until they were acetylated. The present results agree with this observation of Thakker et al. No significant stereoselectivity is observed.

The acid-catalyzed reaction of BAP-4,5-diol in methanol gave the two methoxy derivatives **25** and **26**. However, the acid-catalyzed rearrangement of BAP 4,5-oxide in methanol gave the phenols and the four hydroxy ether adducts (Scheme II). Since reflux conditions are required to carry out the reaction with BAP-4,5-diol, it is conceivable that the reaction initially produces phenols and hydroxy ether adducts which subsequently give rise to the two methoxy compounds **25** and **26**. Phenols of BAP are known to be converted into the methoxy derivatives by heating in acidic methanol.[20]

Hydration of the carbocation from the acid-catalyzed ring opening of BAP 4,5-oxide was observed if the concentration of water was sufficiently high. Thus, use of 20:1 dioxane/water produced measurable amounts of trans dihydro diol in addition to the two phenols. If less water was present, only the phenols were formed. These studies used methanesulfonic acid. However, if HCl was used instead of methanesulfonic acid, trans dihydro diol in addition to the phenols was formed even in 100:1 dioxane/water along with two chlorohydrins. Likewise, if LiCl was present with methanesulfonic acid, all of these products were formed even when the water concentration was low. Chlorohydrins were always present unless the sample workup was altered to include an aqueous washing step, in which case the chlorohydrins disappeared. Generally, loss of chlorohydrins resulted in the formation of increased amounts of trans dihydro diol, although the phenols may also be increased somewhat. Clearly, chloride is an exceptionally good nucleophile in trapping the carbocation generated from BAP 4,5-oxide under acidic conditions. Since only trans products are observed, it appears that trapping of the carbocation by water or chloride must involve some type of concerted process rather than the trapping of an essentially free carbocation. However, this statement is only tentative. The conclusion that only trans chlorhydrins form is based only upon the observation of two products rather than four products by ^{13}C NMR and the observation that trans dihydro diol is formed from these two products. The chlorohydrins were never isolated. They were only observed in the NMR analysis.

Although the formation of chlorohydrins as intermediates as reported here represents the first demonstration of their existence in solvolytic reactions of arene oxides, they have been proposed previously from kinetic studies. Whalen, Jerina, and co-workers observed specific chloride effects in their pH–rate studies of the solvolysis of phenanthrene 9,10-oxide.[22] In this study, however, the product composition which included the 9-phenol and cis and trans dihydro diols did not change in the presence of chloride. It was suggested by these workers that the intermediate trans chlorohydrin could solvolyze to the same carbocation formed by acid-catalyzed or spontaneous ring opening and then subsequently give the same distribution of products as are formed in the absence of chloride. Unlike BAP 4,5-oxide, phenanthrene 9,10-oxide shows a significant spontaneous ring opening. It appears that chloride may be generally reactive toward arene oxides. Clearly, caution should be used in using chloride to adjust the ionic strength of a given solution which is to be used for reactivity studies.

Registry No. 1, 71719-12-5; 2, 71719-13-6; 3, 71719-14-7; **3a**, 60692-95-7; 4, 71719-15-8; **4a**, 60692-96-8; 5, isomer 1, 71719-16-9; 5, isomer 2, 71771-92-1; **5a**, isomer 1, 71719-17-0; **5a**, isomer 2, 71869-28-8; 6, isomer 1, 71719-18-1; 6, isomer 2, 71771-93-2; **6a**, isomer 1, 71719-19-2; **6a**, isomer 2, 71771-94-3; 7, isomer 1, 71719-20-5; 7, isomer 2, 71771-95-4; 8, isomer 1, 71719-21-6; 8, isomer 2, 71771-96-5; 9, 71719-22-7; 10, 71719-23-8; 11, 71719-24-9; 12, 71719-25-0; 13, 71719-26-1; 14, 71719-27-2; 15, 71719-28-3; 16, 71719-29-4; 17, 71719-30-7; 18, 71719-31-8; 19, 71719-32-9; 20, 71719-33-0; 21, 71719-34-1; 22, 71719-35-2; 23, 71719-36-3; 24, 71719-37-4; 25, 71719-38-5; 26, 71719-39-6; BAP 4,5-oxide, 37574-47-3; *tert*-butylmercaptan, 75-66-1; glutathione, 70-18-8; *N*-acetylcysteine, 616-91-1; sodium azide, 26628-22-8; methanesulfonic acid, 75-75-2; cyclohexene, 110-83-8; 2-chlorocyclohexanone, 822-87-7.

(20) H. Yagi, G. M. Holder, P. M. Dansette, O. Hernandez, H. J. C. Yeh, R. A. LeMahieu, and D. M. Jerina, *J. Org. Chem.*, **41**, 977 (1976).

(21) P. Y. Bruice, T. C. Bruice, H. Yagi, and D. M. Jerina, *J. Am. Chem. Soc.*, **98**, 2973 (1976).

(22) D. L. Whalen, A. M. Ross, P. M. Dansette, and D. M. Jerina, *J. Am. Chem. Soc.*, **99**, 5672 (1977).

An Approach to the Linear Representation of Reaction Mechanisms

John S. Littler

School of Chemistry, University of Bristol, Bristol BS8 1TS, England

Received December 7, 1978

The objectives to be achieved by a satisfactory linear notation for the symbolization of reaction mechanisms are discussed. A scheme is proposed, based on simple valence-bond representations of the elementary processes, which is designed to be suitable for computer storage and search, yet is reducible in simple cases to a few symbols which can be easily related to the conventional diagrammatic representation of a reaction mechanism.

A number of proposals have been made recently for the notation of reaction mechanisms in a systematic manner, with a view to improving on the proliferating nonsystematic usage of symbols modeled on Ingold's original S_N1 and S_N2 abbreviations and with a number of other objectives.[1-3] These include an easier indexing and computer retrieval of related reactions, the clear illustration of interreaction relationships, and the use of symbolization as a pedagogical tool. The present proposals sprang from a critical consideration of the proposals of Guthrie.[2] This has resulted in an extended scheme, addressed more di-

(1) J. Mathieu, A. Allais, and J. Valls, *Angew. Chem.*, **72**, 71 (1960).

(2) R. D. Guthrie, *J. Org. Chem.*, **40**, 402 (1975).

(3) D. C. Roberts, *J. Org. Chem.*, **43**, 1473 (1978).

0022-3263/79/1944-4657$01.00/0

rectly to certain identified objectives, which are discussed in the first section of this paper. The proposed notation scheme is presented in the second part, after which follow detailed considerations of particular aspects and examples of its use.

Objectives

There are two main groups of objectives which must be considered, which can be broadly described as defining first the information content which we wish to record and then the constraints on its mode of presentation.

(1) **Information Content.** Our knowledge about a reaction is hierarchical in nature, starting at the top with the fact that a transformation from reagents to products can take place. Classification of transformations, without regard to mechanisms or intermediates, has been considered by others,[4] and such classifications are not the concern of this paper, though a representation of a mechanism must contain the information which enables us to deduce what overall transformation is achieved, as well as the nature of the intermediates involved.

On the next lower level we can use experimental data to form a model of the reaction mechanism. In the absence of discriminating experimental evidence it may well be possible to form a number of different models. The symbolization system should be able to represent any of these alternatives clearly and distinguishably. For any one model we can proceed by dissecting the process into more elementary processes (EP). These are, in principle, observable as separate reactions, have no identifiable intermediates, and lead from one state to another distinguishable in some way from it.

Most such processes have a potential energy maximum or other barrier to pass over. There are a few processes where there may be no enthalpy barrier, e.g., a spontaneous dissociation of a metastable state, or no energy maximum at all, e.g., a bimolecular association at collision rate. The microscopic reversal of a spontaneous dissociation has a large energy hill to climb, so the states are clearly distinguishable. The atomic and electronic movements within an elementary process are concerted (though not necessarily proceeding to the same extent at any given point within the process); if two parts of a process are not concerted they can be represented as separate EPs.

The sequence of such elementary processes must be specified in some way, and some provision must be made for the representation of chain reactions etc. unless the description is to be limited essentially to descriptions of elementary processes.

