Enumeration, Coding, and Complexity of Linear Reaction Mechanisms

Ekaterina Gordeeva

Molecular Design Ltd., San Leandro, California 94577

Danail Bonchev*,† and Dimitar Kamenski

Higher Institute of Chemical Technology, Burgas, Bulgaria

Oleg N. Temkin

Lomonosov Institute of Fine Chemical Technology, Moscow, Russia

Received May 18, 1993®

All topologically distinct linear mechanisms (i.e., mechanisms containing one reaction intermediate on both the left-hand and right-hand sides) of chemical reactions involving up to 16 elementary steps, up to 12 intermediates, and up to 6 reaction routes were computer enumerated and generated. The methodology previously developed for classifying, coding, and analyzing the complexity of such mechanisms was further developed. The complexity analysis of the topological structure of these mechanisms is extended here to all 390 mechanisms that incorporate 4 reaction routes and up to 6 reaction intermediates; these mechanisms are presented with their kinetic graphs, codes, and complexity indexes. Topological patterns that increase or preserve complexity were analyzed in detail and generalized in a complexity flow chart of potential use in the computerized elucidation of reaction mechanisms.

I. INTRODUCTION

The rapid increase in the mechanistic complexity of chemical reactions during the past few decades has led to numerous attempts to systematize or classify reaction mechanisms. Empirical schemes are of limited importance for such aims. It is not surprising then that the first studies along this avenue have been based on more rigorous mathematical formalisms. Sellers[1–4] used group theory to enumerate and generate the mechanisms that emerge for synthesis and substitution reactions. The stability approach of Clarke,[5,6] the works of Snagovskii and Ostrovskii,[7] Barone et al.,[8–10] Zefirov and Trach,[11,12] Brouk and Temkin,[13] and others contributed to these developments. Sinanoğlu and Lee[14–16] proposed network-based methodology for computer-assisted synthesis design. Very recently, Sinanoğlu[17] made use of general networks and topology for the systematic generation of mechanisms and reaction pathways.

Another approach based on graph theory has been developed by Temkin, Bonchev, and others.[18–26] Unlike most of the above-mentioned studies, which proceed from the ***chemical*** information on reaction mechanism (types of reactions, number and type of reactants, etc.), this approach introduced the concept for ***reaction mechanism topology***. The topological component of a reaction mechanism mirrors the interrelations within the space of reaction intermediates, including the number and kind of reaction-route interconnections. This formalism makes use of the cyclic graphs introduced in 1965 by Temkin;[27,28] we have termed these graphs ***kinetic graphs*** **(*KGs*)**. The KG vertexes represent reaction intermediates only, while edges represent the intermediate interconversions (elementary reactions). Cycles in KGs correspond to reaction routes (independent stoichiometric equations). When all reaction steps are reversible, the KGs are simple graphs. Digraphs are useful when irreversible reaction steps take place.

Any number of reagents or products may be associated with any of the KGs. Then, by a systematic increase in the number of reactants, one can combine the topological and chemical information on reaction mechanisms and make their enumeration complete.

This methodology was intensively used for ***linear reaction mechanisms***, which incorporate one intermediate left-hand side and right-hand side of each reaction step: $X_i \rightleftarrows X_j$. Hierarchical classification and code[18,20] has been developed for this large group of reaction mechanisms, as well as methods for evaluating their complexity.[19,21] The analysis was recently extended to nonlinear mechanisms.[23,25] In the present paper, we conclude our studies on linear mechanisms by discussing their computer enumeration, an improvement in their classification and coding, and further complexity analysis.

II. CLASSIFICATION, CODING, AND ENUMERATION OF LINEAR MECHANISMS

Proceeding from the one-to-one correspondence between linear mechanisms and KGs, we proposed a hierarchical classification of these reaction mechanisms in an earlier paper.[20] However, the enumeration of the linear mechanisms and their computer storage and retrieval indicated the need for some changes in both the classification and coding systems. The final hierarchical set of classification criteria is as follows:

(i) number of reaction routes (KG cycles), $M = 1, 2, 3, \ldots$

(ii) number of intermediates (KG vertexes), $N = 2, 3, 4, \ldots$

(iii) types of interconnection of a pair of KG cycles (classes of two-route mechanisms)
- class A = bridging of cycles
- class B = cycles sharing a common vertex
- class C = cycles sharing a common edge
- class Z = disjoint cycles (linkage via other cycles)
- prefix n = number of KG vertexes with degree $a \geq 3$

* To whom correspondence should be addressed.

† On a sabbatical leave at The University of Texas M. D. Anderson Cancer Center, Houston, HMB, Box 151, 1100 Holcombe Blvd., TX 77030.

® Abstract published in *Advance ACS Abstracts*, February 15, 1994.

0095-2338/94/1634-0436$04.50/0

Table 1. Total Number of KGs for M = 2–6 and N = 2–12[a]

M	$N=2$	$N=3$	$N=4$	$N=5$	$N=6$	$N=7$	$N=8$	$N=9$	$N=10$	$N=11$	$N=12$
2	1	2	4	7	10	14	19	24	30	37	44
3	1	3	12	27	65	129	245	422	710	1 113	1 710
4	1	5	23	85	276	764	1 935	4 466	9 583	19 291	36 859
5	1	6	43	210	924	3 403	11 242	33 156	89 789	224 621	526 346
6	1	8	72	469	2652	12 644	52 727	194 909	651 008	CE	CE

[a] CE = combinatorial explosion.

Table 2. Total Number of Classes for M = 2–6 and N = 2–12[a]

M	$N=2$	$N=3$	$N=4$	$N=5$	$N=6$	$N=7$	$N=8$	$N=9$	$N=10$	$N=11$	$N=12$
2	1	1	1	0	0	0	0	0	0	0	0
3	1	2	6	3	2	1	0	0	0	0	0
4	1	4	14	24	33	19	11	4	1	0	0
5	1	5	30	85	192	249	250	153	77	26	7
6	1	7	55	239	798	1746	2800	3082	2576	CE	CE

[a] N for a class includes vertices with $a_i \geq 2$, as well as all loops. CE = combinatorial explosion.

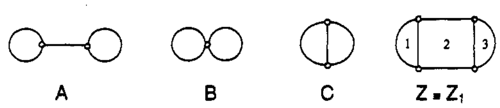


Figure 1. Four basic classes of linear mechanisms. Class Z refers to the nonadjacent pair of cycles 1 and 3. Substituting any loop for a cycle of arbitrary size preserves the class.

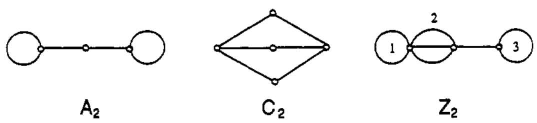


Figure 2. Examples of the A_I, C_K, and Z_V subclasses: there is a two-edge bridge in A_2, two edges shared by the two cycles in C_2, and two-edge distance between cycles 1 and 3 in Z_2 (the complete code for the last KG is ABZ_2).

(iv) subclasses of mechanism (number of elements connecting a pair of KG cycles)
 subclasses A, A_2, A_3, ... (the length of a bridge, I)
 subclasses C, C_2, C_3, ... (the number of common edges, K)
 subclasses Z_0, Z_1, Z_2, ... (the number of edges V separating a pair of cycles lacking connections of type A, B, or C)

(v) number of vertexes in each cycle, N_i

The linear code that results from the above classification criteria is

$$M - N - n - A_I^{\,i} B^j C_K^{\,k} Z_V^{\,v} - N_1, N_2, ..., N_M$$

It describes simple (nondirected) graphs. For digraphs, which refer to mechanisms containing irreversible steps, the code is supplemented by the list of all edge types.[20] The class notation in the linear code is abbreviated; it stands for the ***generalized*** classes and contains superscripts that show the number of times this particular type of cycle linkage occurs. Instead, one can use ***specific*** class notation, which is not shortened, and list all pairwise cycle linkages (A, B, C, or Z) following their canonical numbering (see Table III, *vide infra*). Kinetic ***supergraphs*** (***KSGs***) are used to facilitate the canonical numbering of KG cycles, vertexes, and edges.[20] Each vertex in the KSG represents a cycle in the initial KG, while a KSG edge represents a KG cycle linkage of type A, B, or C. The lack of an edge between two KSG vertexes means no A, B, or C type of linkage for the respective pair of cycles in KG (class Z).

The modifications of our previously adopted classification and coding systems include the ***type*** of reaction mechanism, which was previously denoted in the code by the serial number introduced for each KSG. The computer elucidation of the linear mechanisms, however, would require that standard tables be stored with the serial numbers of all KSGs, whose number increases rapidly for more complex reactions. The retrieval of the mechanisms coded is facilitated by the use of the new class Z introduced in the foregoing, and the class prefix n, which is equal to the number of vertexes in the smallest homeomorphic image of all KGs of the class under consideration. The new code does not contain any symbol for the mechanism type. Yet, the defining of the latter makes sense from the viewpoint of classification. Types of KGs with increased complexity may be denoted by L = 1, 2, 3, 4, ..., an integer indicating the total number of pairwise cycle linkages of type A, B, or C in the KG (see Table III, *vide infra*). The upper limit of the L value is the number of edges in the complete KSG.

