Bull. Chem. Soc. Jpn., 62, 662—667 (1989)

Formulation and Enumeration of Five-Center Organic Reactions

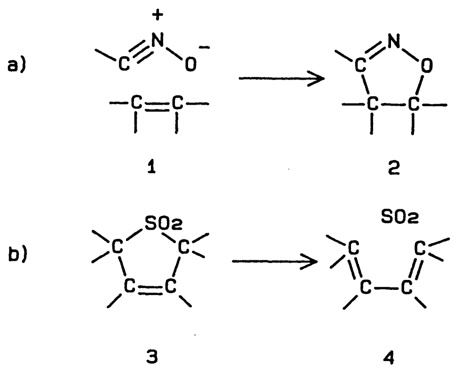
Shinsaku FUJITA
Research Laboratories, Ashigara, Fuji Photo Film Co., Ltd.,
Minami-Ashigara, Kanagawa-ken 250-01
(Received July 30, 1988)

Five-center organic reactions are formulated in the form of pentagonal reaction-center graphs (RCC's) that are subgraphs of imaginary transition structures. This formulation is based on the concept of a charge space that affords rational representations of coordinate bonds. Pólya's theorem is applied to the enumeration of these RCG's, in which the obligatory minimum valencies of vertices restrict the mode of substitution.

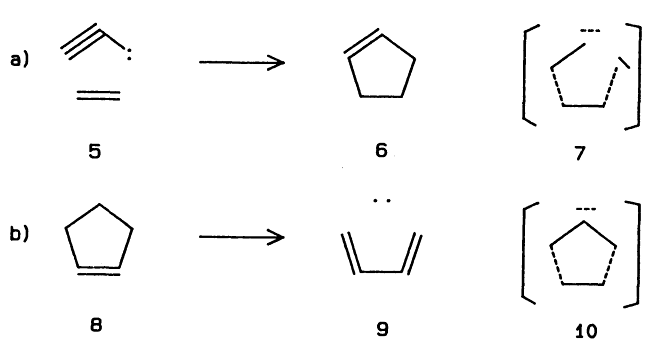
The enumeration of reaction types is important for any computer manipulation of organic reactions. In previous papers,[1] we have reported on the enumeration of generic reaction types in the form of the counting of reaction graphs (RG's) that are defined as subgraphs of imaginary transition structures (ITS's).[1b,c] Thus, we have enumerated pentagonal RG's, which represent five-center organic reactions, such as 1,3-dipolar additions and cheletropic reactions.[1c] However, the enumeration of RG's has disregarded the variation of atoms that occupy the vertices of the RG's and has, thus, been limited within the counting of modes of bond-switchings. For an effective manipulation of organic reactions, it is necessary to count more specific reaction types that define the variation of the atoms as well as the modes of bond-switching. This can be accomplished by counting the reaction-center graphs (RCG's) of the pentagonal class. In a continuation of the enumeration of hexagonal RCG's,[1a] this paper will discuss that of pentagonal RCG's.

A Charge Space for Representing Coordinate Bonds in the Imaginary Transition Structure Approach. In an enumeration of such five-center organic reactions, one of the most troublesome problems is in the treatment of coordinate bonds (Scheme 1). Although many approaches have been devoted to this problem, none of them have ever resulted in a comprehensive solution. For example, Ugi et al.[2] have expressed reactions (1a) and (1b) with graphs (**7** and **10**, respectively) in the most abstract fashion (Scheme 2). They have ascribed the two graphs (**7** and **10**) to irreducible reaction matrices (R-matrices) that are presumed to represent five-membered cyclic processes. However, expression (**7**) disregards the presence of a coordinate bond, which is in turn replaced by lone pairs on the terminal atoms. In other words, graph (**7**) and the corresponding R-matrix indicate an acyclic process in their strict sense. Although they had intended to classify **7** and **10** into the same class, this classification requires the presupposition of an implicit invariant bond. If the implicit invariant bond is neglected, just as the graph (**7**) indicates, there would be no good reason to classify reactions **7** and **10** into the same class.