As the potential complexity of an elementary process is large, we need to describe it in terms of a lower level, more detailed, model. However, our direct knowledge is as yet limited, being restricted largely to the overall changes in interatom connections which the EP achieves, although if we have intermediates or even transition states on which structural studies can be made, we may also have direct evidence as to atom positions and electron distributions. There is a choice, therefore, as to how we represent the EP. It can be considered as a "black box" into which enter molecules of a certain structure, while the products, having a new structure, emerge. The change could be represented as a transformation of a connectivity table from that of the reagents to that of the products. This would contain no information about the way in which an EP occurs. An opposite extreme approach would be to give a complete theoretical description of the multidimensional energy surface leading from reagents to products. More usefully, we can take a middle path and use the traditional terminology of bond formation, breakage, etc. to subdivide the EP into a hypothetical series of "primitive changes" (PC) which are not in themselves observable, though inferred from the experimental evidence. This coincides with the common way of representing such processes on paper and provides a satisfactory framework for the entry of information at a level of detail not inconsistent with the amount of experimental knowledge we actually have about many reactions, from kinetic and product studies, isotopic labeling, etc.

The necessary test of the success of the representation of an elementary process as a group of PCs is that one should be able on paper to generate the product uniquely from the reactant with the aid of this information, as does the actual reaction in the flask (with the hope, but without the certainty, that the paper route is a correct model within the limits imposed by its representation). An alternative way of expressing this is to say that the group of PCs which comprise an elementary process acts as a group of operators which transform the formulas of the reagents or intermediates they act on into the formulas of the products. It is essential to understand that PCs do not exist in isolation but depend on each other, although an elementary process consisting of a single PC is, of course, possible.

Unlike earlier schemes, the present one sets out not to restrict PCs to those involved in simple bond formation (association) or breakage (dissociation), though these are often the predominating ones. There must be representations for π-complex formation, conformational changes, electronic excitation, electron transfer, etc. when these are either the sole primitive change within an elementary process or a significant feature of it.

Roberts[3] has pointed out that it is possible to describe concerted processes at an even lower level in terms of the essential electronic movements (without regard to the chemical natures of the atoms concerned) so that all isoelectronic processes are represented in the same way. This degree of abstraction is certainly useful for analysis of related reactions, but it does not seem to provide an immediately acceptable basis for encoding mechanisms and for decoding the linear representation to a normal diagram. However, the scheme proposed here is capable of being denuded of information as to which element is in which place if we require that the basic isoelectronic relationships be emphasized. It should be emphasized that a minimum of structural information is essential whenever the relative positions of the atoms are a crucial feature of a mechanism. This situation prevails in most EPs which are comprised of more than one PC.

There are various other pieces of information which it may be desirable to include in a linear notation. At a high level the relative yields of competitive reactions, e.g., isomer distributions, may be useful. When specifying the EP the individual PC may need amplifying with stereochemical information, and the existence of resonance structures (i.e., states in which bonds are not localized) may be important. If possible the representation should be capable of including all the potentially verifiable features of a reaction mechanism but should not require the inclusion of speculative processes or details of probable but unobservable precursor and successor reactions.

If rate-limiting steps are known, they are obviously of prime importance, and it may well be desirable to designate them to assist in retrieving the information. It will also be useful to be able to specify that the quantity of

(4) J. Blair, J. Gasteiger, C. Gillespie, and I. Ugi, *Tetrahedron*, **30**, 1845 (1974); J. C. Bart and E. Garagnari, *Z. Naturforsch., B*, **32**, 678 (1977); D. P. N. Satchell, *Naturwissenschaften*, **64**, 113 (1977).

material passing through an EP (the chemical flux) in the retrograde direction is a significant fraction of the forward flux, i.e., that an EP is what is loosely called reversible. In fact, all processes are reversible, no matter how unfavorable the reverse process may be, and the numerical value to be attached to the word "significant" in the previous sentence may vary with experimental circumstances. The use of the term "reversible" is therefore based on a practical or operational definition, rather than an absolute one.

(2) Mode of Presentation. It is essential that the growth of knowledge about a reaction mechanism should not totally alter its representation (unless, of course, the previous mechanism is proved fallacious), and the detection and specification even of new types of interaction should not be excluded. Briefly, it should be open-ended so as not to provide an artificial constraint for our ideas about mechanism.

A primary objective in selecting the mode of presentation of the information is to ensure that the translation process between the linear representation and the conventional diagram or "arrow-pushing" picture is relatively straightforward. Inevitably the linear representation will not be so immediately comprehensible, just as a full IUPAC name or a Wiswesser line formula takes longer to comprehend than does a conventional structural diagram; but if the representation contains a wide variety of easily recognizable symbols structured in a simple way, it will be easier to use than if it has few symbols in a highly complex structure. Prominence should not be given to symbols which convey very little information (e.g., the D and A of Guthrie's scheme where there are only two alternative PCs used, which provide only 1 "bit" of information). A scheme requiring careful counting of identical symbols should not be invoked (Roberts' 1,0 sequences), nor should unconventional uses of well-known symbols (Roberts uses intersecting, rather than nested, parentheses). It should be possible to use normal chemical symbols for the elements.

Another objective is to facilitate indexing and file searching. Clearly, the representation is likely to be too complex to be indexed by a simple alphanumeric sequencing method, and we must envisage that retrieval from a computer file is to be used. It is therefore possible to plan for more elaborate searching and indexing processes. As an example, one might search for all reactions containing a certain elementary process, whether this occurred early or late in a reaction sequence, and present the results in the form of a KWIC index where the elementary process is the keyword. It should also be possible to structure the representation so that optional information fields are included, which can be ignored or considered, as required by the searching program. As computer storage and information retrieval is inevitably the best way of handling representations of this sort, the symbols used must be restricted to the standard printable ASCII (American Standard Codes for Information Interchange) character set, and although it would be useful and improve legibility to use both lower and upper case forms of the alphabet, the absence of lower case on many printers etc. suggests that no distinction should be made between the meanings of the upper and lower case forms. Similarly, subscripts and superscripts must not be used, and the insertion of spaces and new lines into the representation should be optional. A third objective, essential in any scheme of nomenclature or representation, is that the representation should not only be unambiguous but also unique. This means introducing additional rules when working out the correct notation for a reaction, designed to ensure that the result does not depend on whomever works out the notation. This extra complication is not entirely without compensatory advantages, since it enables a certain degree of verification to be applied automatically to detect and eliminate errors.

(3) Further Objectives. It is desirable that simple processes (e.g., S_N1, S_N2) should be represented by short groups of symbols, which might even become familiar if used frequently, and that where a mechanism is such that the overall transformation is effected by one elementary process the description of the mechanism should not be unlike the description of the transformation using the currently acceptable, though not yet standardized, descriptions such as "hydroxy deiodination" for the S_N2 hydrolysis of an alkyl iodide. Neither of these objectives has been fully met, the former because we are encoding more information than is conventionally implied by a symbol such as S_N1 and the latter because the descriptions of transformations deliberately avoid including more than the minimum of mechanistic information, so although one might tailor the representation of a particular mechanistic route closely to the name of the transformation, the alternative mechanistic routes would give very different representations.

It is also desirable that any notation scheme should be easy to understand and to teach to students and new users. The close correlation between this scheme and a conventional diagrammatic representation should help in achieving this objective, and inevitably the actual process of encoding a diagram will ensure that the student looks closely at just what is actually implied by the conventional picture. This can do nothing but good.

It is also necessary to ensure that a scheme of this nature is not totally in conflict with work on representing, classifying, indexing, and naming transformations. In general, there should be no problem in spite of the very different representation used, since the completion of the steps of a reaction mechanism should result in the completion of a specific transformation, which can be expressed as a change of formula, or of connectivity table, or classified as a reaction matrix type, as required.

Main Features of the Notation Scheme

The notation scheme must reflect the hierarchical structure of the information being stored. We therefore first consider its overall features, the notation to be used for PCs, and the problems of resolving ambiguities. Further rules are then introduced to restrict the possible representations to a unique form and at the same time to assist in the checking of the correctness of the representation.