An example illustrating the use of KGs and their coding is given as follows with the catalytic reaction of methanol synthesis. One of the mechanisms proposed[29] incorporates two reaction routes with a total of five reaction steps and four intermediates. Hence, it is represented by a KG containing two cycles, four vertexes, and five edges. The mechanism code includes the class prefix $n = 2$ (the two vertexes of degree 2 are omitted).

$$Z^{*}H_2O + CO_2 \rightleftarrows Z^{*}H_2O^{*}CO_2 \qquad (1)$$

$$Z^{*}H_2O^{*}CO_2 \rightleftarrows Z^{*}CO_2 + H_2O \qquad (2)$$

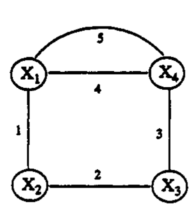
$$Z^{*}CO_2 + H_2 \rightleftarrows Z^{*}CO_2{}^{*}H_2 \qquad (3)$$

$$Z^{*}CO_2{}^{*}H_2 + 2H_2 \rightleftarrows Z^{*}H_2O + CH_3OH \qquad (4)$$

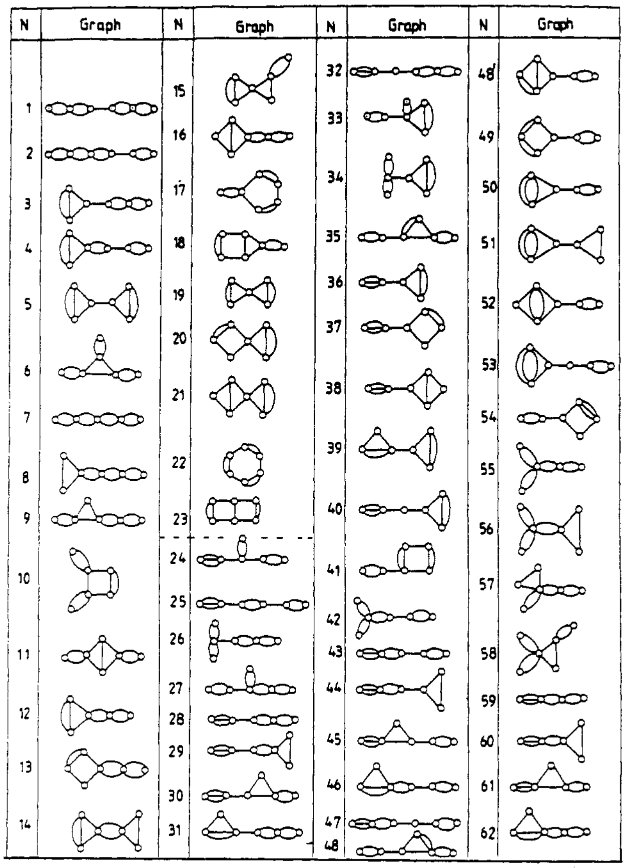
$$Z^{*}CO_2\,H_2 \rightleftarrows Z^{*}H_2O + CO \qquad (5)$$

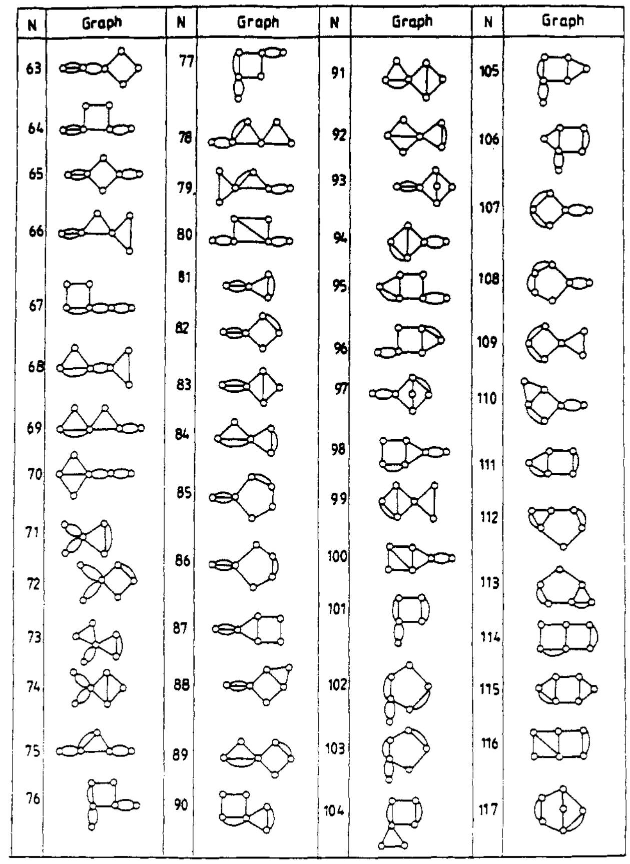
net: $$CO_2 + 3H_2 \rightleftarrows CH_3OH + H_2O$$

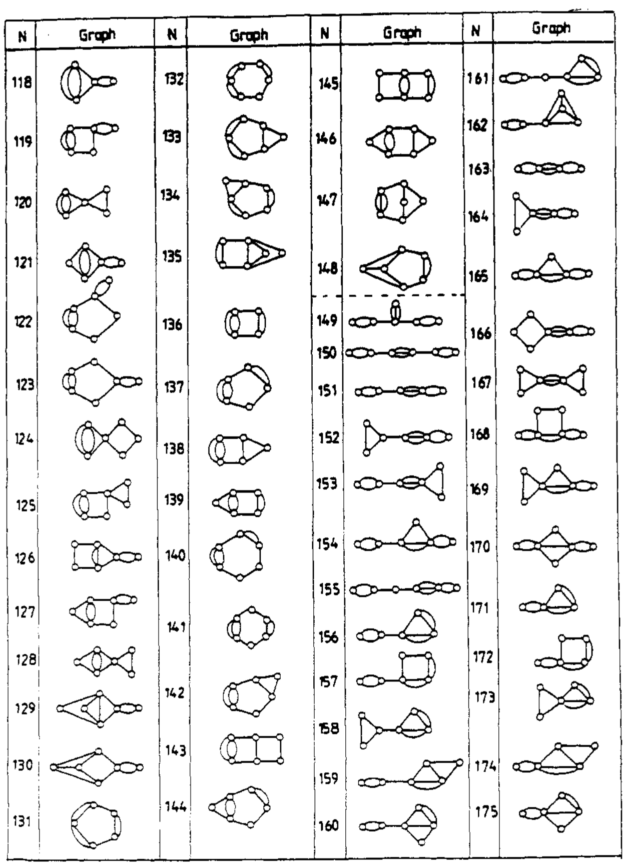
$$CO + 2H_2 \rightleftarrows CH_3OH$$

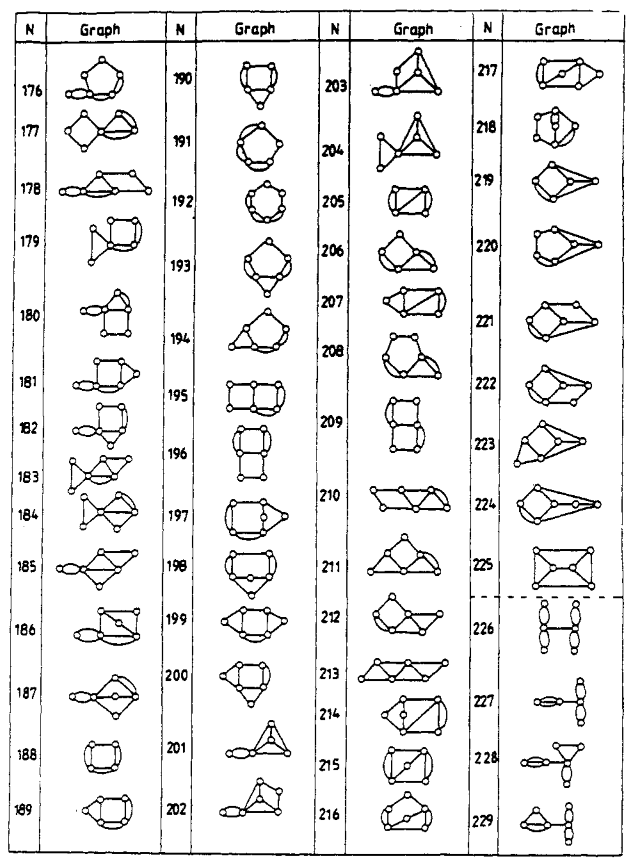
KG 

CODE: 2-4-2-C-2,4


N Graph N Graph N Graph N Graph
1 2 3 4 5 6 7 8 9 10 11 12 13 14
15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48
48' 49 50 51 52 53 54 55 56 57 58 59 60 61 62


N Graph N Graph N Graph N Graph
63 64 65 66 67 68 69 70 71 72 73 74 75 76
77 78 79 80 81 82 83 84 85 86 87 88 89 90
91 92 93 94 95 96 97 98 99 100 101 102 103 104
105 106 107 108 109 110 111 112 113 114 115 116 117


N Graph N Graph N Graph N Graph
118 119 120 121 122 123 124 125 126 127 128 129 130 131
132 133 134 135 136 137 138 139 140 141 142 143 144
145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160
161 162 163 164 165 166 167 168 169 170 171 172 173 174 175


N Graph N Graph N Graph N Graph
176 177 178 179 180 181 182 183 184 185 186 187 188 189
190 191 192 193 194 195 196 197 198 199 200 201 202
203 204 205 206 207 208 209 210 211 212 213 214 215 216
217 218 219 220 221 222 223 224 225 226 227 228 229

