

Description of Organic Reactions Based on Imaginary Transition Structures. 1. Introduction of New Concepts

SHINSAKU FUJITA

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

Received May 27, 1986

The concept of imaginary transition structure (ITS) is proposed for the description of organic reactions. The related concepts derived from the ITS are also introduced: out-bonds; in-bonds; par-bonds; ITS bonds; complex bond numbers; 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, or multistring reactions; ITS rings; and bridges of ring opening, ring closure, or rearrangement.

Computer systems for retrieval of chemical reactions and for design of synthetic pathways have grown separately, although both are concerned with the same organic reactions (Figure 1). For the purpose of retrieval, data of individual reactions are stored and manipulated in machine-readable files.^{1,2} On the other hand, the systems for the latter purpose are based upon data of reaction types rather than those of individual reactions, and they construct synthetic pathways by the logic of retrosynthesis, etc. The reaction type files contain generic schemes abstracted manually from individual reaction data by expert organic chemists (LHASA, etc.)³ or generic reaction matrices derived from mathematical formulation (EROS, etc.).⁴ If the reaction types are abstracted automatically from the individual reaction data, both the approaches can be integrated to help organic chemists more effectively. I intend constructing such an integrated computer system, FORTUNITS (Fuji Organic Reaction Treating Unity based on Imaginary Transition Structures), in which both the retrieval and design of reactions can be supported. For this purpose, a common language or representation is necessary to develop which is applicable to both individual reactions and reaction types.

I will propose here "imaginary transition structures" (ITS), "reaction strings" (RS), and the related new concepts to describe organic reactions and synthetic pathways.⁵

DEFINITION OF IMAGINARY TRANSITION STRUCTURES AND COLORED BONDS

The *imaginary transition structure* (ITS) of a given reaction is a structural formula in which molecules of the starting stage are superposed topologically upon those of the product stage, and the bonds are distinguished and classified into three categories (colors), i.e., *out-bonds* (bonds appearing only in the starting stage), *in-bonds* (bonds appearing only in the product stage), and *par-bonds* (invariant bonds appearing in both the starting and product stages).⁶ In other words, the ITS has all nodes contained in both the starting and the product stages, and said nodes are connected with colored bonds (out-, in-, and par-bonds). For example, the acidic hydrolysis of ethyl acetate (entry 1 in Table I) is represented by the ITS shown in Figure 2. In this paper, out-bonds (*bond color: out*) are denoted by solid lines with a double bar ($\overline{\overline{\text{---}}}$), in-bonds (*bond color: in*) by solid lines with a small circle ($\text{---}\odot\text{---}$), and par-bonds (*bond color: par*) by solid lines (---). These three kinds of bonds may be distinguished by colors when a color graphic display is used.

The ITS shown in Figure 2 is written completely enough to contain hydrochloric acid as a catalyst. However, the corresponding *abbreviated ITS*, in which a chlorine atom is neglected, is also permitted (Table I). Sometimes, the abbreviated ITS's are convenient when one does not refer to the kind of catalysts; compare entries 1 and 2 of Table I.

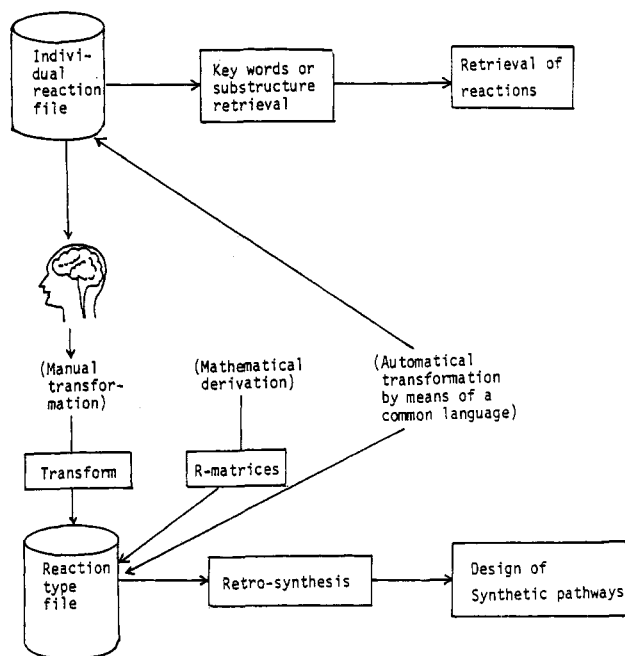


Figure 1. Retrieval of reactions and design of synthetic pathways.

