



are referred to as *bridges of rearrangement* (BR). For example, entry 6 or 7 has a BR (1-2-6-1), which corresponds to the Wagner-Meerwein rearrangement. The Beckmann (entry 21), Pinacol (entry 20), and Claisen rearrangements (entry 22) are characterized by BR's 1-6-7-1 of 21, 2-3-6-2 of 20, and 1-2-3-4-5-6-1 of 22.

Various ITS rings may appear simultaneously in a given ITS. ITS 6 or 7 has a BO1 (1-2-3-4-5-1), a BC1 (1-6-2-3-4-5-1), and a BR (1-2-6-1). This abstraction of ITS rings documents well the fact that entry 6 or 7 is a complex reaction involving ring opening of a tetrahydrofuran, closure to a dihydropyran, and the total Wagner-Meerwein rearrangement. The reaction of entry 21 consists of ring opening of cyclohexane, ring closure to ϵ -caprolactam, and the total Beckmann rearrangement. These reaction features are characterized as BO1 (1-2-3-4-5-6-1), BC1 (1-2-3-4-5-6-7-1), and BR (1-6-7-1), which appear in ITS 21. ITS 24 of the Eschenmoser epoxy-hydrazone fragmentation has BO1 (1-2-3-4-5-6-1), BO1 (2-7-3-4-5-6-1-2), BO2 (2-3-7-2), and BO2 (9-11-10-9).

CONCLUSION

I have introduced the concept of imaginary transition structures (ITS) for description of organic reactions. The ITS is a complete representation of the structural information of a given reaction, from which several important features of the reaction can be abstracted effectively. Since the ITS is a kind of structural formula, manipulation techniques developed for usual chemical structures, e.g., ring perception technique, are also applicable to the present ITS. The treasury of related

concepts introduced in this paper is as follows: abbreviated ITS's; out-bonds; in-bonds; par-bonds; bond colors; ITS bonds; complex bond numbers (*a b*); connection tables of ITS's; imaginary valency; projection to starting stage and to product stage; graphs of reaction centers; reaction strings (RS); connection tables of RS; stringity; one-string, two-string, and multistring reactions; entanglement of RS; primary reactions of a multistring reaction; ITS rings; bridges of ring opening of order *m* (BOM); bridges of ring closure of order *m* (BC_{*m*}); and bridges of rearrangement (BR).

REFERENCES AND NOTES

- (1) Ash, J. E.; Chubb, P. A.; Ward, S. E.; Welford, S. M.; Willett, P. *Communication, Storage, and Retrieval of Chemical Information*; Ellis Horwood: Chichester, UK, 1985; Chapter 8.
- (2) Sello, G. *J. Chem. Inf. Comput. Sci.* **1984**, *24*, 249.
- (3) (a) Corey, E. J.; Wipke, W. T. *Science (Washington, D.C.)* **1969**, *166*, 178. (b) Corey, E. J.; Long, A. K.; Rubenstein, S. D. *Science (Washington, D.C.)* **1985**, *228*, 408.
- (4) Gasteiger, J.; Jachum, C. *Top. Curr. Chem.* **1978**, *74*, 93.
- (5) Presented in part at the 52nd Annual Meeting of the Chemical Society of Japan, Kyoto, April 1986, and at the 192nd National Meeting of the American Chemical Society, Anaheim, CA, September 1986.
- (6) For ionic expressions in reaction diagrams, see part 3 of this series.
- (7) RC graphs of level 0, 1, 2, etc. will be defined elsewhere.
- (8) An equivalent set of reaction strings, 10+7-1+9-12 and 8-7+6-1+9-11, can be extracted since hydrogens (nodes 11 and 12) are equivalent to one another. But, for the purpose of retrieval of organic reactions, it is sufficient to select either of them in the present case. In the case that two terminal atoms are different from one another, the two sets of reaction strings are unequivalent and designate different reactions. A simple graph theoretical approach is insufficient to settle the problem of which we should select from the two, and hence an algorithm considering other chemical factors in addition to graph theoretical ones will be necessary to develop in such a case.
- (9) Arens, J. F. *Recl. Trav. Chim. Pays-Bas* **1979**, *98*, 155, 395, 471.
- (10) (a) Ugi, I.; Bauer, J.; Brandt, J.; Friedrich, J.; Gasteiger, J.; Jochum, C.; Schubert, W. *Angew. Chem., Int. Ed. Engl.* **1979**, *18*, 111. (b) Brandt, J.; von Sholley, A. *Comput. Chem.* **1983**, *7*, 51.
- (11) The recombination of two or more reaction strings will be discussed for the purpose of describing synthetic pathways. See part 5 of this series.

Description of Organic Reactions Based on Imaginary Transition Structures. 2. Classification of One-String Reactions Having an Even-Membered Cyclic Reaction Graph

SHINSAKU FUJITA

Research Laboratories, Ashigara, Fuji Photo Film Co., Ltd., Minami-Ashigara, Kanagawa, Japan 250-01

Received May 27, 1986



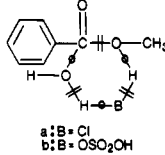


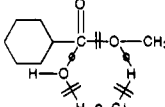
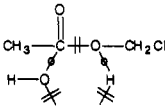
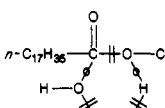


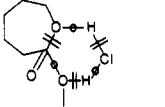



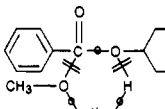


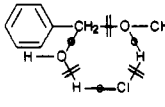
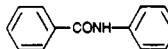


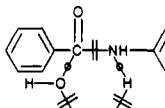
The concept of reaction graphs is derived from that of imaginary transition structures and is found to be useful in the classification of organic reactions. One-string reactions having a hexagonal, a tetragonal, or an octagonal reaction graph are classified and enumerated.

The classification of individual reactions into reaction types is important in the construction of a computer system for retrieval or design of organic reactions. Several critical reviews^{1,2} have discussed many methods to classify organic reactions, but some recent attempts in this field should be mentioned here. In Ugi's comprehensive system,³ reactants and products are denoted in the form of BE (bond/electron) matrices, and a reaction is represented by an operator (*R*) matrix, which is the difference between the product BE matrix and the reactant BE matrix. The *R* matrix describes only a net change in each bond, which is oversimplified from the viewpoint of retrieval of organic reactions. Thus, the Diels-

Alder reaction, the Cope rearrangement, the Claisen rearrangement, the 1,4-addition of Grignard reagents to α,β -unsaturated carbonyl compounds, and some other reactions all have the same *R* matrix⁴ and cannot be discriminated from each other, unless one refers to the reactants and/or the products.