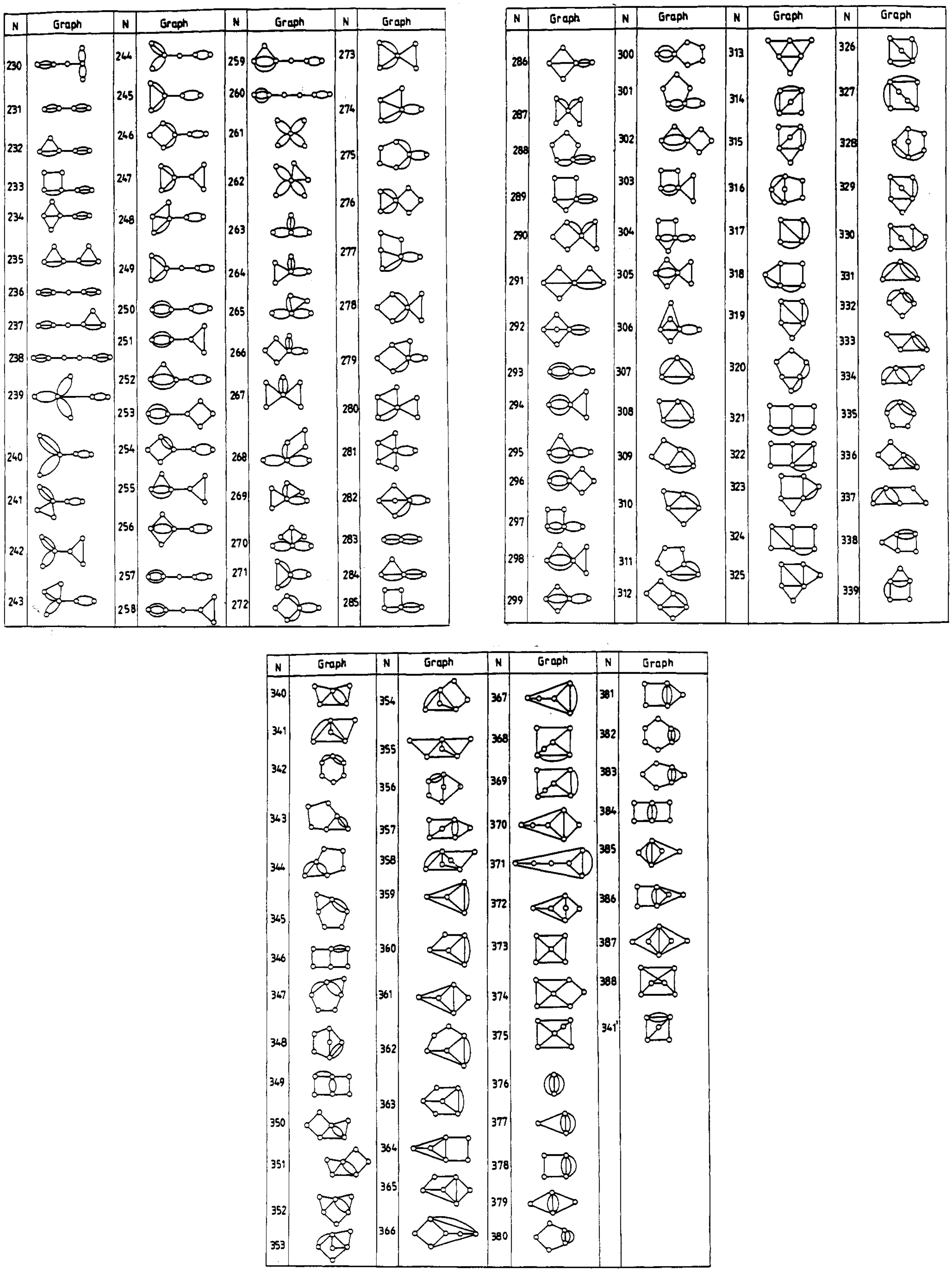


Figure 3. Four-route mechanisms having two to six intermediates.

The linear mechanisms were enumerated by our original program KING (KINetic Graphs), which generates exhaustively all nonredundant KGs for a given number of cycles and vertexes. The KING program is written in C language, under MS-DOS. It runs on an IBM PC or compatible machine and is very inexpensive since it requires relatively little RAM and hard drive space. The combinatorial algorithm used for KG enumeration is similar to that used in the GENESIS program,[30] and it employs an approach to graph enumeration developed by Faradzhev et al.[31]

Table 3. Classification, Codes, and Complexity Indexes of Linear Four-Route Mechanisms with Two to Six Intermediates

	L = 3	
	Generalized Class AB^2	
	Class *4-*ABZ^2BZ	
1	4-6-4 ABZ^2BZ_2-2,2,2,2	2304
	Class *4-*ABZ^3B	
2	4-6-4-ABZ^2Z_2B-2,2,2,2	2304
	Generalized Class ABC	
	Class *5-*ABZ^2CZ	
3	4-6-5-ABZ^2CZ_2-2,3,2,2	2880
	Class *5-*ABZ^3C	
4	4-6-5-ABZ^2Z_2C-2,2,3,2	2880
	Generalized Class AC^2	
	Class *6-*ACZ^2CZ	
5	4-6-6-ACZ^2CZ_2-3,3,2,2	3600
	Generalized Class B^3	
	Class *3-*B^3Z^3	
6	4-6-3-B^3Z^3-3,2,2,2	3408
	Class *3-*B^2Z^2BZ	
7	4-5-3-$B^2Z^2BZ_2$-2,2,2,2	1600
8	4-6-3-$B^2Z^2BZ_2$-2,2,2,3	3408
9	4-6-3-$B^2Z^2BZ_2$-2,3,2,2	3408
	Generalized Class B^2C	
	Class *4-*B^2CZ^3	
10	4-6-4-B^2CZ^3-4,2,2,2	3984
	Class *4-*BCZ^3B	
11	4-6-4-BCZ^2Z_2B-3,2,3,2	4512
	Class *4-*B^2Z^2CZ	
12	4-5-4-$B^2Z^2CZ_2$-2,3,2,2	2000
13	4-6-4-$B^2Z^2CZ_2$-2,4,2,2	3984
14	4-6-4-$B^2Z^2CZ_2$-2,3,3,2	4260
15	4-6-4-$B^2Z^2CZ_2$-3,3,2,2	4260
16	4-6-4-$B^2Z^2CZ_2$-2,3,2,3	4512
	Generalized Class BC^2	
	Class *5-*BC^2Z^3	
17	4-6-5-BC^2Z^3-5,2,2,2	4560
	Class *5-*BCZ^3C	
18	4-6-5-BCZ^2Z_2C-3,2,4,2	5364
	Class *5-*BCZ^2CZ	
19	4-5-5-BCZ^2CZ_2-3,3,2,2	2500
20	4-6-5-BCZ^2CZ_2-3,4,2,2	4980
21	4-6-5-BCZ^2CZ_2-3,3,2,3	5640
	Generalized Class C^3	
	Class *6-*C^3Z^3	
22	4-6-6-C^3Z^3-6,2,2,2	5136
	Class *6-*C^2Z^2CZ	
23	4-6-6-$C^2Z^2CZ_2$-4,4,2,2	6360
	L = 4	
24	4-6-4-$A^3CZ^2_0$-2,2,2,2	1776
	Class *5-*A^3CZ^2	
25	4-6-5-Z^3CZ^2-2,2,2,2	1776
	Generalized Class A^2B^2	
	Class *3-*$A^2B^2Z^2$	
26	4-6-3-$A^2B^2Z^2$-2,2,2,2	2304
	Class *3-*A^2ZB^2Z	
27	4-6-3-A^2ZB^2Z-2,2,2,2	2304
	Generalized Class A^2BC	
	Class *4-*A^2BCZ^2	
28	4-5-4-A^2BCZ^2-2,2,2,2	1240
29	4-6-4-A^2BCZ^2-2,2,2,2	2628
30	4-6-4-A^2BCZ^2-3,2,2,2	2628
31	4-6-4-A^2BCZ^2-2,2,3,2	2880
32	4-6-4-$A^2_2BCZ^2$-2,2,2,2	1776
	Class *4-*A^2ZBCZ	
33	4-6-4-A^2ZBCZ-2,3,2,2	2880
	Class *4-*A^2CBZ^2	
34	4-6-4-A^2CBZ^2-3,2,2,2	2880
	Class *4-*A^2ZCBZ	
35	4-6-4-A^2ZCBZ-2,3,2,2	2880
	Generalized Class A^2C^2	
	Class *5-*$A^2C^2Z^2$	
36	4-5-5-$A^2C^2Z^2$-3,2,2,2	1550
37	4-6-5-$A^2C^2Z^2$-4,2,2,2	3072
38	4-6-5-$A^2C^2Z^2$-3,2,2,3	3480
39	4-6-5-$A^2C^2Z^2$-3,2,3,2	3600
40	4-6-5-$A^2_2C^2Z^2$-3,2,2,2	2220
	Class *5-*A^2ZC^2Z	
41	4-6-5-A^2ZC^2Z-2,4,2,2	3456

	L = 4	
	Generalized Class AB^3	
	Class *3-*AB^2Z^2B	
42	4-6-3-AB^2Z^2B-2,2,2,2	2304
	Generalized Class AB^2C	
	Class *4-*AB^2Z^2C	
43	4-5-4-AB^2Z^2C-2,2,2,2	1240
44	4-6-4-AB^2Z^2C-2,3,2,2	2628
45	4-6-4-AB^2Z^2C-3,2,2,2	2628
46	4-6-4-AB^2Z^2C-2,2,2,3	2880
47	4-6-4-$A_2B^2Z^2C$-2,2,2,2	1776
	Class *4-*$ABCZ^2B$	
48	4-6-4-$ABCZ^2B$-3,2,2,2	2880
	Generalized Class ABC^2	
	Class *5-*$ABCZ^2C$	
48′	4-6-5-$ABCZ^2$-3,2,2,3	3732
	Class *5-*AC^2Z^2B	
49	4-6-5-AC^2Z^2B-4,2,2,2	3456
	Generalized Class AC^3	
	Class *4-*AC^2Z^2C	
50	4-5-4-AC^2Z^2C-3,2,2,2	1470
51	4-6-4-AC^2Z^2C-3,3,2,2	3108
52	4-6-4-AC^2Z^2C-3,2,2,3	3480
53	4-6-4-$A_2C^2Z^2C$-3,2,2,2	2100
54	4-6-4-AC^2Z^2C-4,2,2,2	2952
	Generalized Class B^4	
	Class *2-*B^4Z^2	
55	4-5-2-B^4Z^2-2,2,2,2	1600
56	4-6-2-B^4Z^2-2,2,2,3	3408
57	4-6-2-B^4Z^2-2,2,3,2	3408
58	4-6-2-B^4Z^2-3,2,2,2	3408
	Generalized Class B^3C	
	Class *3-*B^3CZ^2	
59	4-4-3-B^3CZ^2-2,2,2,2	800
60	4-5-3-B^3CZ^2-2,2,2,3	1830
61	4-5-3-B^3CZ^2-3,2,2,2	1830
62	4-5-3-B^3CZ^2-2,2,3,2	2000
63	4-6-3-B^3CZ^2-2,2,2,4	3480
64	4-6-3-B^3CZ^2-4,2,2,2	3480
65	4-6-3-B^3CZ_2-4,2,2,2	3480
66	4-6-3-B^3CZ^2-3,2,2,3	3888
67	4-6-3-B^3CZ^2-2,2,4,2	3984
68	4-6-3-B^3CZ^2-2,2,3,3	4260
69	4-6-3-B^3CZ^2-3,2,3,2	4260
70	4-6-3-B^3CZ^2-2,3,3,2	4512
	Class *3-*B^2CBZ^2	
71	4-5-3-B^2CBZ^2-3,2,2,2	2000
72	4-6-3-B^2CBZ^2-4,2,2,2	3984
73	4-6-3-B^2CBZ^2-3,2,3,2	4260
74	4-6-3-B^2CBZ^2-3,2,2,3	4512
	Class *3-*B^2CZBZ	
75	4-5-3-B^2CZBZ-3,2,2,2	2000
76	4-6-3-B^2-$CZBZ$-4,2,2,2	3984
77	4-6-3-B^2-CZ_2BZ-4,2,2,2	3984
78	4-6-3-B^2CZBZ-3,2,2,2	4260
79	4-6-3-B^2CZBZ-3,3,2,2	4260
80	4-6-3-B^2CZBZ-3,2,2,3	4512
	Generalized Class B^2C^2	
	Class *4-*$B^2C^2Z^2$	
81	4-4-4-$B^2C^2Z^2$-3,2,2,2	1000
82	4-5-4-$B^2C^2Z^2$-4,2,2,2	2140
83	4-5-4-$B^2C^2Z^2$-3,2,2,3	2420
84	4-5-4-$B^2C^2Z^2$-3,2,3,2	2500
85	4-6-4-$B^2C^2Z^2$-5,2,2,2	3924
86	4-6-4-$B^2C^2Z^2_2$-5,2,2,2	3924
87	4-6-4-$B^2C^2Z^2$-3,2,2,4	4740
88	4-6-4-$B^2C^2Z^2$-4,2,2,3	4740
89	4-6-4-$B^2C^2Z^2$-4,2,3,2	4980
90	4-6-4-$B^2C^2Z^2$-3,2,4,2	4980
91	4-6-4-$B^2C^2Z^2$-3,2,3,3	5640
92	4-6-4-$B^2C^2Z^2$-3,3,3,2	5640
93	4-6-4-$B^2C_2CZ^2$-4,2,2,4	5184
	Class *4-*B^2CZCZ	
94	4-5-4-B^2CZCZ-3,2,2,3	2590
95	4-6-4-B^2CZCZ-4,2,2,3	5112
96	4-6-4-B^2CZ_2CZ-4,2,2,3	5112
97	4-6-4-B^2C_2ZCZ-4,2,2,4	5592

