Topological Ligand Design. I. Definition and Derivation of Simple Ligand Graph*

Susumu Такамото

Department of Chemistry, Faculty of Sciences, Gakushuin University,
Mejiro, Toshima-ku, Tokyo 171
(Received December 18, 1978)

A simple ligand graph and a pattern graph are defined topologically in order to design the patterns of the chelating ligands. This ligand graph, which comprises m donor atoms and n carbon chains capable of forming medium-sized chelate rings, is an irregular graph containing neither multiple lines nor self-loops. All the ligand graphs with $m \le 8$ can be systematically obtained from fifty pattern graphs. These are derived from fifteen basic pattern graphs, which are cubic general graphs containing some multiple lines and self-loops. A pattern code and a set of series indices are tabulated for each ligand graph with $m \le 8$. For the ligand graphs with $m \le 6$, all the geometric isomers of their octahedral chelates are also illustrated.

The chemistry of metal chelates has been remarkably developed and is widely applied. The properties and the reactions of the metal chelates are closely related to their structure. When one seeks the most suitable new chelating reagents for a certain purpose, it would be of great use to look over all the patterns of ligand structures. In their reviews on the ligand design, Goodwin¹⁾ and Black and Hartshorn²⁾ were chiefly concerned with the systematic outline and classification of the chelate agents already prepared, but their topological diagrams were mathematically incomplete. Thus, purely geometrical investigations are not available for the sysematic derivation of ligand patterns or for the design of new ligands. The graph theory and computer studies have been successfully applied to the numeration and the documentation of the chemical isomers or the homologs for some special kinds of organic compounds.3) In order to cover all the ligands, one has to treat all the organic compounds, since every compound containing some oxygen- or nitrogen-group elements has more or less ability to coordinate to a metal ion. This designing work is limited to those practical ligands which can form five- or six-membered chelate rings with a metal ion.

Theory and Results

A ligand can be considered as a set of some donor atoms joined by several carbon chains capable of forming medium-sized chelate rings. Then a simple ligand graph L(m, n) can be defined for each ligand, where m is the number of points which correspond to the donor atoms, and n is the number of lines which correspond to the carbon chains. The ligand graph must meet the following requirements: (1) The graph is connected. (2) It is planar. (3) The degree of each point is one, two, or three (The degree of a point is the number of lines incident with it). The first condition is self-evident. A planar graph can be drawn on a plane so that no two lines cross each other. According to the graph theory such a graph is identical to a spherical graph. To be spherical is needed for two carbon chains not to cross each other on a coordination sphere, because these chains are too short

to detour. Three would be the maximum degree of a point in any kind of donor atom. Then, for instance, ethylenediamine, 1,3-propylenediamine, 2,3-butenediamine, N,N,N',N'-tetramethylethylenediamine, 1,2-phenylenediamine, and 1,2-cyclohexenediamine are represented by the same ligand graph L(2, 1), since these ligands contain two donor atoms and only one carbon chain which can form a chelate ring. The graph L(2, 1) is unique, but it should be noted that a graph L(m, n) is not generally determined only by a set of m and n.

A graph in the narrow sense of the word does not include multiple lines; that is, two points may be joined by only one line. Molecules such 1,4-diazacycloheptane, for example, can form metal chelates, so some double lines may be included in the ligand graph. But relatively few such ligands form stable chelates because of the steric hindrance and the bond strains. Moreover, this limited number of double lines can be easily derived by joining a pair of adjacent points of degree one or two in a simple ligand graph. Therefore, no multiple lines are included in this simple ligand graph. Also, no self-loop is needed in this graph, since it forms no chelate ring. Thus the simple ligand graph is a normal irregular graph.

Some new terms should be defined. When all the points of degree two in a ligand graph are neglected, the remainder is named a pattern graph P(p, q), and each line in the pattern graph is called a series. Although an isolated circle without a point may not be generally accepted as a graph, in this study it is by exception admitted as a pattern graph. The degree of every point in the pattern graph is one (end point) or three (branch point). A pattern graph may contain some triple and double lines and some self-loops. This value of p is always an even number because the degree of every point is odd.

When all side-chains are removed from a pattern graph, the remainder is called a basic pattern graph B(r,s). All basic pattern graphs but the isolated circle B(0,1) and an isolated line B(2,1) are cubic general graphs; that is, all their points have the same degree of three, and they also may have multiple lines and self-loops. Examples of each graph are shown in Fig. 1.