Hendrickson's representations for five-center reactions[3] are equivalent to the formulas (**7** and **10**) in an abstract fashion. Hence, his method has the same drawback as that mentioned as above, although it claims that both of the two formulas represent pericyclic five-center reactions. Vladutz[4] has presented a participating charge with a special symbol placed on an atom. Zefirov et al.[5] have treated a charge as an independent node (vertex). These methods permit an ionic representation as well as the covalent one, where both of the two representations are legitimately alternative forms and, in other words, are not integrated. Brandt el al.[6] have adopted Ugi's R matirices along with the same representation of charges as Vladutz's one. Although these approaches have described the appearance and disappearance of charges, they have paid little attention to the difficulty concerned with coordinate bonds and the dual character of a bond (i.e. the covalent vs. ionic character). Dubois et al.[7] have reported the concept of chromatic transfographs,



Scheme 1.



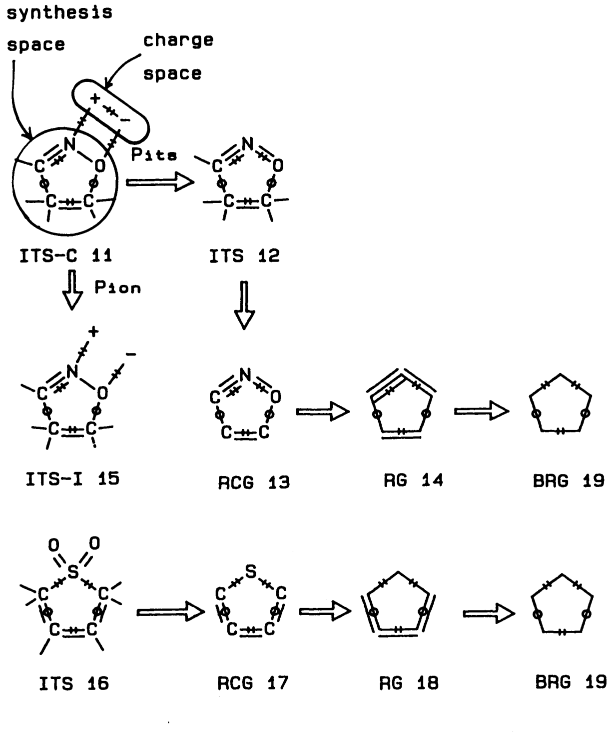
Scheme 2.

which also pays no attention to charges and abnormal valencies.

In order to settle the problem, the ITS approach has introduced a charge space that is attached to a synthesis space (Scheme 3).[1h] The synthesis space contains all nodes that represent atoms or groups participating in a reaction. The charge space consists of charges that are regarded as independent nodes. These nodes (atoms and charges) are linked by par-bonds (——), in-bonds (—⊖—), and out-bonds (—╫—) in accord with any structural changes during a reaction. The resulting structural formula is called a *full ITS* or an *ITS with charges.* The concept of charge space thus allows us to described a 1,3-dipolar reaction (Scheme 1a) in light of a full ITS (**11**). This type of representation is an integrated representation of covalent and ionic characters (Scheme 3). Thus, the ionic representation (e.g. N^+-O^-) of a coordinate bond and the covalent counterpart (e.g. N=O) are interconvertible to one another via the full ITS (e.g. **11**).

The corresponding ITS (**12**) is obtained by projecting to an ITS (P_{its}), which is a deletion of interspace bonds[8] and of charges and a simultaneous attachment of the intraspace bonds[9] to the respective edges. A corresponding ionic representation (**15**) can be generated by projecting to an ionic ITS (P_{ion}) that deletes the bonds of a charge space.

The graphic representations collected in Scheme 3 may be unfamiliar to organic chemists who used to discuss organic reactions in terms of the usual reaction diagrams. To convert an ITS representation into a reaction diagram, projections to the starting stage and to the product stage (PS and PP) are introduced. The PS operation represents a deletion of all in-bonds. This operation regenerates the starting stage. The PP operation, on the other hand, represents a deletion of all out-bonds, the operation of which creates the product stage.