(1) Overall Structure. In line with the above considerations the reaction can be described as a series of elementary processes (EPs), each subdivided into primitive changes (PCs), and the line notation can be divided up into fields by punctuation marks; i.e., using semicolons to separate the EPs, we can represent a reaction as a list of the reagents followed by

EP no. 1; EP no. 2; EP no. 3

Each EP can be expanded as

process header: PC no. 2.1, PC no. 2.2..., PC no. 2.*n*;

The process header is a field which can contain optional extra information about the elementary process to which it is prefixed. The primitive change is then represented as an interaction between two atoms or groups of atoms, e.g.

CldeC

meaning the loss of Cl^- with breakage (represented as de) of the C–Cl bond. Thus the archetypal S_N1 process in an alcohol or water can be represented simply as two EPs, each consisting of one PC

XdeC; OasC

where as means association and the dissociation precedes the solvation of the carbonium ion, while the S_N2 process is OasC, XdeC where the primitive changes are concerted, rather than consecutive, and thus are combined to form a single EP.

(2) Notation of Primitive Changes. The two interacting atoms are written on either side of the symbol chosen for the interaction. The symbols used in the examples are:[5]

AasB—formation of a bond between A and B (heterolytic), with A bringing the electron pair

$$BF_3 + NH_3 \rightarrow BF_3NH_3, \qquad \text{NasB}$$

AdeB—breakage of a bond between A and B (heterolytic) where A carries away the electron pair

$$C(CH_3)_3Cl \rightarrow C(CH_3)_3^+ + Cl^-, \qquad \text{CldeC}$$

AhaB—homolytic association where A initially is a radical (colligation, if B is also a free radical)

$$CH_3\cdot + CH_3\cdot \rightarrow C_2H_5, \qquad \text{ChaC}$$

$$CH_3\cdot + C_6H_6 \rightarrow \cdot C_6H_6CH_3, \qquad \text{C<1>haC<2>}$$

AhdeB—homolytic dissociation (if only one product is a radical it is A)

$$C(CH_3)_3OOC(CH_3)_3 \rightarrow 2C(CH_3)_3O\cdot, \qquad \text{OhdeO}$$

ApaB—formation of a π bond between A and B (which are already σ bonded)

$$C_6H_5N_2^+ \rightarrow C_6H_5^+ + N_2, \qquad \text{NdeC, NpaN}$$

Note here and in most π-bond reactions it is difficult or impossible to find examples in which a single PC can exist as an EP.

ApdeB—breakage of a π bond between A and B (which are still σ bonded); electron pair remains on A

$$H^+ + \underset{1}{CH_2}{=}\underset{2}{CH_2} \rightarrow HCH_2CH_2^+, \qquad \text{C1pdeC2, C1asH}$$

AhpdeB—as previous, but homolytic; odd electron remains on A

$$R\cdot + \underset{1}{CH_2}{=}\underset{2}{CH_2} \rightarrow RCH_2\dot{C}H_2, \qquad \text{RhaCl, C2hpdeC1}$$

ApiB—formation of a π bond, without a σ bond (e.g., π association between halogens and olefins or aromatics)

ApideB—breakage of a π bond, not associated with a σ bond (e.g., reverse of previous reaction)

AwaB—formation of a weak association (e.g., ion pair or hydrogen bond)

$$(C_2H_5)_2O + HOC_2H_5 \rightarrow (C_2C_5)O\cdots HOC_2H_5,$$
$$O\langle 1\rangle \text{ waH}(O\langle 2\rangle)$$

AwdeB—breakage of a weak association

AwwB—formation of a very weak association (e.g., solvent-separated ion pair)

Generally the dissociation symbol is like the association one, with the extra letters "de".

A further selection of primitive changes are perhaps less well defined and may not often be needed but are included to illustrate the flexibility and extensibility of the system:

AvsB—van der Waals association (if necessary to distinguish it from ww)

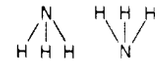
AanhyB—"hydrophobic bond"

AaxA—conformation change, A moves to axial position

AeqA—similarly to equatorial position

ArotB—rotation about the A–B bond

AinvA—conversion of molecule to a new form by inversion of atom A, e.g.



ApipiexB—$\pi \rightarrow \pi^*$ transition in A–B double bond

AnpiexB—n $\rightarrow \pi^*$ transition, n orbital was on A

Further photophysical changes could be included in a similar manner but there is no need to suggest a detailed symbolism here and now, particularly as a glossary of suitable terms and symbols is being prepared. Photochemical changes, in general, can be subdivided into a photophysical PC or EP (excitation etc.) followed by a normal chemical PC or EP.

Reactions such as $2CH_2 \rightarrow C_2H_4$ and $2N \rightarrow N_2$ can be made up of more than one PC, e.g., C<1>asC<2>, C<2->paC<1>. Electron transfers can be included by replacing B by a symbol (e) for the electron; a reduction is then, formally, bond formation to e. For example, in the $S_{RN}1$ reaction the initial step (chain initiation) consists of the transfer of an electron from the free radical initiator A to a C-based π orbital. Bearing in mind that these reactions are radical reactions (i.e., the symbol h is needed), that if only one product is a radical it must be written on the left (and in this case is the electron), and that spin should move from left to right we get

ehdeA, ehpiC;

This is followed by a conventional S_N1 substitution of X by Y on the radical

XdeC; YasC;

followed by a chain transfer from molecule n to the next in the chain $(n + 1)$

ehpideC<n>, ehpiC<$n + 1$>;

In all cases where primitive changes are combined the replacement of one association by another involves the formal breakage of the first and formation of the latter; e.g., if an ion pair collapses, it is represented AwdeB, AasB. Unlike previous suggestions, shifts involving π bonds must be spelled out in detail (unless an acceptable abbreviation can be found). This lengthens the sequence of PCs in, for example, the representation of ester hydrolyses, but if they are omitted, the intermediates appear to include compounds containing pentavalent C and the vital role of the C=O double bond is ignored. Any omission of π-bond shifts violates the principle noted previously that the EP representation must be sufficiently detailed to enable us to reconstruct the structure of the correct products or intermediates, knowing only the precursors of the step concerned.

(5) The precise choice of symbols is irrelevant to the format of the representation, though it is vital to its acceptability to chemists. Originally, punctuation marks were used (e.g., · for ha, * for as, / for weak associations) to avoid confusion with element symbols when only upper case characters are available, but as a result of comments from referees and others, these have been replaced entirely by short sequences of lower case letters, which are chosen to be mnemonics for the PC concerned. When lower case is not available, it will be necessary to enclose the symbols in quotes or use some other means of distinguishing them. It could be advantageous to choose pronounceable mnemonics so that the designation could be more easily used in speech and be more easily read, but I have not attempted this since it would be necessary to consider vocalization in a variety of languages.

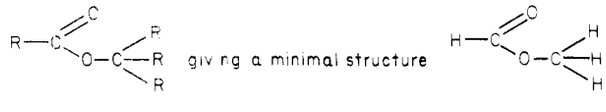
If a conformational change which is thought to be a precursor to a reaction generates a conformation which is a conformational energy maximum and which may be further excited and react, it is not possible experimentally to draw a distinction between the two processes anymore than in, say, the vibrational preactivation of an inorganic complex to increase its ease of electron transfer. Such subdivisions are, however, useful in discussion or for calculating the parameters of the reaction path. The representation of the conformational excitation step then becomes a conformational PC preceding any bonding changes, and any subsequent conformational relaxation of the products can be included after the bonding PCs.

(3) Designation of the Site of Reaction. It would clearly be ludicrous to represent in different ways two reactions such as the addition of bromine to a 9,10 double bond or a 10,11 double bond in a C_{20} unsaturated acid, yet on the other hand it is essential to identify the atoms individually in all but the simplest processes. Designating the carbon atom (for organic substrates) which is first attacked as α, its neighbor as β, etc. leads to difficulties even in an "E2" situation and becomes very ambiguous in any rearrangement. It is therefore necessary to determine a structure containing the minimum features required for a mechanism, the "*minimal structure*", number that structure according to conventional rules, and keep the same designation for each atom throughout, since redesignating atoms for each step is liable to lead to confusion.

