

The Systematic Definition of Organic Reactions

James B. Hendrickson* and Thomas L. Sander

Abstract: A logical system is presented to describe organic reactions. It is based on the *net structural change* in a reaction, rather than on substructure searching. This system rigorously provides a place for any reaction. It is very simple to use by hand, but, since it is also numerical, it provides a basis for organizing and retrieving reactions in a database by computer. The effectiveness of a program, COGNOS, written to implement and test the system on a large database, lends confidence in this logic for organizing reactions.

Keywords

computer chemistry · databases · indexing · reaction organization

Introduction

Organic chemistry has created a systematic nomenclature (IUPAC) and hierarchic indexing (Beilstein) for organic *structures*, but their parallel for organic *reactions* (i.e., a comprehensive, rigorous system for classification and annotation) has not yet been developed. Now that reaction databases have appeared, requiring computerized retrieval, the need for such a system of organization has become acute.

The organization of reactions should mirror the Beilstein system for organizing compounds in having a limited number of main general categories and a defined hierarchic, or taxonomic, nesting of subfamilies within them to further refine lesser distinctions. The main criterion is that the system must be capable of describing all possible reactions, whether currently known or unknown, so that any reaction has a clear place in the organizational scheme. Building on an older abstraction for generalizing organic structure,^[1] we develop here a rigorous description of organic reactions that meets these criteria and affords, as well, unique numerical identifications of reactions for indexing purposes and retrieval from reaction databases.

A reaction will be described as the *net structural change* from the substrate to the product: what bonds are made and what bonds are broken. Such a procedure allows overall generalization of these changing bonds first and then successive refinement in detail. This is not a mechanism-based approach since not all mechanisms are known, nor is it based on comparing substructures, but rather it describes just the bond changes at those atoms that change.

Discussion

Generalization of Structure: Generalization may begin with a focus on carbon as the keystone atom in organic structures, by defining the major kinds of bonds that carbon atoms form and the changes these undergo in reactions. In any structure it is important to distinguish clearly between the *skeleton* and the *functional groups*: a structure is composed of a skeleton of linked carbon atoms, with functional groups located at specific sites on that skeleton. The σ bonds between carbons constitute the skeleton, but the carbon-carbon π bonds are functional groups since they may be broken without altering the skeleton. The rest of the bonds to the carbons may simply be divided into bonds to more electronegative or to more electropositive atoms, since their balance determines the oxidation state and hence its change in a reaction.

The four kinds of bonds on any carbon may be labeled as R, Π , Z, and H, respectively, as summarized in Figure 1, and the number of each bond type at any carbon atom (σ , π , z , and h , respectively) will add up to four. The values of σ are the familiar

	No.
R = σ bond to another carbon (skeletal)	σ
Π = π bond to another carbon (functional)	π
Z = bond to electronegative atom (N, O, S, halogen, etc.)	z
H = bond to electropositive atom (H, B, Al, Si, Sn, metals, etc.)	h
	$\Sigma = 4$

Oxidation state: $x = z - h$

Fig. 1. Four generalized kinds of bonds to any carbon atom.

descriptors of skeletal carbons: primary ($\sigma = 1$), secondary ($\sigma = 2$), tertiary ($\sigma = 3$), and quaternary ($\sigma = 4$). Only three values of π are possible for carbon atoms, namely, $\pi = 0, 1$, and 2 , distinguishing their attached single, double, and triple bonds to other carbons, respectively.

Values of z (0–4) generalize the common electronegative heteroatom functional groups as families of groups interchange-

[*] J. B. Hendrickson, T. L. Sander⁽⁺⁾
 Department of Chemistry, Brandeis University
 Waltham, MA 02254 (USA)
 Telefax: Int. code + (617) 736-2516
 e-mail: bitnet:hendrickson@binah.cc.brandeis.edu

[*] Present address: Hoffmann-La Roche, Basel (Switzerland)

able among themselves by substitutions of one kind of heteroatom for another. Thus $z = 3$ unites all derivatives of the carboxylic acid/nitrile family, and $z = 2$ all derivatives of aldehydes and ketones (themselves distinguished by $\sigma = 1$ or 2), such as acetals, imines, oximes, etc. The z values do not distinguish σ bonds and π bonds to the heteroatoms. Values of h (0–4) refer usually to hydrogens, but also include the other electropositive atoms, or indeed simple unshared electron pairs on carbon, seen as the conjugate base of a bond to hydrogen.

