Description of Organic Reactions Based on Imaginary Transition Structures. 8. Synthesis Space Attached by a Charge Space and Three-Dimensional Imaginary Transition Structures with Charges

SHINSAKU FUJITA

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

Imaginary transition structures with charges are proposed in order to overcome the duality of expression concerning bond character. Covalent and ionic expression are now integrated into unitary representations. Stereochemistry and ionic characters of bonds in organic reactions are manipulated in the form of three-dimensional imaginary transition structures, which are defined in three-dimensional synthesis spaces attached by charge spaces.

The duality of expression concerning bond character is a problem to be solved for computational manipulation of organic reactions. Let us examine a simple case, i.e., formation of an ammonium salt:

$$Me_3N + HCl \rightarrow Me_3NH^+Cl^- \quad (entry\ 1)$$

In this expression, hydrochloric acid is represented by a covalent formula HCl, but it can be denoted H^+Cl^-. Similarly, the ammonium salt ($Me_3NH^+Cl^-$) of ionic character can also be written Me_3NHCl. Moreover, manipulation of stereochemical information is still open to further discussion.

We have previously proposed the concept of imaginary transition structures (ITS's) as a versatile methodology for description of organic reactions.[1] We will extend the concept here in order to overcome the duality associated with the system of organic chemistry. We will also propose a three-dimensional ITS (3D-ITS) and with charges (3D-ITS-C) as full expressions of stereochemistry of organic reactions.

SYNTHESIS SPACE AND CHARGE SPACE

A synthesis space is defined as a set of all nodes that appear in a given reaction (or a synthetic pathway). A synthesis space for entry 1 contains six nodes (1N, 2CH₃, 3CH₃, 4CH₃, 5H, and 6Cl) as shown in Scheme I. In order to describe ionic character of the reaction, the synthesis space is attached by a charge space, which contains plus and minus charges as nodes (7+, 8−, and 9+). These nodes in both spaces are connected by one of three colored bonds (in-, out-, and par-bonds)[2] in accordance with structural changes during the reaction. The resulting imaginary transition structure with charges (ITS-C 1) is the full expression of structural changes of the ammonium salt formation (entry 1). ITS bonds appearing in ITS-C's are classified into two types, i.e., *inter*space bonds (linking between the synthesis and the charge spaces) and *intra*space bonds (contained in either space).

ITS-C's are stored in a computer in the form of connection tables (Table I) in which all ITS bonds[2] are represented by complex bond numbers. Each ITS bond with ionic character is denoted by an asterisk (*), and the information of the charge space is given at the bottom of Table I.

