by synthesis reaction graphs, such as those that are numerous illustrated in Corey and Cheng's (1989) seminal book [6], *The Logic of Chemical Synthesis*. These synthesis graphs are differentiated [7] as to being "linear" or "convergent", and there has been some modest degree of formal characterization (e.g., by Hendrickson and by Bertz [8, 9]). 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", with preferred directions on the edges. These reaction networks have been much studied, even in a formal sense, e.g., as in Eigen [10] or Hill [11], though in these cases the possibility of a posetic interpretation is complicated by their cyclic nature. Yet another type of reaction graph is that for degenerate rearrangements (of non-rigid molecules), which have also been formally examined and reviewed [12]. However, these graphs not only contain cycles, but the edges are best taken as undirected. General complex sets of chemical reactions often similarly exhibit cycles – see, e.g., Temkin *et al.* [13] or Ross *et al.* [14].

But there is a general class of reaction networks which are directed, are generally acyclic, and are somewhat akin to the synthesis graphs, though of a more nicely defined posetic type. 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 Fig. (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  $\alpha$  to a second  $\beta$ , if  $\beta$  can be obtained from  $\alpha$  by a single minimal step of chlorination (i.e. by the replacement of one H-atom by one Cl-atom, without moving around any other Cl-atoms which might already be attached). Note that in general not all the  $n$ -substituted isomers so arise from a particular isomer with substituents. In the mathematical theory, the reaction-network diagram as in Fig. (1) is termed a *Hasse diagram* which has elements  $\alpha \succ \beta$  if  $\alpha$  lies higher in the diagram than  $\beta$  and if also there is a downward directed path (of one or more links) from  $\alpha$  to  $\beta$ . As such this relation satisfies the conditions:

$$\begin{aligned} \alpha \succ \beta &\Rightarrow \beta \not\succ \alpha \\ \alpha \succ \beta \text{ and } \beta \succ \gamma &\Rightarrow \alpha \succ \gamma \end{aligned} \quad (1)$$



**Fig. (1).** The posetic reaction diagram for successive chlorination of the benzene skeleton, with black dots identifying sites where H ligands are substituted by Cl.

Such a collection is termed a "partially ordered set" or *poset*. There is also the relation  $\alpha \succeq \beta$  which often is used to mean that  $\alpha$  and  $\beta$  satisfy either  $\alpha \succ \beta$  or  $\alpha = \beta$ . Indeed there is a large mathematical literature on posets, including: Birkhoff [15], Alberti and Uhlmann [16], Trotter [17], Pecaric and Proschan [18], and Neggers and Kim [19]. These mathematical works having developed a fair amount of general theory, noting that posets occur in many diverse forms in mathematics. Posets also occur quite generally in science as reviewed in Rival [20] (in 1983), though the areas (from behavioral science to physics) considered there somehow miss chemistry. A special (2001) issue [21] of *MATCH Communications in Mathematical and in Computer Chemistry* focuses on posets in chemistry, and a recent book [22] focuses jointly on chemistry and environmental science. These provide many further chemical examples of posets. For the present substitution-reaction posets, there is a direction, though it is not precluded that the reactions occur only to higher degrees of substitution. Rather it is assumed that

this is just an *a priori* direction of "positive" reaction, while reverse "negative" reaction is also possible. Still the partial ordering remains (chemically) relevant – just as for particle motion in 1-dimension, the ability of the particle to move in both directions does not mean the numerical ordering of coordinates is unimportant.

Substitution-reaction posets have special features. In particular, they have:

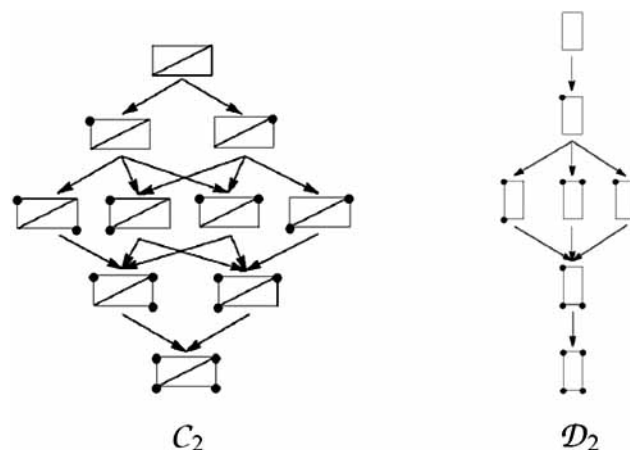
- \* a rank (corresponding to the number of substitutions made);
- \* a unique maximum (being the parent unsubstituted skeleton);
- \* a grading with the number of reaction steps between any two comparable species being independent of directed path (or chain) between the two species;
- \* metric regularity properties (as described elsewhere [23]); and
- \* some symmetry-mediated characteristics (related to the symmetry of the skeleton on which substitution is made).

Moreover, if different substitutions at fixed available sites are independent, then these posets are *self-dual* (involving the interchange of substituted and unsubstituted sites). But it seems that the various mathematical works [15-20] do not address this particular type of poset, with most of the extensive mathematical work on posets being either more generally focused or else directed to other special classes of posets. This lack of mathematical recognition of reaction-network posets is rather surprising, because they can be regarded as complicit in Polya's widely renowned theory [24-26] of enumeration under group action. Even more notably, these posets can be interpreted as embodying the finite geometric manifestation of Felix Klein's renowned [27, 28] "Erlangen program" – where a geometry is characterized (or even defined) by the action of a group on a chosen "space".

Notwithstanding the lack of mathematical attention to such progressive-reaction posets, it should be emphasized that there are many chemical examples of this particular type of reaction network, and that they can be rather large and challenging even to diagram in a clear fashion. As already noted (in conjunction with Fig. (1)) the benzene substitution network provides a prototypical example with 13 members and seven ranks. This poset is complicit in early work (in 1874) by Körner [29, 30] as a summary of an extensive set of experiments seeking to characterize the molecular structure of benzene – and occasional summaries of Körner's work have diagrammatically displayed this reaction network – as in references [31, 32]. Also, this particular reaction network seemingly has in fact been independently rediscovered a few times [33-36]. An interesting such rediscovery of this poset is that by Zivkovic, who shows that if the substitutions are of H by D, then under the assumption of a common potential-energy surface, the (lower-lying) harmonically approximated vibrational frequencies manifest a type of numerical ordering consonant with this poset. The structure of this poset *also* turns out [35,37] to be the same as that for a trigonal prism.

An even simpler posetic reaction network is that involved with the oxygenation of hemoglobin – where O<sub>2</sub> molecules may be adsorbed to any one (or more) of the four component

myoglobin units. For hemoglobin there is [38] a strict 2-fold skeletal point-group symmetry or a more approximate dihedral symmetry which is sufficient for some purposes. Here, the two consequent reaction networks which are essentially posetic Hasse diagrams appear as in Fig. (2), where the great bulk of the structure of the myoglobin units is suppressed to retain iconic skeletal frameworks of the appropriate symmetries. These posets make mostly implicit appearances in the vast body of work on hemoglobin – occasionally with nearly explicit diagrams [39-41] of the reaction network – say with the same spatial arrangement as in Fig. (2), but without the interconnecting lines, even-though there is clear focus on the reactions and associated equilibria. In fact, sometimes even more complicated diagrams are used, in seeking to account for different conformational forms of hemoglobin each of which is favored to different extents by different extents of oxygenation, and possibly also by binding of H<sup>+</sup> or CO<sub>2</sub>. With the summary diagram of Fig. (2), each position displayed there may be interpreted as entailing a mix of conformations.