Arens' coding system⁵ is based upon numerical treatment and describes reactions by operators (\rightarrow), etc., which are also oversimplified so as to be unable to distinguish the above-described reactions. Roberts⁶ has reported a graphic approach in which he treated organic reactions in terms of CP (concerted process) skeletons. This method has abstracted essential

Table I. ITS's of Ester Hydrolyses and of the Related Reactions

entry	starting stage	product stage	ITS
1	 + H ₂ O + HB	 + CH ₃ OH + HB	 a:B=Cl b:B=OSO ₂ OH
2	 + H ₂ O + HCl	 + CH ₃ OH + HCl	
3	CH ₃ COOC ₂ H ₅ + H ₂ O + HCl	CH ₃ COOH + C ₂ H ₅ OH + HCl	
4	<i>n</i> -C ₁₇ H ₃₅ COOCH ₃ + H ₂ O + HCl	<i>n</i> -C ₁₇ H ₃₅ COOH + CH ₃ OH + HCl	
5	 + H ₂ O + HCl	 + HCl	
6	 +  + HCl	 + CH ₃ OH + HCl	
7	 + H ₂ O + HCl	 + CH ₃ OH + HCl	
8	 + H ₂ O + HCl	 +  + HCl	

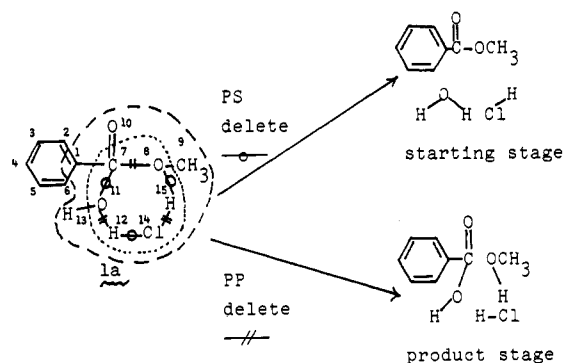
features of organic reactions. However, the description of a reaction is displaced by parallel reference to the starting and product stages.

Hendrickson's approach⁷ to systematize and enumerate thermal pericyclic reactions has been able to classify them into several broad isomorphic classes. His method focused on invariant backbones of reacting species and opened broad prospects to classify organic reactions. However, the representation is not unitary and is essentially parallel reference to the starting and product stages.

In the preceding paper,⁸ I have proposed imaginary transition structures (ITS's), a unitary representation of organic reactions. In this paper, I will report that a graph of reaction centers, especially, a reaction graph, which appears in the ITS as a substructure and contains several reaction strings, can be a descriptor of reaction types.

IMAGINARY TRANSITION STRUCTURES, GRAPHS OF REACTION CENTERS OF VARIOUS LEVELS, REACTION GRAPHS, AND REACTION STRINGS

An imaginary transition structure (ITS) of a given reaction is defined as a structural formula in which all nodes appearing both in the starting and in the product stage are connected



with out-, in-, and/or par-bonds according to the change of bonds during the reaction. For example, hydrolysis of methyl benzoate catalyzed by hydrochloric acid (entry 1 of Table I) is represented by the corresponding ITS 1a, wherein the out-bonds are bonds appearing only in the starting stage and are denoted by the symbol --- , the in-bonds (---) are those appearing only in the product stage, and the par-bonds (---) are invariant bonds during the reaction. Since the structural information of individual reactions is fully documented by their ITS's, various reaction features can be ab-

Table II. Hierarchy of RC Graphs of Various Levels

basic graph	RC graph of level 0 (reaction graph)	RC graph of level 1	RC graph of level 2	examples of ITS
				1a, 2, 3, 4, 5
				6
				7
				1b
				8
				9, 10, 11, 12, 13

stracted from the ITS's by appropriate algorithms. *Projection to the starting stage (PS)* is defined as the deletion of in-bonds ($\text{---}\bigcirc\text{---}$), which produces the starting stage from the ITS. *Projection to the product stage (PP)* is the deletion of out-bonds ($\text{---}\bigcirc\text{---}$) from the ITS, resulting in the formation of the product stage.

Reaction centers of an ITS are nodes that are incident to in- and/or out-bonds. In ITS 1a, nodes 7, 8, 11, 12, 14, and 15 are the reaction centers. A figure of the combined reaction centers with in-, out-, and/or par-bonds is defined as a *graph of reaction centers* (an RC graph). I define an *RC graph of level 1* as an RC graph containing only reaction centers, e.g., a subgraph of 1a encircled by a dotted line. Similarly, an *RC graph of level 2* is one with reaction centers and next-neighbor nodes, e.g., a subgraph of 1a encircled by a broken line. A *reaction graph* (an RC graph of level 0) is defined as a graph in which every node is regarded as a ball in an abstract fashion and combined by in-, out-, and/or par-bonds in accordance with the corresponding ITS. The reaction graph of ITS 1a is a hexagon, the edges of which are out- and in-bonds alternately.

An RC graph of each level consists of strings possessing alternate out- and in-bonds. For example, in the RC graph of level 1 of ITS 1a, a hexagon (7-8+15-14+12-11+7) is such a string. I define such a string as a *reaction string*. When the number of reaction strings in an RC graph is referred to as stringity, all reactions are classified into one-, two-, and three-string and multistring reactions by their stringity. In the case of one-string reactions, their RC graphs (and reaction graphs) are identical with the reaction strings.

An *ITS bond* is an extended bond that consists of out-, in-, and/or par-bonds and connects each pair of adjacent nodes

Table III. ITS's of the Diels-Alder Reactions

entry	reaction	ITS
9		
10		
11		
12		
13		

of a given ITS. The ITS bond is represented by a complex bond number ($a\ b$), where a is the bond multiplicity of the starting stage and b is the difference in bond multiplicity between the product and the starting stages. By means of complex bond numbers, the above-described reaction string can be coded as $7C(1-1)8C(0+1)15H(1-1)14Cl(0+1)12H(1-1)11O(0+1)7C$ in the level of the RC graph of level 1. The reaction string(s) can be stored and retrieved in the form of

Table IV. Reaction Graphs of Hexagonal Class ($m = 0, n = 0-6$)^a

m	n	no. of reacn graphs ^b	no. of reacn pairs ^c	reaction graphs ^d
0	0	1	1	
0	1	2	1	
0	2	4	3	
0	3	6	3	
0	4	4	3	
0	5	2	1	
0	6	1	1	

^aTables IV-VII may contain several reaction graphs that are not reasonable chemically. But they are not omitted for completeness of the list.

^bThe numbers of reaction graphs are the coefficients of $x^m y^n$ in $G(x, y)$ (eq 3). ^cThe numbers of reaction pairs are the coefficients of $x^m y^n$ in $P(x, y)$ (eq 5). ^dEach couple of reaction graphs linked with a bracket is a reaction pair. A reaction graph without a bracket is a self-reaction pair.

such letter string(s) as above or of a connection table having the additional information of neighboring atoms.