All the cubic general graphs for $m \le 8$ can be constructed from the two simplest cubic graphs B(2,3)

^{*} A preliminary report of this work was presented at the 28th Symposium on Coordination Chemisrty, Matsuyama, October 1978.

$$^{\text{H}_2\text{N-C}_2\text{H}_4}$$
 $^{\text{N-C}_2\text{H}_4}$ $^{\text{N-C}_2\text{H}_4}$ $^{\text{N-C}_2\text{H}_4}$ $^{\text{C}_3\text{H}_6}$ $^{\text{C}_3\text{H}_6}$ $^{\text{N-C}_2\text{H}_4}$ $^{\text{N-C}_2\text{H}_4}$ $^{\text{N-C}_2\text{H}_4}$ $^{\text{N-C}_2\text{H}_4}$ $^{\text{N-C}_2\text{H}_4}$

Ligand

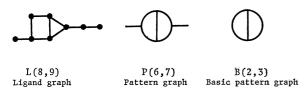


Fig. 1. Reduction of a ligand structure into the basic pattern graph.

by adding new branch lines one by one avoiding the duplication of patterns, up the maximum values of r:

$$r \le m - d - 2t - 2l - 3c \tag{1}$$

where d, t, l, and c are the numbers of double lines, triple lines, self-loops, and isolated circes. This restriction is needed to exclude all the multiple lines and self-loops from the ligand graphs which are derived from the pattern graphs. Naturally r is even number.

All the basic pattern graphs for $m \le 8$ are shown in Fig. 2. The values of t and c are only applicable in the case of (2-1) and (1-1) in Fig. 2 respectively. Then, for example, (4-2) and (4-4) are not the basic pattern graphs for $m \le 7$, because d of (4-2) is 2, and l of (4-4) is 1.

The sum of the degree of the points of a graph G(x, y) is twice the number of lines, and the number of independent cycle i is equal to y-x+1. Then in any cubic graph B(r, s) the following relationships are obtained:

$$s = 3r/2, (2)$$

$$i = 1 + r/2. \tag{3}$$

Fig. 2. Basic pattern graphs $(m \leq 8)$.

The pattern graphs are systematically derived from the basic pattern graphs by adding some side-chains one by one, while basic pattern graphs themselves belong to the set of pattern graphs. The maximum values of p is also restricted by the right hand side of Eq. 1, although p is not less than r of the original basic pattern graph. Fifty pattern graphs thus obtained for $m \le 8$ are illustrated in Fig. 3. As the numbers of points and lines are incressed by the same integer by adding a side-chain to a graph, q-p is always equal to s-r of its original basic pattern graph. However, the number of independent circles i is unchanged by any addition of side-chains.

Similarly, all the ligand graphs are derivable by allotting a definite number of points of degree two to each series of the pattern graph. Ligand graphs thus derived from the same pattern graph are topologically homeomorphic to one another. In this study a pattern graph is simply denoted by a pattern code classified by the basic pattern groups, and alphabetical labels a to h are assigned to each series of the pattern graphs, as shown in Fig. 3. Then all the simple ligand graphs can be characterized by a pattern code and a set of eight series indices, which express the number of the points of degree two allotted to each labeled series of the pattern graph. These allotted numbers are properly restricted so that the numerating repetition which is caused by the symmetry of the ligand graph may be avoided, and so that every cycle in ligand graphs may contain at least three points, which are needed to exclude any multiple lines and self-loops. In the case of Fig. 1, two points of degree two are allotted for series a and h of the pattern graph (24), because series α has the highest priority among the four equivalent series a, b, c, and d in order to avoid the repetition. All the ligand graphs with $m \leq 8$ and their pattern codes and series indices are listed in Table 1. The numbers of the simple ligand graphs thus designed and of their double-line derivatives are summarized in Table 2.

Each ligand may form some configurationally isomeric chelates with a metal ion in a certain coordination structure. As an example, all the geometric configurations in the octahedral coordination for every simple ligand graph with $m \le 6$ are illustrated in Table 3. They were confirmed by use of molecular models.

Discussion

In this study double lines are not included in the simple ligand graph, but double-line derivatives have been derived from the simple ligand graph. All of them, as well as the simple ligand graphs, can directly be derived from the pattern graphs if the restriction 1 is moderated as follows:

$$r \leq p \leq m - t - l - 2c. \tag{4}$$

Then the numbers of the pattern graphs and the basic pattern graphs are increased from 50 and 15 to 112 and 34 respectively. There is little merit in complicating the pattern graphs so much, since the number of such less important ligands is limited.