In this system any carbon can now be described with three variables: σ for its skeletal level, and $z\pi$ for its functional nature, with h as the difference of $\sigma + z + \pi$ from 4. The three digits σ, z, π may be defined as the *character* of any carbon. In this generalization only 24 characters are possible, a number small enough to allow facile coverage of the great variety of organic structures, while detailed enough to retain chemical significance. An important product of this abstraction is that the oxidation state (x) of any carbon in a structure is given by $x = z - h$, with $-4 < x < +4$. Therefore, the change in oxidation state in a reaction will be $\sum \Delta x_i$ over those carbons that change. The three variables describing any carbon may also be σ, π, x , which are equivalent to σ, z, π above, and these are more convenient to describe reaction changes in the following discussion, as $\Delta\sigma$, $\Delta\pi$, and Δx .

If structures can be described with simple numbers for each carbon, then reactions can be described as the change in those numbers on passing from substrate to product. This constitutes the net structural change that characterizes any reaction family. The units that make up these changes may be logically derived.

Definition of Reactions: The essence of an organic reaction is the exchange of bond types on each changing carbon of the skeleton. The simplest reaction change is a single exchange of one bond for another on one carbon, that is, one bond made and one bond broken. On any one carbon there are 16 possible such single exchanges, which derive from these four generalized bond types. We may label these changes at any carbon with a simple notation of two letters, the first for the bond made, the second for the bond broken, as shown in Table 1.

Table 1. The 16 possible unit exchanges at any skeletal carbon.

		$\Delta\sigma$	$\Delta\pi$	Δx
Substitution	HH, ZZ, RR, $\Pi\Pi$	0	0	0
	Oxidation	ZH	0	+2
	Reduction	HZ	0	-2
Elimination	ΠH	0	+1	+1
	ΠZ	0	+1	-1
Addition	$H\Pi$	0	-1	-1
	$Z\Pi$	0	-1	+1
Construction	RH	+1	0	+1
	RZ	+1	0	-1
	$R\Pi$	+1	-1	0
Fragmentation	HR	-1	0	-1
	ZR	-1	0	+1
	ΠR	-1	+1	0

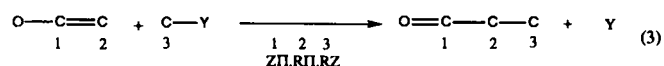
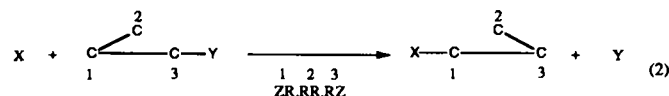
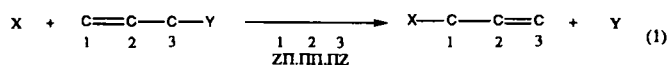
These single exchanges at one carbon are organized in Table 1 into familiar reaction types, namely, substitution, elimination/addition, and the skeletal alterations construction/fragmentation. Each has an oxidative and reductive variant: any carbon changed by +H or -Z is reduced, and any by -H or +Z is oxidized, each by one level of oxidation state, as shown in the Δx column. The simple redox substitutions themselves, HZ and ZH (i.e., +H -Z and +Z -H, respectively), change the oxidation

state by $\Delta x = \pm 2$. The former, HZ, is a reductive substitution of a heteroatom bond by H, as in the reduction of ketones or halides by hydride; the latter, ZH, is the reverse, an oxidative substitution of H by heteroatom Z, as in alcohol oxidation or aromatic nitration. The other substitutions are those of electrophiles (proton or metal exchange, HH) and of nucleophiles, as in S_N2 displacements (ZZ), both with $\Delta x = 0$.