We must therefore decide which are the atoms or groups involved in the reactions. These are those to which a bond is broken, made, or significantly changed (e.g., by resonance with a transition state or intermediate) during a reaction or which are necessary to hold other groups in the correct stereochemical relationship (e.g., in a transannular reaction). This, somewhat arbitrarily, excludes hyperconjugation, inductive, and solvation effects.

All other (noninvolved) groups in the formula of the reagent can then be formally replaced by hydrogen atoms. This gives a "minimal structure". If this resulting compound is numbered according to normal IUPAC rules, an unambiguous numbering of the reactant atoms should often result. The minimal structure(s) should be named explicitly as part of the designation of the mechanism—this is essential if the representations are to be indexed so as to find all examples of a given reaction. If a substituent group is first attacked, and it is not clear to which carbon atom it is attached, that can be specified in parentheses; e.g., O(C1) as H; = protonation of the oxygen attached to C1.

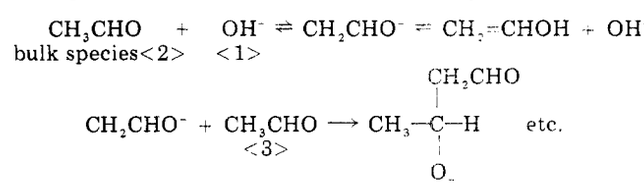
If the minimal structure approach is applied to a reaction such as the hydrolysis of an ester, the structure of this substrate, writing R for hydrogen or alkyl groups not involved in the reaction, is



i.e., the minimal structure is that of methyl formate.

The minimal structures (with also, if required, the solvent) form the list of reagents which precedes the detailed representation of the mechanism. It should be made quite clear that the description of a reaction operating on a minimal structure defined as above does not imply that an actual substance of the same formula as the minimal structure will undergo reaction in the manner described. For example, not all ester hydrolysis mechanisms are applicable to methyl formate. Similarly the S_N1 reactions (Table I) operate on the minimal structure CH_3X, and the pinacol rearrangement (Table I) operates on the minimal structure $CH_2OHCHOHCH_3$, since only one methyl group is directly involved. The normal numbering system in the ester example above gives us two carbon atoms labeled C1, one being in the alkyl moiety and the other being the carbonyl C, and two oxygen atoms, without numbers, which are both involved in the reaction, i.e., the carbonyl and the "ether" oxygen. This difficulty is removed if, in a reaction where two reactant centers are linked by a heteroatom, we apply "replacement nomenclature" (Section C-0.6 of ref 6) so that the substrate is correctly described as 2-oxapropanal, with the acyl C as C1, the "ether" O as O2, and the alkyl C as C3, while the ketone oxygen is O(C1). The numerical locants (positional designations) are then unambiguous.

(4) Designating the Molecules Involved. A familiar problem, that of logically distinguishing the substrate from the reagent, arises when any EP involves a bimolecular process. Each molecule can be named as above, and locants determined for the reacting atoms. We must also designate which atom belongs to which molecule by appending as an arabic numeral a "reagent number" (or in a chain reaction (Table I) the general letter n or m) in distinctive brackets to the atomic symbol, e.g., Cl<2>, where each molecule in turn entering a mechanism is assigned the next higher reagent number. In a sequence of EPs the numbers must be carried forward to ensure that the correct product from the first EP is used in the second, and once labeled by its origin, the atom should keep the same label throughout. For an example, see the Cannizzaro reaction, Table I. However, the actual value of the label may be assigned differently by different people when encoding the same mechanism or when the same reaction is considered under different conditions. For example, the base-promoted condensation of an aldehyde anion



is the same process whether the process is regarded as starting from the aldehyde, its isomer the enol, or the enol anion, so the nucleophilic reagent could be labeled <1> and the electrophilic one <2>, rather than as in the diagram. Here an attempt can be made to generate a unique representation by complicated rules, but, more simply, the ambiguity can be accepted since a search program examining a list of representations could easily be made to take note only whether these numerals were the same or not, without regard to their actual numerical value and sequence. The problem of the substrate/reagent dichotomy then is seen to be a trivial consequence of our assigning to the molecules active and passive roles which are not physically significant. Ambiguity can be avoided by simultaneously specifying the necessary reduced structures and their assigned reagent number. Normally the species specified should be those present in the reaction mixture in bulk even if they are not the actual reactant species, though it is reasonable to include the products of autoprotolysis of the solvent as starting materials. Similarly detailed descriptions of the subsequent equilibria of the products could be omitted.

(5) Sequence of PCs in an EP. The sequence of primitive changes in a reaction is, by definition, arbitrary

(6) "IUPAC Nomenclature of Organic Chemistry, Section C", Butterworths, London, 1965, p 49.

Table I

S_N1 (ref M270) (dissociation limiting, rapid capture of the carbonium ion from either side
CH$_3$X, Y #:XdeC; ran:YasC

S_N1 (ref M271) (dissociation reversible, rate-limiting capture of the carbonium ion—note that the symbol # points a useful distinction here)
CH$_3$X, Y XdeC; #, ran:YasC

S_N1' (ref M303) (allylic—a case where the atoms must be distinguished by locants)
CH$_2$=CHCH$_2$X, Y #, $:C2pdeC3, C2paC1, XdeC1; YasC3

S_N1cB (ref M330) (first step is a bimolecular substitution on hydrogen, followed by carbene formation by loss of X from anion)
CH$_3$X, Y YasH, CdeH; #:XdeC

S_N1cA, A1 (ref M325) (acid- or electrophile-promoted displacement, e.g., Ag^+ + *sec*-alkyl halide)
CH$_3$X, N, Y XasN; XdeC; YasC

S_Ni (ref M302) (OSX leaves, X returns)
CH$_3$OSX OdeC; OpaS, XdeS; XasC

S_N2 (ref M266)
CH$_3$X, Y inv:YasC, XdeC

S_N2' (ref M304) (allylic—note differences from S_N1')
CH$_2$=CHCH$_2$X, Y YasC3, C2pdeC3, C2paC1, XdeC1

S_N2cA, A2 (ref M325) (cf. S_N1cA)
CH$_3$X, M, Y XasM; YasC, XdeC

S_NAr, S_N2Ar (ref M584) (nucleophilic aromatic substitution—only the local C=C is specified, but $ indicates that this is not the whole story)
H$_2$C=CHX, Y $:YasC1, C2pdeC1, C2paC1, XdeC1 (either step rate limiting)

E1 (ref M903) (two-step process, M usually is H (written with first step rate limiting); note that this and the next two examples use the same PCs arranged differently)
MCH$_2$CH$_2$X, Y #:XdeC1; YasM, C2deM, C2paC1

E2 (ref M896) (one-step process)
MCH$_2$CH$_2$X, Y YasM, C2deM, C2paC1, XdeC1

E1cB (ref M904) (two step, via carbanion)
MCH$_2$CH$_2$X, Y YasM, C2deM; C2paC1, XdeC1 (either step rate limiting)

E1cA (ref M924) (when N = H, preprotonation of leaving group)
MCH$_2$CH$_2$X, N XasN; C2deM, C2paC1, XdeC1

S_E1 (ref M459) (via free carbanion, often aromatic)
H$_3$CM, N CdeM; CasN

$S_E1(N)$ (ref M530) (nucleophile-assisted electrophilic substitution)
H$_3$CM, N, X XasM, CdeM; CasN

S_E2 (front) (ref M521)
H$_3$CM, N ret:CasN, CdeM

S_E2 (back) (ref M521)
H$_3$CM, N inv:CasN, CdeM

S_E2Ar (ref M453) (two-step aromatic electrophilic substitution)
H$_2$C=CHM, N $, #:C1pdeC2, C1asN; C1deM, C1paC2

S_E2Ar with conjugative group in ring (e.g., ortho substitution of a phenol)
HOCH=CHM, N $, #:O(C1)paC1, C2pdeC1, C2asN; C2deM, C2paC1, O(C1)pdeC1 (here the PCs show the involvement of the OH groups)

S_E2' (ref M529) (with allylic shift)
H$_2$C=CHCH$_2$M, NC1deM, C1paC2, C3pdeC2, C3asN