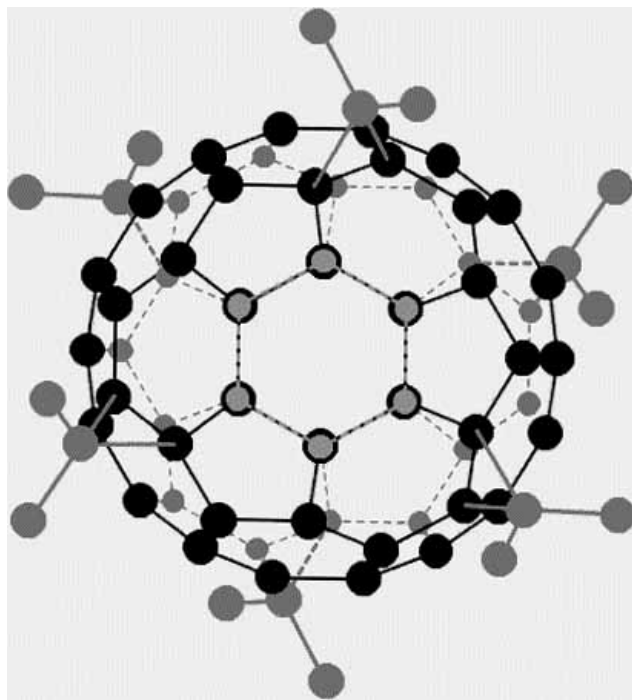


**Fig. (2).** Addition-reaction poset for oxygenation of hemoglobin. The first presumes a skeletal symmetry group of  $C_2$  and the second presumes  $D_2$ .

A 28-member poset with six ranks arises for H-substitution (for H bonded to C) in phenol and is described (and utilized) elsewhere [42]. Another reaction-network poset, now of 34 members and nine ranks, is that for H-substitution in cyclo-butane C<sub>4</sub>H<sub>8</sub>, as described (and extensively utilized) elsewhere [43]. Clearly, there are many more examples – a few further having been elsewhere indicated. And in view of the hemoglobin example, there may be many molecular biological examples involving (allosteric) proteins.

Yet another (larger) progressive reaction poset arises in dealing with a hexamalonate-buckminsterfullerene species C<sub>60</sub>(:C(CO<sub>2</sub>H)<sub>2</sub>)<sub>6</sub>, and various decarboxylation products. Here each malonic acid moiety :C(CO<sub>2</sub>H)<sub>2</sub> is bonded to two adjacent C-atoms of the underlying C<sub>60</sub> fullerene, in such a way that this parent species has a tetrahedral T<sub>h</sub> symmetry as in Fig. (3), where the H and O atoms have been suppressed. Under mildly basic conditions, Cerar *et al.* [44] find that this water soluble species can be decarboxylated to leave different numbers of :CH(CO<sub>2</sub>H) moieties still bonded to the same pairs of C atoms in the underlying C<sub>60</sub> core. Here just as for malonic acid, decarboxylation at one end of a malonic acid moiety makes the second decarboxylation at the other end much more difficult, so that a second decarboxylation does

not occur, under the considered conditions. Thus, the reaction sites are strongly correlated (and not independent), whence only a sub-poset survives from the full poset which would occur with independent substitution. The consequent 40-member poset with seven ranks is represented as in Fig. (4), where we have abbreviated each possible isomeric species by a 6-digit trinary number with the  $i$ th "digit" indicating the state of the  $i$ th malonic acid moiety when numbered as in Fig. (3). There are just three "digits"  $-$ ,  $0$ ,  $+$  with the  $i$ th digit being  $0$  if the  $i$ th malonic acid moiety is undecarboxylated, and  $+$  or  $-$  if (singly) decarboxylated with the  $+$  choice taken if the missing  $\text{CO}_2$  is at the malonic acid end nearer the viewer in the representation of Fig. (3). Notably, this substitution reaction poset of Fig. (4) does not have a unique minimal element due to the correlative exclusionary condition by which no more than one of the two carboxylic acid groups at any malonic acid location can be removed. Also, this poset does not show the self-duality evident in our earlier examples.



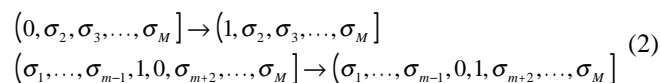
**Fig. (3).** The hexamalonated-buckminsterfullerene species of  $T_h$  symmetry.

Another even larger poset is that for Cl-substitutions in biphenyl, leading to a 210-member poset, which has indeed already been studied [45,46], though the reaction network has not been explicitly displayed. Indeed, it becomes somewhat a challenge to represent this in complete detail, though the beginning and ending portions of this 11-rank poset are displayed in Fig. (5). If one forgoes the display of the molecular structures of the different members of this poset, then the whole poset can be displayed, as in Fig. (6). The numbering of the EPA-set [47] congener numbers for these PCBs proceed in order (left-to-right and then top-to-bottom, with it being understood that the top member of biphenyl is number 0). It is seen from this figure that the overall structure is rather "complicated", evidently embodying much information. Its algebraic representation may be made in terms of a  $210 \times 210$  adjacency matrix for the directed graph of the reaction network (or Hasse diagram) – or alternatively as the set of (840) pairs of structures corresponding to the links in

this network. Moreover, this poset is of great interest as its members, the polychlorinated biphenyls, or PCBs, are carcinogens which seem to occur rather ubiquitously as residuals from various industrial and agricultural activities.

There are numerous posetic reaction networks even more difficult to diagram. For substitution on Paquette's dodecahedrane [48] molecule  $\text{C}_{20}\text{H}_{20}$ , there are [49] 20024 chlorinated isomers  $\text{C}_{20}\text{H}_{20-n}\text{Cl}_n$  (distinguishing enantiomers) which are to be placed into the consequent poset. But if one considers (additive) hydrogenation of buckminsterfullerene ( $\text{C}_{60}$ ), the great majority of such structures are eliminated [50] (as radicaloid), through the restriction that the  $\text{C}_{60}\text{H}_n$  isomers to be those which admit a Kekule structure (whence, e.g.,  $n$  is even). Even then, there remain  $\sim 10^{15}$  members in the poset.

But there are even larger such posets, if one slightly relaxes the fixity of the skeleton or the restriction to just one type of reaction. For instance, consider diffusion down a (polymer-like) chain [11,51] of sites each of which are to be occupied or unoccupied, say by a type of ligand which is being transported from one end of the chain down toward the other end which is a long distance away, so that we denote the states of the length- $M$  such chain by  $(\sigma_1, \sigma_2, \sigma_3, \sigma_4, \dots, \sigma_M)$  with  $\sigma_i = 1$  or  $0$  as the  $i$ th site is occupied or not. (We use a non-symmetric notation (...) for these states, since the absorption is to occur for only one end, the "open" one next to the parenthesis). Then, consider two reactions, first a reaction to adsorb a ligand at the initial site from the bath, and second the reaction to transport a ligand one site along the chain:



Then, e.g., for  $M = 4$ , one obtains the (16-member) reaction network of Fig. (7), where an arrow angled to the left entails the first of the reactions of Eq. (2) and one angled to the right entails the second reaction of Eq. (2). But if  $M$  is of some more reasonable size for a polymer, of say  $M \approx 100$  monomers, then the poset explodes to have millions of moles (i.e.  $>10^{30}$ ) of members. A process somewhat like this might also represent [52] the decalcification of tooth enamel, leading to tooth decay – here the  $\text{Ca}^{2+}$  diffuses out to leave a depleted material.

Yet another example of a progressively ordered reaction poset is that for the methylation of alkanes (replacing a H attached to a C by a  $\text{CH}_3$  group). In this case, the basic skeleton grows without end, so that an infinite poset is obtained – namely that for the subgraph partial ordering of "trees", with degrees limited to  $\leq 4$ .