HIERARCHY OF RC GRAPHS OF VARIOUS LEVELS

In general, each RC graph describes a reaction type, whereas an ITS is a description of an individual reaction. Let us examine the ester hydrolyses (entries 1-5) collected in Table I. The corresponding ITS's (1a-5, except 1b) involve a common substructure in level 1 (A_1) or level 2 (A_{11}) of the RC graphs (Table II). The RC graph of level 1 (A_1) describes the essential features of ester hydrolyses, but not completely: an ester exchange (entry 6) and a benzyl ether cleavage (entry 7) have the same RC graph of level 1 (A_1). At the representation of level 2, the three types of reactions are distinguished from each other. Compare A_{11} , A_{12} , and A_{13} . It should be noted that ITS 1b gives another RC graph of level 1 (A_2), which corresponds to the difference in catalysts. A common reaction graph (A) is abstracted as a superior concept from ITS (1a-7). The ITS of HCl-catalyzed hydrolysis of amides (entry 8) contains the same reaction graph (A) but can be discriminated at level 1 representation from the above-described reactions.

Another set of examples is Diels-Alder reactions (Table III), which give the corresponding ITS's (9-13) for individual reactions and an RC graph (B_1) of level 1 for the reaction type. The reaction graph of the Diels-Alder reactions is a hexagonal graph (B) modified by four single par-bonds.

Hierarchies of RC graphs of various levels such as shown in Table II are useful in the classification of organic reactions. In particular, reaction graphs (RC graphs of level 0) such as A and B are versatile tools for the generic description of reaction types. In the following sections, various types of reaction graphs will be enumerated and discussed.

It should be emphasized that an RC graph representing a reaction type is a substructure (subgraph) of the corresponding ITS representing an individual reaction. This implies that reaction type data are produced automatically rather than manually in the form of RC graphs from individual reaction data, which are stored as a connection table of the ITS.

Further common subgraphs, $H-O-C-O-H$ or $O-C-O$, can be extracted from A_1 and A_2 (for the reactions 1a, 1b, 2, 3, 4, 5, 6, and 7). They are also good descriptors for organic reactions. This type of subgraph will be discussed in part 4 of this series.

ENUMERATION OF HEXAGONAL REACTION GRAPHS

Manipulation of reaction data in the form of ITS's provides various merits, which can be ascribed to the fact that ITS's are regarded as extended structural formulas having three-colored bonds. One of these merits is that enumeration of reaction types is transformed to counting reaction graphs. Reaction graphs of the hexagonal class can be considered to be hexagons that have alternate in-bonds and out-bonds on their edges and said in- and out-bonds may be modified by a number m of double par-bonds (\equiv) and a number n of single par-bonds ($-$). For enumeration of hexagonal reaction graphs, I can formulate this problem as counting isomers of a basic reaction graph (A) with m double par-bonds and n

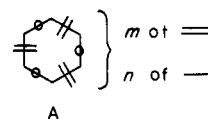


Table V. Reaction Graphs of Hexagonal Class ($m = 1, n = 0-5$)^a

m	n	no. of reacn graphs	no. of reacn pairs	reaction graphs
1	0	2	1	
1	1	6	3	
1	2	12	6	
1	3	12	6	
1	4	6	3	
1	5	2	1	

^aSee footnotes of Table IV.

single par-bonds on the edges.⁹ This problem of counting isomers can be solved by straightforward use of Polya's theorem.¹⁰

The basic reaction graph A is superposed on itself by six operations of a permutation group denoted by D_3 . The cycle index $Z(D_3)$ is represented by eq 1. Then, according to Polya's

$$Z(D_3) = (1/6)(s_1^6 + 3s_1^2s_2^2 + 2s_3^2) \quad (1)$$

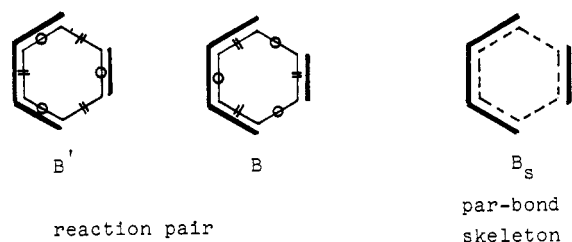
theorem, each variable s_k in cycle index Z is substituted by a figure-counting series:

$$s_k = 1 + x^k + y^k \quad (2)$$

The resulting polynomial series $G(x, y)$ for counting reaction graphs is the generating function, wherein the coefficient of $x^m y^n$ is the number of reaction graphs:

$$\begin{aligned} G(x, y) &= Z(D_3, 1 + x^k + y^k) \\ &= 1 + 2x + 2y + 4x^2 + 6xy + 4y^2 + \\ &\quad 6x^3 + 12x^2y + 12xy^2 + 6y^3 + 4x^4 + 12x^3y + \\ &\quad 18x^2y^2 + 12xy^3 + 4y^4 + 2x^5 + 6x^4y + \\ &\quad 12x^3y^2 + 12x^2y^3 + 6xy^4 + 2y^5 + x^6 + 2x^5y + \\ &\quad 4x^4y^2 + 6x^3y^3 + 4x^2y^4 + 2xy^5 + y^6 \end{aligned} \quad (3)$$

When all in-bonds and out-bonds of a reaction graph are exchanged with each other, the reaction graph of the corresponding reverse reaction is obtained. For example, the reaction graph B of the Diels-Alder reaction is converted by this operation into the graph (B') of the retro Diels-Alder reaction.



I call this operation a *transformation to the reverse reaction*. The resulting pair of reaction graphs, e.g., B and B', is referred to as a *reaction pair*. The two reaction graphs of a reaction pair have a common skeleton of par-bonds, e.g., B_s , for the pair of B and B'.

There are some special cases in which the two reaction graphs of a reaction pair are identical. These cases are called *self-reaction pairs*. Thus, the basic graph A is an example of a self-reaction pair. The reaction graphs of hexagonal class enumerated by eq 3 constitute reaction pairs (or self-reaction pairs). Enumeration of all reaction pairs of hexagonal class is the next problem.¹¹ This enumeration is ascribed to that of par-bond skeletons such as B_s , which is, in turn, equivalent to counting of benzene derivatives ($C_6H_{6-m-n}X_mY_n$).

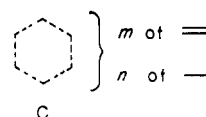


Table VI. Reaction Graphs of Hexagonal Class ($m = 2, n = 0-4$)^a

m	n	no. of reacn graphs	no. of reacn pairs	reaction graphs
2	0	4	3	
2	1	12	6	
2	2	18	11	
2	3	12	6	
2	4	4	3	

^a See footnotes of Table IV.