S_H1 (ref M621) (spontaneous cleavage and then attack on A-B)
H$_3$CJ, A-B ChdeJ; ChaB, AhdeB

S_H2 (ref M621) (A as chain carrier)
H$_3$CJ, A-B AhaJ, ChdeJ; ChaB, AhdeB

S_HAr (ref M622) (A attacking aromatic ring)
H$_2$C=CHJ, A-B AhaC1, C2hdeC1; JhdeC1, JhaB, AhdeB (or other alternative second processes)

$A_{Ac}1$ (ref M350) (a special case of S_N1cA (O2 is ether oxygen; O<2> is that of water)—see text)
HCOOCH$_3$, H$_2$O, H$^+$ O2asH; #:O2deC1; O<2>asC1; O<2>deH

$A_{Al}1$ (ref M350) (alkyl-oxygen fission)
HCOOCH$_3$, H$_2$O, H$^+$ O2asH; #:O2deC3; O<2>asC3; O<2>deH

$A_{Ac}2$ (ref M350) (via tetrahedral intermediate)
HCOOCH$_3$, H$_2$O, H$^+$ O(C1)pdeC1, O(C1)asH; O<2>asC1; O<2>deH; O2asH; O2deC1; OdeH, OpaC1

$A_{Al}2$ (ref M350) (a special case of S_N2cA)
HCOOCH$_3$, H$_2$O, H$^+$ O2asH; O<2>asC3, O2deC3; O<2>deH

$B_{Ac}1$ (ref M350) (Y is normally HO$^-$)
HCOOCH$_3$, Y O2deC1; YasC1

$B_{Al}1$ (ref M350) (a special case of S_N1)
HCOOCH$_3$, Y O2deC3; YasC3

$B_{Ac}2$ (ref M350) (via tetrahedral intermediate)
HCOOCH$_3$, Y YasC1, O(C1)pdeC1; O(C1)paC1, O2deC1 (subsequent proton transfers omitted)

$B_{Al}2$ (ref M350) (a special case of S_N2)
HCOOCH$_3$, Y YasC3, O2deC3

Decarboxylation of a β-Keto Acid (ref M571) (cyclic mechanisms involving proton transfer to the carbonyl group)
HCOCH$_2$COOH C1paO, HdeO, HasO(C3), C3pdeO(C3), C3paC2, C1deC2

Pinacol Rearrangement (shift of C3 from C2 to C1)
H$_2$C(OH)CH(OH)CH$_3$, H$^+$ O(C1)asH; #, trans:O(C2)paC2, C3deC2, C3asC1, OdeC1; O(C2)deH

Cannizzaro Reaction
OH$^-$HCHO, HCHO rev:O<1>asC<2>, OpdeC<2>; #:OpaC<2>, HdeC<2>, HasC<3>, OpdeC<3>; O<3>asH

Halogenation of Methane
CH$_4$<*n*>, Cl$_2$<*m*>
#, photo:ClhdeCl (initiation)
[1:ClhaH, C<*n*>hdeH; 2:C<*n*>haCl, C<*m*>hdeCl<*m*>;] (chain)
1:ClhaCl;
2:ChaC;
1, 2:ChaCl;
epi, 1:ClhaCl;
epi, 2:ChaC
} chain terminations—sequence numbers show which processes are in competition with each other

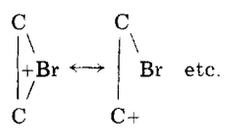
Free Radical Polymerization of Ethenes (Chain Step Only)
HCH$_2$CH$_2$· CH$_2$=CH$_2$ (minimal structures)
<*n*> <*n* + 1>
[C2<*n*>haC1<*n* + 1>, C2<*n* + 1>hpdeC1<*n* + 1>;] (the repeated chain step is enclosed in square brackets)

Cationic Route for Polymerization
[C1<*n* + 1>pdeC2<*n* + 1>, C1<*n* + 1>asC2<*n*>;]

2-Hydroxy-(2-1 methyl)-1-dechlorination
CH$_3$CH$_2$CH$_2$Cl + OH$^-$ (minimal structures)
OasC2, C3deC2, C3asC1, CldeC1; (one step)
(Note: the possible confusion between 1 and l does not arise in ASCII codes)
C3deC2, C3asC1, CldeC1; OasC2 (two step)

Radical Electrodeprotiation (E.g., Ph· + Ferricinium → Phenylferrocene + H$^+$)
C<1>haC<2>; addition of radical (molecule <1>)
C<2>deH; loss of H$^+$

Bromination of a Double Bond in the Presence of Cl$^-$ and N_3^-
CH$_2$=CH$_2$, Br$_2$<1>, Br$^-$<2>, Cl$^-$, N_3^-
$:CpiBr1; π association with one end of Br$_2$
#, $:CpideBr1, Br1asC2, C2pdeC1, C2asBr1, Br2deBr1; formation of



1, inv:Br<2>asC, Br1deC; opening by new Br$^-$

Table I (*Continued*)

1, inv:ClasC, BrdeC; opening by Cl^-
1, inv:NasC, BrdeC; opening by N_3^-

Bromination of Ketone under Acid-Catalyzed Enolization-Limited Conditions
CH_3CHO, Br_2, H^+
OasH; reversible protonation
#, trans:C2deH, C2paC1, OpdeC1; loss of α-H^+
OdeH, OpaC1, C2pdeC1, C2asBr1, Br2deBr1

The $S_{RN}1$ system (ref M597)
ArX, $A\cdot$, Y
ehdeA, ehpiC; transfer of electron to ArX (initiation)
[XdeC<*n*K; loss of X^- from the *n*th molecule in the chain
YasC<*n*>; nucleophilic addition by Y
ehpideC<*n*>, ehpiC<*n* + 1>] transfer of electron to new molecule; square brackets indicate these three EPs repeat (termination unspecified)

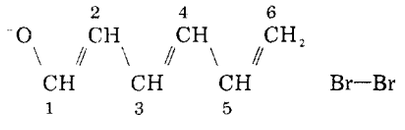
Diels-Alder
ethene<1>, butadiene<2> (minimal structures)
C1<1>asC1<2>, C2<2>pdeC1<2>,
C2<2>paC3<2>, C4<2>pdeC3<2>,
C4<2>asC2<1>, C1<1>pdeC2<1>

Favorskii Rearrangement
$CH_3C(O)CH_2Cl$, OH^- (minimal structure to give numbers to the cations, so that the H atoms can be identified)

Cyclopropanone Route
$:0<2>asH(C3), C3deH; C3asC1, CldeC1; (proton loss and cyclization)
1:0<2>asC2, C3deC2, C3asH; } two equivalent
1:0<2>asC2, C1deC2, C1asH } ring openings

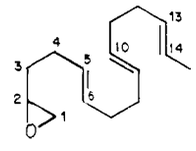
Benzylic-Type Route
O<2>asC2, OpdeC2; (addition of OH^- to carbonyl group)
OpaC2, C3deC2, C3asC1, CldeC1

Halogenation of a Conjugated Enol Anion



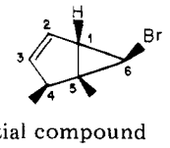
OpaC1, C2pdeC1, C2paC3, C4pdeC3,C4paC5, C6pdeC5, C6asBr1, Br2deBr1

Terpene Cyclization by Acid Attack on Epoxide, As a Succession of Steps



OasH; OdeC1; C6asC1, C6pdeC5; C10asC5, C10pdeC9; etc. (a synchronous mechanism would have the same PCs but in the reverse order and joined by commas into a single EP)

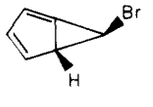
Some Bicyclo[3.1.0]hexene Rearrangements.[16] An Attempt to Apply the Scheme to a Particularly Complicated System



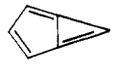
+ an oxygen base

initial compound

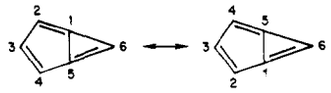
#, 1:C4deH, C4paC3, C2pdeC3, C2paC1, C5paC6, BrdeC6; (leading to bromobenzene)
#, 1:C1deH, C1paC2, C3pdeC2, C3paC4, BrdeC4; (leading to the intermediate