We might return to the traditional formulation of combinatorial chemistry, in the light of these new concepts. The bare traditional view contemplates a rather simplistic reaction graph with the  $i$ th molecule of  $M = (M_1, M_2, \dots, M_p)$  reacting with the  $j$ th ligand of  $L = (L_1, L_2, \dots, L_Q)$  to give a product  $M_i L_j$ , so that the bare reaction graph is naught but a collection of  $P \times Q$  single-step reactions. But there are different ways in which this can be granted further reaction-network "structure". For instance, the molecules of  $M$  could all be related to one another in some reaction network themselves, and multiple types of reactions might be included with the reaction (or reactions) to generate the molecules  $M_i$  and the reactions for the ligand additions. For instance, the set of  $M_i$

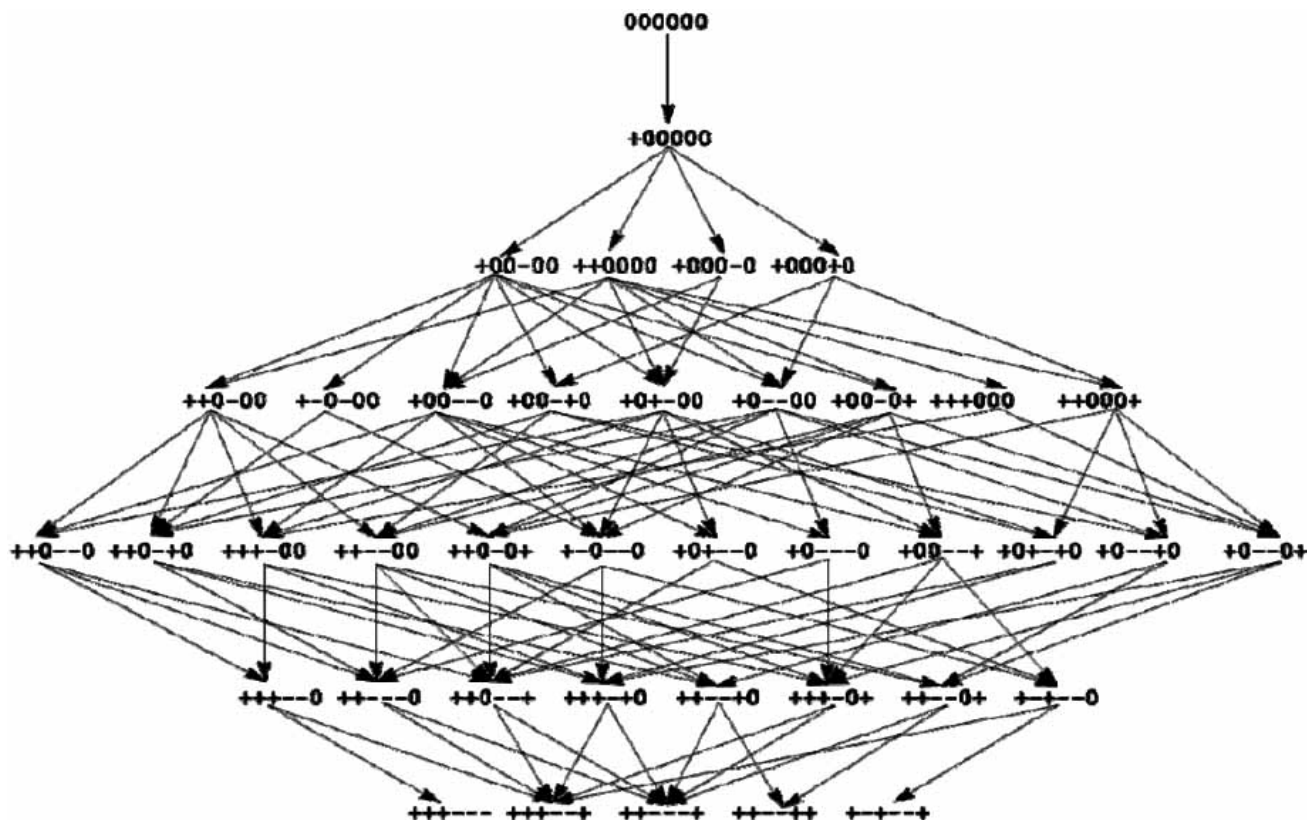


Fig. (4). The decarboxylation poset for the hexamalonate- $C_{60}$  derivative.

could be from one of our "nice" substitution-reaction posets, and the  $L_j$  could be something which further substitutes for the substituents appearing in these  $M_i$  – though in this case

there would be more than one  $M_i, L_j$ -product if there is more than one location at which to substitute.