The projection to starting stage (PS)[1] is now extended to mean deletion of in-bonds of ITS-C (more conveniently followed by deletion of interspace bonds and closing up each pair of nodes connected by said interspace bonds). With respect to such a connection table as Table I, the operation (PS) is the cancellation of all intraspace bonds (denoted by a dollar sign) and the following displacement of each ITS bond (*a b*) of the synthesis space by the corresponding usual bond of multiplicity *a*.
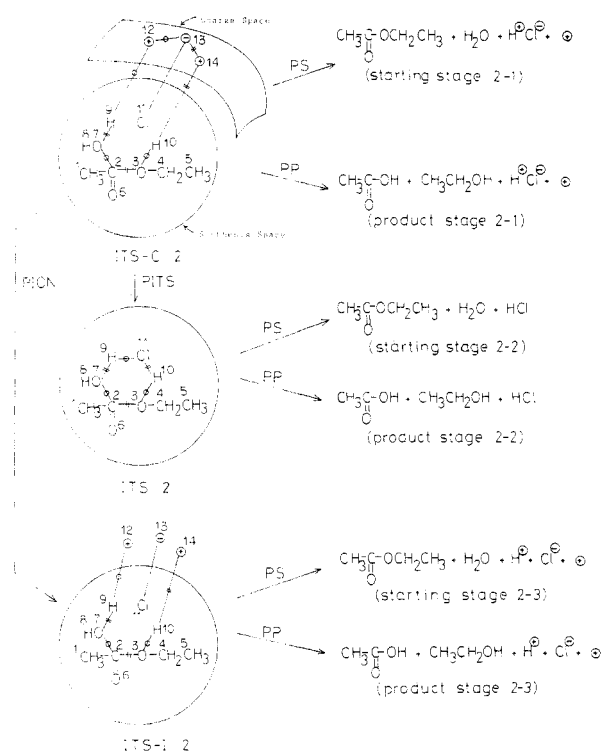
Scheme I
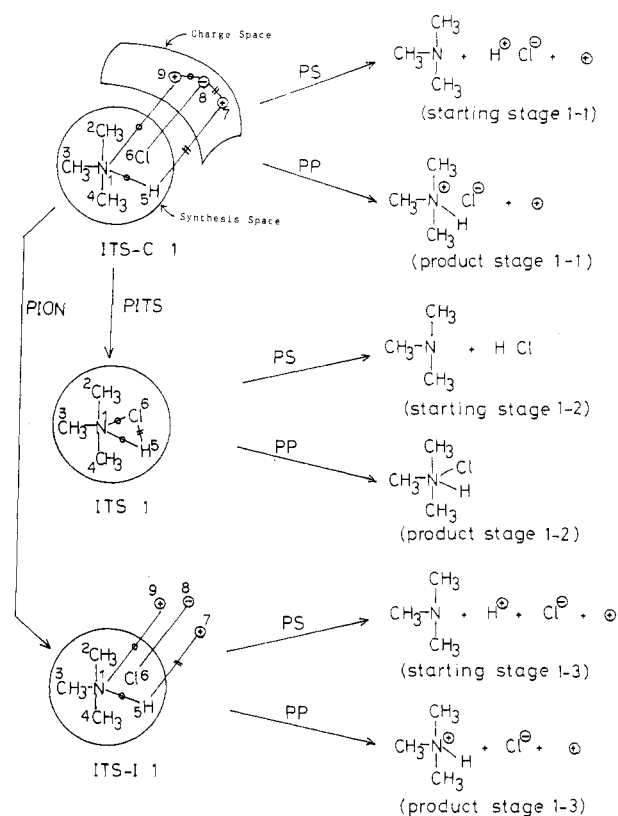
The projection to product stage (PP) is defined as deletion of out-bonds of ITS-C (also followed by deletion of interspace bonds and closing up each pair of nodes connected by said interspace bonds). The operation (PP) is the cancellation of all intraspace bonds (denoted by a dollar sign) and the following displacement of each ITS bond of (*a b*) by the corresponding usual bond multiplicity *a + b*. The corresponding starting and product stages are easily reproduced from ITS-C 1 by the extended PS and PP operations, respectively, as shown in Scheme I.

We introduce the *projection into ITS* (PITS) for the purpose of correlating the ITS-C with the ITS defined previously.[1] This operation is deletion of a charge space, wherein each ITS bond linking adjacent plus and minus charges is mapped on an ITS bond between the nodes said charges are incident to and the charges and interspace bonds are omitted. This operation is illustrated in Scheme I. The operation (PITS) is taking-up of the part of the synthesis space from such a connection table as Table I.

Scheme II

Scheme III

We also define *projection into an ionic ITS* (PION) as deletion of ITS bonds contained in a charge space (intraspace bonds). This operation affords an ionic expression of ITS (ITS-I) as shown in Scheme I. The operation (PION) is the cancellation of ITS bonds denoted by the asterisk and dollar sign.

When PS and PP operations are applied to ITS-C, ITS, and ITS-I, various expressions of starting and product stages are obtained as illustrated in Scheme I. Since they involve ionic and covalent expressions of organic reactions, ITS-C's can be regarded as integrated representation of organic reactions. Thus, the expression of entry 1 consists of the starting stage (1–2) derived from ITS 1 and the product stage (1–1) from ITS-C 1.