Scheme 3.

In the enumeration procedures of the ITS approach, the covalent representation of a bond is preferred rather than the ionic one.[1c] This methodology can be rationalized by the indroduction of a charge space and by the defining of the P_{its} and P_{ion} operations.[1h] The preference of the covalent character has an advantage in that this overcomes an apparent difference provided by the dual representation of the bond character. Note, for example, that though the ionic repesentation ($N^+(=O)-O^-$) of a nitro group indicates the unequality of the two oxygen atoms in the literal sense, the covalent representation ($N(=O)_2$) equalifies them.

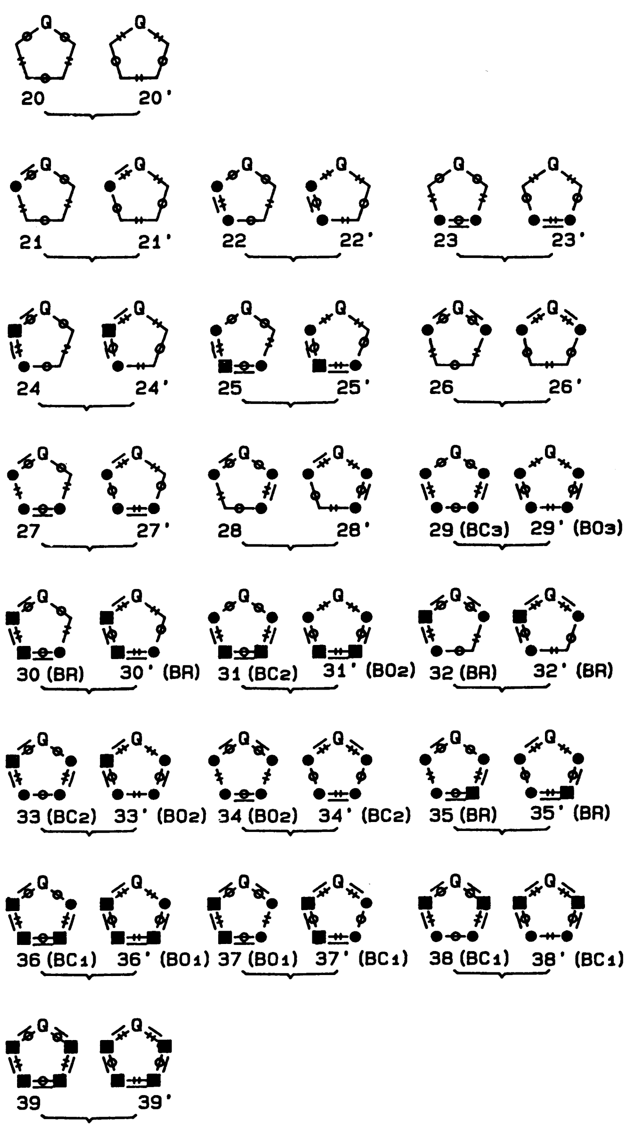
In light of the ITS's extended in this section, a general treatment of five-center reactions is available. As shown in Scheme 3, reaction (1b) is represented by an ITS (**16**) in the usual fashion, since it contains no charge space. Reactions (1a) and (1b) (Scheme 1) produce the corresponding RCG's (**13** and **17**) and RG's (**14** and **18**). The latter pair affords a common basic reaction graph (BRG **19**), which logically indicates that the two reactions belong to the same category of pentagonal class.[10]

Enumeration of Pentagonal RCG's. The present method that enumerates RCG's consists of the following steps:[11,12] (1) finding the permutation group of a parent RG, (2) considering the obligatory minimum valency (OMV) of each vertex (node), (3) obtaining the corresponding cycle index, and (4) substituting the variables of the cycle index by appropriate atom-figure inventories to afford an RCG-counting polynomial as a generating function.