Several examples of reactions and the corresponding ITS's are collected in Table I. It is to be emphasized that the reaction centers can be clarified easily by the glimpse of the out- and in-bonds in the ITS's. Entry 3 is an acid-catalyzed ester formation, which is a reverse reaction of entries 1 and 2. Thus, they have reversely colored ITS's except for the catalysts. Dehydration of cyclohexanol to cyclohexene (entry 4) and dehydrative alkylation of phenol (entry 5) resemble one another if the attention is given to their reaction centers. Dehydration of furfuryl alcohol accompanied by the Wagner-Meerwein rearrangement (entries 6 and 7) needs aluminum oxide or *p*-toluenesulfonic acid as a catalyst, the participation of which is shown in the ITS.

Entries 8 and 9 are the Friedel-Crafts acylations of aromatic compounds, whose ITS's are slightly different from one another. The reaction of entry 9 evolves hydrogen chloride gas, and thus aluminum chloride is present as an independent part in the ITS. On the other hand, boron trifluoride is a component of the ITS of entry 8. The difference between them is presumed to come from the fact that AlCl_3 of the former case is unstable so as to evolve hydrogen chloride.

Entries 11-14 are related to amide formations, in which basic substances are varied. The ITS of entry 11 corresponds to evolution of hydrogen chloride. On the other hand, the hydrogen chloride formed is trapped by triethylamine in entry 12.

Entries 15-18 are examples of more complex reactions, all of which are accompanied with dehydration. Entries 20-23

Table I. ITS's and Abbreviated ITS's of Various Reactions











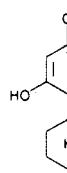
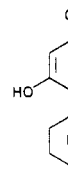
















entry	starting stage	product state	ITS	abbrev ITS
1	$\text{CH}_3\text{COOC}_2\text{H}_5 + \text{H}_2\text{O} + \text{HCl}$	$\text{CH}_3\text{COOH} + \text{C}_2\text{H}_5\text{OH} + \text{HCl}$		
2	$\text{CH}_3\text{COOC}_2\text{H}_5 + \text{H}_2\text{O} + \text{H}_2\text{SO}_4$	$\text{CH}_3\text{COOH} + \text{C}_2\text{H}_5\text{OH} + \text{H}_2\text{SO}_4$		
3	$\text{CH}_3\text{COOH} + n\text{-C}_5\text{H}_{11}\text{OH} + \text{TsOH}$	$\text{CH}_3\text{COOC}_5\text{H}_{11-n} + \text{H}_2\text{O} + \text{TsOH}$		
4	 + H_2SO_4	 + $\text{H}_2\text{O} + \text{H}_2\text{SO}_4$		
5	 +  + H_3PO_4	 + H_2O  + H_3PO_4		
6	 + Al_2O_3	 + $\text{H}_2\text{O} + \text{Al}_2\text{O}_3$		
7	 + TsOH	 + $\text{H}_2\text{O} + \text{TsOH}$		
8	 + CH_3COOH + BF_3	 + COCH_3 + HBF_3OH		
9	 +  + AlCl_3	 + $\text{HCl} + \text{AlCl}_3$		
10	 + HCl	 + H_2O		
11	 + 	 + HCl		
12	 +  + Et_3N	 + Et_3NHCl		
13	 + $(\text{CH}_3\text{CO})_2\text{O} + \text{CH}_3\text{COONa}$	 + CH_3COOH + CH_3COONa		
14	 +  + Py	 + PyHCl		

Table I (Continued)
















entry	starting stage	product state	ITS	abbrev ITS
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				

are typical rearrangements, the features of which are described later. Entries 22 and 23 are the Claisen rearrangement without and with the subsequent hydrogen shift, respectively. Entry 24 is the Eschenmoser epoxy-hydrazone fragmentation.

