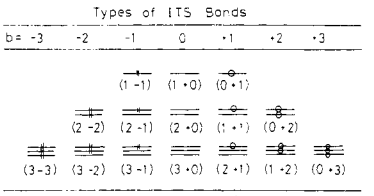
quately by this extended structural formula.

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.



(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

Received October 6, 1986

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.

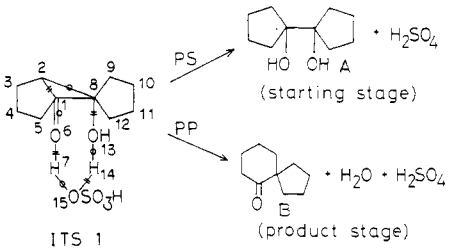
How does a computer recognize rearrangement reactions? Although many computer systems have been reported for retrieval of organic reactions and for design of synthetic pathways, perception of rearrangements has not been discussed so rigorously from the viewpoint of computer manipulation. In conventional 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



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* (*BO*$_1$), which has one ITS bond (*a b*) of $a + b = 0$ and all other bonds of $a + b \neq 0$ and $a \neq 0$. 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* (*BC*$_1$), in which one ITS bond has a complex bond number (*a b*) of $a = 0$ and all other ITS bonds have (*a b*) of $a + b \neq$ and $a \neq 0$. 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 $a + b \neq 0$ and $a \neq 0$. This ring

0095-2338/87/1627-0115$01.50/0

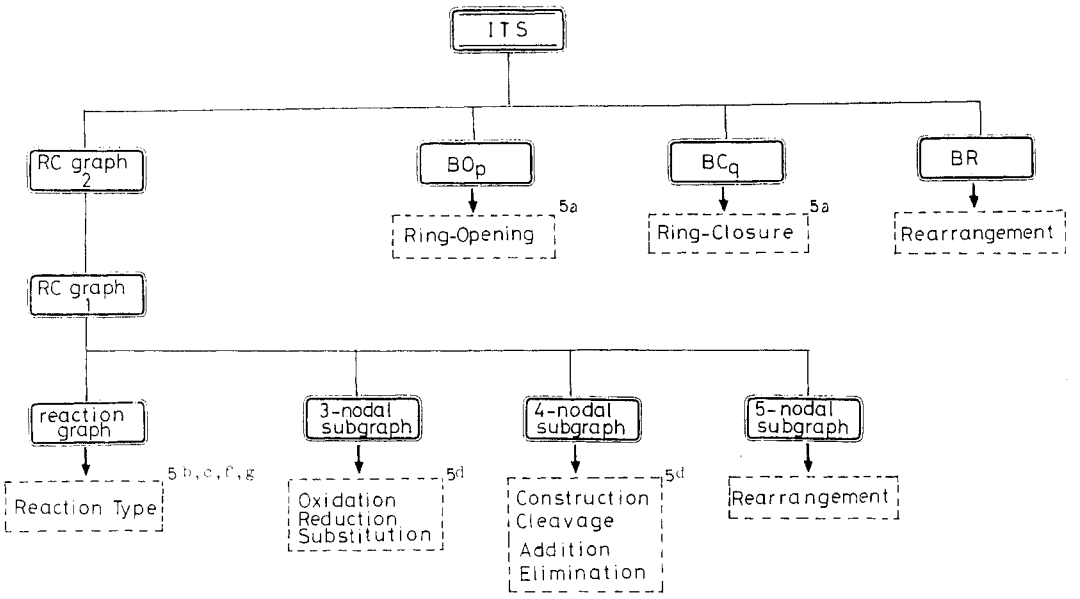


Figure 1. Various subgraphs of ITS's for characterization of organic reactions: solid-line boxes indicate subgraphs; dashed-line boxes indicate perceived features.