The edges of a regular hexagon are superimposable by 12 operations of a permutation group denoted by D_6 . For counting isomers of this skeleton having m double par-bonds and n single par-bonds, the cycle index $Z(D_6)$ and the corresponding polynomial series $P(x, y)$ for counting reaction pairs are shown as follows:

$$Z(D_6) = (1/12)(s_1^6 + 3s_1^2s_2^2 + 4s_2^3 + 2s_3^2 + 2s_6) \quad (4)$$

$$\begin{aligned}
 P(x, y) &= Z(D_6, 1 + x^k + y^k) \\
 &= 1 + x + y + 3x^2 + 3xy + 3y^2 + 3x^3 + \\
 &\quad 6x^2y + 6xy^2 + 3y^3 + 3x^4 + 6x^3y + 11x^2y^2 + \\
 &\quad 6xy^3 + 3y^4 + x^5 + 3x^4y + 6x^3y^2 + 6x^2y^3 + \\
 &\quad 3xy^4 + y^5 + x^6 + x^5y + 3x^4y^2 + 3x^3y^3 + \\
 &\quad 3x^2y^4 + xy^5 + y^6 \quad (5)
 \end{aligned}$$

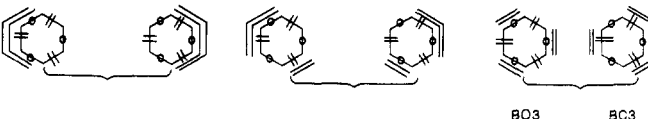
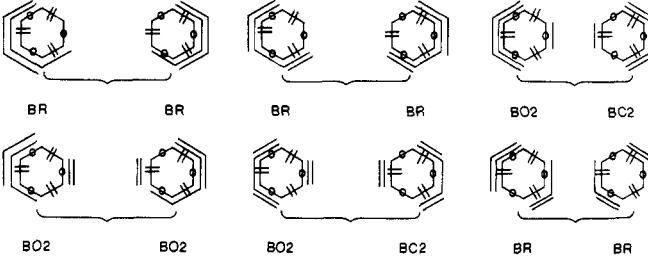
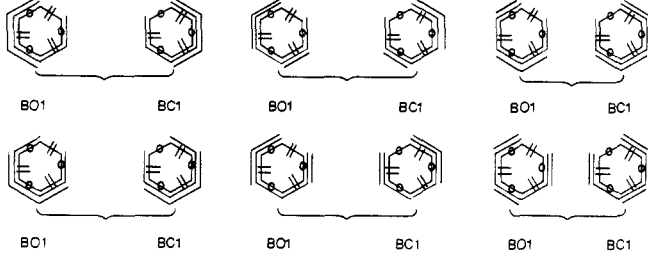

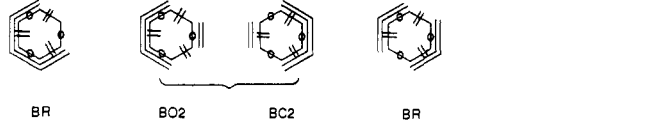
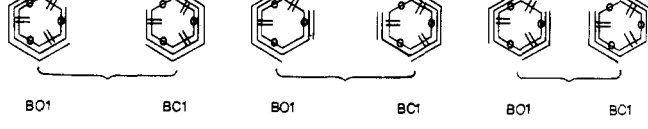

The coefficient of $x^m y^n$ in $P(x, y)$ (eq 5) is the number of reaction pairs with m double par-bonds and n single par-bonds.

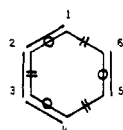
The complete lists of all reaction graphs ($m = 0-4$) of

hexagonal class are found in Tables IV-VII. For example, in the case of $m = 0$ and $n = 4$ (Table IV), there are four reaction graphs, two of which appear as a reaction pair corresponding to the Diels-Alder/retro Diels-Alder reaction pair. The remaining two reaction graphs construct self-reaction pairs individually. This result reflects the fact that the coefficients of $x^0 y^4$ in $G(x, y)$ and in $P(x, y)$ are 4 and 3, respectively.

A coded name of each hexagonal graph is desirable for assigning an individual reaction to a reaction type. The reaction graph can be regarded as an imaginary cyclic compound, each ITS bond of which is represented by a complex bond number ($a b$). If each ITS bond of ($a b$) is displaced by a bond of multiplicity $a + (b - |b|)/2$, said imaginary compound is converted to an enyne, the nodes of which can be numbered by the IUPAC rules. The latter numbering is also used for the parent reaction graph. In other words, the location of each double (i.e., $a + (b - |b|)/2 = 2$) or single par-bond (i.e., $a + (b - |b|)/2 = 1$) is specified by giving the lower number of the two nodes joined by each double or single par-bond. The nodes (vertices) are numbered to give double and single par-bonds the lowest possible set of locants. A couple of examples are shown as follows.

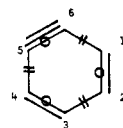
Table VII. Reaction Graphs of Hexagonal Class ($m = 3, n = 0-3; m = 4, n = 0-2$)^a

m	n	no. of reacn graphs	no. of reacn pairs	reaction graphs
3	0	6	3	
3	1	12	6	
3	2	12	6	
3	3	6	3	
4	0	4	3	
4	1	6	3	
4	2	4	3	

^a See footnotes of Table IV.

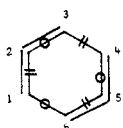
1(1+1)2(2-1)3(1+1)4(1-1)5(1+1)6(1-1)1
 $m = 0, n = 4$; BO2

When there is a choice in numbering, single par-bonds are given the lowest numbers.



1(1+1)2(1-1)3(1+1)4(1-1)5(2+1)6(1-1)1
 $m = 1, n = 2$; BO3

When a choice in numbering is concerned with the selection of in- or out-bonds, an out-bond is selected preferentially.



1(2-1)2(1+1)3(1-1)4(1+1)5(2-1)6(0+1)1
 $m = 0, n = 4$; BR

The present method of coding includes canonization of Ugi's *R* matrices¹² and Arens' coding system,⁵ since each code above-described has integers *b* of alternate signs. But it excels them in holding the data of par-bonds as well. Hendrickson's formulation clarifies the mode of par-bonds modifying the basic reaction graph but lacks such a unitary representation as proposed here. Since the present coding system is based on graphical representation and correlated with the nomenclature of hydrocarbons, it is more friendly to organic chemists than the previous systems are.

The present ITS approach has a further advantage over the systems previously proposed, since additional features of reactions are abstracted graphically. When, in the reaction string of a hexagonal reaction graph, *p* ITS bonds have complex bond numbers (*a b*) of $a + b = 0$ and all other ITS bonds have (*a b*) of $a + b \neq 0$ and $a \neq 0$, said hexagonal graph is a bridge of ring opening of order *p* (BO*p*). The presence of BO*p* corresponds to a ring-opening reaction in which *p* bonds are cleaved.