ITS BONDS AND ITS CONNECTION TABLES

The ITS's collected in Table I contain several different kinds of bonds linking two adjacent nodes. The bonds appearing in the ITS's are referred to as *ITS bonds* (Table II). For example, an ITS bond represented by --- is a single out-bond, which is a single bond in the starting stage but which cleaves after the reaction. Similarly, an ITS bond represented by --- consists of an out- and a par-bond and is referred

Table II. Types of ITS Bonds

						
	(1 - 1)	(1 + 0)	(0 + 1)			
						
	(2 - 2)	(2 - 1)	(2 + 0)	(1 + 1)	(0 + 2)	
						
(3 - 3)	(3 - 2)	(3 - 1)	(3 + 0)	(2 + 1)	(1 + 2)	(0 + 3)

to as a double bond singly cleaved, which is a double bond in the starting stage but which changes into a single bond after the reaction. In some usage, this ITS bond is called an out-bond modified by a par-bond.

Table III. Connection Table of ITS (Entry 1)^a

node	atom or group	coordinate		neighbor 1		neighbor 2		neighbor 3		neighbor 4	
		x	y	node	(a b)	node	(a b)	node	(a b)	node	(a b)
1	CH ₃	0	0	2	(1+0)						
2	C	200	0	1	(1+0)	3	(1-1)	6	(2+0)	7	(0+1)
3	O	400	0	2	(1-1)	4	(1+0)	10	(0+1)		
4	CH ₂	600	0	3	(1+0)	5	(1+0)				
5	CH ₃	800	0	4	(1+0)						
6	O	200	200	2	(2+0)						
7	O	200	-200	2	(0+1)	8	(1-1)	9	(1+0)		
8	H	200	-400	7	(1-1)	11	(0+1)				
9	H	0	-200	7	(1+0)						
10	H	400	-200	3	(0+1)	11	(1-1)				
11	Cl	400	-400	8	(0+1)	10	(1-1)				

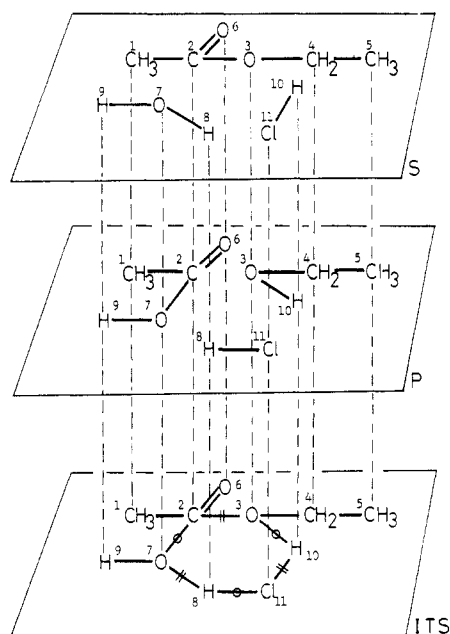
^a The numbering of nodes is given in Figure 2.

Figure 2. Constitution of imaginary transition structures (ITS's). This example is the hydrolysis of ethyl acetate catalyzed by hydrochloric acid. The abbreviations S and P correspond to the starting and product stages, respectively.

Each ITS bond is denoted by a pair of integers ($a b$), wherein the integer a is the bond multiplicity of the corresponding bond of the starting molecule and the integer b is the difference in the bond multiplicity between the product and the starting material. I refer to the pair ($a b$) as a *complex bond number* (Table II). The sum, $a + b$, is the bond multiplicity of the corresponding bond of the product.

A *connection table of ITS* (abbreviated as ITS-CT) can be written easily by using complex bond numbers. For example, the ITS-CT of the reaction (entry 1) is shown in Table III. By extension of the concept of chemical bonds to ITS bonds (Table II), imaginary transition structures can be stored and manipulated as usual structural formulas. It should be emphasized that the *description of organic reactions is now displaced by the manipulation of ITS's as an extended kind of chemical structure*.

I define valency appearing in the ITS as imaginary valency (IV). Thus, the ITS of entry 1 has a node (no. 2, C atom) of IV 5, a node (no. 3, O atom) of IV 3, and so on.

PROJECTION TO STARTING STAGE AND TO PRODUCT STAGE

An ITS contains the whole structural information concerned with a given reaction. When in-bonds in the ITS are canceled, the starting stage can be regenerated as exemplified in Figure 3. This operation is called *projection to starting stage* (PS).