Table 1 lists parent RG's of the pentagonal class.[1c] There are no self-reaction pairs[13] in this series. Each of the RG's has one vertex to which two out-bonds (or two in-bonds) are incident. We call this a valency-changed position (or vertex) because of its chemical meaning. For simplicity of discussions, we suppose that an atom Q always occupies the valency-changed position of each parent RG (Table 1). Several examples for atoms Q are collected in Table 2. The expressions of expanded valencies listed in Table 2 have a sound basis by means of a charge space. The present approach is capable of producing both the covalent representation and the ionic one with the appropriate projections (Scheme 3). The ITS approach permits a quinquivalent nitrogen, a quadrivalent oxygen and so on in the level of ITS's projected from full ITS's.

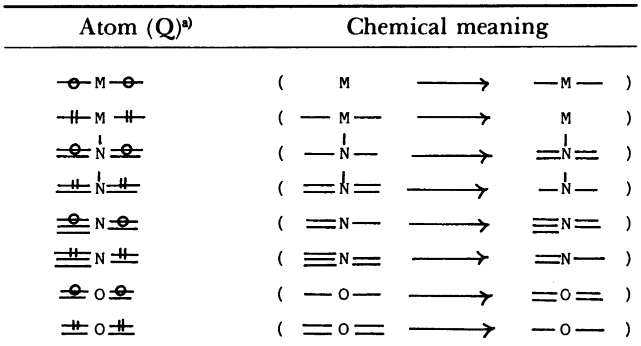
Positions other than a valency-changed atom undergo ordinary interconversions in which one in-bond and one out-bond are attached to each of the positions.

Table 1. Parent Reaction Graphs of Pentagonal Class That Have 0 to 5 Single Par-Bonds and Are Occupied by an Atom Q[a–c)]



a) The positions marked by ■, ●, and none have minimum valencies of v=3, 2, and 1, respectively. b) Each couple linked with a brace is a reaction pair (an RG pair). There are no self-reaction pairs. c) BO_p is a bridge of ring opening of order p. BC_q is a bridge of ring closure of order q. BR is a bridge of rearrangement.

Table 2. Atoms (Q) Capable of Valency Change



a) The symbol (M) indicates a metal.

Table 3. Cycle Indices for the Parent RG's of Pentagonal Class Preoccupied by an Atom Q

RG[a)]	Symmetry	Cycle index[b)]	Eq.
20,20′	C_2	G: $(1/2)(c_1^4+c_2^2)$	1
21,21′	C_1	G: $b_1c_1^3$	2
22,22′	C_1	G: $b_1^2c_1^2$	3
23,23′	C_2	G: $(1/2)(b_1^2c_1^2+b_2c_2)$	4
24,24′	C_1	G: $a_1b_1c_1^2$	5
25,25′	C_1	G: $a_1b_1^2c^1$	6
26,26′	C_2	G: $(1/2)(b_1^2c_1^2+b_2c_2)$	7
27,27′	C_1	G: $b_1^3c_1$	8
28,28′	C_1	G: $b_1^3c_1$	9
29,29′	C_2	G: $(1/2)(b_1^4+b_2^2)$	10
30,30′	C_1	G: $a_1^2b_1c_1$	11
31,31′	C_2	G: $(1/2)(a_1^2b_1^2+a_2b_2)$	12
32,32′	C_1	G: $a_1b_1^2c_1$	13
33,33′	C_1	G: $a_1b_1^3$	14
34,34′	C_2	G: $(1/2)(b_1^4+b_2^2)$	15
35,35′	C_1	G: $a_1b_1^3$	16
36,36′	C_1	G: $a_1^3b_1$	17
37,37′	C_1	G: $a_1^2b_1^2$	18
38,38′	C_2	G: $(1/2)(a_1^2b_1^2+a_2b_2)$	19
39,39′	C_2	G: $(1/2)(a_1^4+a_2^2)$	20

a) The parent RG's are collected in Table 1. b) G: a cycle index for counting RCG's.