Table VIII. One-String Reactions Having Hexagonal Reaction Graphs

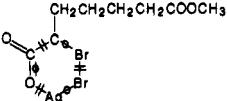

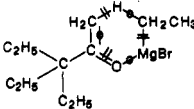

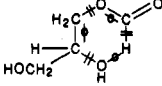

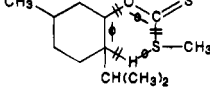

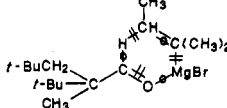

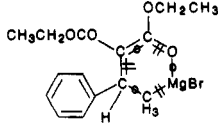



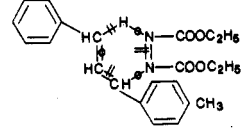

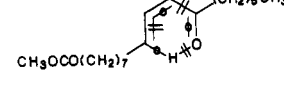

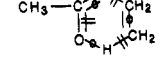

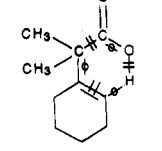

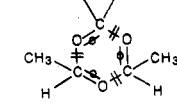

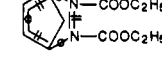

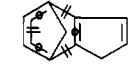

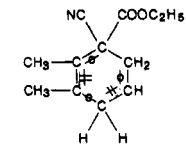

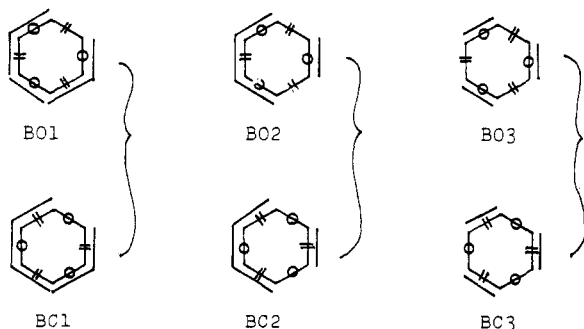
entry	reaction name	ITS	code for RC graph of level 1	reaction graph	ref
14	Hunsdieker-Borodin reaction		1O(1+1)2C(1-1)3C(0+1)4Br(1-1)5Br(0+1)6Ag(1-1)1O $m = 0, n = 1$		13
15	enolization of a ketone by a Grignard reagent		1O(2-1)2C(1+1)3C(1-1)4C(0+1)5C(1-1)6Mg(0+1)1O $m = 0, n = 2$		14
16	decomposition of glycerol formate		1C(1+1)2C(1-1)3O(1+1)4C(1-1)5H(0+1)6O(1-1)1C $m = 0, n = 2$		15
17	Tschugaeff reaction		1C(1+1)2C(1-1)3O(1+1)4C(1-1)5S(1-1)6H(1-1)1C $m = 0, n = 2$		16
18	reduction of a ketone by sterically hindered Grignard compound		1C(1+1)2C(1-1)3H(0+1)4C(2-1)5O(0+1)6Mg(1-1)1C $m = 0, n = 2$		17
19	1,4-addition of a Grignard reagent		1C(2-1)2C(1+1)3C(2-1)4O(0+1)5Mg(1-1)6C(0+1)1C $m = 0, n = 3$		18
20	1,4-addition of bromine to butadiene		1C(2-1)2C(1+1)3C(2-1)4C(0+1)5Br(1-1)6Br(0+1)1C $m = 0, n = 3$		19
21	ene reaction		1C(1+1)2C(2-1)3C(0+1)4N(2-1)5N(0+1)6H(1-1)1C $m = 0, n = 3$		20
22	thermal decomposition of methyl ricinolate		1C(2-1)2C(1+1)3C(1-1)4C(1+1)5O(1-1)6H(0+1)1C $m = 0, n = 3$		21
23	pyrolysis of ethyl acetate		1O(2-1)2C(1+1)3O(1-1)4C(1+1)5C(1+1)6H(0+1)1O $m = 0, n = 3$		22
24	decarboxylation of a β,γ -unsaturated carboxylic acid		1C(2-1)2C(1+1)3C(1-1)4C(1+1)5O(1-1)6H(0+1)1C $m = 0, n = 3$		23
25	decomposition of paraldehyde		1C(1+1)2O(1-1)3C(1+1)4O(1-1)5C(1+1)6O(1-1)1C $m = 0, n = 3; BO3$		24
26	Diels-Alder reaction		1C(2-1)2C(1+1)3C(2-1)4C(0+1)5N(2-1)6N(0+1)1C $m = 0, n = 4; BC2$		25
27	retro Diels-Alder reaction		1C(2-1)2C(1+1)3C(2-1)4C(0+1)5C(2-1)6C(0+1)1C $m = 0, n = 4; BO2$		26
28	Cope rearrangement		1C(2-1)2C(1+1)3C(1-1)4C(1+1)5C(2-1)6C(0+1)1C $m = 0, n = 4; BR$		27

Table VIII (Continued)

entry	reaction name	ITS	code for RC graph of level 1	reaction graph	ref
29	oxy-Cope rearrangement		1C(2-1)2C(1+1)3C(1-1)4C(1+1)5C(2-1)6C(0+1)1C $m = 0, n = 4$; BR		28
30	Claisen rearrangement		1C(2-1)2C(1+1)3C(1-1)4O(1+1)5C(2-1)6C(0+1)1C $m = 0, n = 4$; BR		29
31	photochemical formation of a ketone		1C(1+1)2C(2-1)3C(1+1)4C(2-1)5C(1+1)6C(1-1)1C $m = 0, n = 5$; BO1		30
32	ring closure of an isocyanate		1C(2-1)2C(1+1)3C(2-1)4C(1+1)5C(2-1)6O(0+1)1C $m = 0, n = 5$; BC1		31
33	thermolysis of an acetylenic ether		1C(1+1)2C(1-1)3O(1+1)4C(3-1)5C(0+1)6H(1-1)1C $m = 1, n = 2$		32
34	formation of benzonitrile		1O(2-1)2C(1+1)3O(1-1)4N(2+1)5C(1-1)6H(0+1)1O $m = 1, n = 2$		33
35	decomposition of an azo compound		1C(1+1)2C(1-1)3C(1+1)4C(1-1)5N(1+1)6N(1-1)1C $m = 1, n = 2$; BO3		34
36	Diels-Alder reaction		1C(2-1)2C(1+1)3C(2-1)4C(0+1)5C(3-1)6C(0+1)1C $m = 1, n = 3$; BC2		35
37	trimerization of an acetylene		1C(3-1)2C(0+1)3C(3-1)4C(0+1)5C(3-1)6C(0+1)1C $m = 3, n = 0$; BC3		36



For example, in the case of $m = 0, n = 4$, the following two graphs are characterized by BR.



It should be emphasized that perception of ring opening, ring closure, and rearrangement is transformed to examination of ring structures in ITS's (or in RC graphs of various levels).