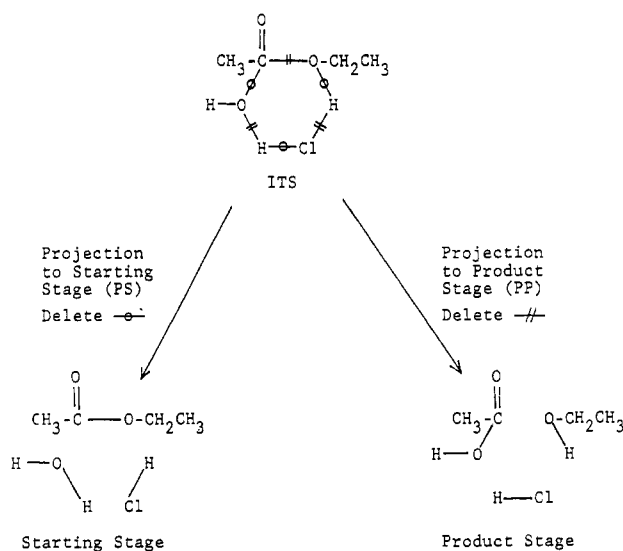


Figure 3. Projection to starting stage and to product stage.

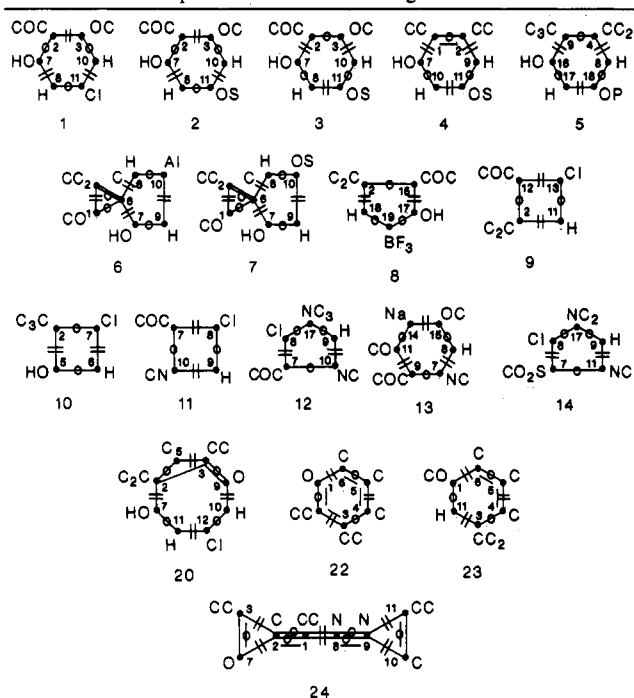
With respect to the ITS-CT (e.g., Table III), PS operation is the displacement of each ITS bond ($a b$) by the corresponding usual bond of multiplicity a .

Similarly, *projection to product stage* (PP) is defined as the deletion of out-bonds, which provides the product stage of a given reaction (Figure 3). The operation PP is the displacement of each ITS bond of ($a b$) by the corresponding usual bond of multiplicity $a + b$.

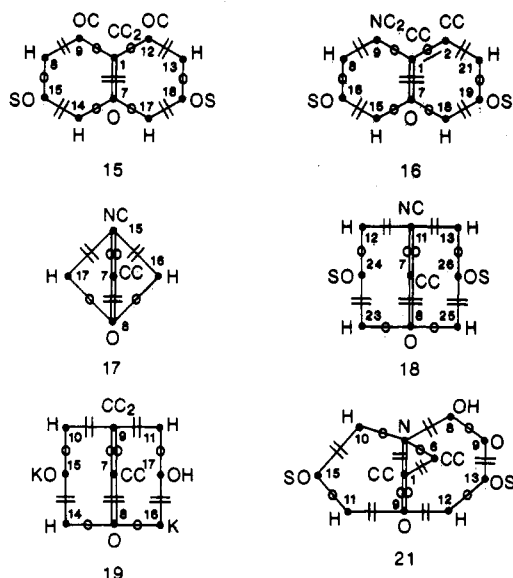
In the same way, each ITS collected in Table I provides the corresponding starting and product stage by PS and PP, respectively. When an abbreviated ITS is adopted, information about a catalyst is missed, but the other essential features of the reaction can be obtained by the operations PS and PP.