Mechanical derivations of compounds often have

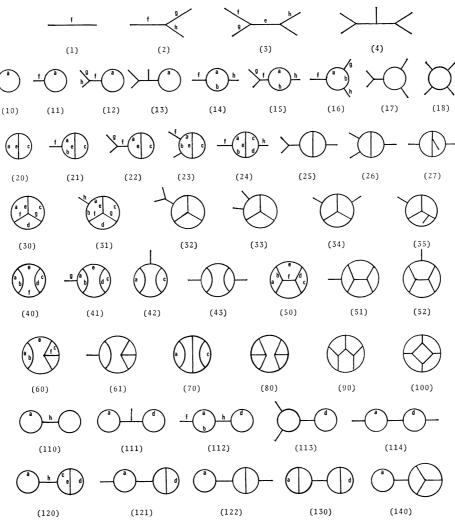


Fig. 3. Pattern graphs and labeling of series $(m \le 8)$.

a tendency to introduce some useless ligands which may offer some difficulties for coordinating. The present manner of deriving the ligand graphs is sufficiently systematic that it produces no such worthless ligands.

Every ligand graph with m=5 can cover all the trigonalbipyramidally or square-pyramidally disposed positions. Every planar graph with m=6 can also cover all the octahedral points, but not all the points of a trigonal prism. The planar m-polygonal corners can be covered with only two ligand graphs of order m which have a linear or a circular shape.

Since any donor atoms of nitrogen or oxygen group elements may be held at any points of a ligand graph, ligands of mixed donor atoms may also be expressed by the same ligand graphs. Then, for example, most of the simple cryptands belong to No. 8-055, and edta or its analogs to No. 6-04 in Table 1. The lines in a simple ligand graph might represent not only a carbon chain but also boron-, silicon-, or any other chains.

Previously, the author⁴⁾ reported that 1,4,7-triazacyclononane (a type of No. 3-02) forms the most stable metal chelates in all the triamines, and that 1,2-di(1,4,7-triazacyclononan-1-yl)ethane, which is

composed of two such triazacycloalkanes connected by a ethylene group (a type of No. 6-28) is also a good chelating agent.⁵⁾ Since cyclam (a type of No.4-03) is well known for the extreme stability of its chelates, a ligand of type No. 6-26, which is composed of two triazacycloalkanes and one tetraazacycloalkane, should form the most stable chelates.

Ligands of type No.4-06 or No.6-27 have a basketlike structure. All the lone-pair electrons of their donor atoms are directed to the center of the polyhedron, so no metal ion might enter inside or be coordinated. But if these ligands could be synthesized around a metal ion by some means, the metal atom should be confined completely within their network, and a hydrophobic spherical cation would be produced.

It may be seen from the above lists that only a few types of ligands have ever been prepared; many types with interesting structures still remain unknown. For example, most studies on sexidentate ligands were focused upon only five types of No.6-01 to No.6-05 among the twenty eight ligand graphs with $m \le 6$; therefore a number of new ligands with various patterns can be synthesized and investigated in the near future. Showing their possibilities is the most important aim and the successful fruit of this work.