Table I. ITS Rings Appearing in ITS **1**

ring	nodes	ring size	$a + b \neq 0$; $a \neq 0$	$a + b = 0$	$a = 0$	ring type
1	1-2+8-13+14-15+7-6+1	8	1	4	3	RS
2	1-2-3-4-5-1	5	4	1	0	BO_1
3	1-8-2-3-4-5-1	6	5	0	1	BC_1
4	1-2-8-1	3	1	1	1	BR
5	8-9-10-11-12-8	5	5	0	0	invariant
6	1-8-13-14-15-7-6-1	7	2	3	2	trivial

Table II. Types of ITS Rings

ring type	ring size	type of ITS bond (a b)		
		$a + b \neq 0$; $a \neq 0$	$a + b = 0$	$a = 0$
BO_p	m	$m - p$	p	0
BC_q	m	$m - q$	0	q
BR	m	$m - 2$	1	1
invariant	m	m	0	0

indicated that the reaction is a rearrangement, i.e., 1,2-migration.

Rings 2–4 reveal that the pinacol rearrangement represented by ITS **1** implies a ring opening of a five-membered ring, a ring closure to a six-membered ring, and total rearrangement reaction. Ring 5 is an invariant ring, and ring 6 is a trivial one.

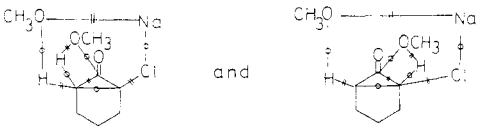
In general, BO_p, BC_q, and BR are defined as shown in Table II, where integer m is ring size, integer p is the number of ITS bonds (a b) of $a + b = 0$, and integer q is the number of ITS bonds (a b) of $a = 0$. Hereafter, we define type I rearrangements as reactions characterized by their BR's. In the following section, we discuss various types of BR's.

BRIDGE OF REARRANGEMENT: PERCEPTION OF TYPE I REARRANGEMENT

A. Rearrangement Reaction Characterized by Three-Membered BR'S. Table III collects representative 1,2-migration reactions, each of which involves a carbon center shifted from one to another carbon atom. All of them are characterized by BR's that are the subgraphs of the respective ITS's. The three-membered BR's are represented, in a more abstract fashion, by a general bridge of rearrangement (GBR **1**), in which the (1 −1) bond may be replaced by a (2 −2) or (3 −3) bond, and the (0 +1) bond may be substituted by a (0 +2) or (0 +3) bond.

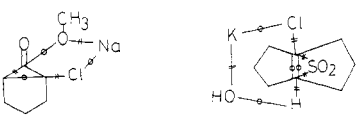
GBR 1

The Favorskii rearrangement is characterized by a cyclopropane intermediate.[25] This feature is represented by ITS **20a** and **20b**, which are in accordance with scrambling of ^{14}C



20a 20b

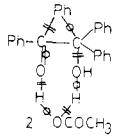
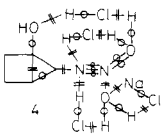
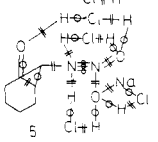
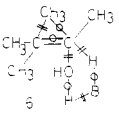
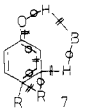
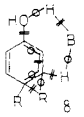
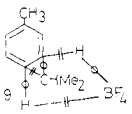
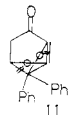
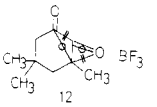
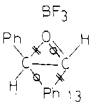
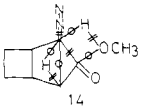
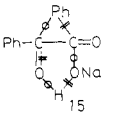
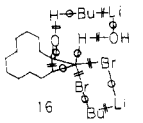
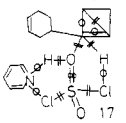
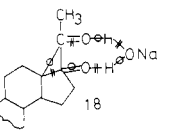
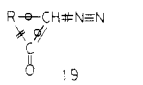

21 22


20' 23

at the chlorinated carbon.[26] A three-membered BR appears in both ITS's (**20a** and **20b**) as a subgraph. Since ITS **20a** (or **20b**) is of two-strings, it can be divided into two primary ITS's (**21** and **22**). Although ITS **21** is characterized by BC_1 and ITS **22** by BO_1, there appear to be no BR's. Thus, the Favorskii rearrangement is a rearrangement which is divided by a ring closure followed by ring opening.