Their obligtory minimum valencies (OMV) are v=1 (unmarked), v=2 (marked with ●), or v=3 (marked with ■) in accord with the substitution of par-bonds. The unmarked positions are substituted by any atoms that have a valency of larger than 1. We choose carbon, hydrogen, oxygen, and nitrogen atoms as substituents in order to simplify the following discussions. Any position marked with ● (v=2) is capable of taking C, N, or O, but incapable of taking a hydrogen atom. Each position of v=3 (■) is capable of having C or N as a substituent.

The OMV requires an extension of Pólya's theorem, in which positions of different OMV's take different atom-figure inventories.[1n)] Table 3 summarizes the cycle indices corresponding to the parent RG's of Table 1, where variables a_r, b_r, and c_r are assigned to the vertices marked by ■, ●, and none, respectively.[14)]

If the variables of the cycle indices (Table 3) are substituted by atom-figure inventories (Eqs. 21 to 23), RCG-counting polynomials are obtained as generating functions:[15)]

$$a_r = 1 + u^r, \tag{21}$$

$$b_r = 1 + u^r + v^r, \tag{22}$$

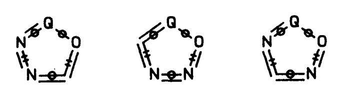
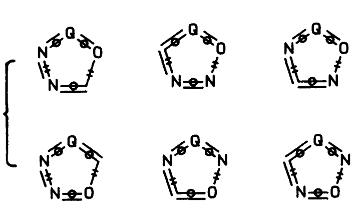
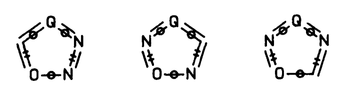
and

$$c_r = 1 + u^r + v^r + w^r. \tag{23}$$

Table 4. Enumeration of Pentagonal RCG's with an Atom Q

Parent RG			**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**
			20′	**21′**	**22′**	**23′**	**24′**	**25′**	**26′**	**27′**	**28′**	**29′**	**30′**	**31′**	**32′**	**33′**	**34′**	**35′**	**36′**	**37′**	**38′**	**39′**
i	*j*	*k*	Number of RCG's																			
0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	0	0	2	4	4	2	4	4	2	4	4	2	4	2	4	4	2	4	4	4	2	4
0	1	0	2	4	4	2	3	3	2	4	4	2	2	1	3	3	2	3	1	2	1	0
2	0	0	4	6	6	4	6	6	4	6	6	4	6	4	6	6	4	6	6	6	4	6
1	1	0	6	12	12	6	9	9	6	12	12	6	6	3	9	9	6	9	3	6	3	0
0	2	0	4	6	6	4	3	3	4	6	6	4	1	1	3	3	4	3	0	1	1	0
3	0	0	2	4	4	2	4	4	2	4	4	2	4	2	4	4	2	4	4	4	2	4
2	1	0	6	12	12	6	9	9	6	12	12	6	6	3	9	9	6	9	3	6	3	0
1	2	0	6	12	12	6	6	6	6	12	12	6	2	1	6	6	6	6	0	2	1	0
0	3	0	2	4	4	2	1	1	2	4	4	2	0	0	1	1	2	1	0	0	0	0
4	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	1	0	2	4	4	2	3	3	2	4	4	2	2	1	3	3	2	3	1	2	1	0
2	2	0	4	6	6	4	3	3	4	6	6	4	1	1	3	3	4	3	0	1	1	0
1	3	0	2	4	4	2	1	1	2	4	4	2	0	0	1	1	2	1	0	0	0	0
0	4	0	1	1	1	1	0	0	1	1	1	1	0	0	0	0	1	0	0	0	0	0
0	0	1	2	3	2	1	2	1	1	1	1	0	1	0	1	0	0	0	0	0	0	0
1	0	1	6	9	6	3	6	3	3	3	3	0	3	0	3	0	0	0	0	0	0	0
0	1	1	6	9	6	3	4	2	3	3	3	0	1	0	2	0	0	0	0	0	0	0
2	0	1	6	9	6	3	6	3	3	3	3	0	3	0	3	0	0	0	0	0	0	0
1	1	1	12	18	12	6	8	2	6	6	6	0	2	0	4	0	0	0	0	0	0	0
0	2	1	6	9	6	3	2	1	3	3	3	0	0	0	1	0	0	0	0	0	0	0
3	0	1	2	3	2	1	2	1	1	1	1	0	1	0	1	0	0	0	0	0	0	0
2	1	1	6	9	6	3	4	2	3	3	3	0	1	0	2	0	0	0	0	0	0	0
1	2	1	6	9	6	3	2	1	3	3	3	0	0	0	1	0	0	0	0	0	0	0
0	3	1	2	3	2	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0
0	0	2	2	3	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	2	3	6	1	1	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	2	3	6	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	2	2	3	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	2	3	6	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
0	2	2	2	3	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 5. RCG Isomers with N_2O Based on **36**, **37**, and **38**