Scheme II is another example of ITS-C, which corresponds to acid-catalyzed hydrolysis of ethyl acetate. The above-defined operations can be applied to ITS-C of Scheme II to result in regeneration of reaction features in various levels.

Hydrolysis of a nitrile to a carboxylic acid affords ITS-C 3, in which participation of sulfuric acid is partially considered (Scheme III). Such an abbreviated ITS is convenient to describe essential features of the reaction. In this case, four reaction strings are obtained by considering only intraspace bonds or by extracting from the corresponding ITS 3 projected.

With the present extension, minor change or extension of concepts related to ITS is necessary and will be discussed elsewhere. It should be noted that the present method of description can be applied to describe coordinate bonds in such compounds as sulfoxide, nitro compounds, etc. as shown in Scheme IV. The PITS and PION operations are useful to correlate the present description to the conventional ones. In light of the present representation, organic reactions involving sulfoxide, nitro compounds, phosphonium ylides, diazonium salts, etc. are described successfully.
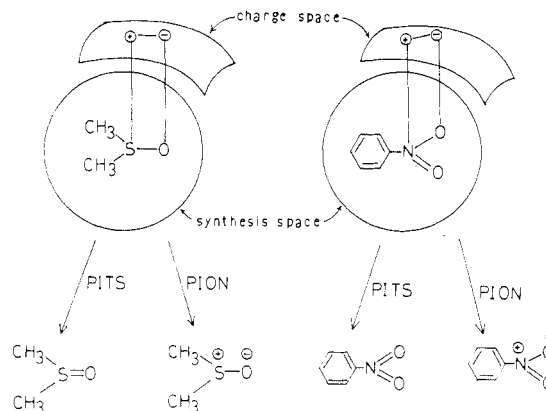
THREE-DIMENSIONAL IMAGINARY TRANSITION STRUCTURE WITH CHARGES

For the description of stereochemical aspects, all nodes appearing in a given reaction (or a synthetic pathway) are assigned to appropriate three-dimensional coordinates in a
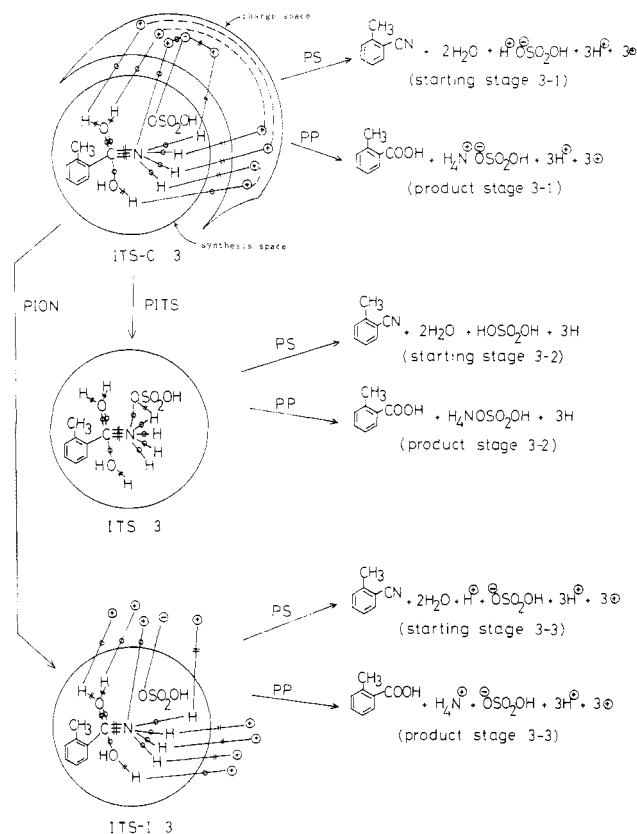
Scheme IV

synthesis space. We attach a charge space to the synthesis space in order to describe ionic character. A three-dimensional imaginary transition structure (3D-ITS) is obtained by writing its ITS bond[2] between said nodes in accordance with a structural change during the reaction.