2, dis:C1paC6, C5deC6, C5paC4, C3pdeC4, C3paC2, C1pdeC2; (disrotatory ring opening, leading to bromofulvene)
$, 2,cis:C5deH, C5paC6, BrdeC6; (formation of bicyclic triene



$:C1paC6, C5pdeC6, C5paC4, C3pdeC4, C3paC2, C1pdeC2; (resonance process which interconverts certain atom labels)



In *t*-BuOH
aequo:OasC6, C5pdeC6, C5asH; addition of *t*-BuOH
dis:C1paC6, C5deC6, C5paC4, C3pdeC4, C3paC2, C1pdeC2; (disrotatory ring opening giving butoxyfulvene)

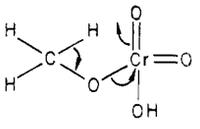
In Me_2NH
aequo, trans:NasC6, C5pdeC6, C5paC1, C2pdeC1, C2asH;
cis:C1asH(C2), C2deH(C2), C2paC3, C4pdeC3, C4paC5, C1pdeC5; H shift
dis:C1paC2, C3pdeC2, C3paC4, C5pdeC4, C5paC6, C1deC6 (ring opening to (dimethylamino)fulvene)

Note, this implies
(1) that all reactions specifically written for one canonical form can apply to the other (identical) form:
(2) that attack by *t*-BuOH and Me_2NH can occur either side;
(3) cis: for the H shift is freshly invoked to notate the observed stereospecificity.

The numbering of these transformations gives rise to the eventual products showing immediately where carbon isotopes should appear and also where hydrogen isotopes should come if not shifted or exchanged. This is quite useful in the context of these particular papers where it is easy to get confused when the authors change from bicylic numbering to fulvene numbering.

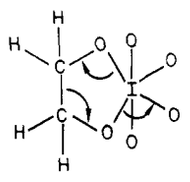
Oxidation State or Coordination Number Changes

Chromic Acid Oxidation of an Alcohol, a Decomposition Process of a Coordination Complex



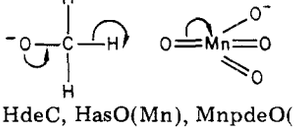
There appears to be no inherent limitation to organic or organometallic reactions, though extra PCs may need to be defined.
CpO(C), CrdeO(C), O(Cr)pdeCr, O(Cr)asH, CdeH

Glycol + Periodate



C1pO(C1), IdeO(C1), O(C2)deI, O(C2)paC2, C1deC2;

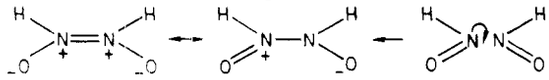
Alcohol Anion + Permanganate



O(C)paC, HdeC, HasO(Mn), MnpdeO(Mn);

Table I (*Continued*)

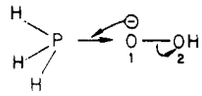
Dimerization of Nitroso Compounds



$:N<1>asN<2>, O<2>depN<2>, N<2>paN<1>, O<1>depN<1>;

Oxidation of R_3P by ^1O_2H

PasO1, O2deO1, O1paP;



The Claisen Rearrangement

CH_2=CHOCH$_2$CH=CH_2 (3-oxahexa-1,5-diene is the minimal structure)

C1asC6, C5pdeC6, C5paC4, OdeC4, OpaC2, C1pdeC2;

C1deH, C1paC2, OpdeC2, OasH (including subsequent keto-enol shift if the substrate is aromatic)

Hydroboration (four-center mechanism)

ethene, BH_3

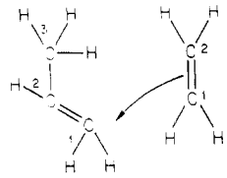
C1asB, HdeB, HasC2, C1pdeC2

since they are concerted, but there are advantages in defining the order in which they are to be written down. In the general case of a sequence of electronic shifts with bond formations and breakages a simple rule to get a unique representation is to write the PCs from left to right in the same sequence as is followed by the formal movement of the electrons or (if a radical process) of the unpaired spin. This rule cannot, of course, apply to electrocyclic and similar processes, so an arbitrary decision must be taken. A possible solution, illustrated for the "ene" reaction, is as follows:

(i) Write down and number the minimal structures of the reagents.

(ii) Consider the bonds which are to be formed between these and choose the one between the lowest numbered atoms (or if the reaction involves transfer of a H atom choose the transfer between the lowest numbered atoms).

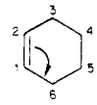
(iii) Consider the two atoms concerned from the viewpoint of the Cahn–Ingold–Prelog sequence rule; assume arbitrarily that the atom lowest in priority is the one providing the electrons for the new bond



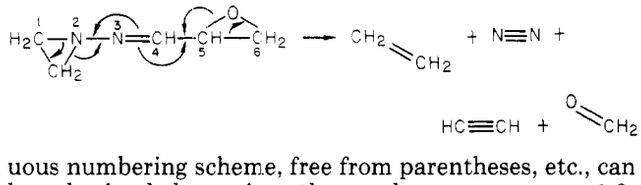
C1 of ethene has lower priority than the C1 of propene

(iv) Write this bond formation down first and follow in sequence round the cyclic transition state; i.e., start with the bond formation shown by the head of the arrow.

In a reaction such as the retro-Diels–Alder a simple rule is to take the lowest numbered atom in the minimal structure (normally C1) and write the sequence from there assuming that this atom is the electron-rich partner in the bond changes which affect it, giving the representation C1paC6, C5deC6, C5paC4, C3deC4, C3paC2, C1pdeC2.



This rule also provides an unambiguous description of the Eschenmoser fragmentation reaction.[7] An unambiguous numbering scheme, free from parentheses, etc., can be obtained by using the replacement name 1,2-methylene-5,6-epoxy-2,3-diazahex-3-ene with the electron movements, assumed to be concerted, being described as C1paC(C1), N2deC(C1), N2paN3, C4pdeN3, C4paC5, C6deC5, C6paO, C5deO, C5paC4, N3deC4, N3paN2, C1deN2. Clearly, a major problem here is not the definition of the electron movement but the implied requirement that we should use our already established numbering systems to get a unique numeric identification for each atom which is involved in the reaction. The Wiswesser line notation[8] for structures, which is designed for computer storage, includes a potentially useful system of locants for ring compounds, but it does not provide locants to refer to individual atoms or groups in a chain. Other systems have been proposed,[8] but no generally accepted alternatives are available.



(6) Sequence of EPs. A mechanism is written as a sequence of EPs which are normally assumed to proceed from left to right as read. The only simple way to represent a complex network of reactions including different paths to the same intermediates or products is to affix a sequence number to each branching or rejoining point; all EPs starting at the same branching point are prefixed by the same number; EPs arriving at the same confluence point (other than a final product) are all suffixed by another number. If an EP or EPs are repeated (e.g., in a chain reaction), the repeated portion can be enclosed in brackets, i.e. [ ].

The sequence number can be included in the process header of an EP, the confluence number at its end (vide infra).

Optional Features of the Representation

An essential feature of the scheme is that a place is provided in the representation for the insertion of optional extra information, which may be used, if required, when a list of representations is searched for particular processes. The process header is the most appropriate point to include this information, so that it is not confused with the essential PC and EP details.

(1) Representation of the Stereochemistry. Each EP has characteristic stereochemical properties, which may be ignored for simplicity if it is sufficiently well-known (e.g., "S_N2" inversion), unimportant (E2 reactions to give ethene), indeterminate (equally likely capture of carbonium ions from either side), or unknown. The stereochemical properties may in some reactions be largely attributable to a single PC (e.g., the π-bond formation in the elimination), but this seems rather artificial and clearly is not true in all cases (e.g., suprafacial and antarafacial electrocyclic addition). Not all PCs in context need stereochemical specifications; if such specifications were included, it might force us to use elaborate orbital descriptions. If possible, therefore, the stereochemical information

(7) D. Felix, R. Muller, U. Horn, R. Joos, J. Schreiber, and E. Eschenmoser, *Helv. Chim. Acta*, **55**, 1276 (1972).