An alternative ITS (**20'**) is rejected by mechanistic studies of the Favorskii rearrangement. The rearrangement of entry

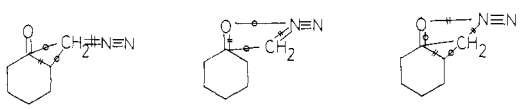
Table III. Rearrangement Reactions Characterized by Three-Membered BR's

entry	ITS	BR	name	ref
2			pinacol rearrangement	8
3			ring enlargement of cyclopropanol	9
4			Demjanov rearrangement	10
5			Tiffeneu–Demjanov rearrangement	11
6			retro-pinacol rearrangement	11
7			dienone–phenol rearrangement	12
8			dienol–benzene rearrangement	13
9			isomerization of aromatics	14
10			photochemical enone rearrangement	15
11			photochemical dienone rearrangement	16
12			rearrangement of an epoxide	17
13			rearrangement of an epoxide	18
14			rearrangement of an diazoketone	19
15			benzil–benzilic acid rearrangement	20
16			ring enlargement of a macrocycle	21
17			cyclopropane to cyclobutane rearrangement	22
18			D-homosteroid rearrangement	23
19			Wolff rearrangement	24

14 is a two-string reaction and resembles the case of the ITS (**20a** or **20b**).

In analogy to the Favorskii rearrangement, the reaction represented by ITS **23** is called the Ramberg–Buchlung "rearrangement".[27] However, no BR's appear in ITS **23** or the corresponding primary ITS's. Hence, this reaction is not a rearrangement reaction from the criterion of the present ITS approach.

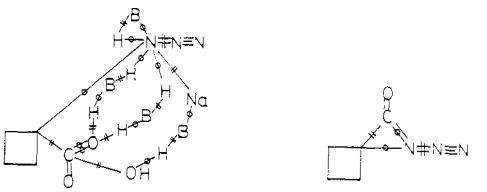
The ring enlargement of cyclohexanone with diazomethane[28] is denoted by ITS **24**, which has no BR's as subgraphs. But



24 25 26

ITS **24** is constructed by recombination of ITS's **25** and **26**, which are the corresponding primary reactions. ITS **26** contains a BR, and the ring enlargement implies a hidden rearrangement.

Table IV shows three-membered BR's containing heteroatoms. Entries 27 and 28 are two types of the Beckmann rearrangement. Entries 29–31 are the Curtius and related rearrangements. The Schmitt reaction[41] represented by ITS **39** is totally a substitution at the ring carbon of a cyclobutane. However, the essential step is the Curtius rearrangement represented by ITS **40**.



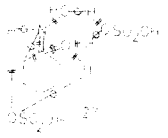
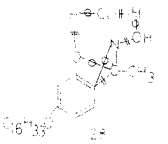
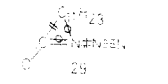
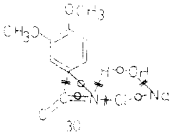
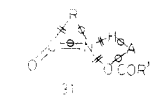
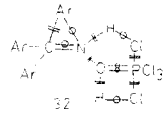
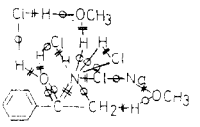
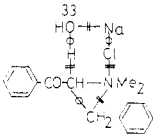
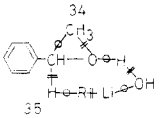
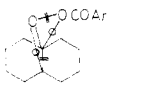
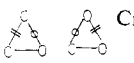
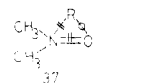
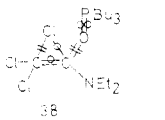
39 B = OSO_2OH 40

The Neber rearrangement denoted by ITS **33** has a BR. However, this reaction is divided into BC_1 and BO_1 in a way similar to that described in the case of the Favorskii rearrangement.

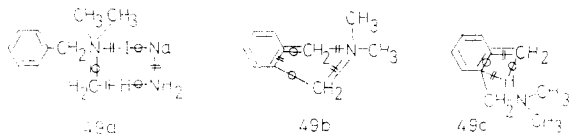
B. Rearrangement Reactions Characterized by Four-Membered BR'S. Table V summarizes rearrangement reactions characterized by four-membered BR's. Entry 41 is the Chapman rearrangement, whose BR represents migration from the oxygen to the nitrogen terminal. Entry 44 should be followed by aromatization, which is omitted for briefness.