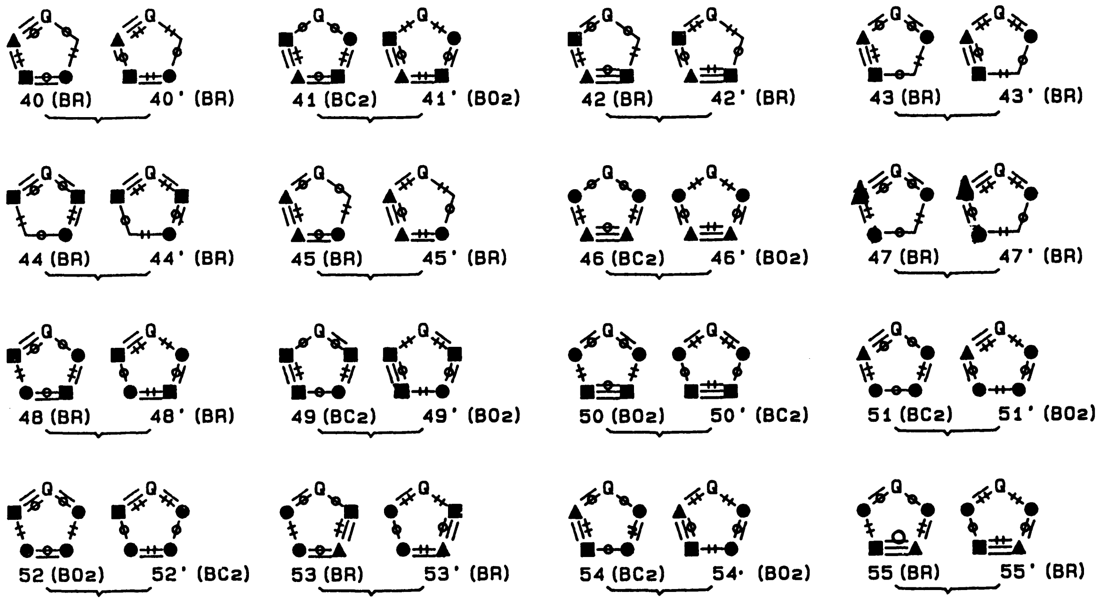
Parent RG	RCG with N_2O
36	
37	
38	

Equation 21 is an atom-figure inventory for counting the modes of C and N substitution, where a variable for C is not expressed explicitly. Equation 22 enumerates the modes of C, N, and O substitution. The modes of C, N, O, and H substitution are counted by Eq. 23. The coefficients of the RCG-counting polynomial series are collected in Table 4, in which each column is in accord with a given parent RG, each line indicates the powers (i, j, and k) and their intersection is the coefficient of the $u^i v^j w^k$. This value is the number of RCG isomers with i nitrogens, j oxygens, and k hydrogens ($N_iO_jH_k$).

Table 5 shows RCG isomers derived form parent RG's (**36**, **37**, and **38**) by N_2O substitution. The number of RCG isomers from **36** is 3, which is found at column **36** and line i=2, j=1, and k=0 in Table 4. The number of RCG isomers derived from **37** and **38** are also listed in Table 4.