Table 1. Catalog of ligand graphs $(m \leq 8)$

		g :					`				
Graph	Pat- tern	Series indices,	Graph	Pat- tern	Series indices,	Graph	Pat- tern	Series indices,	Graph	Pat- tern	Series indices,
No.	code	adcdef gh	No.	code	abcdef gh	No.	code	abcdef gh	No.	code	abcdefgh
9.01		0	7-01		5	8-001		6	8-064		103-00
2-01	1	0	7-01 7-02	$\frac{1}{2}$	300	8-001 8-002	1 2	400	8-065	21 21	103-00
			7-02 7-03	2	210	8-003	2	310	8-066	21	102-10
			7-03 7-04	2	111	8-004	2	220	8-067	21	101-11
3-01	1	1	7-05	3	1000	8-005	2	211	8-068	21	101-02
3-02	10	3	7-06	3	0100	8-006	3	2000	8-069	21	004-00
			7-07	10	7	8-007	3	1100	8-070	21	003-10
			7-08	11	50	8-008	3	0200	8-071	21	003-01
			7-09	11	41	8-009	3	0110	8-072	21	002-20
4-01	1	2	7-10	11	32	8-010	3	0101	8-073	21	002-11
4-02	2	000	7-11	11	23	8-011	4		8-074	21	002-02
4-03	10	4	7-12	12	3000	8-012	10	8	8-075	21	001-12
4-04	11	20	7-13	12	2100	8-013	11	60	8-076	21	001-03
4-05	20	1-1-0	7-14	12	2010	8-014	11	51	8-077	22	1-1-000-
4-06	30	0-00000-	7-15 7-16	14 14	300-0 210-0	8-015 8-016	11 11	42 33	8-078	22	0-2-000-
			7-16 7-17	14 14	201-0	8-016 8-017	11	33 24	8-079 8-080	22 22	0-1-100- 0-1-010-
			7-17 7-18	14	111-0	8-017	12	4000	8-081	22	0-1-010-
5-01	1	3	7-10 7-19	14	102-0	8-019	12	3100	8-082	23	101-00
5-02	$\overset{1}{2}$	100	7-20	14	101-1	8-020	12	3010	8-083	23	011-00
5-03	10	5	7-21	15	10000	8-021	12	2200	8-084	23	002-00
5-04	11	30	7-22	16	10000	8-022	12	2110	8-085	23	001-10
5-05	11	21	7-23	16	0100	8-023	12	2020	8-086	23	001-01
5-06	14	100-0	7-24	20	4-1-0	8-024	12	2011	8-087	24	200000-0
5-07	20	2-1-0	7-25	20	3-2-0	8-025	13	2	8-088	24	110000-0
5-08	20	1-1-1	7-26	20	3-1-1	8-026	14	400-0	8-089	24	101000-0
5-09	21	001-00	7-27	20	2-2-1	8-027	14	310-0	8-090	24	100100-0
5-10	30	1-00000-	7-28	21	201-00	8-028	14	301-0	8-091	24	100010-0
			7-29	21	111-00	8-029	14	220-0	8-092	24	100001-0
			7-30 7-31	$\frac{21}{21}$	102-00 101-10	8-030 8-031	14	211-0	8-093	24	100000-1
6-01	1	4	7-31 7-32	21	101-10	8-032	14 14	202-0 201-1	8-094 8-095	$\frac{24}{24}$	000020-0 000011-0
6-02	2	200	7-32 7-33	21	003-00	8-033	14	112-0	8-096	24	000011-0
6-03	2	110	7-34	21	002-10	8-034	14	111-1	8-097	24	000001-1
6-04	3	000	7-35	21	002-01	8-035	14	203-0	8-098	25	
6-05	10	6	7-36	21	001-11	8-036	14	102-1	8-099	26	
6-06	11	40	7-37	12	000-12	8-037	15	20000	8-100	27	
6-07	11	31	7-38	22	0-1-000-	8-038	15	11000	8-101	30	4-00000-
6-08	11	22	7-39	23	001-00	8-039	15	10100	8-102	30	3-10000-
6-09	12	2000	7-40	24	100000-0	8-040	15	10010	8-103	30	3-00001-
6-10	14	200-0	7-41	24	000010-0	8-041	16	10001	8-104	30	2-20000-
6-11	14	110-0	7-42	24	000001-0	8-042	16	20000	8-105	30	2-11000-
6-12	14	101-0 00000	7-43 7-44	30 30	3-00000- 2-10000-	8-043 8-044	16	11000	8-106	30	2-10100-
6-13	16 20	3-1-0	7-44 7-45	30 30	2-10000-	8-044 8-045	16 16	10100 10001	8-107 8-108	30 30	2-10010-
6-14 6-15	20 20	3-1-0 2-2-0	7-43 7-46	30 30	1-11000-	8-045	16	00200	8-108 8-109	30 30	2-10001- 2-00002-
6-15 6-16	20	2-1-1	7-40 7-47	30	1-10100-	8-047	16	00200	8-110	30	1-11100-
6-17	21	1-1-00	7-48	30	1-10010-	8-048	17		8-111	30	1-11100-
6-18	21	002-00	7-49	31	10000000	8-049	18		8-112	31	20000000
6-19	21	001-10	7-50	31	00100000	8-050	20	5-1-0	8-113	31	11000000
6-20	21	001-01	7-51	31	00000010	8-051	20	4-2-0	8-114	31	10100000
6- 21	24	000000-0	7-52	31	00000001	8-052	20	4-1-1	8-115	31	10010000
6- 22	30	2-00000-	7-53	40	201000	8-053	20	3-3-0	8-116	31	10000010
6-23	30	1-10000-	7-54	40	111000	8-054	20	3-2-1	8-117	31	10000001
6-24	30	1-00001-	7-55	40	110010	8-055	20	2-2-2	8-118	31	00200000
6-25	31	00000000	7-56	41	00100-00	8-056	21	301-00	8-119	31	00110000
6-26	40 50	101000 000000	7-57 7-58	50 50	100000 000010	8-057 8-058	21 21	211-00 202-00	8-120 8-121	31	00101000
6-27 6-28	30 110	220	7-36 7-59	60	100-00	8-059	21	201-10	8-121 8-122	31 31	00100100 00100010
U-40	110	440	7-39 7-60	110	320	8-060	21	201-10	8-123	31	00100010
			7-61	110	221	8-061	21	112-00	8-124	31	00000020
			7-62	112	10-2-0-0	8-062	21	111-10	8-125	31	00000011
			7-63	120	2-0100	8-063	21	111-01	8-126	31	00000002