(8) W. J. Wiswesser, "A Line-Formula Chemical Notation", Thomas Y. Cowell Co., New York, 1954; R. G. Dromey, *J. Chem. Inf. Comput. Sci.*, **18**, 255 (1978).

on the elementary processes themselves should be carried in the process header in the form of a symbol or symbols which are similar to those conventionally used. Like the list of PCs, this list is essentially open-ended to allow for new stereochemical insights. Designations which come to mind are:

inversion of a potentially chiral or similar center	inv:
retention of configuration at a similar center	ret:
randomization of configuration at a similar center	ran:
cis addition/elimination to form a double bond	cis:
trans addition/elimination to form a double bond	trans:
suprafacial (needs specifying for both reagent <1> and <2>)	sup:
antarafacial	ant:
disrotatory (with groups of highest priority at opposite ends approaching each other)	disz:
conrotatory (with groups of highest priority at opposite ends approaching each other	conz:
disrotatory (with groups of highest priority receding)	dise:
conrotatory (with groups of highest priority receding)	cone:
attack from less sterically hindered side	open:
attack from more sterically hindered side	hind:
attack from either side	aequo:

Anomeric effects and the choice of stereoelectronically perferred protons in enolization can probably be included with cis and trans above.[9]

(2) Additional Kinetic Information. Designation of the rate-limiting EP, or jointly rate-limiting EPs, of a reaction sequence is an optional feature, but it is clearly useful when indexing. In addition, work on a reaction normally provides the most information about the rate-limiting EP, and some hypothesis about which EP is rate limiting must be part of our model of a reaction. We choose the symbol ‡ to indicate the rate-limiting EPs under the conditions used for the study of the reaction, bearing in mind that a change of reaction conditions may result in a different EP becoming rate limiting.[10-12]

(3) Other Information. It may be useful to include other information at this point if it is potentially useful for distinguishing between otherwise similarly represented processes. Thus if an EP is dependent for its existence on being in a certain phase or a specified solvent or at a phase boundary or within a crystal, then it might be desirable to indicate it here. Alternatively, the solvent might with advantage be specified in the list of the reduced structures of reagents. One possibility is to include a code word, e.g., "epi" to indicate a reaction at a phase boundary. Similarly specific photoexcitation should be included as the appropriate PC, but if the mechanism of excitation is not known, the photochemically induced bond changes could be preceded by the word "photo" in the process header. Similarly "radio" could be used for excitation by ionizing radiation. A percentage could be included to show the partition of material between competing EPs.

Some Further Considerations

(1) Verification of the Accuracy of a Representation. There is a measure of redundancy in the representation in that in a sequence of PCs each atom is likely to occur twice, both times on the same side of the PC symbol, once with and once without the "de" flag, in successive PC terms. This provides a quick check for accuracy of electron counting etc. The only exceptions to this are where an atom changes its coordination number or oxidation level (e.g., carbene reactions, pinacol + HIO_4 etc.; see examples in Table I). It is also clear as a consequence of the rules recommended for writing PCs and EPs that Br^+, H^+, carbonium ions, and other species which enter or leave without the electron pair must always appear on the right-hand side of a PC, while halide ions, carbanions, H^-, etc. must always appear on the left-hand side. Also in an EP any attacking nucleophiles will enter in the first (left hand) bond-forming PC and break away in the last bond-breaking PC whereas electrophiles will enter in the right-hand PC, and leave on the left. This is a consequence of the rule that electrons flow from left to right and that the left-hand partner in a heterolytic PC carries the electron pair. Another check is that for each "curly arrow" in a conventional diagram there must be two PCs, one (dissociative) for its tail and one (bond forming) for its head. If additional information were packed into the PC, e.g., by writing it in the reverse order if the stereochemistry were abnormal,[2] the redundancy might be eliminated, but the ease of verification and interpretation would be lost.

(2) Indexing and File Searching. The obvious unit by which a reaction mechanism should be indexed is the EP, and, in general, it will be that of the rate-limiting process which will be most important. A file search by a suitably sophisticated computer program could, on request, either look for or ignore the rate-limiting-step symbol and so find all the reactions in a list which include a requested sequence of concerted PCs. Sequence numbers and other header information could be used or ignored as desired, and the only complexity would be that mentioned above in the section on the reagent numbers. PCs not involving bond changes could also be ignored on request. An alternative or additional criterion would be the name or formula of the minimal structure, since we may only be interested in the reaction of particular substrates. If the file search program is set to disregard even the nature of the atoms involved and whether the bonds involved are

(9) A referee points out that the topological equivalence of trans, inv, ant, and con means that one term (? ant) will suffice, modified by z or e, together with its converse (supz, supe). I concur but do not feel that this is the place to attempt to systematize the whole of stereochemical nomenclature.

(10) Arguments have been adduced[11] that kinetic information such as the specification of the rate-limiting step is irrelevant, since the representation only describes the fate of one molecule and that it can only be defined meaningfully if a reaction is carried out under standard state conditions. The first argument could also be used against any attempt to represent chain or competitive reactions and would thereby exclude many processes for which a representation could be useful. The second objection is a real one but it results from the common, rather lax use of the phrases "rate limiting" and "rate determining" which tends to imply that they are synonymous. A brief definition of the rate-limiting EP can be achieved if we use the concept of chemical flux (quantity of material being converted from reagent to product per unit volume per unit time). The rate-limiting EP is then the earliest EP in which the forward flux is least in excess of the overall rate of reaction under the prevailing conditions.[12] If the forward and backward fluxes from a particular intermediate are of the same order of magnitude, then the EP forming that intermediate, and the one by which it proceeds toward the product, can be described as jointly rate limiting. The definition of "jointly rate limiting", like the earlier one of "reversible", is by intention operational rather than absolute.

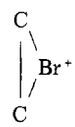
(11) R. D. Guthrie, Report for IUPAC Commission III.2, 1978.

(12) V. Gold, "A Glossary of Terms for Use in Physical Organic Chemistry" (IUPAC Commission III.2) in press.

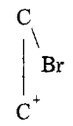
σ or π, it should separate mechanisms into categories like those used in Roberts' classification. It should therefore be possible to sort a series of mechanisms into a table in which isoelectronic processes are listed together and in which all EPs with the same number and types of bond-changing PCs are together. All that is necessary to decide the final order in the table is to specify a sequence of priorities of the PC operations which preferably makes some chemical sense. The list of the main PC types given above is already arranged in a possible sequence.

(3) Abbreviations. The full representation of any reaction, including precursor association, ion pairs, stereochemistry, etc., may be very lengthy, but it should be clear that much of this material can be omitted without violating the structure of the representation and without preventing the use of this type of system in indexing. This is a particularly valuable feature since it means that any details which might be guessed at, but for which there is little experimental evidence, can be safely omitted. There are, however, a number of reactions where the representation in terms of bond shifts becomes rather unwieldly. The first example is the ester hydrolysis, e.g., $A_{Ac}2$ (Table I). Here there undoubtedly are a large number of separately identifiable elementary processes, and the only way of simplifying the processes themselves is to omit the internal rearrangement of electrons (i.e., in the conversion of $HO^- + >C{=}O$ to $>C(OH)O^-$, the breakage of the π bond), but this is to make the representation a travesty since the product the EP generates would have pentavalent C. An alternative might be to define a new PC to indicate bond formation to one end of a π-bond system (together with the converse dissociative PC).

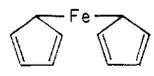
A closely related problem occurs if a reagent or the product has a delocalized structure. As we are describing the reaction in essentially valence-bond terms, we might wish to start an EP afresh with a particular canonical form, and, of course, the electron movements which give rise to this form can be detailed. Similarly the electron excursions to give other canonical forms could be included, but this is pointless unless the ultimate product structure is closely related to one of the other canonical forms. However, it may be useful to indicate that the product of an EP is not solely represented by the single structure the EP description generates—here a suitably noncommittal printed sign such as \$ could be inserted in the process header. It should be noted that inclusion of \$ enables us to write only one canonical form in such reactions as the bromination of a C–C double bond, where the bromonium ion can be represented by the form



without specifying alternative structures such as



or the ferrocene molecule where the canonical form



can do duty for the sandwich complex.

