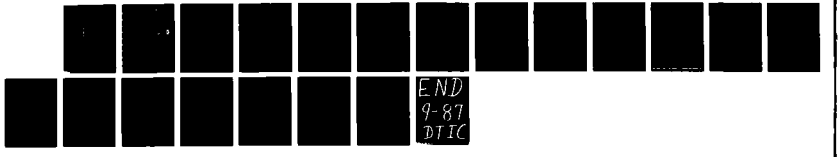
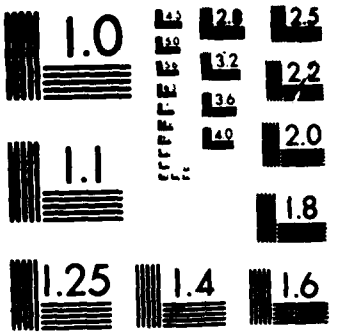
AD-A183 699 NOVEL APPLICATION OF RANDIC MOLECULAR CONNECTIVITY INDEX ON DATA REDUCTION OF CHEMICAL GRAPHS(U) GEORGIA UNIV ATHENS DEPT OF CHEMISTRY S EL-BASIL 11 AUG 87 TR-53 N00014-84-K-0365 F/G 7/3 1/1

UNCLASSIFIED NL



END
9-87
DTIC



OFFICE OF NAVAL RESEARCH

Contract N00014-84-K-0365

R&T Code 4007001-6

Technical Report No. 53

Novel Application of Randić Molecular Connectivity Index On Data Reduction of Chemical Graphs

by

Sherif El-Basil

Prepared for Publication

in the

Chem. Phys. Letters

University of Georgia
Department of Chemistry
Athens, Georgia 30602

August 11, 1987

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

A183 699

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Report No. 53	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Novel Application of Randić Molecular Connectivity Index on Date Reduction of Chemical Graphs		5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Sherif El-Basil		8. CONTRACT OR GRANT NUMBER(s) N00014-84-K-0365
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Georgia Department of Chemistry Athens, GA 30602		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 4007001-6
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Department of the Navy Arlington, VA 22217		12. REPORT DATE August 11, 1987 13. NUMBER OF PAGES 16
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) 15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

This document has been approved for public release and sale; its distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

To be published in Chem. Phys. Letters

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Graph Theory
Randić Molecular Connectivity Index
Data Reduction
Caterpillar Trees
Benzenoid Hydrocarbons

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Benzenoid hydrocarbons are studied in terms of the much simpler caterpillar trees. Using molecular connectivity indices of the latter almost exact linear relations are obtained with natural logarithms of five properties of benzenoid hydrocarbons including all self-avoiding paths, conjugated circuits, number of Kekulé structures, electronic absorption spectra, and heats of atomizations.

(Key words)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

137 (6), 543 - 547 (1987

Novel Applications of Randić Molecular Connectivity Index

On data reduction of chemical graphs

Sherif El-Basil

Department of Chemistry, University of Georgia

Athens, GA 30606 U.S.A.

Benzenoid hydrocarbons are studied in terms of the much simpler caterpillar trees. Using molecular connectivity indices of the latter almost exact linear relations are obtained with natural logarithms of five properties of benzenoid hydrocarbons including all self-avoiding paths, conjugated circuits, number of Kekulé structures, electronic absorption spectra, and heats of atomizations.

Key words:

- Graph Theory ,
- Randić Molecular Connectivity Index
- Data Reduction
- Caterpillar trees
- Benzenoid Hydrocarbons

1. Introduction and definitions

One of the main goals of every science is the endeavoring to arrange and collate the numerous individual observations and details which present themselves, in order that they become part of one comprehensive picture.[1] Such an objective may be achieved through the use of data reduction which is an essential part of chemical and physical data analysis. In the past the task of data reduction involved mainly curve fitting procedures which have recently become more efficient due to increased availability of computers. Topological approaches which consider additivities of molecular properties have been almost ignored in the past. However some recognition of the role of graph theory became apparent soon after the (important) work of Smolenskii[2] and (to a lesser degree) the work of Gordon and Kennedy[3] who formalized their schemes of general approaches to molecular additivities. Revival of interest in chemical graph theory is probably due to Balaban via his (editing) of a book[4] which provided a source of problems and numerous other papers. Several counting polynomials were defined which act as descriptors of the connectivity pattern of various types of molecular graphs. The latters are well-defined mathematical objects which describe the bonding relation between atoms in the original molecules. Such counting polynomials are then a combinatorial form of data reduction. Equation 1 describes the general form of a graph-theoretical counting polynomial[5], $F(G;x)$, for an arbitrary graph G, containing n vertices:

$$F(G;x) = \sum_{k=0}^{M} \rho\ \theta(G;k)\ x^{f(k,n)} \qquad (1)$$

where ρ is either 1 or $(-1)^k$, $\theta(G;k)$ enumerates certain selected graph-invariants in G taken in k independent tuples (i.e. no two of them are adjacent) and M is

the maximal value of k. For the purpose of this paper we consider three polynomials for all of which $\rho = 1$ and $f(k,n) = n$. These are:

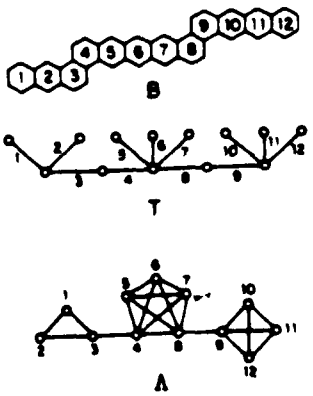
i) The counting polynomial (of Hosoya)[6], H(G;x) for which $\theta(G;k) = p(G;k)$ = the number of k-matchings in G.

ii) The sextet polynomial (of Hosoya and Yamaguchi)[7], $\sigma(B;x)$ of which $\theta(G;k) = r(B;k)$ = the number of selections of k mutually resonant but nonadjacent aromatic sextet in a benzenoid system (≡ polyhex graph[8]), B.

iii) The independence polynomial (of Gutman)[9], $\omega(G;x)$, for which $\theta(G;k) = 0(G;k)$ = the number of selections of k independent vertices of G.