Table 1. (Continued)

Graph No.	Pat- tern code	Series indices, abcdefgh									
8-127	28		8-143	41	00110-0-	8-159	60	101-00	8-175	112	10-2-1-0
8-128	34		8-144	41	00101-0-	8-160	60	100-10	8-176	112	10-2-0-1
8-129	34		8-145	41	00100-0-	8-161	60	100-01	8-177	113	2
8-130	35		8-146	42	1-1	8-162	61		8-178	114	11
8-131	40	301000	8-147	43		8-163	70	1-1	8-179	120	3-0100
8-132	40	211000	8-148	50	200000	8-164	80		8-180	120	2-1100
8-133	40	202000	8-149	50	110000	8-165	90		8-181	120	2-0200
8-134	40	201100	8-150	50	101000	8-166	100		8-182	120	2-0110
8-135	40	201010	8-151	50	100100	8-167	110	420	8-183	120	2-0101
8-136	40	111100	8-152	50	100010	8-168	110	330	8-184	121	11
8-137	40	111010	8-153	50	100001	8-169	110	321	8-185	122	2
8-138	40	101020	8-154	50	000020	8-170	110	222	8-186	130	11
8-139	40	101011	8-155	50	000011	8-171	111	22	8-187	140	2
8-140	41	10100-0-	8-156	51		8-172	112	20-2-0-0			
8-141	41	01100-0-	8-157	52		8-173	112	11-2-0-0			
8-142	41	00200-0-	8-158	60	110-00	8-174	112	10-3-0-0			

Table 2. Number of Ligand graphs $(m \leq 8)$

Donor atoms	2	3	4	5	6	7	8	
Basic pattern graphs	1	2	4	4	7	9	15	
Pattern graphs	1	2	6	8	16	24	50	
Ligand graphs	1	2	6	10	28	63	187	
Double-line derivatives	1	2	6	12	37	102	318	

Table 3. Octahedral configurations ($m \leq 6$)

Graph Configura- number tions	Graph Configura- number tions	Graph Configura- number tions		Graph Configura- number tions
2-01	5-01	6-01 () () 😂 🐳	6-11 🖒 🚫 🛇	6-21
3-01	5-02	6-02	6-12 🖒 🔾	6-22
3-02	5-03	6-03	6-13	6-23
4-01	5-04 🖒 🚫	6-04	6-14 🔷 🕏	6-24
4-02	5-05 💍 🧓	6-05 🔷 💢	6-15	6-25
4-03 .	5-06 🔾 🔾	6-06 🔘 🖒	6-16	6-26
4-04	5-07	6-07 🖒 🖒	6-17 🔘 🚫 🚫	6-27
4-05 .	5-08	6-08 🗘 🔾	6-18	6-28
4-06 (🛆)	5-09	6-09	6-19	
	5-10	6-10	6-20	

A ligand containing some special branched carbon atoms, where the branches might form several chelate rings, cannot be represented by a simple ligand graph, for a line of a graph can branch only at the point representing a donor atom. In order to express such complicated ligands the definition of the ligand graph

must further be extended. These studies will be reported in a later paper.

The author wishes to express his thanks to Prof. Takeshi Iizuka of Gunma University for valuable suggestions. The present work was partially supported by a Grant-in-Aid for Scientific Research from the Ministry of Education.

References

1) H. A. Goodwin, "Chelating Agents and Metal Chelates," ed by F. P. Dwyer and D. P. Mellor, Academic Press, New York (1964), pp. 143—181.

- 2) D. St. C. Black and A. J. Hartshorn, Coord. Chem.
- Rev., 9, 219 (1973).
 3) "Chemical Applications of Graph Theory," ed by A. T. Balaban, Academic Press, London (1976).
- 4) T. Arishima, K. Hamada, and S. Takamoto, Nippon Kagaku Kaishi, 1973, 1119.
- 5) N. Tanaka, Y. Kobayashi, and S. Takamoto, Chem. Lett., 1977, 101.