	L = 4	
	Class *4-*B^2CZCZ	
98	4-6-4-B^2CZCZ-3,2,2,4	5364
99	4-6-4-B^2CZCZ-3,2,3,3	5520
100	4-6-4-B^2CZCZ-3,3,2,3	5892
	Class *4-*B^2ZC^2Z	
101	4-5-4-B^2ZC^2Z-2,4,2,2	2400
102	4-6-4-B^2ZC^2Z-2,5,2,2	4560
103	4-6-4-$B^2Z_2C^2Z$-2,5,2,2	4560
104	4-6-4-B^2ZC^2Z-3,4,2,2	5112
105	4-6-4-B^2ZC^2Z-2,4,2,3	5364
106	4-6-4-B^2ZC^2Z-2,4,3,2	5364
	Class *4-*BC^2Z^2B	
107	4-5-4-BC^2Z^2B-4,2,2,2	2400
108	4-6-4-BC^2ZZ_2B-5,2,2,2	4560
109	4-6-4-BC^2Z^2B-4,3,2,2	5112
110	4-6-4-BC^2Z^2B-4,2,2,3	5364
	Generalized Class BC^3	
	Class *5-*BC^3Z^2	
111	4-5-5-BC^3Z^2-4,2,3,2	3090
112	4-6-5-BC^3Z_2Z-5,2,3,2	5832
113	4-6-5-BC^3Z^3-5,2,3,2	5832
114	4-6-5-BC^2Z^2-4,2,4,2	6360
115	4-6-5-BC^3Z^2-4,2,3,3	6900
116	4-6-5-BC^3Z^2-4,3,3,2	7020
117	4-6-5-$BC_2C^2Z^2$-5,2,4,2	6684
	Class *3-*BC^2Z^2C	
118	4-4-3-BC^2Z^2C-3,2,2,2	952
119	4-5-3-BC^2Z^2C-4,2,2,2	2060
120	4-5-3-BC^2Z^2C-3,3,2,2	2170
121	4-5-3-BC^2Z^2C-3,2,2,3	2420
122	4-6-3-BC^2Z^2C-5,2,2,2	3804
123	4-6-3-$BC^2Z^2_2C$-5,2,2,2	3804
124	4-6-3-BC^2Z^2C-3,4,2,2	4116
125	4-6-3-BC^2Z^2C-4,3,2,2	4368
126	4-6-3-BC^2Z^2C-3,2,2,4	4860
127	4-6-3-BC^2Z^2C-4,2,2,3	4860
128	4-6-3-BC^2Z^2C-3,3,2,3	5148
129	4-6-3-BC^2Z^2C-3,2,3,3	5712
	Class *5-*BC^2Z^2C	
130	4-6-5-BC^2Z^2C-4,2,3,3	6768
	Class *5-*$BCZCZC$	
131	4-5-5-$BCZCZC$-2,2,5,2	2800
132	4-6-5-$BCZCZ_2C$-2,2,6,2	5136
133	4-6-5-$BCZCZC$-2,2,5,3	6216
134	4-6-5-$BCZCZC$-2,3,5,2	6216
	Generalized Class C^4	
	Class *4-*C^4Z^2	
135	4-6-4-C^4Z^2-4,3,3,2	6852
136	4-4-4-C^4Z^2-4,2,2,2	1152
137	4-5-4-C^4Z^2-5,2,2,2	2370
138	4-5-4-C^4Z^2-4,2,2,3	2760
139	4-5-4-C^4Z^2-4,2,3,2	2920
140	4-6-4-C^4Z^2-6,2,2,2	4248
141	4-6-4-$C^4Z^2_2$-6,2,2,2	4248
142	4-6-4-C^4Z^2-5,2,2,3	5220
143	4-6-4-C^4Z^2-4,2,2,4	5376
144	4-6-4-C^4Z^2-5,2,3,2	5580
145	4-6-4-C^4Z^2-4,2,4,2	5856
146	4-6-4-C^4Z^2-4,2,3,3	6528
147	4-6-4-$C^2C_2CZ^2$-5,2,2,4	5820
	Class *6-*C^4Z^2	
148	4-6-6-C^4Z^2-5,3,3,2	7896
	L = 5	
	Generalized Class A^4C	
	Class *4-*A^2CZA^2	
149	4-6-4-$A^2CZ_0A^2$-2,2,2,2	1776
150	4-6-4-A^2CZA^2-2,2,2,2	1776
	Generalized Class A^2B^2C	
	Class *3-*A^2ZCB^2	
151	4-5-3-A^2ZCB^2-2,2,2,2	1240
152	4-6-3-A^2ZCB^2-3,2,2,2	2628
153	4-6-3-A^2ZCB^2-2,2,2,3	2628
154	4-6-3-A^2ZCB^2-2,2,3,2	2880
155	4-6-3-$A^2_2ZCB^2$-2,2,2,2	1776

Table 3. (Continued)

	L = 5	
	Generalized Class A^{2}BC^2	
	Class *4-A^{2}ZCBC*	
156	4-5-4-A^{2}ZCBC-2,2,3,2	1640
157	4-6-4-A^{2}ZCBC-2,2,4,2	3456
158	4-6-4-A^{2}ZCBC-3,2,3,2	3480
159	4-6-4-A^{2}ZCBC-2,2,3,3	3732
160	4-6-4-A^{2}ZCBC-2,3,3,2	3732
161	4-6-4-A$^2_2$ZCBC-2,2,3,2	2352
	Generalized Class A^{2}C^3	
	Class *5-A^{2}ZC^3*	
162	4-6-5-A^{2}ZC^3-2,3,3,3	4608
	Generalized Class B^{4}C	
	Class *2-B^{2}CZB^2*	
163	4-4-2-B^{2}CZB^2-2,2,2,2	800
164	4-5-2-B^{2}CZB^2-2,2,3,2	1830
165	4-5-2-B^{2}CZB^2-2,2,2,3	2000
166	4-6-2-B^{2}CZB^2-2,2,4,2	3480
167	4-6-2-B^{2}CZB^2-2,3,3,2	3888
168	4-6-2-B^{2}CZB^2-2,2,2,4	3984
169	4-6-2-B^{2}CZB^2-2,2,3,3	4260
170	4-6-2-B^{2}CZB^2-3,2,2,3	4512
	Generalized Class B^{3}C^2	
	Class *3-B^{2}CZBC*	
171	4-4-3-B^{2}CZBC-2,2,2,3	1056
172	4-5-3-B^{2}CZBC-2,2,2,4	2400
173	4-5-3-B^{2}CZBC-2,3,2,3	2420
174	4-5-3-B^{2}CZBC-2,2,3,3	2590
175	4-5-3-B^{2}CZBC-3,2,2,3	2590
176	4-6-3-B^{2}CZBC-2,2,2,5	4560
177	4-6-3-B^{2}CZBC-2,4,2,3	4608
178	4-6-3-B^{2}CZBC-2,2,4,3	5112
179	4-6-3-B^{2}CZBC-2,3,2,4	5112
180	4-6-3-B^{2}CZBC-4,2,2,3	5112
181	4-6-3-B^{2}CZBC-2,2,3,4	5364
182	4-6-3-B^{2}CZBC-3,2,2,4	5364
183	4-6-3-B^{2}CZBC-2,3,3,3	5520
184	4-6-3-B^{2}CZBC-3,3,2,3	5520
185	4-6-3-B^{2}CZBC-3,2,3,3	5892
186	4-6-3-B^{2}CZBC$_2$-2,2,4,4	5592
187	4-6-3-B^{2}C$_2$ZBC-4,2,2,4	5592
	Generalized Class B^{2}C^3	
	Class *4-B^{2}CZC^2*	
188	4-4-4-B^{2}CZC^2-2,2,2,4	1312
189	4-5-4-B^{2}CZC^2-2,2,3,4	3180
190	4-5-4-B^{2}CZC^2-3,2,2,4	3180
191	4-5-4-B^{2}CZC^2-2,2,2,5	2800
192	4-6-4-B^{2}CZC^2-2,2,26	5136
193	4-6-4-B^{2}CZC^2-3,2,2,5	6216
194	4-6-4-B^{2}CZC^2-2,2,3,5	6216
195	4-6-4-B^{2}CZC^2-2,2,4,4	6240
196	4-6-4-B^{2}CZC^2-4,2,2,4	6240
197	4-6-4-B^{2}CZCC$_2$-2,2,4,5	6816
198	4-6-4-B^{2}C$_2$ZC^2-4,2,2,5	6816
199	4-6-4-B^{2}CZC^2-2,3,3,4	7152
200	4-6-4-B^{2}CZC^2-3,2,3,4	7152
	Class *4-B^{2}ZC^3*	
201	4-5-4-B^{2}ZC^3-2,3,3,3	3200
202	4-6-4-B^{2}ZC^3-2,3,3,4	6768
203	4-6-4-B^{2}ZC^3-2,3,4,3	6768
204	4-6-4-B^{2}ZC^3-3,3,3,3	6816
	Class *4-BC^{3}ZB*	
205	4-4-4-BC^{3}ZB-3,2,3,2	1376
206	4-5-4-BC^{3}ZB-3,2,4,2	3090
207	4-5-4-BC^{3}ZB-3,2,3,3	3370
208	4-6-4-BC^{3}ZB-3,2,5,2	5832
209	4-6-4-BC^{3}ZB-4,2,4,2	6360
210	4-6-4-BC^{3}ZB-3,2,3,4	6648
211	4-6-4-BC^{3}ZB-3,3,4,2	6900
212	4-6-4-BC^{3}ZB-3,2,4,3	7020
213	4-6-4-BC^{3}ZB-3,3,3,3	7680
214	4-6-4-BC^{2}C$_2$ZB-3,4,4,2	7368
215	4-5-4-BC$_2$C^{2}ZB-4,2,4,2	3320
216	4-6-4-BC$_2$C^{2}ZB-4,2,5,2	6684
217	4-6-4-BC$_2$C^{2}ZB-4,2,4,3	7300
218	4-6-4-BC$_3$C^{2}ZB-5,2,5,2	6480