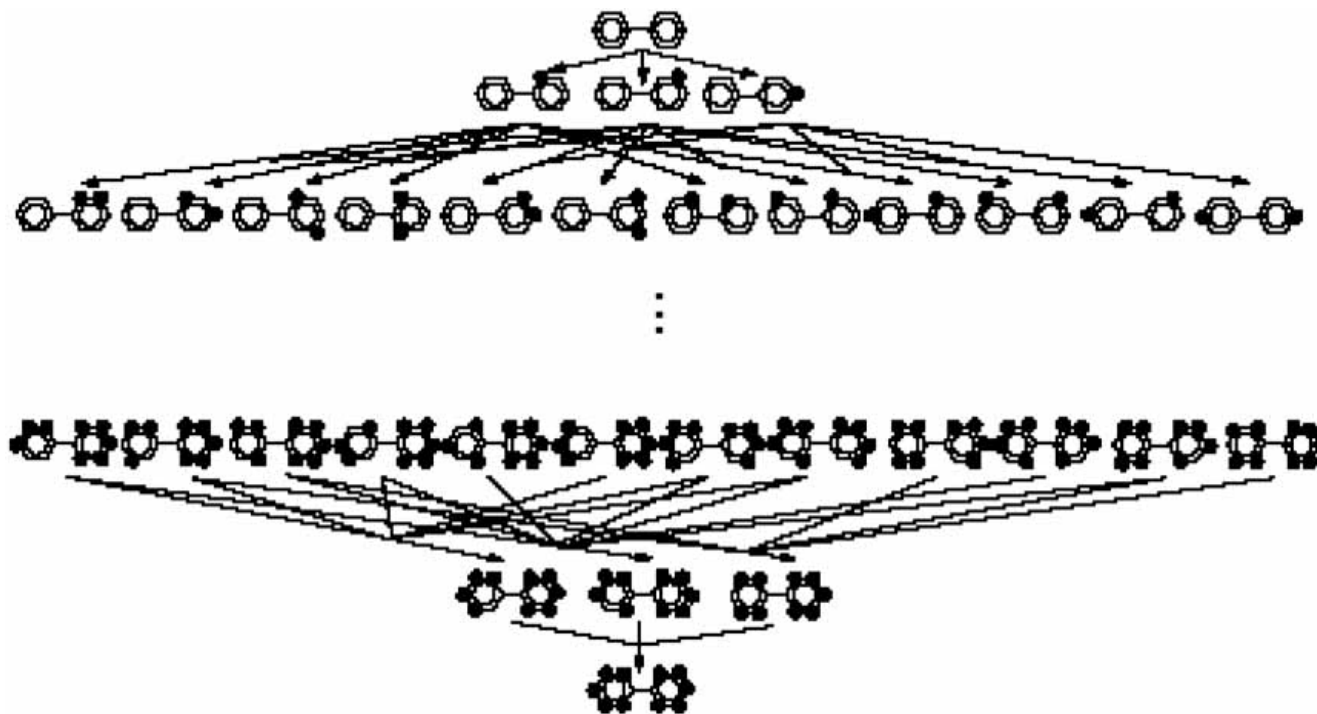
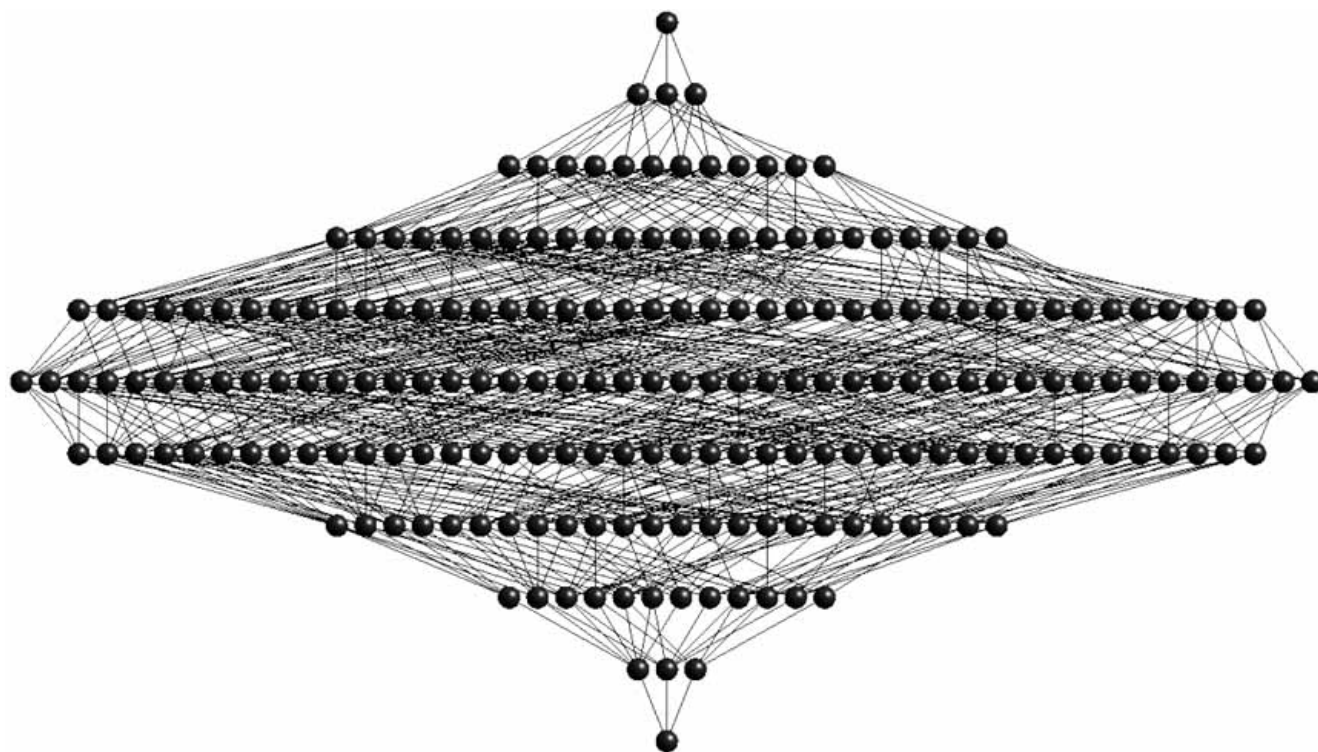
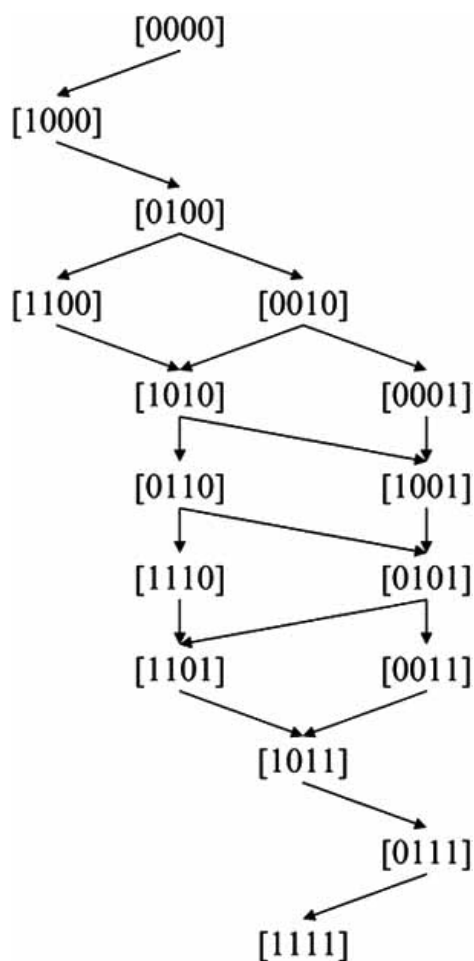


Fig. (5). The initial and latter parts of the biphenyl substitution network.



**Fig. (6).** The full biphenyl substitution network, with explicit structures suppressed, but implicit in using a conventional IUPAC congener ordering.



**Fig. (7).** The progressive reaction poset for diffusion down a length  $M = 4$  chain of sites.

Evidently, different sorts of methods to deal with such posetic reaction networks are of interest to offer novel insights, though there may be substantive mathematical or computational challenges. Mathematical formalization can extend stereochemical theory (where our posetic structures are implicit), and a further conceivable use could be to interpolate/extrapolate properties for the different species arising in the network.

### 3. FORMALIZATION FOR REACTION POSETS

There are a fair number of formal mathematical results which apply for the case of a substitution-reaction network, especially if applied to a fixed finite set  $S$  of substitution sites, say on an underlying molecular *skeleton*. The possibilities of the reaction are characterized in terms of a permutational symmetry group  $G$  for the skeleton. That is, the permutations  $P$  in  $G$  are such that each site  $i$  (in  $S$ ) is symmetry equivalent to site  $Pi$ . For a set  $S$  of potential substitution sites on the molecular skeleton, a *configuration* with  $n$  substituents at positions  $\{i_1, i_2, \dots, i_n\} \equiv C \subseteq S$  is deemed *symmetry equivalent* to other configurations  $\{i_{p_1}, i_{p_2}, \dots, i_{p_n}\} \equiv PC$  for  $P \in G$ . These equivalence classes of configurations correspond to chemical isomers and then comprise the elements of the substitution-reaction poset  $P$  – and for  $\xi, \zeta \in P$  one has  $\xi \succ \zeta$  if there are configurations  $C \in \xi$  and  $C' \in \zeta$  such that  $C \subset C'$ . Thence, the poset depends on the set  $S$  of available skeletal substitution sites and the group  $G$  drawing the equivalences amongst subsets  $C$  of  $S$ .

As an example, consider the hemoglobin molecule of Fig. (2), for which  $S$  is a set of four substitution positions  $\{1, 2, 3, 4\}$ , locating the positions just as the quadrants in a

coordinate system. Then, the group  $C_2$  is  $\{I, (13)(24)\}$ , where  $(ij)$  denotes the permutation interchanging sites  $i$  and  $j$ , while leaving all others fixed. The less discerning group  $D_2 = C_2 \cup \{(13)(24), (14)(23)\}$  gives rise to fewer equivalence classes of arrangements or configurations  $C \subseteq S$ . For instance, the conformation  $\{1,2\}$  (with oxygenation in positions 1 and 2) occurs in the equivalence class

$$D_2\{1,2\} = \{P1, P2\}; P \in D_2 = \{\{1,2\}, \{3,4\}\} \quad (3)$$

and the configuration  $\{1,2,3\}$  occurs in the equivalence class

$$D_2\{1,2,3\} = \{\{1,2,3\}, \{2,3,4\}, \{1,3,4\}, \{1,2,4\}\} \quad (4)$$

The overall reaction network may then be re-expressed in terms of this (more algebraic) labeling of equivalence classes, as in Fig. (8).

Different choices of group  $G$  yield different posets, even when  $S$  is fixed. Indeed, this has already been illustrated in Fig. (2) for two choices of the point group ( $C_2$  and  $D_2$ ) for the case of hemoglobin. For the  $C_{60}$ -based example of Fig. (3) and Fig. (4), we chose the point group to be  $T_h$ , though if we had wished to presume just  $T$ , then a distinction between stereoisomers would have been made (and the poset would appear somewhat more elaborate than as depicted in Fig. (4)). For biphenyl, the group used is denotable as  $(C_{2a} \times C_{2b})C_{2ab}$ , where  $C_{2a}$  and  $C_{2b}$  denote independent 2-fold rotations for the two biphenyls ( $a$  and  $b$ ) while  $C_{2ab}$  interchanges  $a$  and  $b$  ends of the molecule. This group is isomorphic to  $D_{2h}$ , but more rigorously (at least at lower temperatures), the group should be  $D_2$  with the rotations on the two phenyl ends coordinated. Similarly, the quoted count of 20024 for dodecahedrane substitution is with the use of the point group  $I_h$ , whereas a higher count is obtained with the enantiomers-distinguishing proper icosahedral group  $I$ . In general, different choices for the skeletal group correspond to different choices for the type of isomer to be considered. A large group  $G$  and a smaller more rigorously realized one  $G'$  can both be relevant – say using  $G$  and only including the effect of  $G'$  as a perturbation. In fact, often even the minimal group of  $\{I\}$  is relevant in that then it is just primitive conformations which are identified. Presumably, the idea of overall group with a second perturbative group may be relevant in dealing with hemoglobin, and other allometric enzyme systems.