GRAPH OF REACTION CENTERS

Reaction centers of a given ITS are the nodes to which in- or out-bonds are incident. A figure of the combined reaction centers is defined as a *graph of reaction centers* (RC graph). The nodes of the RC graph can be balls, atoms, or groups according to the levels of description.⁵ Tables IV and V show the RC graphs of the reaction collected in Table I, wherein the nodes are represented by balls with several pieces of structural information. If two nonadjacent nodes of the reaction centers are linked directly by a par-bond, said par-bond is also drawn in the RC graph (e.g., entry 20 in Table IV). The RC graph is a generic descriptor of the respective reaction types. It should be emphasized that an RC graph describing a reaction type is a substructure (or subgraph) of the corresponding ITS describing an individual reaction. Thus, data of reaction types and those of individual reactions are now closely combined.

Table IV. RC Graphs of Several One-String Reactions^a

^aEach graph number corresponds to the entry number of Table I. The numbering of the nodes is identical with that of Table I. Each node is characterized by the atom at said node and the neighboring atoms.

Table V. RC Graphs of Several Two-String Reactions^a

^aEach graph number corresponds to the entry number of Table I. The numbering of the nodes is identical with that of Table I. Each node is characterized by the atom at said node and the neighboring atoms.

REACTION STRINGS

When the ITS's (Table I) or the corresponding RC graphs of Tables IV and V are examined carefully, the reaction centers of each reaction are found to be connected with alternate in-bonds and out-bonds in the manner of a picture drawn with a single stroke of the brush. For example, the reaction center of entry 1 appears as a string looped, 8-7+2-3+10-11+8 (the numbering of the nodes is shown in Figure 2), wherein - and + represent an out-bond and an in-bond, respectively.

I define a *reaction string* (RS) as a two- or three-colored string of various length, which has alternate in-bonds (—○—) and out-bonds (—/—) and can be modified by par-bonds (—).

Table VI. Connection Table of the Reaction String for the Reaction of Entry 1

node	atom	bond (a b)	neighbor 1		neighbor 2	
			node	atom	node	atom
8	H	(1-1)				
7	O	(0+1)	9	H	6	O
2	C	(1-1)	1	C		
3	O	(0+1)	4	C		
10	H	(1-1)				
11	Cl	(0+1)				
8	H					

The nodes that appear in the reaction string can be balls in an abstract fashion, atoms, or groups according to levels of description.⁷ In general, each reaction string of the complete ITS's or of the corresponding RC graphs is closed to form a loop, but that of the abbreviated ITS's or of the corresponding RC graph is not always closed.

Each of the reactions in Table IV contains only a single reaction string, which can be easily picked up. On the other hand, each of the reactions collected in Table V possesses two reaction strings. To prevent some ambiguity in adoption of reaction strings, I decided on a criterion that the same colored bonds incident to a node are preferred to form different reaction strings if possible. In other words, each reaction string is preferred, if possible, not to contain each node doubly. This criterion is reasonable chemically since a double or triple bond is considered to form or to cleave stepwise in realistic organic reactions.

In the case of the Beckmann rearrangement (entry 21), I adopt two RS, i.e., 1-6+7-8+14-13+12-9+1 and 1-7+10-15+11-9+1. But I do not take up a combination of the above-described RS's (i.e., 1-6+7-8+14-13+12-9+1-7+10-15+11-9+1) nor other possible RS's (1-7+6-1+9-11+15-10+7-8+14-13+12-9+1, etc.). In abbreviated representations of ITS's (Table I), the same criterion is also effective. Thus, I abstract two RS's for the abbreviated ITS of entry 21, i.e., 10+7-1+9-11 and 8-7+6-1+9-12.⁸ Another set of RS's (11-9+1-6+7-1+9-12 and 8-7+10) is not permitted in the light of the criterion described above. The reactions collected in Table V are two-string reactions, the strings of which are entangled or, in other words, share two or more nodes. The last points will be discussed later.

In several reactions such as entry 24, double and triple bonds are formed or cleaved at the same time, and hence only a single reaction string is possible.

Each reaction string can be stored as a string (e.g., 8H(1-1)7O(0+1)2C(1-1)3O(0+1)10H(1-1)11Cl(0+1)8H) or as a connection table shown in Table VI. They contain at least the adjacent node of RS and the complex bond number (a b) for each ITS bond. It should be noted that integers b appear in plus and minus value alternately in such representations of reaction strings.