C. Rearrangement Reactions Characterized by Five-Membered BR'S. Table VI shows three examples of this category.

Table IV. Rearrangement Reactions Characterized by Three-Membered BR's Containing Heteroatoms

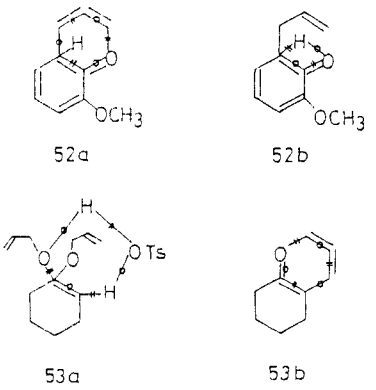
entry	ITS	BR	name	ref
27			Beckmann rearrangement	29
28			Beckmann rearrangement	30
29			Curtius rearrangement	31
30			Hoffmann rearrangement	32
31			Lossen rearrangement	33
32			Stieglitz rearrangement	34
33			Neber rearrangement	35
34			Stevens rearrangement	36
35			Wittig rearrangement	37
36			Criegee rearrangement	38
37			Meisenheimer rearrangement	39
38				40

The Sommlet rearrangement (entry 49) can be divided into three steps, i.e., N-ylide formation (ITS **49a**), rearrangement



to the exomethylene derivative (ITS **49b**), and subsequent aromatization (ITS **49c**). The recombination of reaction strings appearing in **49a**, **49b**, and **49c** affords ITS **49** of overall Sommlet rearrangement.

D. Rearrangement Reactions Characterized by Six-Membered BR'S. Table VII summarizes this category of rearrangement reactions. The typical Claisen rearrangement represented by ITS **52** can be divided into two primary reactions, i.e., rearrangement to the keto derivative (ITS **52a**) and subsequent aromatization (ITS **52b**). These ITS's contain BR's in accord with their structural changes.



ITS **53** shows the partial deacetalization of cyclohexanone diallyl acetal (ITS **53a**) followed by the Claisen rearrangement (ITS **53b**). In this case, the keto form is stable and undergoes no further reactions. ITS's **52a**, **53b**, and **54**, have the same reaction graph, which is identical with the BR's. The Cope and oxy-Cope rearrangements (**55** and **56**) are characterized by six-membered BR's as shown in Table VII.

Table V. Rearrangement Reactions Characterized by Four-Membered BR's

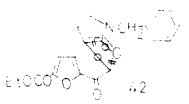
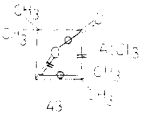
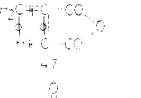
entry	ITS	BR	name	ref
41			Chapman rearrangement	42
42			spiro compound formation	43
43			rearrangement of a ketene dimer	44
44			rearrangement of a silyl compound	45
45			ring enlargement of a silyl compound	46
46			1,3-sigmatropic reaction	47
47			hydrogen migration	48
48			photochemical ring contraction	49

Table VI. Rearrangement Reactions Characterized by Five-Membered BR's

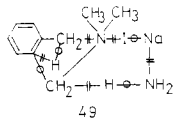
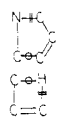
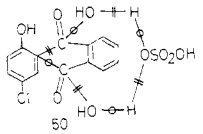
entry	ITS	BR	name	ref
49			Sommlet rearrangement	50
50			Hayashi rearrangement	51
51			ring enlargement of an epoxy ketone	52

Table VII. Rearrangement Reactions Characterized by Six-Membered BR's

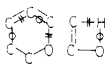
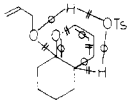
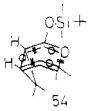
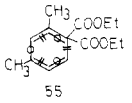
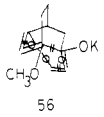
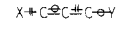
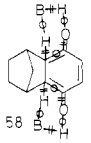
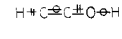
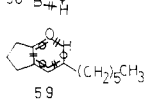
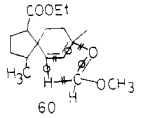
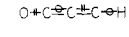
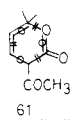
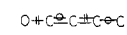
entry	ITS	BR	name	ref
52			Claisen rearrangement	53
53			Claisen rearrangement	54
54			Claisen rearrangement	55
55			Cope rearrangement	56
56			oxy-Cope rearrangement	57

Table VIII. Rearrangement Reactions Characterized by Five-Nodal Subgraphs

entry	ITS	five-nodal subgraph	name	ref
57			allylic rearrangement	58
58			keto–enol tautomerization	59
59			π-rearrangement	60
60			reductive π-rearrangement	61
61			Carroll reaction	62

EXAMINATION OF FIVE-NODAL SUBGRAPHS: PERCEPTION OF TYPE II REARRANGEMENT

There is another type of rearrangement reaction. For example, allylic rearrangement (entry 57) cannot be characterized by BR. We call these reactions type II rearrangements. However, five-nodal subgraphs collected in Table VIII are clues for the perception of this category of reactions.