In general, one sees that each member of the poset is represented by a set of configurations obtainable from any member configuration  $C$  comprised from sites  $i_1, i_2, \dots, i_n$  by the action of  $P \in G$  on each of these sites to obtain a configuration  $PC$ . Each of these configurations itself has a symmetry group which is the intersection of  $G$  with the subgroup leaving the configuration invariant – here the group which is a product of the group  $\Gamma_C$  of all permutations on the set  $\{i_1, i_2, \dots, i_n\} = C$  of substituted sites, and the complementary group  $\Gamma_{C^*}$  where  $C^*$  is the *dual* set of skeletal sites without substituents. That is, the symmetry group of the configuration  $C$  is

$$G \cap (\Gamma_C \times \Gamma_{C^*}) \equiv G_C \quad (5)$$

and the symmetry groups for all other configurations equivalent to  $C$  are isomorphic. Each isomer  $\xi \in P$  then is labelled (up to isomorphism) by a generic permutation group  $G_\xi (\approx G_C$  for  $C \in \xi$ ), which corresponds to a molecular point group for  $\xi$  when  $G$  itself does. Examples of these isomer symmetry groups  $G_\xi$  are indicated for  $G = D_2$  hemoglobin in the second part of Fig. (8). And also shown in the third part, different "symmetry numbers" are given: first, at each poset member the numbers of arrangements constituting that isomer; and second, the number of ways a given arrangement of one isomer results in the product isomers with one more substituent site. At each poset location, one could also append the symmetry numbers, involving counts of the "types" of substituted and of unsubstituted sites, where sites  $i$  and  $j$  in a configuration  $C$  are of the same *type* iff there is a  $P \in G_C$  such that  $Pi = j$ . One may say also that such  $i$  and  $j$  are "symmetry equivalent", or that  $i$  and  $j$  are in the same "orbit" of sites with regard to the skeletal symmetry group  $G$ . Notably the number of orbits of substituted sites for  $\xi \in P$  identifies the number  $d_{\xi \rightarrow}$  of successors in the Hasse diagram, and the number of orbits of substituted sites identifies the number  $d_{\rightarrow \xi}$  of predecessors in the Hasse diagram. Further, at a given rank, these numbers evidently decrease as the symmetry group "increases". That is, if  $G_\xi$  is a subgroup of  $G_\zeta$  for  $\xi$  and  $\zeta$  at the same rank, then  $d_{\zeta \rightarrow} \geq d_{\xi \rightarrow}$  and  $d_{\rightarrow \zeta} \geq d_{\rightarrow \xi}$ . Also to deal with the point groups  $G_\xi$  there is the method of marks [53-57], which could perhaps be integrated with the current reaction-poset approach. Amusingly this would couple our

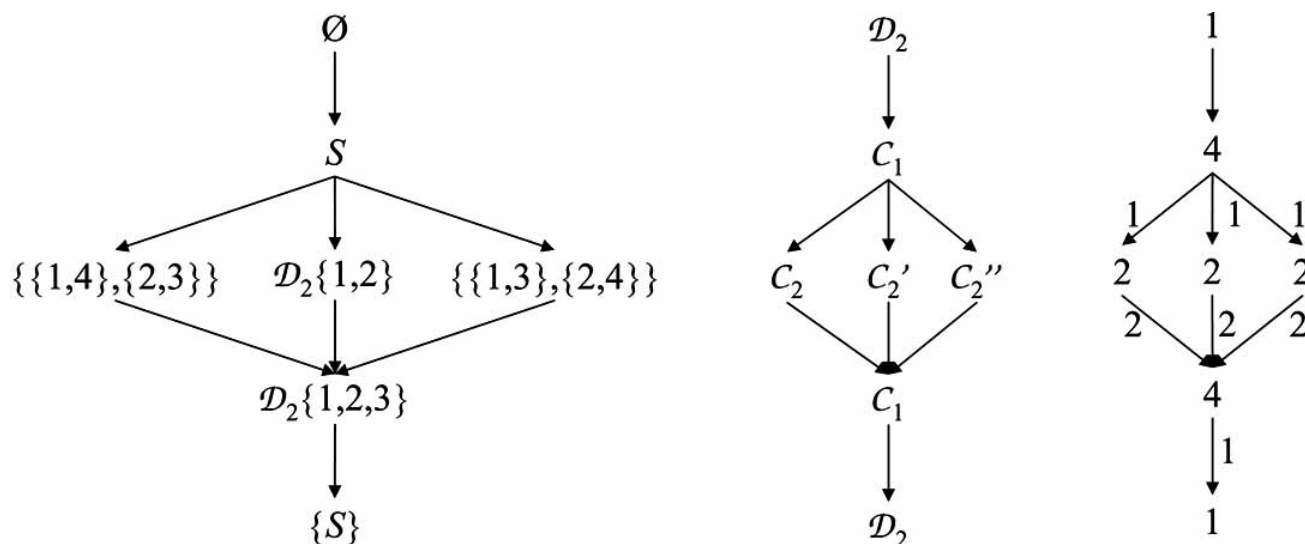


Fig. (8). The  $D_2$  hemoglobin-oxygenation network, with the algebraic representation of the poset members. Then the same  $D_2$  network, first with symmetries of its entries, and associated "symmetry numbers".

reaction poset with the lattice of subgroups of  $\Gamma_S$ , such a lattice being but another special type of poset.

Yet further, the different isomers in the poset are each in correspondence with different formal group-theoretic structures, called  $G, (\Gamma_C \times \Gamma_{C^*})$  double cosets in  $\Gamma_S$ . This is established in the fundamental article of Ruch *et al.* [58] and later reviewed [59] in a wider context, though without explicit reference to our current posetic structure. This posetic dimension (and the relation to "double cosets") deserves further attention, as all this could lead to further fundamental development of stereochemical theory.

Evidently, there is an intimate relation between the permutation-group-theoretic representation of isomers and the current poset. It seems that the poset expressing the isomers in terms of sets of configurations has embedded within it different kinds of information, possibly of some chemical relevance. One is tempted to seek to utilize the so-represented formal poset for further chemical applications.

#### 4. MOLECULAR PROPERTIES AND QSSAR

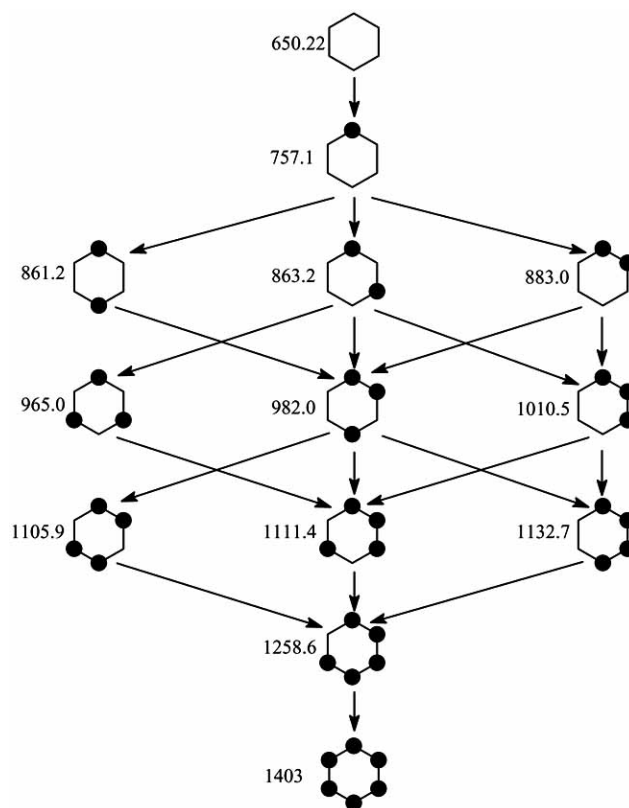
One sees that the structure of a posetic reaction diagram retains some information about molecular structure. Certainly, the investigation of the information inherent in a reaction diagram should be developed in a theoretical format and different possible uses explored. To proceed beyond the local structure to placement in a whole super-structural network, one can speak of a quantitative super-structure/activity relationship: QSSAR. In dealing with more conventional physical or chemical properties, similarly "quantitative super-structure/property relationship" would be QSSPR.