The most attractive area for abbreviations is that of electrocyclic, sigmatropic, and other pericyclic reactions. If the chemist appreciates the number of electrons involved and the σ bonds formed or borken, he can usually work out satisfactorily the necessary bond shifts. It is easy to think of simple abbreviations for frequently met reactions, but a study of Hendrickson's summary of the possible variety of reaction results gives rise to some doubts whether a simple general scheme could be devised.[13] It would, however, be easy to introduce ad hoc abbreviations based on existing nomenclature for a few frequently met processes.

Some Examples of the Use of the Representation

(1) Example of the Format of the EP. An elementary process can be written down as a sequence of characters which may be as short as four (the two atoms concerned and the PC by which they interact) or, if all the options are included and there are several PCs, quite lengthy. The following list gives examples of the possibilities that have been mentioned in sequence with comments.[14]

Symbol	Comment
#,	include if a rate-limiting process (separate by commas)
rev,	include if reversible under experimental conditions
epi,	type of reaction symbol (phase, etc.)
trans,	stereochemical information
20%,	proportion of material taking this route, if there are alternatives and if it is useful to specify it
3	sequence number if needed
\$	include if the product has a delocalized structure
:	termination of header (omit if no header details)

The above form the *process header* and are all optional.

C	element symbol (atom bearing electron pair or radical)
5	position of atom in reduced structure (optional) (or, in parentheses, symbol and number of atom to which it is attached)
<1>	molecule <1> is involved here; not necessary for many unimolecular processes
as	nature of primitive change
O(C1)<2>	second atom involved (details optional)
,	PC terminator unless this is the end of the process
C1pdeC2	further PC if needed
:7	confluence sequence number if needed
;	end of process (if end of mechanism, use full stop followed by a new line)

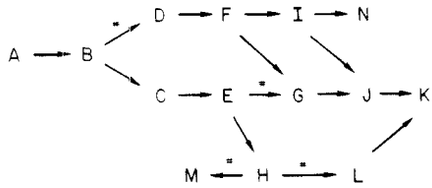
New lines should be inserted only after PC or EP terminator marks, spaces are optional between any of the above symbols.

(2) Example of the Combination of EPs in an Overall Reaction Scheme. A simple sequence of such elementary processes can be listed in order, each being marked off from the next by a semicolon. If a more complicated sequence is involved, sequence numbers can be used as follows: the chain which most rapidly reaches the rate-limiting step is written down first. For example, for

(13) J. B. Hendrickson, *Angew. Chem., Int. Ed. Engl.*, 13, 47 (1974).

(14) It should be noted that the first five or even six symbols reflect experimental details to varying degrees, and hence these could be considered strictly irrelevant to the representation of a model of a reaction mechanism. This is, of course, why they are included in the process header, rather than in the process representation proper, and why they are optional. There is a close parallel here with the insertion of comments into a computer program which are ignored by the compiler but which are extremely valuable to the programmer or to others looking at the program.

the hypothetical set of reactions



where the rate-limiting step for each path is marked #. Using X → Y as the representation of a step which converts X to Y

A → B; #,1:B → D; D → F; 2:F → I; 3:I → N;
1:B → C; C → E; #,4:E → G:5;
G → J:6; J → K;
4:E → H; #,7:H → L; L → K;
#,7:H → M;
2:F → G:5;
3:I → J:6;

The numbering of intermediates at the confluence of paths (G,J) is not strictly necessary but helps to indicate that products arrived at by both paths are the same and that they are not the final products.

(3) Examples of the Application of the Notation to Known Reactions (See Table I). The applicability of this scheme has been illustrated by notating a number of reaction mechanisms with the suggested rules. The first section consists of some of the commoner "Ingold-type" mnemonics for particular processes. A simple one-to-one translation of the Ingold-type symbols is not possible, and situations occur where either several symbols refer to essentially identical mechanisms or one symbol encompasses a variety of situations where there are differences in the mechanism or, at least, the rate-limiting step. A somewhat random selection of other reactions follows. Reference numbers prefaced by M refer to the relevant pages in ref 15. The general symbols used are X or Y for leaving groups or reagents of nucleophilic type, M or N for electrophilic types (including H^+), and J for a substituent subject to radical displacement; the molecule A–B is assumed capable of giving rise to radicals A and B. The minimal structures of reagents are listed before the representation of these mechanisms, with H replacing any unimportant groups. Optional information is included in some examples.

Acknowledgment. I thank Professor D. P. N. Satchell, King's College, London, and colleagues at Bristol University for helpful discussions of some of the ideas in this paper and various members of the IUPAC Commission in Physical Organic Chemistry (III.2) for encouraging me to develop them. I also thank the Royal Society for a Wates-IUPAC Bursary which has enabled me to meet Commission members for discussions.

(15) J. March, "Advanced Organic Chemistry", 2nd ed., McGraw-Hill, New York, 1977.
(16) W. N. Washburn and R. Zahler, *J. Am. Chem. Soc.*, **98**, 7827, 7828 (1976).

Photostimulated $S_{RN}1$ Reactions of Phenyl Selenide and Phenyl Telluride Ions with Halo- and Dihaloarenes in Liquid Ammonia[1]

Adriana B. Pierini[2] and Roberto A. Rossi*

Departamento de Química Orgánica, Facultad de Ciencias Químicas, Universidad Nacional de Córdoba, Est. 32, 5000 Córdoba, Argentina

Received April 30, 1979

The photostimulated reactions of haloarenes with phenyl selenide **2** and phenyl telluride **6** ions were studied in liquid ammonia. In the reactions of **2** with chloro- and bromobenzene, the yields were low, but with iodobenzene and halogen derivatives of naphthalene, phenanthrene, and quinoline, the yields of substitution products were good. With **6** and iodobenzene, a very good yield of diphenyl telluride was obtained, but with 1-chloro- or 1-bromonaphthalene, besides the unsymmetrical 1-naphthyl phenyl telluride, the symmetrical diphenyl telluride and bis(1-naphthyl) telluride were formed. In the photostimulated reaction of **6** and *p*-iodoanisole, both the unsymmetrical and the symmetrical substitution products were obtained. A reversible coupling of aryl radicals with aryl telluride ion is suggested in terms of the $S_{RN}1$ mechanism. There is no dark reaction with either nucleophile. **2** and **6** also react with dihaloarenes under photostimulation to give disubstitution products in fair to good yields.

Unactivated aromatic substrates bearing suitable leaving groups react under photostimulation with several nucleophiles by the $S_{RN}1$ mechanism.[3] The steps comprising a typical photostimulated $S_{RN}1$ reaction are outlined in Scheme I.

Photons probably stimulate electron transfer from the nucleophile to the substrate, forming a radical anion and a residue (step 1). This radical anion decomposes into an aryl radical and the leaving group (step 2). The radical

Scheme I

$$ArX + Nu^- \xrightarrow{h\nu} (ArX)^{\cdot -} + \text{residue} \quad (1)$$

$$(ArX)^{\cdot -} \longrightarrow Ar\cdot + X^- \quad (2)$$

$$Ar\cdot + Nu^- \longrightarrow (ArNu)^{\cdot -} \quad (3)$$

$$(ArNu)^{\cdot -} + ArX \longrightarrow ArNu + (ArX)^{\cdot -} \quad (4)$$

then reacts with the nucleophile to give a new radical anion (step 3). This radical anion can transfer its extra electron to the substrate (step 4), but in some systems it can decompose in other ways, depending on the identities of the aromatic moiety and the nucleophile.[3] Steps 2–4 are the propagation steps of a chain mechanism.

Phenoxide ion does not react under photostimulation by this mechanism in liquid ammonia, although it has been

(1) Research supported in part by the Consejo Nacional de Investigaciones Cientificas y Tēcnicas, Argentina.
(2) Based in part on the Ph.D. Thesis of Adriana B. Pierini, Universidad Nacional de Córdoba, Argentina, April 1979.
(3) (a) J. F. Bunnett, *Acc. Chem. Res.*, **11**, 413 (1978); (b) J. F. Wolfe and D. R. Carver, *Org. Prep. Proced.*, **10**, 225 (1978).

0022-3263/79/1944-4667$01.00/0