CONCLUSION

The perception of type I rearrangement reactions is formulated as detection of bridges of rearrangement (BR). The type II rearrangements can be recognized by examining five-nodal subgraphs. Thus, the perception of rearrangement requires single access to the ITS file in the present ITS approach.

REFERENCES AND NOTES

(1) Dubois, J.-E. *Pure Appl. Chem.* **1981**, *53*, 1313.
(2) Roberts, D. C. *J. Org. Chem.* **1978**, *43*, 1473.
(3) Ugi, I.; Bauer, J.; Brandt, J.; Friedrich, J.; Gasteiger, J.; Jochum, C.; Schubert, W. *Angew. Chem., Int. Ed. Engl.* **1979**, *19*, 111.
(4) Hendrickson, J. B. *Angew. Chem., Int. Ed. Engl.* **1974**, *13*, 47.
(5) For the terminology, see the previous papers of this series: (a) Part 1. Fujita, S. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 205. (b) Part 2. Fujita, S. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 212. (c) Part 3. Fujita, S. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 224. (d) Part 4. Fujita, S. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 231. (e) Part 5. Fujita, S. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 238. Parts 6–8 of this series appear as the preceding papers in this issue.
(6) Deletion of in-bonds (⊖—).
(7) Deletion of out-bonds (—‖—).
(8) Bachmann, W. E. *Organic Syntheses*; Wiley: New York, 1943; Coll. Vol. 2, p 73.
(9) Denis, J. N.; Conia, J. M. *Tetrahedron Lett.* **1972**, 4593.
(10) Smith, P. A. S.; Baer, D. R. *Org. React.* **1960**, *11*, 157.
(11) Zelinsky, N.; Zelikow, J. *Ber.* **1901**, *34*, 3249.
(12) von Auwers, K.; Ziegler, K. *Justus Liebigs Ann. Chem.* **1921**, *425*, 217.
(13) Plieninger, H.; Keilich, G. *Angew. Chem.* **1956**, *68*, 618. *Ber.* **1958**, *91*, 1891.
(14) Pearson, E. E.; Wysong, R. D.; Finkel, J. M. *Organic Syntheses*; Wiley: New York, 1973; Coll. Vol. 5, p 332.
(15) Zimmerman, H. E.; Wilson, J. W. *J. Am. Chem. Soc.* **1964**, *86*, 4036.
(16) Zimmerman, H. E.; Schuster, D. I. *J. Am. Chem. Soc.* **1962**, *84*, 4527.
(17) Ryerson, G. D.; Wasson, R. L.; House, H. O. *Organic Syntheses*; Wiley: New York, 1963; Coll. Vol. 4, p 957.
(18) Reif, D. J.; House, H. O. *Organic Syntheses*; Wiley: New York, 1963; Coll. Vol. 4, p 375.
(19) Wiberg, K. B.; Furtek, B. L.; Olli, L. K. *J. Am. Chem. Soc.* **1979**, *101*, 7675.
(20) Ballard, D. S.; Dehn, W. M. *Organic Syntheses*; Wiley: New York, 1941; Coll. Vol. 1, p 89.
(21) Taguchi, H.; Yamamoto, H.; Nozaki, H. *J. Am. Chem. Soc.* **1974**, *96*, 6510.
(22) Trost, B. M.; Keeley, D. E.; Arndt, H. C.; Bogdamowicz, M. J. *J. Am. Chem. Soc.* **1977**, *99*, 3088.
(23) Ruzicka, L.; Meldahl, H. *Helv. Chim. Acta* **1938**, *21*, 1760; **1939**, *22*, 421.
(24) Bachmann, W. E.; Struve, W. S. *Org. React.* **1942**, *1*, 38.
(25) (a) Goheen, D. W.; Vanghan, W. R. *Organic Syntheses*; Wiley: New York, 1963; Coll. Vol. 4, p 594. (b) Kende, A. S. *Org. React.* **1960**, *11*, 261.
(26) Gould, E. S. *Mechanism and Structure in Organic Chemistry*; Henry Holt: New York, 1959; p 642.
(27) Paqutte, L. A. *Org. React.* **1977**, *25*, 1.
(28) de Boer, T. J.; Backer, H. J. *Organic Syntheses*; Wiley: New York, 1963; Coll. Vol. 4, p 225.
(29) Eck, J. C.; Marvel, C. S. *Organic Syntheses*; Wiley: New York, 1943; Coll. Vol. 2, p 76.
(30) Fujita, S.; Koyama, K.; Inagaki, Y. *Synthesis* **1982**, 68.
(31) Allen, C. F. H.; Bell, A. *Organic Syntheses*; Wiley: New York, 1955; Coll. Vol. 3, p 846.
(32) Buck, J. S.; Ide, W. S. *Organic Syntheses*; Wiley: New York, 1943; Coll. Vol. 2, p 44.
(33) Renfrow, W. B., Jr.; Hauser, C. R. *J. Am. Chem. Soc.* **1937**, *59*, 2308.
(34) Stiegliz, J.; Leech, P. N. *Ber.* **1913**, *46*, 2147. *J. Am. Chem. Soc.* **1914**, *36*, 272.
(35) Baumgarten, H. E.; Petersen, J. M. *Organic Syntheses*; Wiley: New York, 1973; Coll. Vol. 5, p 909.
(36) Pine, S. H. *Org. React.* **1970**, *18*, 403.
(37) Wittig, G.; Löhmann, L. *Justus Liebigs Ann. Chem.* **1942**, *550*, 260.
(38) Criegee, R.; Kaspar, R. *Justus Liebigs Ann. Chem.* **1948**, *560*, 127.
(39) Meisenheimer, J. *Ber.* **1919**, *52*, 1667.