It is rather natural to consider that different chemical properties might be ordered in concert with the partial ordering of a progressive reaction poset. That is, for general structures  $\alpha$  and  $\beta$  in the reaction poset such that  $\beta$  follows  $\alpha$ , then property  $X$  might be ordered such that  $X(\alpha) < X(\beta)$ . If this is true for every pair  $\alpha \succ \beta$ , then the ordering for  $X$  is consonant with the partial ordering of the reaction network. Even if whenever  $\beta$  follows  $\alpha$ , it turns out that  $X(\alpha) > X(\beta)$ , the property ordering is still "consonant" with the partial ordering of the reaction network – the sense of the correspondence between the two orderings is just reversed. Such "consonance" is more technically termed *isotonicity*. This isotonicity does happen, e.g., as illustrated in Fig. (9), for chromatographic retention indices (on squalene at 96°C) of methylbenzenes. Isotonicity seems [60,61] to be a frequent occurrence for chlorobenzenes, and for methylbenzenes (from an examination of over two dozen properties and activities).

But there is also the possibility of "imperfect isotonicity". Such indeed does occur, occasionally to some slight extent [60,61] for chlorobenzenes and methylbenzenes, while for fluorobenzenes, "non-isotonicity" seems to be a rather more frequent circumstance. But even without perfect isotonicity, the poset still may be of use, say if there is a gradual non-monotonic change of a property along a posetically monotonic chain through the network.

In any event, a prototypical problem to be addressed might be that of interpolating or extrapolating properties when only a fraction of the property values for the members of the poset are known. Generally, methods which do not presume isotonicity would be desirable, and also it would be nice if the interpolation or extrapolation could be made with

fewer numbers of known values. Thence indeed, there are three or four interpolation/extrapolation schemes available – which we now briefly describe.



**Fig. (9).** The benzene-substitution network, showing chromatographic Kovats retention indices for methyl benzenes. The result is seen to be isotonic.

A first especially simple sort of a scheme to interpolate property values, at least in suitable cases, is the so-called *average-poset* method. In this scheme a prediction  $\bar{X}(\beta)$  is made for a property  $X$  at a position  $\beta$  in the poset if the property values are available for all other positions immediately adjacent to  $\beta$  in the Hasse diagram. One takes  $\bar{X}(\beta)$  as the average of two values: first the average of all  $X(\alpha)$  with  $\alpha$  an immediate predecessor of  $\beta$ , and second the average of all  $X(\gamma)$  with  $\gamma$  an immediate successor of  $\beta$ . Thence for *ortho*-dichloro-benzene (denoted by  $12\text{Cl}_2\phi$  or  $o\text{Cl}_2\phi$ ), the predicted toxicity from the data for chlorobenzene (denoted  $\text{Cl}\phi$ ) and for 1,2,3- and 1,2,4-trichloro-benzenes (denoted  $123\text{Cl}_3\phi$  and  $124\text{Cl}_3\phi$ ) is obtained *via*

$$\bar{X} = (o\text{Cl}_2\phi) = \frac{1}{2} \left\{ X(\text{Cl}\phi) + \frac{1}{2} [X(123\text{Cl}_3\phi) + X(124\text{Cl}_3\phi)] \right\} \quad (6)$$

somewhat similar ideas appear elsewhere [37,62-64]. But also this scheme is akin to the neighbor-averaging scheme originally used by Mendeleev to make predictions for selected properties of then unknown elements (e.g., *eka-silicon*, later named germanium). Indeed, this scheme in application to Mendeleev's periodic table is often described

in texts for introductory chemistry courses. Notably, Mendeleev's scheme evidently appears to have even more in common with our present work when Mendeleev's periodic table is interpreted as a poset (such as indeed has been suggested [65]). This poset-average method has a disadvantage though in requiring a great deal of data, for all positions adjacent to the interpolated one in the poset – and consequently in this simple form, this scheme is not suitable to predict property values for the maximum or minimum members of a poset. A natural way to extend this idea would be to include further than just neighbors, such as the next scheme does.

A second more general interpolation/extrapolation scheme [66] is a posetic splinoid fitting method. This splinoid procedure imagines a function defined on each of the line segments shown in the Hasse diagram (each such segment corresponding to a single step of the reaction here considered), the function is taken to be a cubic polynomial, such that:

- (1) its values (corresponding to property values) at the end points are the same for each line segment incident at a node;
- (2) it is smooth at each node (with continuous derivatives) in the direction of increase for the poset; and
- (3) subject to the two preceding constraints, the total effective "curvature" of the function (integrated over all the line segments) is minimized.

Such a fitting scheme for the special case of a completely ordered poset yields the standard and highly successful spline fit, as in de Boor [67] or Ruitishauser [68]. For more general posets such as the reaction networks here, the technique is more recent. The consequent linear-algebraic formulas involve standard matrix inversions of matrices with sizes no more than the number of members of the poset. With this approach, even when structures at adjacent positions in a reaction network have property values which are unknown and are wished to be predicted, the spline-fitting method is applicable, giving the vector of unknown values in terms of the vector  $\vec{u}$  of known values in terms of the vector  $\vec{k}$  of unknown values *via*

$$\vec{u} = M \vec{k} \quad (7)$$

with M a matrix computed in terms of intrinsic characteristics of the poset (and the partitioning between known and unknown values). Indeed there is initial evidence that the scheme is robust under the circumstance that even sizable fractions of the property values are missing – perhaps even a majority of the values might be missing. The errors may of course be anticipated generally to increase with increasing deficits of known property values. As an example, this spline-fitting scheme has been used to deal with environmental accumulation of polychlorinated biphenyls (PCBs). Here, there are about 40 toxicities available, though there are 210 members of the substitution-reaction poset. Potentially, this scheme is applicable to very large posets, though perhaps not so large as the poset-average method – except that the poset-average method can be viewed as a special case of the present method, where the fitting for any member is done simply in terms of the sub-poset comprised from this member and its adjacent members. In the advent of very large posets (of more than a few thousand members), though ma-

trix inversions for the full poset might become inapplicable, subposets larger than this minimal nontrivial neighbor choice and smaller than the full poset could plausibly be used.

A third general fitting method uses a so-called *cluster expansion* for the reaction poset. This gives a property  $X$  for species  $\beta$  as a linear combination of a corresponding "constituent" property  $x$  for earlier members of the poset

$$X(\beta) = \sum_{\alpha}^{>\beta} c_{\alpha\beta} x(\alpha) \quad (8)$$

with  $\chi_{\alpha\beta}$  intrinsic to the poset and the  $X$ -dependent  $x(\alpha)$  being determined either by "fitting" or by "inversion". Usually, a simplifying approximation is invoked wherein all but a few of the very earliest  $\alpha$  of the poset have  $x(\alpha) = 0$  – the result being somewhat like a truncated Taylor-series expansion for a continuous variable [43]. This is developed in detail elsewhere, but the method is intimately related to more standard (and widely reviewed) substructural cluster expansions (more often called "group-function" expansions, or "group-contribution" methods, or "linear free energy relationships"), where the partial ordering is that of being a subgraph (or substructure) of the molecular graph (or structure). But as emphasized [69] by Rota, such expansions apply for rather general posets. Typical accounts of the substructural expansion in chemistry most often do not explicitly identify the posetic character, but sometimes [70,71] it is emphasized. Like the splinoid method, the cluster-expansion method is typically applicable to large posets, even with much missing data. The cluster-expansion approach is very closely related to the chemical substructural expansions, the least-squares fitting of the parameters  $x(\alpha)$  bears close resemblance to ordinary QSAR usages, and the parameters are often interpretable as representing contributions from different substructures. Applications of the cluster-expansion and splinoid scheme have typically been done [42,43,45,61] together.