In Table 1 the four single exchanges involving just H and Z take place on only one carbon, but those with changes in Π and R must invoke the same change on an adjacent carbon. In Table 1 eliminations will have + Π at each carbon and be either oxidative (ΠH , with $\Delta x = +1$) or reductive (ΠZ ; $\Delta x = -1$), while the additions are the reverse. A complete reaction requires two adjacent carbons, so that a reductive addition is labeled as $H\Pi \cdot H\Pi$ for the two carbons, with $\sum \Delta x = -2$, as in the hydrogenation of a double bond; bromination is an oxidative addition, labeled as $Z\Pi \cdot Z\Pi$, with $\sum \Delta x = +2$ (see Table 1). Similarly, a simple dehydrohalogenation is an *isohypsic*^[2] elimination, $\Pi H \cdot \Pi Z$ ($\sum \Delta x = 0$), forming a π bond between two carbons at the expense of H on one and Z on the other. Since inversion of each two-letter pair in the string yields the reverse reaction, the notation $H\Pi \cdot Z\Pi$ refers to the addition of HX or H_2O to a C-C double bond.

Reactions that alter the skeleton must also involve two carbons, linked in the product (constructions, +R) or in the substrate (fragmentations, -R). A simple construction, such as reaction of an organolithium reagent with a ketone, is $RH \cdot RZ$, linking two adjacent carbons with +R change on each. This is read as a construction reaction, which forms a C-C σ bond (R), with loss of H (i.e., Li) at one carbon and loss of Z (i.e., bond to oxygen) at the other carbon.

The single exchanges $\Pi\Pi$ and RR at one carbon in Table 1 must invoke the same changes in the adjacent carbons *on each side*. The former is characteristic of the double-bond shift at the central carbon of an allylic, or vinylogous, group, like that observed in the allylic substitution in Equation (1), and so must involve three adjacent carbons. Similarly, the exchange RR at one carbon must also involve three adjacent carbons in a parallel way. This is characteristic of the migrating carbon in a simple 1,2-rearrangement, in which one C-C σ bond is formed and another broken (Eq. (2)); the RR exchange is not restricted to rearrangements, but other examples (i.e., without the bond between carbons 1 and 3) are rare.



Unit Reactions: It is not apparent from common usage just what constitutes a "single reaction", for a single reaction in the synthetic laboratory may often consist of several reaction steps in the mechanism. In order to provide rigorous and universal reaction nomenclature, we need a clear definition of a single reaction.

We define a *unit reaction* as just a single exchange of bonds at each changing carbon. Therefore, all carbons in a unit reaction are adjacent, linked in a reacting strand, and bonded either in

substrate or product. The notation for a reacting strand is the string of two-letter exchanges, one for each changing carbon, from Table 1. A strand of only one carbon will have only $\pm H, Z$ exchanges, but in all longer strands the linked carbons of a unit reaction are all coupled by sharing either $\pm R$ or $\pm \Pi$. Linear strands will exhibit $\pm H, Z$ only at the end carbons. Cyclic unit reactions can also exist, in which all carbons are linked only by $\pm R$ or $\pm \Pi$.

These notation strings for the reacting strands constitute a description of the *net structural change* in any unit reaction. Strands of one and two carbons were illustrated above, as simple substitutions on only one carbon, while eliminations or additions (with $\pm \Pi$) and constructions or fragmentations ($\pm R$) require at least two carbons. Three-carbon strands are characteristic of vinylogous substitutions (with the middle $\Pi\Pi$ carbon), as in Equation (1), or rearrangements (with the middle RR carbon), as in Equation (2). The enol alkylation in Equation (3) is an example of the third reacting strand of three carbons, with the middle one now as $R\Pi$. Four linked carbons constitute the strand of a vinylogous addition/elimination, as in 1,4-addition of HBr to a diene, represented as $H\Pi \cdot \Pi\Pi \cdot \Pi\Pi \cdot Z\Pi$, and of course longer vinylogous strands are possible but uncommon.

Cyclic strands may be distinguished by placing the notation string in parentheses. These are exemplified by photolytic cycloaddition of two double bonds as $(R\Pi \cdot R\Pi \cdot R\Pi \cdot R\Pi)$ on the four carbons in a 4-ring cycle, and the Diels–Alder cycloaddition as $(R\Pi \cdot R\Pi \cdot \Pi\Pi \cdot \Pi\Pi \cdot R\Pi \cdot R\Pi)$ in a six-membered ring cycle. (The general representation of all reactions as cycles is discussed in ref. [3].)