For example, the Walden inversion[3] is represented by 3D-ITS with charges (3D-ITS-C) as shown is Scheme V. For simplicity of description or of data input, the z coordinate of each node is given as a discrete quantity.[4] Each ITS bond, which connects two nodes different in z value (for example, nodes 10 and 1 and 1 and 22), is represented by a directed bond (an upward arrow with respect to the z axis). This treatment can simplify the projection of the 3D-ITS into the corresponding (two-dimensional) ITS without missing the stereochemical information.

The formation of a sulfonic acid salt and the disappearance of sodium acetate should have ionic character. These features are represented by plus and minus charges in a charge space, which are connected by in-, out-, and/or par-bonds. In ad-

Table I. Connection Table of ITS-C 1[a]

node	atom or group	x	y	z	neighbor 1 node	(a b)	neighbor 2 node	(a b)	neighbor 3 node	(a b)	neighbor 4 node	(a b)	neighbor 5 node	(a b)
		coordinate												

node	atom or group	x	y	z	node	(a b)	node	(a b)	node	(a b)	node	(a b)	node	(a b)
					Synthesis Space									
1	N	0	0	0	2	(1+0)	3	(1+0)	4	(1+0)	5	(0+1)	6	(0+1)
2	CH₃	0	100	0	1	(1+0)								
3	CH₃	-100	0	0	1	(1+0)								
4	CH₃	0	-100	0	1	(1+0)								
5	H	-50	150	0	1	(0+1)	6	(1-1)*						
6	Cl	50	100	0	1	(0+1)*	5	(1-1)*						
					Charge Space									
7	+	-50	150	∞	5	(1-1)@	8	(1-1)$						
8	-	50	100	∞	6	(1+0)@	7	(1-1)$	9	(0+1)$				
9	+	0	0	∞	1	(0+1)@	8	(0+1)$						

[a] The asterisk (*) represents an ITS bond attached by charge-spatial expressions. The interspace and intraspace bonds are denoted by rate at (@) and dollar sign ($), respectively.
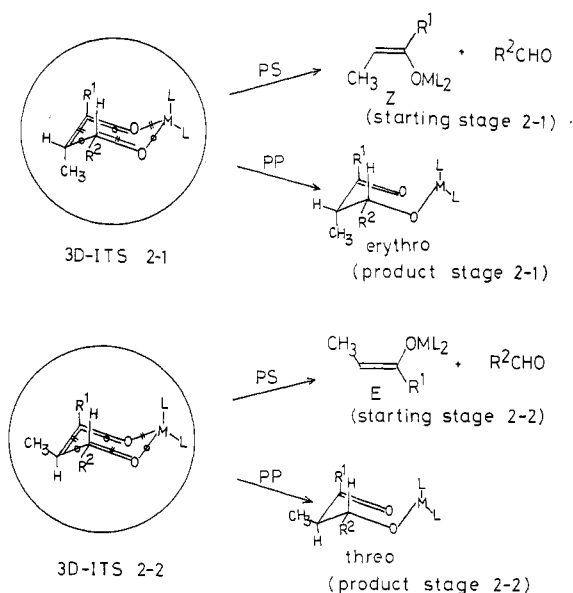
Scheme V