	L = 5	
	Generalized Class BC^4	
	Class *5-BC^{3}ZC*	
219	4-5-5-BC^{3}ZC-3,2,4,3	3980
220	4-6-5-BC^{3}ZC-3,2,5,3	7896
221	4-6-5-BC^{3}ZC-3,2,4,4	8304
222	4-6-5-BC^{3}ZC-4,2,4,3	8304
223	4-6-5-BC^{3}ZC-3,3,4,3	8976
224	4-6-5-BCC$_2$CZC-4,2,4,4	9024
	Generalization Class C^5	
	Class *6-C^{5}Z*	
225	4-6-6-C^{5}Z-4,4,3,3	10560
	L = 6	
	Generalized Class A^{4}B^2	
	Class *2-A^{2}B^{2}A^2*	
226	4-6-2-A^{2}B^{2}A^2-2,2,2,2	2304
	Generalized Class A^{4}BC	
	Class *3-A^{2}BCA^2*	
227	4-5-3-A^{2}BCA^2-2,2,2,2	1240
228	4-6-3-A^{2}BCA^2-2,2,2,3	2628
229	4-6-3-A^{2}BCA^2-2,2,3,2	2880
230	4-6-3-A$^2_2$BCA$^2_2$-2,2,2,2	1776
	Generalized Class A^{4}C^2	
	Class *4-A^{2}C^{2}A^2*	
231	4-4-4-A^{2}C^{2}A^2-2,2,2,2	624
232	4-5-4-A^{2}C^{2}A^2-2,2,2,3	1550
233	4-6-4-A^{2}C^{2}A^2-2,2,2,4	3072
234	4-6-4-A^{2}C^{2}A^2-2,3,3,2	3480
235	4-6-4-A^{2}C^{2}A-2,2,3,3	3600
236	4-5-4-A$^2_2$C^{2}A$^2_2$-2,2,2,2	960
237	4-6-4-A$^2_2$C^{2}A$^2_2$-2,2,2,3	2220
238	4-6-4-A$^2_3$C^{2}A$^2_3$-2,2,2,2	1368
	Generalized Class A^{3}B^3	
	Class *2-A^{3}B^3*	
239	4-6-2-A^{3}B^3-2,2,2,2	2304
	Generalized Class A^{3}B^{2}C	
	Class *3-A^{3}B^{2}C*	
240	4-5-3-A^{3}B^{2}C-2,2,2,2	1240
241	4-6-3-A^{3}B^{2}C-2,3,2,2	2628
242	4-6-3-A^{3}B^{2}C-3,2,2,2	2628
243	4-6-3-A^{3}B^{2}C-2,2,2,3	2880
244	4-6-3-A$^3_2$B^{2}C-2,2,2,2	1776
	Generalized Class A^{3}BC^2	
	Class *4-A^{3}BC^2*	
245	4-5-4-A^{3}BC^2-2,2,2,3	1640
246	4-6-4-A^{3}BC^2-2,2,2,4	3456
247	4-6-4-A^{3}BC^2-3,2,2,3	3480
248	4-6-4-A^{3}BC^2-2,2,3,3	3732
249	4-6-4-A$^3_2$BC^2-2,2,2,3	2352
	Generalized Class A^{3}C^3	
	Class *3-A^{3}C^3*	
250	4-4-3-A^{3}C^3-2,2,2,2	576
251	4-5-3-A^{3}C^3-3,2,2,2	1300
252	4-5-3-A^{3}C^3-2,2,2,3	1470
253	4-6-3-A^{3}C^3-4,2,2,2	2448
254	4-6-3-A^{3}C^3-2,2,2,4	2952
255	4-6-3-A^{3}C^3-3,2,2,3	3108
256	4-6-3-A^{3}C^3-2,2,3,3	3480
257	4-5-3-A$^3_2$C^3-2,2,2,2	880
258	4-6-3-A$^3_2$C^3-3,2,2,2	1848
259	4-6-3-A$^3_2$C^3-2,2,2,3	2100
260	4-6-3-A$^3_3$C^3-2,2,2,2	1248
	Generalized Class B^6	
	Class *1-B^6*	
261	4-5-1-B^6-2,2,2,2	1600
262	4-6-1-B^6-2,2,2,3	3408
	Generalized Class B^{5}C	
	Class *2-B^{5}C*	
263	4-4-2-B^{5}C-2,2,2,2	800
264	4-5-2-B^{5}C-2,3,2,2	1830
265	4-5-2-B^{5}C-2,2,2,3	2000
266	4-6-2-B^{5}C-2,4,2,2	3480
267	4-6-2-B^{5}C-3,3,2,2	3888
268	4-6-2-B^{5}C-2,2,2,4	3984
269	4-6-2-B^{5}C-2,3,2,3	4260
270	4-6-2-B^{5}C-2,2,3,3	4512

	L = 6	
	Generalized Class B^{4}C^2	
	Class *3-B^{4}C^2*	
271	4-4-3-B^{4}C^2-2,2,2,3	1056
272	4-5-3-B^{4}C^2-2,2,2,4	2400
273	4-5-3-B^{4}C^2-3,2,2,3	2420
274	4-5-3-B^{4}C^2-2,2,3,3	2590
275	4-6-3-B^{4}C^2-2,2,2,5	4560
276	4-6-3-B^{4}C^2-4,2,2,3	4608
277	4-6-3-B^{4}C^2-2,2,4,3	5112
278	4-6-3-B^{4}C^2-3,2,2,4	5112
279	4-6-3-B^{4}C^2-2,2,3,4	5364
280	4-6-3-B^{4}C^2-3,2,3,3	5520
281	4-6-3-B^{4}C^2-2,3,3,3	5892
282	4-6-3-B^{2}CC$_2$-2,2,4,4	5592
	Class *3-B^{2}C^{2}B^2*	
283	4-3-3-B^{2}C^{2}B^2-2,2,2,2	360
284	4-4-3-B^{2}C^{2}B^2-2,2,2,3	1000
285	4-5-3-B^{2}C^{2}B^2-2,2,2,4	2140
286	4-5-3-B^{2}C^{2}B^2-2,3,3,2	2420
287	4-5-3-B^{2}C^{2}B^2-2,2,3,3	2500
288	4-6-3-B^{2}C^{2}B^2-2,2,2,5	3924
289	4-6-3-B^{2}C^{2}B^2-2,3,4,2	4740
290	4-6-3-B^{2}C^{2}B^2-2,2,3,4	4980
291	4-6-3-B^{2}C^{2}B^2-2,3,3,3	5640
292	4-6-3-B^{2}CC$_2$B^2-2,4,4,2	5184
	Generalized Class B^{3}C^3	
	Class *2-B^{3}C^3*	
293	4-3-2-B^{3}C^3-2,2,2,2	336
294	4-4-2-B^{3}C^3-3,2,2,2	848
295	4-4-2-B^{3}C^3-2,2,2,3	952
296	4-5-2-B^{3}C^3-4,2,2,2	1720
297	4-5-2-B^{3}C^3-2,2,2,4	2060
298	4-5-2-B^{3}C^3-3,2,2,3	2170
299	4-5-2-B^{3}C^3-2,2,3,3	2420
300	4-6-2-B^{3}C^3-5,2,2,2	3048
301	4-6-2-B^{3}C^3-2,2,2,5	3804
302	4-6-2-B^{3}C^3-4,2,2,3	4116
303	4-6-2-B^{3}C^3-3,2,2,4	4368
304	4-6-2-B^{3}C^3-2,2,3,4	4860
305	4-6-2-B^{3}C^3-3,2,3,3	5148
306	4-6-2-B^{3}C^3-2,3,3,3	5712
	Class *3-B^{2}CBC^2*	
307	4-3-3-B^{2}CBC^2-2,2,2,3	480
308	4-4-3-B^{2}CBC^2-2,2,3,3	1328
309	4-5-3-B^{2}CBC^2-2,2,4,3	2840
310	4-5-3-B^{2}CBC^2-2,3,3,3	3290
311	4-6-3-B^{2}CBC^2-2,2,5,3	5208
312	4-6-3-B^{2}CBC^2-2,3,4,3	6528
313	4-6-3-B^{2}CBC^2-3,3,3,3	7560
314	4-5-3-B^{2}CBCC$_2$-2,2,4,4	3240
315	4-6-3-B^{2}CBCC$_2$-2,3,4,4	7380
316	4-6-3-B^{2}CBCC$_2$-2,2,5,4	6312
	Class *4-B^{2}C^{2}BC*	
317	4-4-4-B^{2}C^{2}BC-2,2,3,3	1376
318	4-5-4-B^{2}C^{2}BC-2,2,3,4	3090
319	4-5-4-B^{2}C^{2}BC-2,3,3,3	3370
320	4-6-4-B^{2}C^{2}BC-2,2,3,5	5832
321	4-6-4-B^{2}C^{2}BC-2,2,4,4	6360
322	4-6-4-B^{2}C^{2}BC-2,4,3,3	6648
323	4-6-4-B^{2}C^{2}BC-2,3,4,3	6900
324	4-6-4-B^{2}C^{2}BC-2,3,3,4	7020
325	4-6-4-B^{2}C^{2}BC-3,3,3,3	7680
326	4-5-4-B^{2}C^{2}BC$_2$-2,2,4,4	3320
327	4-6-4-B^{2}C^{2}BC$_3$-2,2,5,5	6480
328	4-6-4-B^{2}C^{2}BC$_2$-2,2,4,5	6684
329	4-6-4-B^{2}C^{2}BC$_2$-2,3,4,4	7500
330	4-6-4-B^{2}CC$_2$BC-2,4,4,3	7368
	Generalized Class B^{2}C^4	
	Class *3-B^{2}C^4*	
331	4-3-3-B^{2}C^4-2,2,2,3	450
332	4-4-3-B^{2}C^4-2,2,2,4	1152
333	4-4-3-B^{2}C^4-3,2,2,3	1224
334	4-4-3-B^{2}C^4-2,2,3,3	1272
335	4-5-3-B^{2}C^4-2,2,2,5	2370
336	4-5-3-B^{2}C^4-4,2,2,3	2590
337	4-5-3-B^{2}C^4-2,2,4,3	2750