Table 6 collects parent RG's in which a pentagonal basic reaction graph (BRG **19**) has one double par-bond and two single par-bonds. One Q atom preoc-

Table 6. Parent Reaction Graphs of Pentagonal Class That Have One Double Par-Bond and Two Single Par-Bonds and Are Occupied by an Atom Q[a-c)]



a) The positions marked by ▲, ■, ●, and none have minimum valencies of v=4, 3, 2, and 1, respectively. b) Each couple linked with a brace is a reaction pair (an RG pair). c) BO_p is a bridge of ring opening of order p. BC_q is a bridge of ring closure of order q. BR is a bridge of rearrangement.

Table 7. Cycle Indices for the Parent RG's of Pentagonal Class Preoccupied by an Atom Q

RG[a)]	Symmetry	Cycle index[b)]	Eq.
40,40′	C_1	G: $a_1b_1c_1$	24
41,41′	C_1	G: $a_1^2b_1$	25
42,42′	C_1	G: $a_1^2c_1$	26
43,43′	C_1	G: $a_1b_1c_1$	27
44,44′	C_1	G: $a_1^2b_1c_1$	28
45,45′	C_1	G: b_1c_1	29
46,46′	C_1	G: b_1^2	30
47,47′	C_1	G: $b_1^2c_1$	31
48,48′	C_1	G: $a_1^2b_1^2$	32
49,49′	C_1	G: $a_1^3b_1$	33
50,50′	C_2	G: $(1/2)(a_1^2b_1^2+a_2b_2)$	34
51,51′	C_1	G: b_1^3	35
52,52′	C_1	G: $a_1b_1^3$	36
53,53′	C_1	G: $a_1b_1^2$	37
54,54′	C_1	G: $a_1b_1^2$	38
55,35′	C_1	G: $a_1b_1^2$	39

a) The parent RG's are collected in Table 6. b) G: a cycle index for counting RCG's.

cupies the valency-changed position of each of the RG's. The other positions are discriminated by their OMV's (i.e. ▲ for v=4, ■ for v=3, ● for v=2, and none for v=1).

Table 7 summarizes the cycle indices corresponding to the respective RG's of Table 6. Each of the cycle indices contains variables selected from a_r (for v=3), b_r (for v=2), and c_r (for v=1). The position marked by ▲ can be substituted only by a carbon atom and, hence, takes no variables in the present treatment. If the variables (a_r, b_r, and c_r) of each cycle index (Table 7) are substituted by the atom figure inventories (Eqs. 21 to 23), the corresponding RCG-counting polynomial is obtained as a generating function for each case. The coefficient of the term $u^iv^jw^w$ in the RCG-counting polynomial indicates the number of RCG isomers having $N_iO_jH_k$.

Table 8. RCG Isomers with One Oxygen on **34′**, **50′**, and **52′**

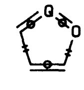
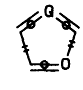
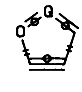
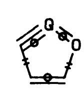
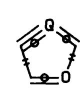
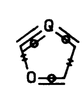
Parent RG	RCG isomers		
34′			
50′			
52′			

Table 8 compares the RG's (**34′**, **50′**, and **52′**) substituted by one oxygen. The number of isomeric RCG's based on each of the RG's emerges as the coefficient of term v in an RCG-counting polynomial derived from Eq. 15, 34, or 36.

Conclusion

Five-center organic reactions are formulated in the form of pentagonal reaction-center graphs (RCG's). These RCG's are enumerated by extending Pólya's theorem, so that different variables are assigened to the vertices of each parent reaction graph (RG) in accord with the obligatory minimum valencies of the vertices.