EXAMPLES OF ONE-STRING REACTIONS HAVING A HEXAGONAL REACTION GRAPH

Many reactions as well as those described in the previous section have hexagonal reaction graphs of various types. Examples reported in Arens' article⁵ are reexamined here by the present method of description. Table VIII shows their ITS's, codes for RC graphs of level 1, and reaction graphs. Although further ambiguity in numbering is present in the codes for RC graphs of level 1, canonization is not attempted here. The numbers of double par-bonds (m) and of single par-bonds (n) are good descriptors to characterize the reaction graphs. These numbers are listed as well as the notations,

A bridge of ring closure of order p (BC p) is also defined as an ITS ring, wherein p ITS bonds have ($a\ b$) of $a = 0$ and all other ITS bonds have ($a\ b$) of $a + b \neq 0$ and $a \neq 0$. It implies a ring closure, in which p bonds are formed during the reaction. The reaction graphs characterized by BO p can be converted to BC p graphs by the above-defined transformation to reverse reactions. Examples cited above show this transformation. Tables IV-VII contain several couples of BO p and BC p as reaction pairs.

If the reaction string of hexagonal class has one ITS bond of $a + b = 0$, another ITS bond of $a = 0$, and other bonds all of $a + b \neq 0$ and $a \neq 0$, it is a bridge of rearrangement (BR).

Table IX. Reaction Graphs of the Tetragonal Class

m	n	no. of reacn graphs ^a	no. of reacn pairs ^a	reaction graphs
0	0	1	1	
0	1	2	1	
0	2	3	2	
0	3	2	1	
0	4	1	1	
1	0	2	1	
1	1	4	2	
1	2	4	2	
1	3	2	1	
2	0	3	2	
2	1	4	2	
2	2	3	2	
3	0	2	1	
3	1	2	1	
4	0	1	1	

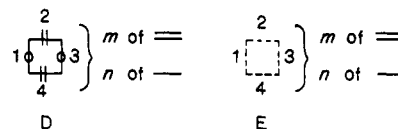
^aThe numbers of reaction graphs are the coefficients of $x^m y^n$ in $G(x, y)$ (eq 7). ^bThe numbers of reaction pairs are the coefficients of $x^m y^n$ in $P(x, y)$ (eq 9).

BOp, BCp, and BR. Since RC graphs of level 1 are subgraphs that are easily abstracted from ITS's, they are omitted from Table VIII. The starting and product stages can be reproduced by the operations PS and PP, respectively, applying to ITS's.

Since the reaction graphs are linked closely to the corresponding ITS as shown in Table VIII, the present ITS approach affords additional merits for the retrieval of organic reactions. In the above-described examination of reaction graphs, I have extracted BOp, BCp, and BR coinciding with said reaction graphs. When attention is expanded to ITS's, which are supergraphs of the reaction graphs, there are other types of BOp, BCp, and BR in the ITS's. For example, the ITS of entry 35 involves a three-membered ITS ring of BO1 and a six-membered ITS ring of BO2 in addition to the reaction graph of BO3. The ITS of entry 37 is another example which has a four-membered ITS ring of BC1.

ENUMERATION OF TETRAGONAL REACTION GRAPHS

In the case of the tetragonal class, the numbers of reaction graphs and of reaction pairs are obtained by counting isomers of a basic graph (D) and a tetragon (E), respectively, with m double par-bonds and n single par-bonds on the edges. When the edges of the basic reaction graph D are numbered as illustrated, graph D can be superposed on itself by permutations (1)(2)(3)(4), (13)(24), (1)(3)(24), and (2)(4)(13) of a



permutation group denoted by D_2 . According to Polya's theorem, the cycle index $Z(D_2)$ for D and the corresponding polynomial series $G(x, y)$ for counting reaction graphs are represented by the following equations:

$$Z(D_2) = (1/4)(s_1^4 + s_2^2 + 2s_1^2 s_2) \quad (6)$$

$$G(x, y) = Z(D_2, 1 + x^k + y^k) \\ = 1 + 2x + 2y + 3x^2 + 4xy + 3y^2 + 2x^3 + 4x^2y \\ + 4xy^2 + 2y^3 + x^4 + 2x^3y + 3x^2y^2 + 2xy^3 + y^4 \quad (7)$$

The coefficients of $x^m y^n$ are the numbers of reaction graphs that have m double par-bonds and n single par-bonds.

For the tetragon E, the cycle index $Z(D_4)$ and the reaction-pair-counting series $P(x, y)$ are obtained similarly:

$$Z(D_4) = (1/8)(s_1^4 + 3s_2^2 + 2s_1^2 s_2 + 2s_4) \quad (8)$$

$$P(x, y) = 1 + x + y + 2x^2 + 2xy + 2y^2 + x^3 + 2x^2y + 2xy^2 + y^3 + x^4 + x^3y + 2x^2y^2 + xy^3 + y^4 \quad (9)$$

Table X. One-String Reactions Having Tetragonal Graphs

entry	reaction name	ITS	code for RC graph of level 1	reaction graph	ref
38	substitution of hydroxyl by chlorine		1C(1-1)2O(0+1)3H(1-1)4Cl(0+1)1C $m = 0, n = 0$		37
39	Friedel-Crafts acylation		1C(0+1)2C(1-1)3Cl(0+1)4H(1-1)1C $m = 0, n = 0$		38
40	amide formation		1C(1-1)2Cl(0+1)3H(1-1)4N(0+1)1C $m = 0, n = 0$		39
41	addition of a Grignard reagent to acetone		1C(2-1)2O(0+1)3Mg(1-1)4C(0+1)1C $m = 0, n = 1$		40
42	dimerization of a fluorinated olefin		1C(2-1)2C(0+1)3C(2-1)4C(0+1)1C $m = 0, n = 2; BC2$		41
43	ring opening of a cyclobutene		1C(1+1)2C(2-1)3C(1+1)4C(1-1)1C $m = 0, n = 3; BO1$		42
44	ring enlargement		1C(1+1)2C(2-1)3C(1+1)4C(1-1)1C $m = 0, n = 3; BO1$		43
45	photochemical ring enlargement		1C(1+1)2C(1-1)3C(2+1)4C(1-1)1C $m = 1, n = 1; BO2$		44
46	decomposition of an azo compound		1N(2+1)2N(1-1)3C(0+1)4C(1-1)1N $m = 1, n = 0$		45

The numbers of reaction pairs are the coefficients of $x^m y^n$ in $P(x, y)$ (eq 9).

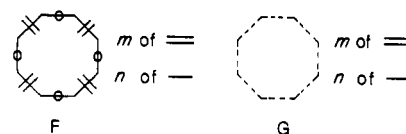
The complete list of reaction graphs of tetragonal class is found in Table IX. The tetragonal reaction graphs can be coded in the same line as described in the case of the hexagonal ones and are also characterized by the concept of ITS rings (i.e., BOp, BCp, and BR). Table IX may contain several unreasonable reaction graphs from the viewpoint of organic chemistry, but they are included to ensure completeness of the list.