Scheme VI
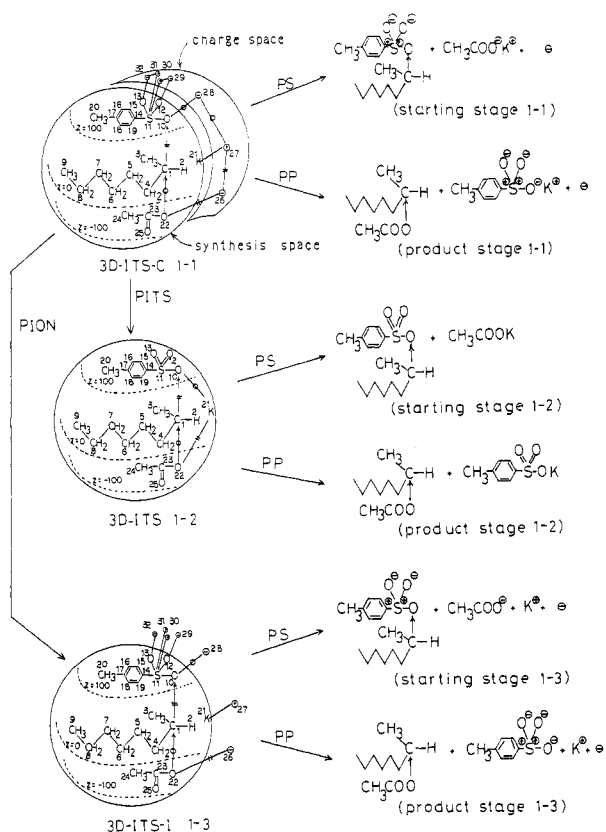
dition, the concept of a charge space settles a problem embraced in the description of sulfonates, which is troublesome to organic chemists. Thus, alternative expressions $ArS^{2+}(O^-)_2OR$ and $ArS(=O)_2OR$ are integrated in the present formulation.

The 3D-ITS-C can be stored in a computer as a connection table such as shown in Table II, in which an arrow is represented by a sign (plus (+) or minus (−)) with respect to its direction. Each ITS bond with ionic character is denoted by an asterisk (*), and the information of the charge space is given at the bottom of this connection table.

The projection into 3D-ITS (PITS) and into ionic 3D-ITS (PION) affords 3D-ITS and 3D-ITS-I, respectively. The characteristics of the three types of ITS's (3D-ITS-C, 3D-ITS, and 3D-ITS-I) become clear when one applies PS and PP operations to these expressions. The starting and product stages from 3D-ITS-I are almost the same as those derived from 3D-ITS-C but are convenient when one wishes to abbreviate information on the gegen cation. All starting and product stages of Scheme I hold information on the stereo-

chemistry of this reaction. Moreover, attachment of a charge space can describe ionic characteristics of this reaction exactly.

Stereoselectivity in aldol reactions of enols has been explained by presuming six-membered cyclic transition states.[5] The stereochemical features can be stored and manipulated in light of the present 3D-ITS. Thus, 3D-ITS's are obtained easily when transient bonds expressed by dotted lines in the presumed transition states are substituted by in-, out-, or par-bonds as illustrated in Scheme VI. The PS and PP operations reveal stereochemical aspects of these reactions, in which erythro and threo isomers are predominant, respectively, from Z and E isomers.

We have previously introduced various concepts related to ITS: graphs of reaction centers; reaction graphs; reaction strings; and bridges of ring opening, ring closure, and rearrangement.[1] Minor modification or extension of these concepts is necessary with the present extension and will be discussed elsewhere.

CONCLUSION

A three-dimensional imaginary transition structure with charges is proposed as a full expression of structural changes during a reaction. Covalent and ionic expression of each bond is integrated into this unitary representation, which is a kind of structural formula with three colored bonds and is defined in a synthesis space attached by a charge space. Stereochemical features of the reaction are also represented ade-