Table 3. (Continued)

	L = 6			L = 6			L = 6	
	Generalized Class B^2C^4			Generalized Class B^2C^4			Generalized Class B^2C^4	
	Class 3-B^2C^4			Class *3-BC^4B*			Class *5-B^2C^4*	
338	4-5-3-B^2C^4-3,2,2,4	2760	356	4-6-3-$B^2C_2C^3$-4,2,2,5	5820	373	4-5-5-BC^4B-3,3,3,3	4440
339	4-5-3-B^2C^4-2,2,3,4	2920	357	4-6-3-$B^2C_2C^3$-4,2,3,4	6876	374	4-6-5-BC^4B-3,3,3,4	9228
340	4-5-3-B^2C^4-3,2,3,3	3120	358	4-6-3-$B^2CC^3_2$-2,4,4,4	7632	375	4-6-5-BC^3C_2B-3,4,3,4	9744
341	4-5-3-B^2C^4-2,3,3,3	3250		Class *4-B^2C^4*			Generalized Class C^6	
341′	4-5-3-$B^2C_2C^3$-4,2,2,4	2900	359	4-4-4-B^2C^4-2,3,3,3	1616		Class *2-C^6*	
342	4-6-3-B^2C^4-2,2,2,6	4248	360	4-5-4-B^2C^4-2,3,4,3	3730	376	4-2-2-C^6-2,2,2,2	120
343	4-6-3-B^2C^4-5,2,2,3	4716	361	4-5-4-B^2C^4-3,3,3,3	4020	377	4-3-2-C^6-2,2,2,3	396
344	4-6-3-B^2C^4-2,2,5,3	5076	362	4-6-4-B^2C^4-2,3,5,3	7152	378	4-4-2-C^6-2,2,2,4	944
345	4-6-3-B^2C^4-3,2,2,5	5220	363	4-6-4-B^2C^4-2,4,4,3	7812	379	4-4-2-C^6-2,2,3,3	1120
346	4-6-3-B^2C^4-4,2,2,4	5376	364	4-6-4-B^2C^4-4,3,3,3	7992	380	4-5-2-C^6-2,2,2,5	1860
347	4-6-3-B^2C^4-2,2,3,5	5580	365	4-6-4-B^2C^4-3,3,4,3	8604	381	4-5-2-C^6-2,2,3,4	2420
348	4-6-3-$B^2C_2C^3$-5,2,2,4	5568	366	4-6-4-$B^2C^3C_2$-2,4,4,4	8700	382	4-6-2-C^6-2,2,2,6	3240
349	4-6-3-B^2C^4-2,2,4,4	5856	367	4-5-4-$B^2CC_2C^2$-2,4,4,3	3700	383	4-6-2-C^6-2,2,3,5	4464
350	4-6-3-B^2C^4-4,2,3,3	6156	368	4-6-4-$B^2C^3C_2$-2,3,5,4	8028	384	4-6-2-C^6-2,2,4,4	4872
351	4-6-3-B^2C^4-3,2,4,3	6276	369	4-6-4-$B^2CC_2C^2$-2,4,5,3	7776	385	4-5-2-C^6-2,3,3,3	2960
352	4-6-3-B^2C^4-3,2,3,4	6528	370	4-6-4-$B^2CC_2C^2$-3,4,4,3	8208	386	4-6-2-C^6-2,3,3,4	5928
353	4-6-3-B^2C^4-2,3,3,4	6852	371	4-6-4-$B^2CC_3C^2$-2,5,5,3	8220	387	4-6-2-C^6-3,3,3,3	7056
354	4-6-3-B^2C^4-2,3,4,3	6840	372	4-6-4-$B^2C_2C^3$-4,3,3,4	9120		Class *6-C^6*	
355	4-6-3-B^2C^4-3,3,3,3	7404				388	4-6-6-CC^4_2C-4,4,4,4	11616

All mechanisms having up to six reaction routes and up to 12 vertexes were enumerated, except in the case of $M = 6$ for $N = 11$, and $N = 12$, for which the computational time was unreasonably high (Table 1). The number of classes was also enumerated (Table 2). We found that, at a constant number of reaction routes and an increasing number of intermediates, the number of classes passes through a maximum and behaves close to the normal distribution. Both tables give evidence for the potential existence of a tremendously large variety of topologically distinct linear mechanisms. This result is in sharp contrast to some estimates based on mechanistic chemical but not topological information.[3,4] Besides the incompleteness of the purely chemical approach, such comparisons may also indicate that some mechanisms that are topologically allowed might be chemically forbidden. The elucidation of this important question needs further studies.

III. COMPLEXITY OF LINEAR MECHANISMS

1. Complexity Index *K*. A quantitative measure for the complexity of reaction mechanisms may play an important role in their computer handling. It allows one to introduce a complexity based mechanistic hierarchy and, therefore, can help in both the generation and the discrimination of the totality of hypotheses for the mechanisms of complex reactions. In previous publications[19,21] we developed such a complexity measure based on the complexity of the steady-state kinetic model, which can readily be obtained for linear mechanisms by making use of graph theory. Calculated by means of the spanning trees of the KG and some of its subgraphs, this complexity index reflects the complexity of kinetic graphs as well and may be of use for complexity analysis of any cyclic graphs.

Our complexity index K is based on the fractional–rational form of the rate laws for reaction routes within the framework of the Vol'kenshtein–Gol'dshtein algorithm.[32,33] More specifically, K is defined as the total number of weights (rate constants) of the elementary steps (KG edges) included in the kinetic laws for all M routes of a multiroute reaction. For mechanisms containing reversible steps only, it is calculated by eq 6, where T_i is the number of spanning trees in vertex i (this number is the same for every vertex in the KG); the double sum counts the number of spanning trees of the KG subgraphs obtained after subsequently contracting each of the graph cycles p and its encompassing cycles pk to a vertex.

$$K = MN(N-1)T_i + 2N\sum_{p=1}^{M}\sum_{k=0}^{k=\max} D_{pk} \tag{6}$$

In a previous publication[21] we discussed the complexity of mechanisms with two and three reaction routes. Here, we extend this analysis to mechanisms incorporating four independent routes. Instead of using general methods for enumerating spanning trees,[34] we made use of an explicit formula derived earlier:[19]

$$\begin{aligned} T_4 = {} & N_1N_2N_3N_4 - (E_{12}^{\,2}N_3N_4 + E_{13}^{\,2}N_2N_4 + E_{14}^{\,2}N_2N_3 + \\ & E_{23}^{\,2}N_1N_4 + E_{24}^{\,2}N_1N_3 + E_{34}^{\,2}N_1N_2) - (2E_{12}E_{13}E_{23}N_4 + \\ & 2E_{12}E_{14}E_{24}N_3 + 2E_{13}E_{14}E_{34}N_2 + 2E_{23}E_{24}E_{34}N_1) - \\ & (2E_{13}E_{14}E_{23}E_{24} + 2E_{12}E_{13}E_{24}E_{34} + 2E_{12}E_{14}E_{23}E_{34}) + \\ & (E_{12}^{\,2}E_{34}^{\,2} + E_{13}^{\,2}E_{24}^{\,2} + E_{14}^{\,2}E_{23}^{\,2}) \end{aligned} \tag{7}$$

Equations 6 and 7 provide fast calculation of the spanning trees and the complexity index directly from the mechanism linear code, where one can find both the cycle size N_p and the number of edges two cycles have in common E_{ij}. The latter is obviously zero for classes A and B, while for class C it is equal to the subclass subscript (1 for $C = C_1$, 2 for C_2, etc.).