The concepts of RC graphs and reaction strings I propose here can be proved to be equivalent in some cases to the numerical representations reported previously.^{9,10} However, the latter mathematical formulations are so oversimplified as to omit par-bonds modifying the reaction strings. Thus, in these approaches, RC graphs of entries 1, 3, and 21 (in Table IV) cannot be discriminated unless one refers to the starting and product stage. They are not so applicable to the reactions having two or more reaction strings as the present concepts are. Moreover, additional structural information around the nodes is discarded during the numerical treatments. It is not the case in the present approach, since the RC graphs with reaction strings are the subgraphs (or substructures) of the corresponding ITS's. The reaction strings of this work are based upon graphical representations (ITS's) and so are easy

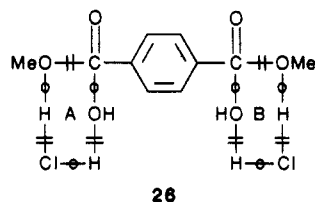
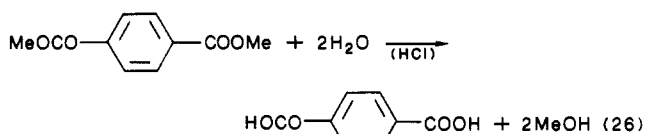
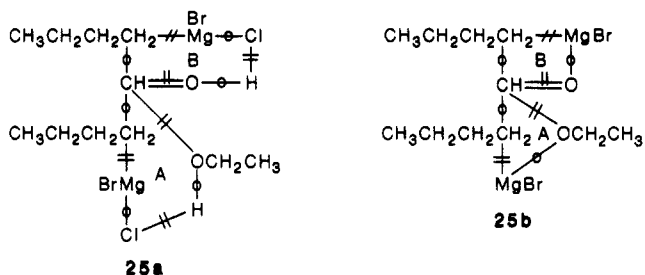
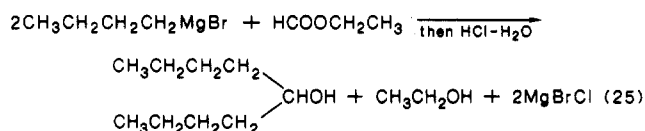
to understand for organic chemists.

CLASSIFICATION OF REACTIONS BY STRINGITY

In the preceding section, I described the reactions of one or two strings. Here, I define *stringity* of the reaction as the number of reaction strings contained in the ITS of the reaction. All reactions can be classified by their stringity. A *one-string reaction* has a single RS (e.g., the examples in Table IV). A *two-string reaction* has two RS's (e.g., the examples in Table V), and so on.

The mode of *entanglement of reaction strings* is also a versatile guide for classification of organic reactions. The two strings of the reaction (entry 15 or 16) share two nodes. Each set of two strings of the reactions (entries 17–19 and 21) shares three nodes.

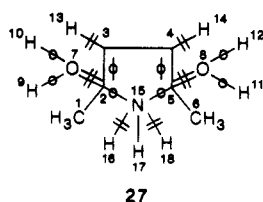
Entry 25 has two reaction strings (A and B) which share one node with one another. Two kinds of ITS's are written



for this reaction. ITS **25a** indicates the total reaction in which acidic workup is also considered. ITS **25b** is the representation of entry 25 without such acidic workup. The stringity is invariant in both of the ITS's.

Two strings of a given reaction are separated in some cases. For example, hydrolysis of dimethyl terephthalate (entry 26) gives an ITS of two strings that share no nodes. This type of reaction is called a *separate two-string reaction*.

A further complex example is a pyrrole formation from a 1,4-diketone (entry 27). The abbreviated ITS **27** has four



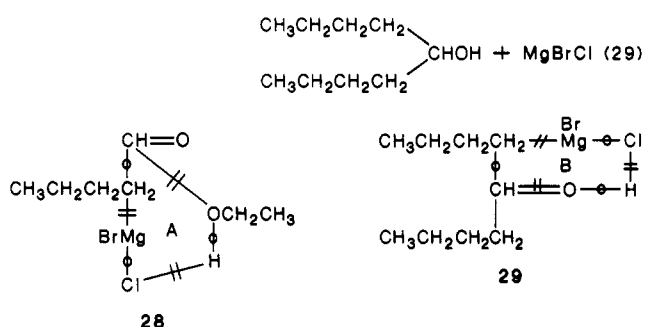
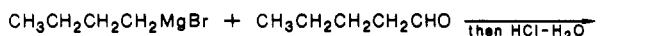
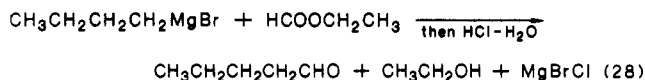
reaction strings, which are 10+7-2+3-13, 9+7-2+15-16,

12+8-5+4-14, and 11+8-5+15-18. I refer to this reaction as an *entangled four-string reaction*.