(40) Speziale, A. J.; Freeman, R. C. *Organic Syntheses*; Wiley: New York, 1973; Coll. Vol. 5, p 387.
(41) Werner, N. W.; Casnova, J., Jr. *Organic Syntheses*; Wiley: New York, 1973; Coll. Vol. 5, p 273.
(42) Shulenberg, J. W.; Archer, S. *Org. React.* **1965**, *14*, 1.
(43) Sundberg, R. J.; Pearce, B. C. *J. Org. Chem.* **1985**, *50*, 425.
(44) Hasek, R. H.; Clark, R. D.; Mayberry, G. L. *Organic Syntheses*; Wiley: New York, 1973; Coll. Vol. 5, p 456.
(45) Kita, M.; Yoshida, H.; Sakurai, H. *J. Am. Chem. Soc.* **1985**, *107*, 7767.
(46) Salaün, J. *Tetrahedron Lett.* **1984**, *25*, 1269, 1273.
(47) Krow, G. R.; Reilley, J. *J. Am. Chem. Soc.* **1975**, *97*, 3837.
(48) Shriner, R. L.; Ford, S. G.; Roll, L. J. *Organic Syntheses*; Wiley: New York, 1943; Coll. Vol. 2, p 140.
(49) Crockett, G. C.; Koch, T. H. *J. Org. Chem.* **1977**, *42*, 272.
(50) Brasen, W. R.; Hauser, C. R. *Organic Syntheses*; Wiley: New York, 1963; Coll. Vol. 4, p 585.
(51) Hayashi, M. *J. Chem. Soc.* **1927**, 2516.
(52) Carlson, R. G. *J. Chem. Soc., Chem. Commun.* **1973**, 223.
(53) Allen, C. F. H.; Gates, J. W., Jr. *Organic Syntheses*; Wiley: New York, 1955; Coll. Vol. 3, p 418.
(54) Howard, W. L.; Lorrette, N. B. *Organic Syntheses*; Wiley: New York, 1961; Coll. Vol. 5, p 25.
(55) Funk, R. L.; Munger, J. D., Jr. *J. Org. Chem.* **1985**, *50*, 707.
(56) Rhods, S. J.; Raulins, N. R. *Org. React.* **1975**, *22*, 1.
(57) Evans, D. A.; Golob, A. M. *J. Am. Chem. Soc.* **1975**, *97*, 4765.
(58) DeVolfe, J. R.; Young, W. G. *Chem. Rev.* **1956**, *56*, 763.
(59) Diels, O.; Alder, K. *Ber.* **1929**, *62*, 2337.
(60) Arnold, R. T.; Smolinsky, G. *J. Am. Chem. Soc.* **1960**, *82*, 4918.
(61) Oppolzer, W. *Helv. Chim. Acta* **1973**, *56*, 1812.
(62) Carroll, M. F. *J. Chem. Soc.* **1940**, 704.