2. Standard Tables with the Complexities of All Topologically Distinct Four-Route Mechanisms Having Two to Six Intermediates. Before proceeding with a complexity analysis, we present here in Figure 3 and Table 3 all 390 four-route mechanisms having two to six intermediates, as generated by the KING program together with their codes and complexity indexes. These are mechanisms containing reversible elementary steps only. However, each of the mechanisms presented can be used to generate a certain number of mechanisms with irreversible steps, as well as an additional number of mechanisms incorporating intermediates that are involved only in an equilibrium elementary step (KGs with pendant vertexes).

3. Trends Increasing Mechanism Complexity. The complexity analysis we performed confirmed the trends toward a higher complexity of linear mechanisms found previously for two-route and three-route mechanisms.[21] Clearly, the complexity index K of the four-route mechanisms is considerably higher than that of the three-route mechanisms with the same number of intermediates. Similarly, at $M = 4 =$ constant, the

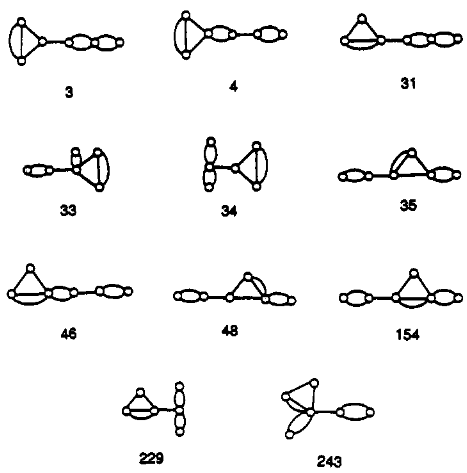


Figure 4. 11 isocomplex KGs with complexity index $K = 2880$. The KG numbers correspond to those in Figure 3 and Table 3.

Table 4. Degeneracy of the Complexity Index of Linear Reaction Mechanisms Having One to Four Reaction Routes

no. of routes	total no. of mechanisms	total no. of the different index values	degree of degeneracy
1	5	5	1
2	24	23	1.04
3	104	65	1.60
4	390	171	2.28

increase in the number of intermediates greatly increases the mechanism complexity. The subtle topological patterns enhancing complexity are reflected by the following series of classes and subclasses ordered with respect to the increase in K:

$$A_3 < A_2 < A < B < C < C_2 < C_3 < \ldots \qquad (8)$$

At a constant number of reaction routes and intermediates, as well as within the same class and subclass, K increases with equalizing cycle sizes, thus manifesting an entropylike behavior. As an illustration, compare KGs 123, 126, and 129 from Figure 3 which have six intermediates and belong to the same class 3-BC^2Z^2C but differ in cycle sizes, which are respectively 5,2,2,2; 3,2,2,4; and 3,2,3,3. The cycle-size equalizing results in an increase in the complexity index from 3804 to 4860 to 5712, respectively.

Another trend of increasing mechanistic complexity is easily proved. It refers to the increase in the newly introduced class prefix n in the mechanism code. Since n, by definition, is equal to the number of vertexes in the smallest homeomorphic image of the KGs from a certain class of mechanisms, then the larger this number, the more complex the mechanism. This trend can be illustrated by comparing in Table 3 the three specific classes belonging to the same generalized class B^3C^3: 2-$B^3C^3 \rightarrow$ 3-$B^2CBC^2 \rightarrow$ 4-B^2C^2BC (graphs 293 to 306, 307 to 316, and 317 to 330, respectively).

IV. ISOCOMPLEXITY

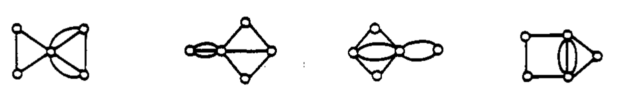
1. Complexity Index Degeneracy. Albeit closely related to the unique linear code, the complexity index of the linear mechanisms is not entirely discriminating. The number of distinct KGs with the same value of the K index increases rapidly with the increase in the number of reaction routes. This is illustrated in Table 4, where the degree of degeneracy of the complexity index is calculated as the ratio of the total number of mechanisms and that of the mechanisms with different K values.

The high degeneracy found for four-route mechanisms reflects the higher degree of similarity of the graphs having four cycles. The difficulties involved in discriminating the highly connected KGs parallel those involved in discriminating the kinetic hypotheses for four-route reaction mechanisms. Thus, the complexity index K helps explain why it is so difficult to discriminate some mechanisms, the reason being the high similarity of topological structure among mechanisms.

2. Isocomplexity Levels. The phenomenon of isocomplexity encompasses not only mechanisms differing in minor structural details but also covers all classification levels of mechanisms: types, generalized classes, specific classes, subclasses, and different distributions of cycle sizes.

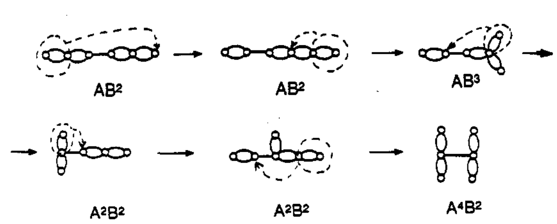
An illustration is presented in Figure 4, where 11 KGs belonging to 11 specific classes, 6 generalized classes, and 4 types of linear mechanism have the complexity index $K = 2880$ showing the highest degeneracy. For example, KGs 31 and 33–35 belong to the specific classes 4-A^2BCZ^2, 4-A^2ZBCZ, 4-A^2CBZ^2, and 4-A^2ZCBZ, respectively, all of which are included in the generalized class A^2BC and type $L = 4$ (four-cycle interconnections). Another generalized class, AB^2C, of the same type $L = 4$, is also represented by KGs 46 and 48 (specific classes 4-AB^2Z^2C and 4-$ABCZ^2B$, respectively). KG 154 is of generalized class A^2B^2C and type $L = 5$, and KGs 229 and 243 are of type $L = 6$ and generalized classes A^4BC and A^3B^2C, respectively.

In addition to the intrinsic mechanism isocomplexity described above, it should be mentioned that 15 cases of accidental degeneracies have been found. These are cases in which the same K index value results by chance from different summands reflecting different mechanistic topology; no systematic graph transformations connect these KGs. An example is presented below, in which four linear mechanisms have the same complexity index ($K = 2420$).

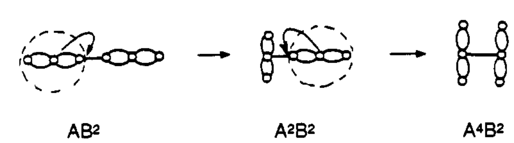


273 286 299 381

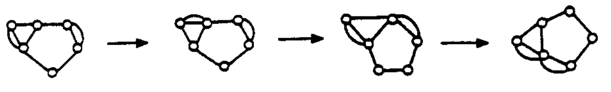
3. Graph Transformations Preserving Complexity. The analysis of eq 7 indicates that graph transformations preserving complexity are all transformations that do not change cycle sizes N_i and the number E_{ij} of the edges common for cycles i and j. Otherwise, these are different cases of "positional isomerism" that deal mainly with A and B classes (weak intercycle linkage). Upon such a graph transformation, a cycle linked by a bridge or by a common vertex is displaced so as to be connected to other cycles by any one of these weak linkages.



In general, the same type of transformation can be performed for subgraphs containing two or more weakly connected cycles:



Some transformations of strongly connected cycles (class C) also produce isocomplexity. These are displacements of an outer cycle sharing a common edge with a large cycle whose sites are nonequivalent:



112 113 208 320

K = 5832

4. Complexity Flow Chart. All isocomplexity relationships found for the classes of four-route mechanisms can be presented in a flow chart (Figure 5). The classes with the same complexity are connected there by vertical lines. The flow chart also shows the relationships of increasing complexity; these are shown by horizontal or diagonal lines for all generalized classes of all four types (L = 3–6) of the four-route mechanisms. From Figure 5 one can see that the KG transformations that increase complexity include all $B \rightarrow C$ transitions, as well as some of the $A \rightarrow B$ ones. The first trend deals with replacing the common vertex between two KG cycles with a common edge (or, otherwise, with replacing a common intermediate with a common elementary step). The second trend, the replacement of a bridge between two KG cycles with a common vertex, is weaker because both are a "weak" type of cycle linkage. The increase in complexity in such cases comes (see eq 7) from the increase by 1 in the size of one of the KG cycles in order to preserve a constant total number of KG vertexes. However, in those cases in which the $A \rightarrow B$ transformation can be performed by cycle displacements only (i.e., without any alteration of the cycle sizes), the complexity index remains unchanged.

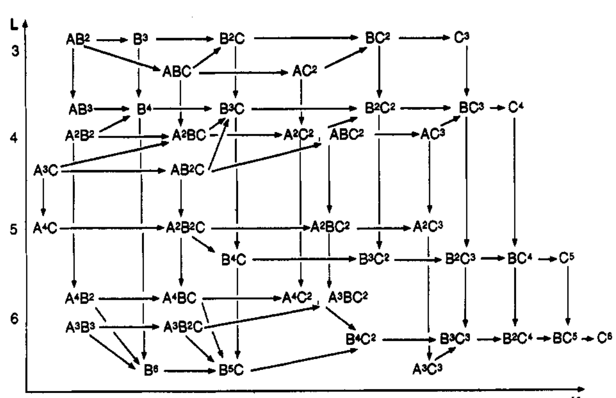


Figure 5. Complexity flow chart for the classes of four-route linear mechanisms.

V. CONCLUDING REMARKS

This study reports the first large-scale enumeration of the theoretically possible linear mechanisms of chemical reactions. On the basis of the novel concept that mechanisms have a topological structure, this is an exhaustive enumeration which indicates that the number of topologically distinct mechanisms of complex chemical reactions could be very large. This finding differs drastically from the few other known attempts[3,4] at mechanism enumeration, which proceed from the chemical information on the reactions and produce a rather limited number of distinct mechanisms. Evidently, in order to be complete, any mechanism enumeration should take into account all possible interrelations of reactants, elementary steps, and reaction routes. The enumeration we report in Table 1 is also incomplete. It refers to mechanisms containing only reversible steps. Indeed, a specified number of mechanisms with irreversible elementary reactions can be deduced for each of the mechanisms counted in Table 1. Graph-theoretically, this is the problem of counting the digraphs that correspond to a certain nondirected graph. A second extension of the enumeration procedure may handle mechanisms with reaction intermediates that are involved in an equilibrium elementary step only. In terms of graph theory, this problem can be reformulated as counting the number of graphs with pendant vertexes that correspond to each of the digraphs of interest. Finally, after the exhaustive topological enumeration described above, one could search for procedures that would produce an even larger number of theoretically possible mechanisms by accounting for their chemical specificity. Different classes of chemical reactions or reactants may be incorporated into our enumeration scheme by regarding graphs with weighted edges and/or vertexes. The results obtained by all these developments will be a subject of a future publication.[35] The large numbers of theoretically possible reaction mechanisms, revealed by our method, however, does not necessarily presuppose their real existence. One may expect some of the mechanisms that are topologically allowed to be forbidden for some chemical reasons. The search for such rules of selection in chemical kinetics might be a real challenge.