Table II. Connection Table of 3D-ITS-C 1ᵃ

node	atom or group	coordinate x	y	z	neighbor 1 node	(a b)C	S	neighbor 2 node	(a b)C	S	neighbor 3 node	(a b)C	S	neighbor 4 node	(a b)C	S	neighbor 5 node	(a b)C	S
								Synthesis Space											
1	C	0	0	0	2	(1+0)		3	(1+0)		4	(1+0)		10	(1-1)	+	22	(0+1)	-
2	H	100	0	0	1	(1+0)													
3	CH₃	-100	50	0	1	(1+0)													
4	CH₂	-100	-150	0	1	(1+0)		5	(1+0)										
5	CH₂	-200	-50	0	4	(1+0)		6	(1+0)										
6	CH₂	-300	-150	0	5	(1+0)		7	(1+0)										
7	CH₂	-400	-50	0	6	(1+0)		8	(1+0)										
8	CH₂	-500	-150	0	7	(1+0)		9	(1+0)										
9	CH₃	-600	-50	0	8	(1+0)													
10	O	0	300	100	1	(1-1)	-	11	(1+0)		21	(0+1)*							
11	S	-100	300	100	10	(1+0)		12	(2+0)*		13	(2+0)*		14	(1+0)				
12	O	-50	400	100	11	(2+0)*													
13	O	-150	400	100	11	(2+0)*													
14	C	-200	300	100	1	(1+0)		15	(2+0)		19	(1+0)							
15	CH	-220	350	100	14	(2+0)		16	(1+0)										
16	CH	-280	350	100	15	(1+0)		17	(2+0)										
17	C	-300	300	100	16	(2+0)		18	(1+0)		20	(1+0)							
18	CH	-280	250	100	17	(1+0)		19	(2+0)										
19	CH	-220	250	100	14	(1+0)		18	(2+0)										
20	CH₃	-400	300	100	17	(1+0)													
21	K	200	50	0	10	(0+1)*	+	22	(1-1)*		23	(1+0)							
22	O	0	-300	-100	1	(0+1)		21	(1-1)*										
23	C	-100	-300	-100	22	(1+0)		24	(1+0)										
24	CH₃	-200	-300	-100	23	(1+0)					25	(2+0)							
25	O	-100	-400	-100	23	(2+0)													
								Charge Space											
26	-	0	-300	∞	22	(1-1)@		27	(1-1)$										
27	+	200	50	∞	21	(1+0)@		26	(1-1)$		28	(0+1)$							
28	-	0	300	∞	10	(0+1)@		27	(0+1)$										
29	-	-50	400	∞	12	(1+0)@		30	(1+0)$										
30	+	-100	300	∞	11	(1+0)@		29	(1+0)$										
31	+	-100	300	∞	11	(1+0)@		32	(1+0)$										
32	-	-150	400	∞	13	(1+0)@		31	(1+0)$										

ᵃThe asterisk (*) in column C represents an ITS bond attached by charge-spatial expressions. The interspace and intraspace bonds are denoted by rate at (@) and dollar sign ($), respectively, in column C. The stereochemical information is represented by plus (+) and minus (−) signs in column S.

quately by this extended structural formula.

REFERENCES AND NOTES

(1) Fujita, S. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 205, 212, 224, 231, 238.
(2) An imaginary transition structure (ITS) for a given reaction is a structural formula which has three colored bonds, i.e., in-bonds (⊖—), out-bonds (—‖—), and par-bonds (—), in accordance with the changes during the reaction. In-bonds are those involved in the product stage. Out-bonds are contained only in the starting stage. Par-bonds are invariant bonds during the reaction. Each ITS bond consists of a combination of the three colored bonds as shown below.

Types of ITS Bonds

(3) Phillips, H. *J. Chem. Soc.* **1925**, *127*, 2552.
(4) The *x*, *y*, and *z* coordinates of a real transition state can be used for more elaborate description.
(5) Murata, S.; Noyori, R. *Kagaku Zokan (Kyoto)* **1981**, *91*, 117.

Description of Organic Reactions Based on Imaginary Transition Structures. 9. Single-Access Perception of Rearrangement Reactions

SHINSAKU FUJITA

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

The perception of rearrangement reactions is replaced by the examination of subgraphs of imaginary transition structures (ITS's). The bridges of rearrangement are a clue for the detection of type I rearrangement. Type II rearrangement can be recognized in light of five-nodal subgraphs.