"Structure–Reaction Type" Paradigm in the Conventional Methods of Describing Organic Reactions and the Concept of Imaginary Transition Structures Overcoming This Paradigm

SHINSAKU FUJITA

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

Received October 6, 1986

The description of organic reactions is discussed from the viewpoint of a structure–reaction type (SRT) paradigm, in which structural information and reaction type are stored and manipulated more or less independently. This paradigm must be overcome in order to construct an integrated system that will support both retrieval of organic reactions and synthetic design. The concept of imaginary transition structures (ITS) is introduced as a unitary representation free from the SRT paradigm.

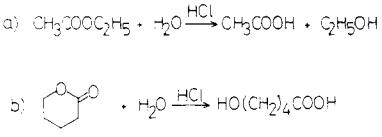
Sellow pointed out in his critical review[1] that "an important amount of success was achieved in the field of codification of structure but not in that of codification of reactions". He ascribed this to two strongly correlated reasons: "the lack of a real and efficient methodology for analyzing the theoretical aspects of reactions and the necessity of translating them from usual language to code". Although his criticism is to the point, the difficulty in describing organic reactions should be discussed more thoroughly from another point of view.

There are two types of computer systems manipulating organic reactions: systems for the retrieval of organic reactions and systems for the design of synthetic pathways. The former systems deal with information on *individual reactions*. The latter concern *reaction types* rather than individual reactions. These two types have grown separately, although both are based on information on organic reactions. In this paper, we discuss why these two types of systems have never been integrated. We reveal a structure–reaction type (SRT) paradigm that restricts the conventional methods of description of organic reactions and, as a result, hinders the integration of the systems.[2]

DUALITY IN THE DESCRIPTION OF ORGANIC REACTIONS: STRUCTURE–REACTION TYPE (SRT) PARADIGM

Let us work out a simple reaction shown in Scheme Ia. How do we describe this reaction? First, we recognize the structure of the substrate as ethyl acetate (or the corresponding coded name). Next, we discern between the substrate and the product and recognize this reaction as hydrolysis (or the corresponding code) from our own knowledge. And then we describe this as a combination of ethyl acetate and hydrolysis. In the process of recognition, we do not describe the individual reaction in itself, but we replace the description of the reaction

Scheme I



by (1) that of the structure of the substrate and (2) that of the reaction type.[3]

Figure 1 summarizes the duality in the conventional description of organic reactions. In general, structural data of a given reaction are divided into information on the structure and that on the reaction type, which are stored separately, more or less independently, in a structure file and in a reaction-type file, respectively. We call this type of duality a structure–reaction type (SRT) paradigm. This paradigm has never been formulated, so that the previous efforts to develop new representations of organic reactions resulted unintentionally in intimate combination of (1) and (2) at the utmost. There have been no attempts to overcome this.

In the SRT paradigm, retrieval of organic reactions requires dual reference or access to the structure file and to the reaction-type file as shown in the right-hand side of Figure 1. In this case, the correspondence between the two files is crucial to construct an effective system of retrieval. However, this is difficult as discussed below so long as we use conventional methods of description.

On the other hand, synthetic design is simpler than retrieval of organic reactions from the viewpoint of the SRT paradigm.[4] Once an expert extracts reaction types from the data of individual reactions, synthetic design requires only single access to the reaction-type file. In other words, structural information is *given as a target molecule* in the case of synthetic design. If our efforts are limited to synthetic designe field, the SRT

0095-2338/87/1627-0120$01.50/0