Finally, error estimates for the fitting schemes might be mentioned. To estimate the accuracy of either the average-poset or splinoid schemes, all the presumed known values are fit exactly. Then by leaving out each one of the known values in succession and predicting it along with the other unknowns, a quality-of-fit statistic is obtained as the standard deviation for the left-out knowns. And this same statistic may be used for the cluster expansion.

## 5. DISCUSSION AND PROGNOSIS

It may be worthwhile to emphasize a little further the generality and ubiquity of the general ideas described here. Indeed, the idea of progressive reaction posets may be viewed to underlie the general idea of periodic tables, with Mendeleev's periodic table of the elements providing the most conspicuous example of such a periodic table. That is, relations down columns of the table may be identified as chains in the poset, also including interconnection from an element in an A-column to the element in the corresponding B-column one row down in the table. Here, the reactions from one element to that below it in this posetic periodic table are hypothetical nuclear reactions, but still there are relations between properties of elements consonant with this partial ordering. In addition to Mendeleev's periodic table, there also is Randić and Wilkins [72-74] "periodic table of alkanes", Dias [75,76] "formula periodic table of benzenoids", and a [77,78] "periodic table of all acyclics". All these periodic tables fall into two-dimensional arrangements

with one type of hypothetical reaction down columns and another type along rows. Most of these periodic-table examples end up with more than one chemical species at each position (or node) of what then is recognized as a reaction network. Especially in Randić and Wilkins periodic-table poset for alkanes, the type of property interpolation techniques developed here should be rather directly applicable. Thence, the substitution-reaction posets might be interpreted as primitive "periodic tables".

Also, it has been noted here that there evidently are many molecular biological occurrences of these reaction posets. And it has been noted that some extensions of our ideas may be relevant to conventional combinatorial chemistry. Overall, it seems that the general ideas underlying our approach are widespread, and perhaps many applications of the general theory of posets may be made, maybe with some of the techniques specifically adapted to the evidently ubiquitous case of progressive reaction posets. It has been suggested that they are a mathematically fundamental special kind of poset. The currently noted interpolative fitting techniques (poset-averaging, splinoid fitting, cluster expansion) go beyond simple ordering and ranking. Moreover, the considered poset is determined *a priori*. This may be contrasted with the environmental-science applications [62-64] of Brüggemann, Carlsen and others where the considered poset is determined *a posteriori*, in terms of the considered property values. The property correlations for substitution-reaction posets as reported in the various quoted articles are encouraging. Especially with larger posets (with small fractions of known data points to interpolate from), the robustness of the splinoid and cluster-expansion fitting methods is often very relevant, e.g., with the chloro-substitution reaction for biphenyl, where out of 210 members one finds only about 40 toxicities reported in the literature.

The "poset-average" and splinoid fitting procedures give novel QSSAR fitting procedures, which are in essence parameter-free. The cluster-expansion approach yields fitting parameters in a more conventional manner, and the possibility of the transfer of these parameters between different reaction posets is a problem for future study. In the (few) cases so far studied to make fits and predictions of properties, the various techniques seem to work quite reasonably. The "poset-average" procedure recommends itself with its great simplicity. The splinoid and cluster-expansion approaches however have an advantage of being tolerant to larger deficits of information, while still seemingly making reasonable predictions. The novelty of the approach in comparison to typical QSAR approaches, and in particular the attention beyond structure to placements in reaction networks, suggests the designation of our approach to be that of QSSAR (quantitative super-structure/activity relationships). Further, it is emphasized that such directed reaction graphs are widespread structure in chemistry, and conceivably even in other sciences and mathematics as well. Sometimes, the posets turn out to be quite complex, as with the briefly noted addition-reaction poset for hydrogenation of buckminsterfullerene – leading to a poset of  $\sim 10^{15}$  members. In some cases, the poset may be infinite, as with the noted alkane poset. More broadly, related types of progressive reaction posets may occur in other areas – for example, the posets of ancestors or food webs in biology could be susceptible to some of the same analyses. Thence, there is much promise for our presently indicated ideas and techniques. Notably, the interpolation techniques considered here seem to be applica-

ble to the somewhat more general periodic-table-like posets mentioned near the beginning of this section.

Overall, notable success is evidenced to date in the few investigations on simple substitution-reaction networks. As such, this indicates a potentially wide-range of chemical applications, and possibly to other fields as well. Further posetic techniques should be developed and explored.

## ACKNOWLEDGEMENTS

The authors acknowledge the support (*via* grant BD-0894) from the Welch Foundation of Houston, Texas.

## REFERENCES

- [1] Ugi, I.; Goebel, M.; Gruber, B.; Heilingbrunner, M.; Heiss, C.; Horl, W.; Kern, O.; Starnecker, M.; Domling, M. *Res. Chem. Intermed.*, **1996**, *22*, 625.
- [2] Wieland, T. *J. Math. Chem.*, **1997**, *21*, 141.
- [3] Stadler, B.M.R.; Stadler, P.F. *J. Chem. Inf. Comp. Sci.*, **2002**, *42*, 577.
- [4] Haggarty, S.J.; Clemons, P.A.; Schreiber, S.L. *J. Am. Chem. Soc.*, **2003**, *125*, 1.
- [5] Ivanciuc, O.; Klein, D.J. *Croat. Chem. Acta.*, **2002**, *75*, 577.
- [6] Corey, E.J.; Choung, X.-M. *The logic of chemical synthesis*, Wiley: New York, **1989**.
- [7] Hendrickson, J.B. *J. Am. Chem. Soc.*, **1977**, *99*, 5439.
- [8] Bertz, S.H. *J. Chem. Soc. Chem. Commun.*, **1984**, *4*, 218.
- [9] Bertz, S.H. *J. Chem. Soc. Chem. Commun.*, **1986**, *22*, 1627.
- [10] Eigen, M. *Naturwissenschaften*, **1971**, *58*, 465.
- [11] Hill, T.L. *Free energy transduction in biology*, Academic Press: New York, **1977**.
- [12] Balaban, A.T. In *Graph theoretical approaches to chemical reactivity*; Bonchev, D.; Mekenyan, O., Eds.; Kluwer Academic Publishers: Dordrecht, **1994**; pp. 137-180.
- [13] Temkin, O.N.; Zeigarnik, A.V.; Bonchev, D. *Chemical reaction networks: a graph-theoretical approach*, CRC Press: Boca Raton, **1996**.
- [14] Ross, J.; Schreiber, I.; Vlad, M.O. *Determination of complex reaction mechanisms*, Oxford University Press: Oxford, **2006**.
- [15] Birkhoff, G. *Lattice theory*, American Mathematical Society: Providence, **1948**.
- [16] Alberti, P.M.; Uhlmann, A. *Stochasticity and partial order*, VEB Deutscher Verlag der Wissenschaften: Berlin, **1981**.
- [17] Trotter, W.T. *Combinatorics and partially ordered sets - dimension theory*, The Johns Hopkins University Press: Baltimore, **1992**.
- [18] Pecaric, J.E.; Proschan, F.; Tong, Y.L. *Convex functions, partial orderings and statistical applications*, Academic Press: Boston, **1992**.
- [19] Neggers, J.; Kim, H.S. *Basic posets*, World Scientific: Singapore, **1998**.
- [20] Rival, I.; Reidel, D. *Ordered sets*, Springer: Dordrecht, **1982**.
- [21] *MATCH Commun. Math. Comput. Chem.*, **2000**, *41*, Spec. Ed. Klein, D.J.; Brickmann, J.
- [22] Brüggemann, R.; Carlson, L. *Partial order in environmental sciences and chemistry*, Springer Verlag: Berlin, **2006**.
- [23] Klein, D.J. "Substitution reaction posets and neatly graded posets", **2007**, in preparation.
- [24] Pólya, G. *Acta Math.*, **1937**, *68*, 145; translation of this classical paper appears in [25].
- [25] Pólya, G.; Read, R.C. *Combinatorial enumeration of groups, graphs and chemical compounds*, Springer-Verlag: New York, **1987**.
- [26] Kerber, A. *Applied finite group actions*, Springer-Verlag: Berlin, **1999**.
- [27] Klein, F. "Vergleichende Betrachtungen über neue geometrische Forschungen" (Erlangen, **1972**); reprinted in *Gesammelte math. Abhand.* I Berlin, **1921**, p. 460.
- [28] Stillwell, J. *The four pillars of geometry*, Springer-Verlag: Berlin, **2000**, chapters 7 and 8.
- [29] Körner, W. *Gazz. Chim. Ital.*, **1874**, *4*, 305.
- [30] Körner, W. *Palermo giornale di scienze naturali ed economiche*, **1869**, *5*, 208, as summarized by Armstrong, H.E. *J. Chem. Soc.*, **1876**, *29*, 204.
- [31] Cohen, J.B. *Theoretical organic chemistry*, Macmillan & Co.: London, **1923**, pp. 395-396.