It should be noted here that the stringity does not strictly correspond to the steps of the reaction. For instance, the stringity of reaction 25 is invariant with or without the additional step of acidic workup. This point will be discussed later.

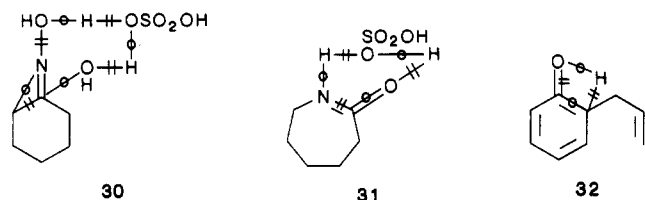
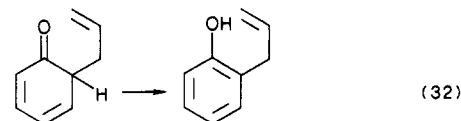
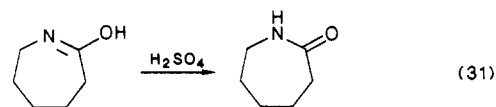
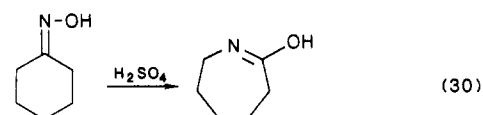
PRIMARY REACTIONS OF MULTISTRING REACTIONS

A reaction having two or more reaction strings can be divided into the corresponding *primary reactions of one string*.¹¹ The Grignard reaction of entry 25 consists of two primary reactions (entries 28 and 29), the ITS's of which are **28** and **29**, respectively. The reaction string of ITS **28** and that of



29 are accumulated to form the two reaction strings of ITS **25a**. This fact corresponds to the fact that reaction 25 consists of the two primary reactions 28 and 29.

The Beckmann rearrangement of entry 21 is divided into two primary reactions, the rearrangement to an enol form (30) and the subsequent proton shift (31). The corresponding ITS's



(30 and 31) are of one string. The two reaction strings are contained in ITS **21**, which corresponds to the total process of the Beckmann rearrangement. This case exemplifies the above-described criterion to pick up reaction strings.

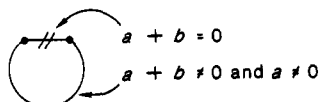
The Claisen rearrangement of entry 24 can be regarded as the sequence of the reaction (entry 23) and the following hydrogen shift (32). The reaction string of ITS **23** and that

of 32 are not accumulated but degenerated to a single RS of ITS 24.

RING STRUCTURES IN ITS'S

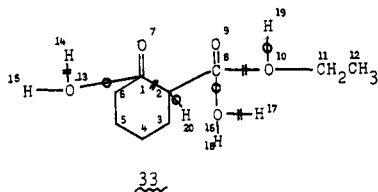
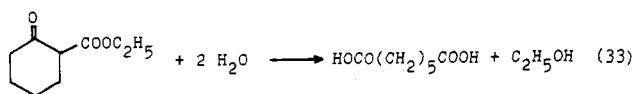
Any ring structures appearing in ITS's are referred to as *ITS rings*. The reaction string in full can form an ITS ring by itself, and other types of ITS rings are present in ITS's. They are important to describe the features of reactions. Suppose that each ITS bond of an ITS ring has a complex number ($a b$), where $a + b \neq 0$ and $a \neq 0$. This ITS ring is invariant during the reaction. Table I contains several examples of this type.

When only one ITS bond of an ITS ring had ($a b$) of $a + b = 0$ and all other bonds have ($a b$) of $a + b \neq 0$ and $a \neq 0$, said ITS ring is called a *bridge of ring opening of order 1*. This case is denoted by BO1 and is shown schematically as



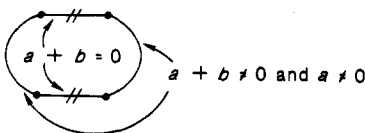
wherein a single out-bond can be displaced by a double or triple out-bond. The order of the bridge of ring opening is the number of ITS bonds of $a + b = 0$ contained in the ITS ring.