How does a computer recognize rearrangement reactions? Although many computer systems have been reported for retrieval of organic reactions and for design of synthetic pathways, perception of rearrangements has not been discussed so rigorously from the viewpoint of computer manipulation. In conventional methods of describing organic reactions, dual access to substrate and to product data is necessary to perceive a rearrangement reaction. For example, Dubois has isolated the pivot focus as a site of adjunction or ablation in order to describe the Favorskii rearrangement.[1] But his treatment does not clarify the reaction as a rearrangement reaction unless one compares the substrate with the product. Moreover, the intermediacy of a cyclopropanone in the reaction has been neglected completely in his approach. Roberts has treated the Claisen rearrangement by means of a CP skeleton, which involves only a net change of dynamic electrons.[2] Hence, his approach cannot discriminate the Claisen rearrangement from the Diels–Alder reaction, the Cope rearrangement, and so on. The situation of Ugi's *R* matrix is the same as described above, since the *R* matrix also represents a net change of the reaction only.[3] Hendrickson[4] has recognized that rearrangement reactions correspond to specific σ-shells (e.g., 4x, 4u, and 5) in thermal pericyclic reactions of a six-atom framework. However, his method, too, requires dual reference to substrate and to product data for this perception.

We have proposed the concept of an imaginary transition structure (ITS) as a unitary representation of respective organic reactions.[5] One of the merits of the ITS approach is that the characterization of organic reactions can be replaced by a subgraph or substructure search of ITS's. Thus, various subgraphs of ITS's are found to be descriptors of reaction features (Figure 1).

This paper deals with specific ring structures of ITS's as clues for the perception of rearrangements. And this will point out another type of rearrangement reaction that can be perceived in light of five-nodal subgraphs.

IMAGINARY TRANSITION STRUCTURE (ITS) AND ITS RINGS

First we discuss the pinacol rearrangement represented by ITS **1**.[5] The ITS consists of various bonds, each of which is a combination of in-bonds (—⊖—), out-bonds (—‖—), and/or

par-bonds (—). Each bond appearing in an ITS is denoted

ITS **1**
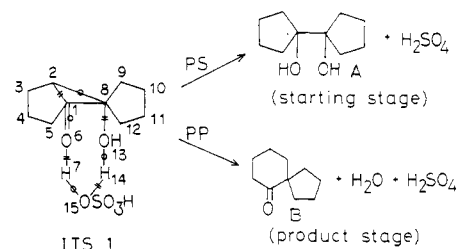
by a complex bond number (*a b*), wherein the first integer, *a*, corresponds to the bond multiplicity of the starting stage and the second one, *b*, is the change of the multiplicity during the reaction. For example, the ITS bond between nodes 1 and 6 is denoted by

and is coded as a (1 +1) bond. The constitution of the ITS is easy to understand when one operates the *projection to starting stage (PS)*[6] and the *projection to product stage (PP)*.[7]

Since the ITS is regarded as a kind of structural formula that has three colored bonds (in-, out-, and par-bonds), the ring perception technique, which has been developed for manipulation of usual structural formulas, is applicable to the ITS. We have defined "ITS rings" as ring structures appearing in a given ITS. At least six ITS rings are detected from ITS **1** as collected in Table I. Ring 1 is a looped reaction string, wherein in-bonds and out-bonds appear alternately. Ring 2 is perceived as a *bridge of ring opening of order 1 (BO₁)*, which has one ITS bond (*a b*) of *a* + *b* = 0 and all other bonds of *a* + *b* ≠ 0 and *a* ≠ 0. This BO_1 corresponds to the fact that the pinacol rearrangement is accompanied by opening of a five-membered ring. Ring 3 is recognized as a *bridge of ring closure of order 1 (BC₁)*, in which one ITS bond has a complex bond number (*a b*) of *a* = 0 and all other ITS bonds have (*a b*) of *a* + *b* ≠ 0 and *a* ≠ 0. The appearance of BC_1 stems from the fact that the reaction is a ring closure to form a six-membered ring.

Ring 4 is characterized as a *bridge of rearrangement (BR)*, which has one ITS bond of *a* + *b* = 0, another ITS bond of *a* = 0, and all others of *a* + *b* ≠ 0 and *a* ≠ 0. This ring