Another essential part of this study deals with the complexity analysis of linear mechanisms. The complexity index K, introduced in our previous publications, proved to be a reliable tool in assessing the complexity of both the kinetic models and cyclic graphs (KGs) used to represent them. Being derivable from the code developed for the computer storage of linear mechanisms, the K index evidences that our hierarchical mechanistic classification is associated with a systematic increase in the complexity of the types, classes, and subclasses of these mechanisms. By examining all 390 generated mechanisms having 4 reaction routes and up to 6 reaction intermediates, we were able to outline the major trends in increasing or preserving mechanistic complexity. The isocomplex mechanisms were treated in detail by specifying the different hierarchical levels of isocomplexity, as well as by determining the type of KG transformations that preserve mechanistic complexity. This analysis sheds some light on why it is so difficult to discriminate mechanisms with a larger number of reaction routes, the answer being that their complexity is frequently the same or very similar. On the other hand, it may be of theoretical interest to treat the isocomplexity problem not by using equations like eq 7 but, more generally, by finding the necessary and sufficient conditions for two cyclic graphs to have the same total number of spanning trees, as well as the same number of spanning trees in the subgraphs corresponding to the algebraic complements of the graph cycles.

Besides being of academic interest, this study is also practical. It is related to the creation in the Lomonosov Institute of Fine Chemical Technology in Moscow of a system for computer-assisted mechanism elucidation with a data base of mechanisms within a large range of reaction intermediates, elementary steps, and reaction routes. The modifications to the previously developed hierarchical classification and code of the linear mechanisms, simplified the computer storage and retrieval of mechanisms. The flow chart developed for the four-route mechanisms, along with those of the two- and three-route mechanisms,[21] reveals the mechanism topological interrelations that increase or preserve complexity will facilitate the planning of kinetic experiments for more effective model discrimination.

REFERENCES AND NOTES

(1) Sellers, P. H. Algebraic Complexes Which Characterize Chemical Networks. *SIAM J. Appl. Math.* **1967**, *15*, 13–68.

(2) Sellers, P. H. Combinatorial Analysys of Chemical Networks. *J. Franklin Inst.* **1970**, *290*, 113–130.

(3) Sellers, P. H. Introduction to a Mathematical Theory of Chemical Reaction Networks, Parts 1 and 2. *Arch. Ration. Mech. Anal.* **1971**, *44*, 23–40, 376–386.

(4) Sellers, P. H. In *Chemical Applications of Graph Theory*; King, R. B., Ed.; Elsevier: Amsterdam, 1983; pp 420–429.

(5) Clarke, B. L. In *Advances in Chemical Physics*; Prigogine, I., Rice, S. A., Eds.; Wiley: New York, 1980; Vol. 43, pp 1–215.

(6) Clarke, B. L. In *Chemical Applications of Topology and Graph Theory*; King, R. B., Ed.; Elsevier: Amsterdam, 1983; pp 322–357.

(7) Snagovskii, Y. S.; Ostrovskii, G. M. *Modeling Kinetics of Heterogeneous Catalytic Processes*; Khimiya: Moscow, 1976 (in Russian).

(8) Barone, R.; Chanon, M.; Green, M. L. H. Possible Mechanisms for Ethylene Dimerization Derived by Use of a New Computer Program. *J. Organomet. Chem.* **1980**, *185*, 86–93.

(9) Theodosiou, I.; Barone, R.; Chanon, M. New Computer Program Able to Propose Mechanistic Schemes and Secondary Products for Transformations Catalyzed by Transition Metal Complexes. Possible Mechanisms for Hydroformulation of Ethylene. *J. Mol. Catal.* **1985**, *32*, 27–50.

(10) Theodosiou, I.; Barone, R.; Chanon, M. Computer Aids for Organometallic Chemistry and Catalysis. *Adv. Organomet. Chem.* **1986**, *26*, 165–216.

(11) Trach, S. S.; Zefirov, N. S. Problems of Molecular Design and Electronic Computers. 7. Formal-Logical Approach to Organic Reactions. Basic Concepts and Terminology. *Zh. Org. Khim.* **1982**, *18*, 1561–1583.

(12) Zefirov, N. S.; Trach, S. S. Symbolic Equations and Their Applications to Reaction Design. *Anal. Chim. Acta* **1990**, *235*, 115–134.

(13) Brouk, L. G.; Temkin, O. N. In *Catalytic Reaction Mechanisms (3rd All-Union Conference)*, Part 2; Nauka: Novosibirsk, 1982; p 101 (in Russian).

(14) Sinanoğlu, O.; Lee, L. S. Finding All Possible a Priori Mechanisms for a Given Type of Overall Reaction. Molecular Rearrangements (A–)B. 2. Molecular Association (A+B–)C Reaction Types. *Theor. Chim. Acta* **1978**, *48*, 287–299. Sinanoğlu, O.; Lee, L. S. Finding the Possible Mechanisms for a Given Type of Overall Reaction. Case of the (A+B) – (C+D) Overall Reaction Types. *Theor. Chim. Acta* **1979**, *51*, 1–9.

(15) Lee, L. S.; Sinanoğlu, O. Reaction Mechanisms and Chemical Networks. Types of Elementary Steps and Generation of Laminar Mechanisms. *Z. Phys. Chem.* **1981**, *125*, 129–160.

(16) Sinanoğlu, O. 1-Topology and 2-Topology of Reaction Networks. *J. Math. Phys.* **1981**, *22*, 1504–1512.

(17) Sinanoğlu, O. Autocatalytic and Other General Networks for Chemical Mechanisms, Pathways, and Cycles: Their Systematic and Topological Generation. *J. Math. Chem.* **1993**, *12*, 319–363.

(18) Bonchev, D.; Temkin, O. N.; Kamenski, D. On the Classification and Coding of Linear Reaction Mechanisms. *React. Kinet. Catal. Lett.* **1980**, *15*, 113–118.

(19) Bonchev, D.; Temkin, O. N.; Kamenski, D. On the Complexity of Linear Reaction Mechanisms. *React. Kinet. Catal. Lett.* **1980**, *15*, 119–124.

(20) Bonchev, D.; Temkin, O. N.; Kamenski, D. Graph Theoretical Classification and Coding of Chemical Reactions with a Linear Mechanism. *J. Comput. Chem.* **1982**, *3*, 95–111.

(21) Bonchev, D.; Kamenski, D.; Temkin, O. N. Complexity Index for the Linear Mechanisms of Chemical Reactions. *J. Math. Chem.* **1987**, *1*, 345–388.

(22) Temkin, O. N.; Bruk, L. G.; Bonchev, D. Topological Structure of Complex Reaction Mechanisms. *Teor. Eksp. Khim.* **1988**, *24*, 282–291.

(23) Temkin, O. N.; Bonchev, D. In *Mathematical Chemistry, Graph Theory. Introduction and Fundamentals*; Bonchev, D., Rouvray, D. H., Eds.; Gordon and Breach: Chichester, U.K., 1991; Vol. II, Chapter 2.

(24) Temkin, O. N.; Bonchev, D. Application of Graph Theory to Chemical Kinetics. Part 1. Kinetics of Complex Reactions. *J. Chem. Educ.* **1992**, *69*, 544–550.

(25) Temkin, O. N.; Zeigarnik, V. A.; Bonchev, D. In: *Graph-Theoretical Approaches to Chemical Reactivity*, Bonchev, D., Mekenyan, O., Eds.; Kluwer: Dordrecht, The Netherlands, in press.

(26) Zeigarnik, A. V.; Temkin, O. N. Formal-Logical Approach to Constructive Enumeration of Elementary Reaction Classes in Catalysis by Transition Metal Complexes. *J. Math. Chem.*, submitted for publication.

(27) Temkin, M. I. Kinetic Equations of Complex Reactions Deduced by Theory of Graphs. *DAN SSSR* **1965**, *165*, 615–618.

(28) Temkin, M. I. In *Mechanism and Kinetics of Complex Reactions*; Roginskii, S. Z., Ed.; Nauka: Moscow, 1970; pp 57–71 (in Russian).

(29) Rozovskii, A. Y.; Lin, G. I. *Theoretical Foundations of the Methanol Synthesis Process*; Khimiya: Moscow, 1990 (in Russian).

(30) Gordeeva, E. V.; Molchanova, M. S.; Zefirov, N. S. General Methodology and Computer Program for the Exhausted Restoring of Chemical Structures by Molecular Connectivity Index. Solution of the Inverse Problem in QSAR/QSPR. *Tetrahedron Comput. Method.* **1990**, *3*, 389–415.

(31) Faradjev, I. A. Generation of Nonisomorphic Graphs with a Given Distribution of Vertex Types. *Algorithmic Investigations in Combinatorics*; Nauka, Moscow, 1978; pp 11–19.

(32) Vol'kenshtein, M. V.; Gol'dshtein, B. N. Application of Graph Theory to Calculation of Complex Reactions. *Dokl. Acad. Nauk SSSR* **1966**, *170*, 963–965.

(33) Vol'kenshtein, M. V. *Physics of Enzymes*; Nauka: Moscow, 1967; p 155 (in Russian).

(34) Harary, F. *Graph Theory*; Addison-Wesley: Reading MA, 1969.

(35) Gordeeva, E.; Bonchev, D.; Temkin, O. N. Work in progress.