The presence of this type of ITS rings (BO1) implies that the reactions are accompanied by ring opening. For example, abbreviated ITS 33 contains a BO1 (1-2-3-4-5-6-1) that

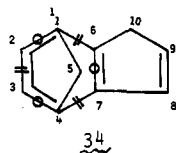
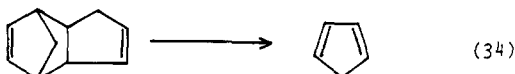


corresponds to the ring opening of cyclohexanone during the reaction (entry 33). ITS rings 1-2-3-4-5-1 of ITS 6 or 7, 1-2-3-4-5-6-1 of ITS 21, and 1-2-3-4-5-6-1 and 1-2-7-3-4-5-6-1 of ITS 24 are bridges of ring opening BO1.

Suppose that two ITS bonds of an ITS ring have ($a b$) of $a + b = 0$ and other bonds have ($a b$) of $a + b \neq 0$ and $a \neq 0$. Said ITS ring is referred to as a *bridge of ring opening of order 2* (BO2). This ITS ring is illustrated as



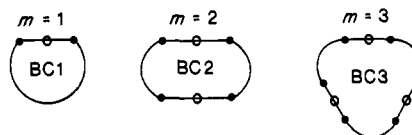
wherein single out-bonds can be displaced by double or triple out-bonds. The ITS ring BO2 indicates a ring opening where two bonds are cleaved. The retro Diels-Alder reaction of entry 34 involves cleavage of two bonds, and hence the corresponding



ITS 34 has ITS rings 1-2-3-4-7-6-1 and 1-5-4-7-6-1, which can be characterized as bridges of ring opening of order

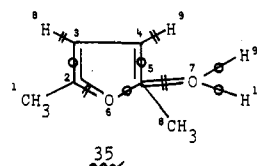
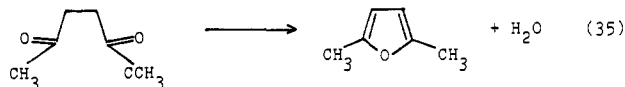
2. In a similar way, a bridge of ring opening of higher order can be defined.

A *bridge of ring closure of order m* (BC m) is defined as an ITS ring that consists of m ITS bonds of $a = 0$ and an arbitrary number of ITS bonds of $a + b \neq 0$ and $a \neq 0$. This is illustrated as



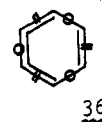
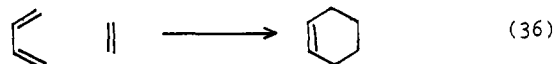
wherein each single in-bond can be substituted by a double in-bond or a triple one.

The furan formation of entry 35 is a ring closure, the abbreviated ITS (35) of which has two reaction strings. This



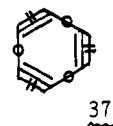
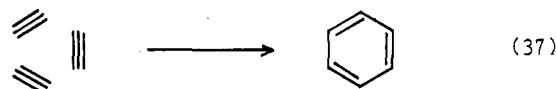
reaction is characterized as BC1, since an ITS ring (6-2-3-4-5-6) is present.

ITS 36 of the Diels-Alder reaction (entry 36) is charac-



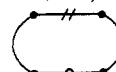
terized by an ITS ring of BC2, which involves the same nodes as the reaction string does. Cyclic acetal formation (entry 15) and pyrrole formation (entry 27) have other types of BC2 (1-9-10-11-12-1 and 2-3-4-5-15, respectively), each ITS ring of which contains two bonds of $a = 0$ belonging to different reaction strings.

If one considers the trimerization of acetylene to benzene (37), BC3 appears in the corresponding ITS 37. The ITS ring



(BC3) overlaps the reaction string of 37.

Suppose that an ITS ring has one ITS bond of $a + b = 0$, another ITS bond of $a = 0$, and an arbitrary number of ITS bonds of $a + b \neq 0$ and $a \neq 0$, i.e.



wherein the ITS bonds $//$ and \bigcirc can be substituted by bonds of types I and II, respectively. These types of ITS rings



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