- [32] Heilbronner, E.; Dunitz, J.D. *Reflections on symmetry*, VCH: New York, **1993**, pp. 46-51.
- [33] Merrifield, R.E.; Simmons, H.E. *Topological methods in chemistry*, John Wiley: New York, **1983**, Chap. 12.
- [34] Dolfing, J.; Harrison, B.K. *REMS microbiology ecology*, **1993**, *13*, 23.
- [35] Shtarev, A.B.; Pinkhassik, E.; Levin, M.D.; Stibor, I.; Michl, J. *J. Am. Chem. Soc.*, **2001**, *123*, 3484.
- [36] Zivkovic, T.P. *Croat. Chem. Acta*, **2007**, *80*, 547.
- [37] Klein, D.J.; Bytautas, L. *MATCH Commun. Math. Comput. Chem.*, **2000**, *42*, 261.
- [38] Dickerson, R.E.; Geis, I. *The structure and action of proteins*, Harper & Row: New York, **1969**.
- [39] Ackers, G.K.; Smith, F.R. *Ann. Rev. Biophys. Biophys. Chem.*, **1987**, *16*, 583.
- [40] Perrella, M.; Colosimo, A.; Benazi, L.; Ripamonti, M.; Rossi-Bernardi, L. *Biophys. Chem.*, **1990**, *37*, 211.
- [41] Ackers, G.K.; Doyle, M.L.; Myers, D.; Daugherty, M.A. *Science*, **1992**, *255*, 54.
- [42] Ivanciuc, T.; Ivanciuc, O.; Klein, D.J. *Intl. J. Molec. Sci.*, **2006**, *7*, 358.
- [43] Ivanciuc, T.; Klein, D.J.; Ivanciuc, O. *J. Math. Chem.*, **2007**, *41*, 355.
- [44] Cerar, J.; Pompe, M.; Gucek, M.; Cerkovnik, J.; Skerjanc, J. *J. Chromatography A*, **2007**, *1169*, 86.
- [45] Ivanciuc, T.; Ivanciuc, O.; Klein, D.J. *Mol. Diversity*, **2006**, *10*, 133.
- [46] Ivanciuc, T.; Ivanciuc, O.; Klein, D.J. "Posetic quantitative superstructure activity relationships (QSSAR) for chromatographic retention indices of polychlorinated biphenyl (PCBs)", **2007** (in preparation).
- [47] This numbering is explained and detailed at: <http://www.epa.gov/toxteam/pcb/pcbtable.htm>
- [48] Paquette, L.A. *Chem. Rev.*, **1989**, *89*, 1051.
- [49] Hosoya, H. *Gendai Kagaku* (in Japanese), **1987**, *201*, 38.
- [50] Babić, D.; Došlić, T.; Klein, D.J.; Misra, A. *Bull. Chem. Soc. Japan*, **2004**, *77*, 2003.
- [51] Shapiro, L.W.; Zeilberger, D. *J. Math. Biol.*, **1982**, *15*, 351.
- [52] Anderson, P.; Elliot, J.C. *Caries Res.*, **2000**, *34*, 33.
- [53] Hasselbarth, W. *MATCH Commun. Math. Comput. Chem.*, **1986**, *20*, 241.
- [54] Brocas, J. *J. Am. Chem. Soc.*, **1986**, *108*, 1135.
- [55] Mead, C.A. *J. Am. Chem. Soc.*, **1987**, *109*, 2130.
- [56] Fujita, S. *Symmetry and combinatorial enumeration in chemistry*, Springer-Verlag: Berlin, **1991**.
- [57] Fujita, S. *Theor. Chim. Acta.*, **1995**, *91*, 291.
- [58] Ruch, E.; Hasselbarth, W.; Richeter, B. *Theor. Chem. Acta.*, **1970**, *19*, 259.
- [59] Ruch, E.; Klein, D.J. *Theor. Chem. Acta.*, **1983**, *63*, 447.
- [60] Ivanciuc, T.; Klein, D.J. *J. Chem. Inf. Comput. Sci.*, **2004**, *44*, 610.
- [61] Ivanciuc, T.; Ivanciuc, O.; Klein, D.J. *J. Chem. Inf. Model.*, **2005**, *45*, 870.
- [62] Brüggemann, R.; Pudenz, S.; Carlsen, L.; Sørensen, P.B.; Thomsen, M.; Mishra, R.K. *SAR QSAR Environ.*, **2001**, *11*, 473.
- [63] Carlsen, L.; Sørensen, P.B.; Thomsen, M.; Brüggemann, R. *SAR QSAR Environ. Res.*, **2002**, *13*, 153.
- [64] Carlsen, L. *Molecules*, **2004**, *9*, 1010.
- [65] Klein, D.J. *J. Math. Chem.*, **1995**, *18*, 321.
- [66] Došlić, T.; Klein, D.J. *J. Comput. Appl. Math.*, **2005**, *177*, 175.
- [67] De Boor, C. *A practical guide to splines*, Springer: New York, **1978**.
- [68] Ruitishauser, H. *Lectures on numerical mathematics*, Birkhauser: Boston, **1990**.
- [69] Rota, G.C. *Zeit. Wahrscheinlichkeitstheorie and Verw. Gebiete*, **1964**, *2*, 340.
- [70] Klein, D.J. *Intl. J. Quantum Chem.*, **1986**, *20*, 153.
- [71] Klein, D.J.; Schmalz, T.G.; Bytautas, L. *SAR QSAR Environ. Res.*, **1999**, *10*, 131.
- [72] Randić, M.; Wilkins, C.L. *J. Phys. Chem.*, **1979**, *83*, 1525.
- [73] Randić, M.; Wilkins, C.L. *Chem. Phys. Lett.*, **1979**, *63*, 332.
- [74] Randić, M. *J. Chem. Ed.*, **1992**, *69*, 713.
- [75] Dias, J.R. *Acc. Chem. Res.*, **1985**, *18*, 241.
- [76] Dias, J.R. *J. Math. Chem.*, **1990**, *4*, 17.
- [77] Bytautas, L.; Klein, D.J. *Croat. Chem. Acta.*, **2000**, *73*, 331.
- [78] Bytautas, L.; Klein, D.J.; Schmalz, T.G. *New J. Chem.*, **2000**, *24*, 329.