EXAMPLES OF ONE-STRING REACTIONS HAVING A TETRAGONAL REACTION GRAPH

One-string reactions of tetragonal class are collected in Table X. Conversion of *tert*-butyl alcohol into *tert*-butyl chloride (entry 38) is represented by an ITS of tetragonal class. This implies that hydrogen chloride acts as a reagent rather than as a catalyst. The Friedel-Crafts acylation (entry 39) and amide formation (entry 40) have the same reaction graph, which is identical with the basic graph of tetragonal class. In the former case, aluminum chloride is a catalyst which constitutes an independent part of the ITS's. Addition of a Grignard reagent to acetone (entry 41) is represented by the ITS that contains a tetragonal reaction graph modified by a single par-bond, when workup with water is not taken into consideration. Entry 42 is an example of a cycloaddition, which provides a tetragonal reaction graph of BC2. The ring-opening reactions (entries 43 and 44) have the same tetragonal reaction graph modified by three single par-bonds, which is characterized as BC1. Entries 45 and 46 indicate modification of the tetragonal reaction graph with at least a double par-bond.

ENUMERATION OF OCTAGONAL REACTION GRAPHS

The methodology for the present enumeration is the same as described above. Basic graphs F and G can be superposed

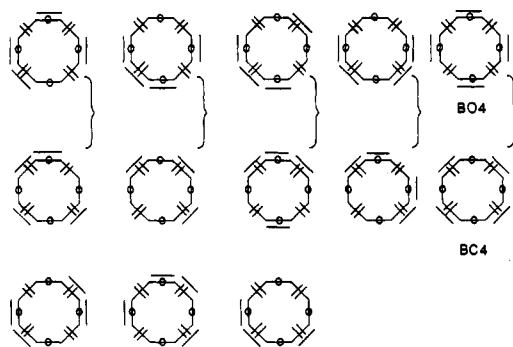


upon themselves by permutations belonging to D_4 and D_8 , respectively. For counting reaction graphs based on $F(D_4)$, the following cycle index and counting series are obtained:

$$Z(D_4) = (1/8)(s_1^8 + 4s_1^2s_2^3 + s_2^4 + 2s_4^2) \quad (10)$$

$$\begin{aligned}
 G(x, y) &= Z(D_4, 1 + x^k + y^k) \\
 &= 1 + 2x + 2y + 6x^2 + 8xy + 6y^2 + 10x^3 + \\
 &\quad 24x^2y + 24xy^2 + 10y^3 + 13x^4 + 38x^3y + \\
 &\quad 60x^2y^2 + 38xy^3 + 13y^4 + 10x^5 + 38x^4y + \\
 &\quad 76x^3y^2 + 76x^2y^3 + 38xy^4 + 10y^5 + \\
 &\quad 6x^6 + 24x^5y + 60x^4y^2 + 76x^3y^3 + 60x^2y^4 + \\
 &\quad 24xy^5 + 10y^6 + 2x^7 + 8x^6y + 24x^5y^2 + \\
 &\quad 38x^4y^3 + 38x^3y^4 + 24x^2y^5 + 8xy^6 + 2y^7 + x^8 + \\
 &\quad 2x^7y + 6x^6y^2 + 10x^5y^3 + 13x^4y^4 + 10x^3y^5 + \\
 &\quad 6x^2y^6 + 2xy^7 + y^8 \quad (11)
 \end{aligned}$$

For counting reaction pairs based on $G(D_8)$, one can obtain cycle index and counting series as follows:

Table XI. Reaction Graphs of the Octagonal Class ($m = 0, n = 4$)^a

^a Each couple of reaction graphs linked with a bracket is a reaction pair.

$$Z(D_8) = (1/16)(s_1^8 + 4s_1^2s_2^3 + 5s_2^4 + 2s_4^2 + 4s_8) \quad (12)$$

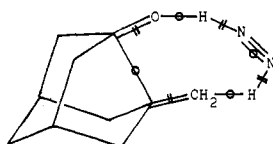
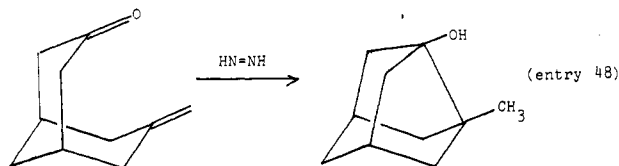
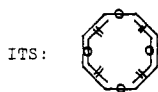
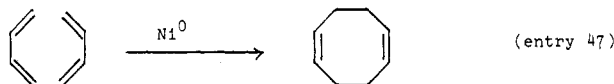
$$P(x,y) = 1 + x + y + 4x^2 + 4xy + 4y^2 + 5x^3 + 12x^2y + 12xy^2 + 5y^3 + 8x^4 + 19x^3y + 33x^2y^2 + 19xy^3 + 8y^4 + 5x^5 + 19x^4y + 38x^3y^2 + 38x^2y^3 + 19xy^4 + 5y^5 + 4x^6 + 12x^5y + 33x^4y^2 + 38x^3y^3 + 33x^2y^4 + 12xy^5 + 4y^6 + x^7 + 4x^6y + 12x^5y^2 + 19x^4y^3 + 19x^3y^4 + 12x^2y^5 + 4xy^6 + y^7 + x^8 + x^7y + 4x^6y^2 + 5x^5y^3 + 8x^4y^4 + 5x^3y^5 + 4x^2y^6 + xy^7 + y^8 \quad (13)$$

The coefficients of $x^m y^n$ in $G(x,y)$ (eq 11) and in $P(x,y)$ (eq 13) are the numbers of reaction graphs and of reaction pairs of octagonal class, respectively.

The full list of octagonal reaction graphs is too large to be recorded. Therefore, I exemplify the case of $m = 0$ and $n = 4$ in Table XI, where there are eight reaction pairs (corresponding to the coefficient $x^0 y^4$ in $P(x,y)$) and thirteen reaction graphs (corresponding to the coefficient $x^0 y^4$ in $G(x,y)$).

EXAMPLES OF ONE-STRING REACTIONS HAVING AN OCTAGONAL REACTION GRAPH

Dimerization of butadiene has an octagonal reaction graph of $m = 0$ and $n = 6$, which is characterized as BC2. The code of the RC graph of level 1 is obtained as follows: 1C(2-1)-



1C(2-1)2C(0+1)3C(2-1)4C(0+1)
5H(1-1)6H(2+1)7H(1-1)8H(0+1)1C

2C(1+1)3C(2-1)4C(0+1)5C(2-1)6C(1+1)7C(2-1)8C(0+1)1C; $m = 0, n = 6$; BC2.