Actual reactions are mostly just unit reactions. An analysis of four major reaction databases (155000 reactions) with this system showed that about 80% are in fact simple unit reactions,^[4] and about two-thirds of the rest are just two successive unit reactions. These transformations incorporating two successive unit reactions are referred to here as *composite* reactions. In composite reactions some or all of the changing carbons exchange more than one bond. As examples, the Wolff–Kishner reduction may be described as $HZ + HZ$ on one carbon, or the Wittig reaction as $RH \cdot RZ + \Pi Z \cdot \Pi Z$ on two carbons. It was also clear from the survey^[4] that the number of linked changing carbons is rarely more than four and almost never more than six, as in the Diels–Alder reaction.

Generation of Reaction Families: We can now logically generate all the possible unit reactions taking place over a strand of adjacent changing carbons. This may be done by stringing together all valid combinations of single carbon exchanges from Table 1: any carbon with $+\Pi$ or $-\Pi$ must have an adjacent carbon with the same change, as must any with $+R$ or $-R$.

The central dichotomy of skeleton and functionality, defined above for structures, is mirrored in their reactions by skeletal alterations with $\pm R$, and refunctionalizations, reactions with no $\pm R$, which do not change the carbon skeleton. Each of these two central classes of reactions will now contain a set of reaction families, which define the possible functional group changes in that class in terms of the changes per carbon (Table 1). We see at work in these combinations of changes the three familiar variables $\Delta\sigma$, $\Delta\pi$, and Δx , that is, construction/fragmentation, addition/elimination, and oxidation/reduction, respectively. These reflect the three variables above for the description of any carbon in a structure.

Refunctionalization Families: The refunctionalization class, with $\Sigma\Delta\sigma = 0$, involves only two of these variables and so may be

organized by $\Sigma\Delta\pi$ and $\Sigma\Delta x$, as in Table 2, which shows the 10 basic families of refunctionalizations. These refunctionalization families are divided into substitutions, in which only one carbon changes, and eliminations and additions, in which two carbons change. The families are simply characterized with familiar labels, shown in brackets for each family in Table 2. These are $[X]$ for oxidative substitution, $[R]$ for reductive substitution, and $[S]$ for substitution of bonds Z ; H substitution is separated as $[H]$. The three eliminations are reductive $[RE]$, isohypsic $[E]$, and oxidative $[XE]$, and the additions are similarly $[RA]$, $[A]$, and $[XA]$. Thus these appended labels for each family are readily recognized as familiar chemical terms.

Table 2. Unit refunctionalization reactions.

$\Sigma\Delta x =$	-2	0	+2	
$\Sigma\Delta\pi =$	$\Pi Z, \Pi Z$	$\Pi H, \Pi Z$	$\Pi H, \Pi H$	Elimination
+2	$[RE]$	$[E]$	$[XE]$	
0	HZ	HH, ZZ	ZH	
0	$[R]$	$[H], [S]$	$[X]$	
-2	$H\Pi, H\Pi$	$H\Pi, Z\Pi$	$Z\Pi, Z\Pi$	Addition
	$[RA]$	$[A]$	$[XA]$	
	Reductive	Isohypsic	Oxidative	

Another 10 families may be created in parallel by using their vinylogues, that is, by adding two more carbons into each strand as a π bond. Thus the simple substitution (ZZ on one carbon) becomes allylic substitution ($Z\Pi \cdot \Pi\Pi \cdot \Pi Z$) on three carbons, as in Equation (1); and the four-carbon vinylogue of reductive addition will be $H\Pi \cdot \Pi\Pi \cdot \Pi\Pi \cdot H\Pi$, a 1,4-addition of hydrogen. These 10 vinylogues are now designated with primes, as $[S']$ for the allylic substitution above ($Z\Pi \cdot \Pi\Pi \cdot \Pi Z$; Eq. (1)). The four-carbon vinylogues of the eliminations/additions are labeled as $[RE']$ and $[XA']$, etc., as with $[RA']$ for the 1,4-hydrogenation ($H\Pi \cdot \Pi\Pi \cdot \Pi\Pi \cdot H\Pi$). Doubly allylic reactions exist also and are all treated the same way, that is, five-carbon doubly allylic substitutions as $[S'']$ and six-carbon doubly allylic eliminations/additions as $[E'']/[A'']$, etc. Therefore, there are overall 30 families of unit refunctionalization reactions including vinylogues and double vinylogues, with reacting strands of up to six carbons. Higher vinylogues can be annotated similarly, but are rare.