References

1) a) S. Fujita, *J. Chem. Inf. Comput. Sci.*, **26**, 205 (1986). See also *Chem. Eng. New.* (*Sept. 29*), 75 (1986). b) S. Fujita, *J. Chem. Inf. Comput. Sci.*, **26**, 221 (1986). c) S. Fujita, *J. Chem. Inf. Comput. Sci.*, **26**, 224 (1986). d) S. Fujita, *J. Chem. Inf. Comput. Sci.*, **26**, 231 (1986). e) S. Fujita, *J. Chem. Inf. Comput. Sci.*, **26**, 238 (1986). f) S. Fujita, *J. Chem. Inf. Comput. Sci.*, **27**, 99 (1987). g) S. Fujita, *J. Chem. Inf. Comput. Sci.*, **27**, 104 (1987). h) S. Fujita, *J. Chem. Inf. Comput. Sci.*, **27**, 111 (1987). i) S. Fujita, *J. Chem. Inf. Comput. Sci.*, **27**, 115 (1987). j) S. Fujita, *J. Chem. Inf. Comput. Sci.*, **27**, 120 (1987). k) S. Fujita, *J. Chem. Inf. Comput. Sci.*, **28**, 1 (1988). l) S. Fujita, *J. Chem. Inf. Comput. Sci.*, **28**, 78 (1988). m) S. Fujita, *J. Chem. Soc., Perkin Trans. 2*, **1988**, 598. n) S. Fujita, *Bull. Chem. Soc. Jpn.*, **61**, 4189 (1988).

2) J. Bauer, R. Herges, E. Fontain, and I. Ugi, *Chimia*, **39**, 43 (1985).

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

4) G. Vladutz, "Modern Approaches to Chemical Reaction Searching," ed by P. Willet, Gower Publishing, England (1986), Chap. 15.

5) a) N. S. Zefirov and S. S. Tratch, *Zhur. Org. Khim.*, **20**, 1121 (1984); b) N. S. Zefirov, S. S. Tratch, and G. A. Gamziani, *Zhur. Org. Khim.*, **22**, 1341 (1986).

6) a) J. Brandt and K. Stadler, "Modern Approaches to Chemical Reaction Searching," ed by Gower Publishing Co., Ltd., England (1986), Chap. 16. b) J. Brandt, J. Bauer, R. M. Frank, A. von Scholley, *Chem. Scripta*, **18**, 53 (1981).

7) a) J.-E. Dubois, G. Sicouri, and R. Ricchiottino, "Modern approaches to Chemical Reaction Seaching," Gower Publishing, ed by P. Willet, England (1986), Chap. 17. b) G. Sicouri, Y. Sobel, R. Picchiottino, and J.-E. Dubois, *C. R. Acad. Sc. Paris*, **299**, 523 (1984).

8) An interspace bond is defined as a linkage between a synthesis space and a charge space. See Ref. 1h.

9) An intraspace bond is defined as a bond which links two nodes of the same space. See Ref. 1h.

10) The RCG's, RG's, and BRG's represent the levels of a reaction hierarchy in ascending order of generality.

11) For the enumeration of organic compounds, see a) D. H. Rouvray, *Chem. Soc. Rev.*, **3**, 355 (1974). b) "Chemical Applications of Graph Theory," ed by A. T. Balaban, Academic Press, London (1976). c) N. L. Biggs, E. K. Lloyd, and R. J. Wilson, "Graph Theory 1736—1936," Clarendon Press, Oxford (1976), Chap. 4.

12) a) G. Pólya, *Acta Math.*, **68**, 145 (1937). b) G. Pólya, R. E. Tarjan, and D. R. Woods, "Notes on Introductory Combinatorics," Birkhaeuser Boston Inc. (1983), Chapt 6. c) C. L. Liu, "Introduction to Combinatorial Mathematics," McGraw-Hill (1969).

13) A. reaction pair is a pair of RG's that are interconvertible by an exchange operation between in-bonds and out-bonds. A self reaction pair is an RG that is invariant in the operation.

14) For the derivation of cycle indices, see Ref. 1n. It should be noted that, since variables a_r, b_r, and c_r are assigned in accord with OMV's, unequivalent vertices can take the same variable.

15) The substitution of carbon atoms is presumed implicitly in all cases.