Diimide reduction accompanied by a transannular reaction (entry 48) has an octagonal reaction graph of $m = 1$ and $n = 2$.⁴⁶

CONCLUSION

I have derived several useful concepts from imaginary transition structures: RC graphs of various levels, reaction graphs, and reaction pairs. One-string reactions have been classified by their reaction graphs. Enumeration of hexagonal, tetragonal, and octagonal reaction graphs has proven that the present ITS approach is a versatile tool for the description and classification of organic reactions.

REFERENCES AND NOTES

- (1) Ash, J. E.; Chubb, P. A.; Ward, S. E.; Welford, S. M.; Willett, P. *Communication, Storage, and Retrieval of Chemical Information*; Ellis Horwood: Chichester, UK, 1985; Chapter 8.
- (2) Sello, G. *J. Chem. Inf. Comput. Sci.* **1984**, *24*, 249.
- (3) Ugi, I.; Bauer, J.; Brandt, J.; Friedrich, J.; Gsteiger, J.; Jochum, L.; Schubert, W. *Angew. Chem., Int. Ed. Engl.* **1979**, *18*, 111.
- (4) Dugundii, J.; Gillespie, P.; Marquarding, D.; Ugi, I.; Ramirez, F. In *Chemical Application of Graph Theory*; Balaban, A. T., Ed.; Academic: London, 1976; Chapter 6.
- (5) Arens, J. F. *Recl. Trav. Chim. Pays-Bas* **1979**, *98*, 155.
- (6) Roberts, D. C. *J. Org. Chem.* **1978**, *43*, 1473.
- (7) Hendrickson, J. B. *Angew. Chem., Int. Ed. Engl.* **1974**, *13*, 47.
- (8) Fujita, S. *J. Chem. Inf. Comput. Sci.*, preceding paper in this issue.
- (9) Hereafter, balls on vertices are omitted in a reaction graph.
- (10) Polya, G. *Acta Math.* **1937**, *68*, 145. See also the other chapters of ref 4.
- (11) Hendrickson has reported an alternative representation for this problem in the case of $m = 0$. See ref 7.
- (12) Brandt, J.; von Sholley, A. *Comput. Chem.* **1983**, *7*, 51.
- (13) Hunsdiecker, H.; Hunsdiecker, C. *Ber. Dtsch. Chem. Ges.* **1942**, *75*, 297.
- (14) Whitmore, F. C.; Lewis, C. E. *J. Am. Chem. Soc.* **1942**, *64*, 1618.
- (15) Kamm, O.; Marvel, C. S. *Org. Synth.* **1941**, *Coll. Vol. I*, 42.
- (16) Tschugaeff, L. *Ber. Dtsch. Chem. Ges.* **1899**, *32*, 3332. See also: Nace, H. R. *Org. React. (N.Y.)* **1962**, *12*, 57.
- (17) Whitmore, F. C.; Whitaker, J. S.; Mosher, W. A.; Breivik, D. N.; Wheeler, W. R.; Miner, C. S., Jr.; Sutherland, L. H.; Wagner, R. B.; Clapper, T. W.; Lewis, C. E.; Lux, A. R.; Popkin, A. H. *J. Am. Chem. Soc.* **1941**, *63*, 643.
- (18) Migridichian, V. *Organic Synthesis*; Reinhold: London, 1957; Vol. 1, p 573.
- (19) Owen, L. N. *J. Chem. Soc.* **1949**, 245.
- (20) Huisgen, R.; Pohl, H. *Chem. Ber.* **1960**, *93*, 527.
- (21) (a) Krafft, F. *Ber. Dtsch. Chem. Ges.* **1877**, *10*, 2035. (b) Barbot, A. *Bull. Soc. Chim. Fr.* **1935**, *2*, 895.
- (22) Oppenheim, A.; Precht, H. *Ber. Dtsch. Chem. Ges.* **1876**, *9*, 325.
- (23) Arnold, R. T.; Elmer, O. C.; Dodson, R. M. *J. Am. Chem. Soc.* **1950**, *72*, 4359.
- (24) Troeger, J. *Ber. Dtsch. Chem. Ges.* **1892**, *25*, 3316.
- (25) Gassman, P. G.; Mansfield, K. T. *Org. Synth.* **1973**, *Coll. Vol. V*, 96.
- (26) Moffett, R. B. *Org. Synth.* **1963**, *Coll. Vol. IV*, 238.
- (27) Cope, A. C.; Hardy, E. M. *J. Am. Chem. Soc.* **1940**, *62*, 441. See also: Rhoads, S. J.; Raulins, N. R. *Org. React. (N.Y.)* **1975**, *22*, 1.
- (28) Viola, A.; Iorio, E. J.; Chen, K. K.; Glover, G. M.; Nayak, U.; Kociecki, P. *J. Am. Chem. Soc.* **1967**, *89*, 3462.
- (29) Claisen, L. *Ber. Dtsch. Chem. Ges.* **1912**, *45*, 3157.
- (30) Barton, D. H. R.; Quinkert, G. *J. Chem. Soc.* **1960**, 1.
- (31) Baydar, A. E.; Boyd, G. V. *J. Chem. Soc., Chem. Commun.* **1976**, 718.
- (32) Nieuvenhuis, J.; Arens, J. F. *Recl. Trav. Chim. Pays-Bas* **1958**, *77*, 761, 1153.
- (33) Lach, B. *Ber. Dtsch. Chem. Ges.* **1884**, *17*, 1571.
- (34) Maerck, R. E.; Battiste, M. A. *J. Chem. Soc., Chem. Commun.* **1972**, 1171.
- (35) Bailey, W. J.; Cummins, E. W. *J. Am. Chem. Soc.* **1954**, *76*, 1940.
- (36) Vollhardt, K. P. C. *Acc. Chem. Res.* **1977**, *10*, 1.
- (37) Norris, J. F.; Olmsted, A. W. *Org. Synth.* **1941**, *Coll. Vol. I*, 144.
- (38) Fujita, S.; Sakanoue, S. U.S. Patent 3999991, 1976.
- (39) Franzen, H. *Ber. Dtsch. Chem. Ges.* **1909**, *42*, 2465.
- (40) Whitmore, F. C.; Woodburn, H. M. *J. Am. Chem. Soc.* **1933**, *55*, 361.
- (41) Lacher, J. R.; Tompkin, G.; Park, J. D. *J. Am. Chem. Soc.* **1952**, *74*, 1693.
- (42) Anderson, J. L.; Putnam, P. E.; Sharkey, W. J. *J. Am. Chem. Soc.* **1961**, *83*, 382.
- (43) Mori, T.; Nakahara, T.; Nozaki, H. *Can. J. Chem.* **1969**, *47*, 4266.
- (44) Coyle, J. D. *Chem. Soc. Rev.* **1974**, *3*, 329.
- (45) Gassman, P. G.; Mansfield, K. T. *Org. Synth.* **1973**, *Coll. Vol. V*, 96.
- (46) Kimoto, K.; Kawanishi, M. *Chem. Ind. (London)* **1971**, 1174.