Skeletal Alteration Families: The skeletal alteration class involves the third variable, $\Sigma\Delta\sigma$. The simple unit reactions of construction/fragmentation are each further divided into two *half-reactions*,^[11] showing the reaction changes on each of the two strands of carbons starting from the two carbons involved in the formation (or cleavage) of the $C-C$ σ bond. In these half-reactions the first carbon in each half-reaction strand will be $\pm R$. The rest of the changes can now be presented as in the refunctionalization table, similarly organized by $\Sigma\Delta\pi$ and $\Sigma\Delta x$ in Table 3.

Table 3. Single construction/fragmentation half-reactions.

$\Sigma\Delta x =$	Reductive		Oxidative		
	-1	-1	+1	+1	
$\Sigma\Delta\pi =$	$\Pi R, \Pi Z$		$\Pi R, \Pi H$		Elimination
+2	$[REF]$	–	$[XEF]$	–	
0	HR	RZ	ZR	RH	Substitution
0	$[RF]$	$[RC]$	$[XF]$	$[XC]$	
-2	–	$R\Pi, H\Pi$	–	$R\Pi, Z\Pi$	Addition
		$[RAC]$		$[XAC]$	
	Fragmentation	Construction	Fragmentation	Construction	

Table 3 now includes all possible unit half-reactions on one or two carbons. Each fragmentation is the reverse of a construction and has an inverted oxidation state change. The eight construction/fragmentation half-reactions are similarly labeled (in brackets) by adding the letter C for construction or F for fragmentation to the reaction label. Comparison with Table 2 shows the parallels, in terms of substitution and elimination/addition, between the skeletal alteration half-reactions and the refunctionalization reactions. As an example, the reductive addition of an alkyl group to an electrophilic double bond (i.e., $R + C=C \rightarrow R-C-CH$) would be labeled an [RAC] half-reaction, annotated as $R\Pi \cdot H\Pi$ with $\sum\Delta x = -1$. If its nucleophilic other half were an RH change (labeled as an [XC] half-reaction), with $\sum\Delta x = +1$, this would be the net change of the Michael addition as a full construction, that is, $RH \cdot R\Pi \cdot H\Pi$, and isohypsic overall ($\sum\Delta x = 0$).

The vinylogues of these eight skeletal half-reactions are similarly labeled with primes and doubly vinylogous ones with double primes, as with the refunctionalizations above, making 24 families of half-reactions overall on half-reaction strands of 1–6 carbons. The full constructions or fragmentations, made by combining two half-reactions, can therefore have strands of 2–12 carbons, but those above five or six carbons are very rare.

Pairing these simple half-reactions allows full unit reactions that only make (or break) one C–C σ bond. There is also a small logical set that makes/breaks two σ bonds in one unit reaction. On two carbons the double construction half-reaction is $R\Pi \cdot R\Pi$, with the vinylogous $R\Pi \cdot \Pi\Pi \cdot \Pi\Pi \cdot R\Pi$ on four carbons. These two can be paired for a six-carbon cyclic strand, the Diels–Alder reaction above; the reverse fragmentations are paired for the retroreaction. For a linear strand a full unit reaction will have three half-reactions: two single construction or fragmentation halves at each side of a double one. The half-reaction will be RR on only one carbon, characteristic of rearrangements, or its vinylogue $R\Pi \cdot \Pi\Pi \cdot \Pi\Pi \cdot R\Pi$ on three carbons. For the full unit reaction these must be flanked by a single fragmentation half-reaction at one side and a single construction at the other as in the notation string $ZR \cdot RR \cdot RZ$ of the unit reaction in Equation (2).

The 54 families of Tables 2 and 3 with their vinylogues could all be merged and presented in a single table of all unit reactions with a third dimension for $\sum\Delta\sigma$, that is, with single construction half-reactions ($\sum\Delta\sigma = +1$) as a plane above refunctionalizations ($\sum\Delta\sigma = 0$) and the fragmentation halves ($\sum\Delta\sigma = -1$) below. The double construction and fragmentation half-reactions ($\sum\Delta\sigma = +2$ and -2 , respectively) would then be above and below the single ones on such a unified table, which represents a kind of "periodic table" of organic reactions.

Notation and Classification Examples: Some illustrative common reactions and their reaction notations are shown in Figure 2. In the first example, three adjacent carbons are seen to change their attachments, and these three are numbered 1, 2, 3. Their changes are characterized below the arrow and are recognized as an allylic reduction, labeled as [R']. The second case is similar, an allylic substitution [S'], but the difference from the first case emphasizes that the classification of reaction families is based strictly on the changes at the reacting carbons and not on the nature of the overall structure with all its unchanging structural parts.

The next two cases in Figure 2 are Michael additions, constructions composed of two half-reactions, the electrophilic half the same in each case ($R\Pi \cdot H\Pi$). In the first, four carbons change attachments, that is, two in each construction half-reaction. The nucleophilic half (carbons 1, 2) is an oxidative addition

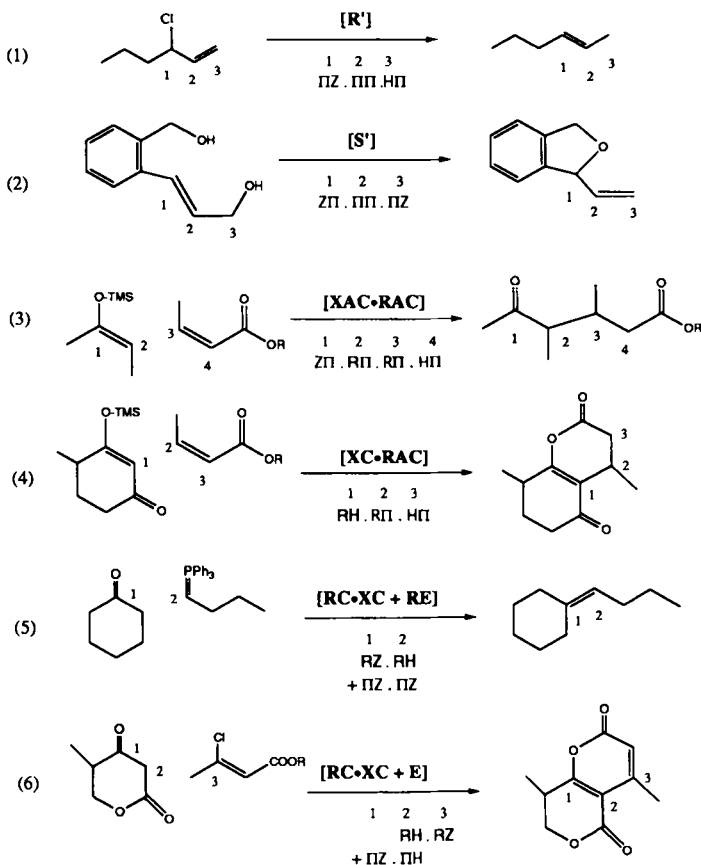


Fig. 2. Examples of reaction classification.

[XAC], the electrophilic half a reductive addition [RAC] at carbons 3, 4. The chemically similar second case (example (4)) has an *overall* net structural change of only RH at C-1 in the nucleophilic half, and the electrophilic half is the same as in the previous case. The generalized attachments at the oxygen-bearing carbons are unchanged: C-1 remains as $z, \pi = 1, 1$ and the ester remains as $z, \pi = 3, 0$.

The last two examples in Figure 2 show the notation for composite reactions. The Wittig reaction in case (5) has two joining half-reactions followed by a reductive elimination and so is labeled as [RC·XC + RE]. Another composite construction, in case (6), is the joining of two simple half-reactions (carbons 2, 3) followed by isohypsic elimination at carbons 1–2 to aromatize the pyrone, hence labeled as [RC·XC + E].

Finally, in reactions that change only the heteroatoms in a functional group, bonds to carbon are unaffected. In these cases the heteroatoms themselves may be treated as if they were carbon atoms with four attachments:^[5] charges are ignored and unshared electron pairs are annotated as H, that is, as the conjugate base of an acid. Thus the common reactions are reduction [R] and oxidation [X], as with carbon, and they are labeled with the symbol of the changing base atom, as [R(N)] for reduction of nitro, or [X(S)] for oxidation of sulfide.

Indexing Reactions: The digital nature of the definitions here makes it possible to create a numerical identification, to order and index organic reactions. For this purpose we may return to the sense of a reaction as the *net structural change* between substrate and product, and define this as the change in the values of σ , z , and π . The class of the reaction is determined first by recognition of any carbons with $\Delta\sigma \neq 0$. Within each class then the net structural change is only a function of the change

in functional groups, that is, the change in the values of z and π between substrate and product.

A list of these $z\pi$ values over the strand of linked changing carbons in the product can be subtracted from the corresponding carbons' $z\pi$ list in the substrate to yield a $\Delta z\pi$ list, a binary number, which serves as an identification number for the reaction family.^[6] Within each class all unit reaction families afford different numbers; in fact it was found that all of the composite reactions had unique identifying numbers as well.^[6]

The number of examples within a reaction family such as described here may still be very large. Closer refinement of the description can now be developed for purposes of focusing a match when searching for similar reactions. Since the reaction family is defined by the $\Delta z\pi$ list of the change from substrate to product, particular cases may be further narrowed down by describing the starting values of σ, π, z on the changing carbons, in the substrate, as in Equation (4).

$$(z\pi \text{ list})_{\text{SUB}} - (\Delta z\pi \text{ list})_{\text{REAX}} = (z\pi \text{ list})_{\text{PROD}} \quad (4)$$

Therefore, detailed matches consist of defining not only the generic reaction change ($\Delta z\pi$ list), but also the starting character (including σ) of the changing carbons; the product character can equally be used for this matching refinement. In practice, even on very large databases, this matching of both reaction family and starting character proves to be a very effective pruning tool.

The COGNOS Program for Reaction Retrieval: The logic of reaction description outlined here can be tested for efficacy by implementing it in a computer program to retrieve from a reaction database those reactions that represent close matches to an input reaction query. The COGNOS program has been developed to utilize these concepts and to test them in practice.^[7]

The program recasts the entries from any database in REACCS or SYNLIB format^[8] using the $\sigma z\pi$ character format described here for the changing carbons. These are then indexed by their reaction family and the character of their starting compounds, and the entries are arranged in numerical order of their family identification numbers and their starting characters. When a query reaction is drawn on the screen, COGNOS classifies it in the same way and directly seeks its matching family and starting character from the database. Because of the imposed ordering of like reaction entries together in the recast database, this retrieval is virtually instantaneous.

Besides the reaction family and the starting character of the entries, the index retains information about carbons adjacent to the reacting site for further pruning of matches if needed. The program also keeps the nature of those atoms or groups that leave the reacting site or that enter it in the product, that is, the actual group represented by $\pm Z$ or $\pm H$ in the generalized reaction notation. When the user chooses among these several pruning options in COGNOS, the screen immediately responds with the number of matches found; thus the user always knows how many entries to expect when he proceeds to examine the individual examples on the screen.

Summary

The COGNOS program has been designed to meet the criteria described at the outset, generalizing organic transformations into unit reaction (or composite reaction) families and indexing database entries into these families for instant retrieval. The system for doing this is fundamental in chemical terms and rigorous in application to digital format for the computer. The success of the program in finding correct matches from the databases lends confidence in this logic for organizing reactions, and the speed with which it finds matches from large databases makes it very practical to use. The literature precedents obtained are sharply defined by the system so that it is always clear just what results one will obtain, and what will be excluded.

Acknowledgment: We wish to thank the Deutsche Forschungsgemeinschaft for a postdoctoral fellowship for Dr. Sander.

Received: January 31, 1995 [F77]

- [1] J. B. Hendrickson, *J. Am. Chem. Soc.* **1971**, *93*, 6847. The development of the system since this original paper has appeared in fragmented form, but not completely described as in this presentation; see J. B. Hendrickson, *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 1286.
- [2] The term *isohypsic*, from the Greek for "equal level", was introduced to mean neither oxidative nor reductive [1].
- [3] J. B. Hendrickson, *Recl Trav. Chim. Pays-Bas* **1992**, *111*, 323.
- [4] J. B. Hendrickson, T. M. Miller, *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 403; *J. Am. Chem. Soc.* **1991**, *113*, 902.
- [5] J. B. Hendrickson, *J. Chem. Educ.* **1965**, *62*, 245.
- [6] The computing details for the reaction identification numbers can be found in J. B. Hendrickson, T. L. Sander, *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 251.
- [7] The COGNOS program is available as a CD-ROM from InfoChem, Munich (Germany), coupled with their ChemReact database of 370 000 reactions.
- [8] The databases are REACCS from Molecular Design, San Leandro, California (USA) and SYNLIB from Distributed Chemical Graphics, Meadowbrook, Pennsylvania (USA).