2. The Caterpillar Tree[10], the Polyhex Graph[11] and the Clar Graph[9,12]

In a given Kekulé structure an (aromatic) sextet is defined as a set of three double bonds circularly conjugated so that no two sextets can have a common bond. When two rings (hexagons) can have aromatic sextets so that all other carbon atoms are spanned by a sextet or by double bond, such two rings are called resonant.[9,12] The individual hexagons of a given benzenoid system may or may not be resonant[7]. Information on such resonance relations (among the individual hexagons) are best "reduced" (i.e. described) using either a caterpillar tree[10] or a Clar graph.[9] To illustrate this fundamental relation[13] we consider the three graphs shown below:



The hexagons of B may be grouped into three subsets, viz., {1,2,3}; {4,5,6,7,8}; {9,10,11,12}. One observes that no two hexagons are resonant if they belong to one subset. Similarly the edges of T may be subdivided into analgous three subsets observing that no two edges of the same subset can be adjacent. A similar description obtains for the vertices of graph Λ. In fact bijective (i.e. one-to-one, onto)[14] mapping might be defined between the hexagons of B, the edges of T and the vertices of Λ: the latter is called the Clar graph[9,12] of B. It is easy to verify the following identity:

$$\sigma(B;x) = H(T;x) = \omega(\Lambda;x) = 1 + 12X + 45X^2 + 53X^3 \qquad (2)$$

A caterpillar tree such as T whose counting polynomial is identical to the sextet polynomial of a benzenoid hydrocarbon B will be called an equivalent caterpillar to the benzenoid hydrocarbon. the set of graphs {B,T,Λ} shown above will be called a set of equivalent graphs. Gutman[15] demonstrated that for every nonbranched[16] benzenoid hydrocarbon there is an "equivalent" caterpillar tree.

3. Conjecture

The physical properties of a benzenoid hydrocarbon which depend on its topological structure may be predicted from graph-theoretical properties of its equivalent graphs.

4. Results and tests of Conjectures

Fig. 1 shows benzenoid hydrocarbons and their equivalent caterpillar trees studied in this work.

4.a Graph theoretical (combinatorial) properties

Three graph-theoretical properties of benzenoid hydrocarbons, viz., self-avoiding paths[19], conjugated circuits[20] and the number of Kekulé structures (i.e. number of perfect matchings) are correlated with the molecular connectivity indices,[21] χ's of the equivalent caterpillar (or pseudocaterpillar) trees. Fig. 2 shows the types of linear correlations obtained. In all cases correlation coefficients $\geq$ 0.999 were obtained.

4b. Physical properties

Two physical properties viz., electronic absorption spectra[22], and heats of atomizations[23] of several homologous series of benzenoid hydrocarbons are linearly correlated with χ (T)'s; the molecular connectivity indices[21] of the corresponding equivalent trees. Details of data are available on request[17] but as an illustraton Table 1 is included which contains β and p bands of the electronic spectra of two families of hydrocarbons of Fig. 3. A separate manuscript is submitted elsewhere[24] dealing with electronic absorption spectra of 27 arene systems all linearly correlated with χ(T) values. The same type of data reduction is obtainable[17] when molecular properties (topological or physical) of benzenoid hydrocarbons are studied in terms of connectivity functions of their equivalent Clar graphs.[25] Fig. 4 is a plot of heats of atomizations of some benzenoid hydrocarbons versus connectivity indices of the corresponding caterpillar trees.

6. Discussion

Caterpillar trees and Clar graphs seem to be quite promising "storage devices" which preserve many properties of (the much larger) equivalent polyhex graphs. In contrast to the Balaban-Harary dualist graphs[26] they do not have a geometric element and as such, connectivity alone best describes them. It is amazing that some of the studied properties such as the total number of paths[19] in a polyhex graph (Fig. 2) take nearly 30 minutes on a Corona computer,[27] yet can be retrieved by very simple hand calculation of the connectivity index of a much simpler graph! Table 2 lists percent retrievals of the total number of self-avoiding paths for a series of nonbranched benzenoid hydrocarbons shown in Fig. 1. All the above mentioned correlations imply classification of molecules into families (i.e. homologous series). The concept of classification is well-known in science: the periodic table of the elements and recently table of alkanes introduced by Randić and Wilkins[28] and more recently table of benzenoid hydrocarbons of Dias[29] are all non-mathematical forms of set theory. However in constrast to the latter two tables, which result from graph theoretical considerations, the "old" periodic table is a result of group theory as it follows from symmetry properties of atoms.

Acknowledgments

I thank the U.S. Office of Naval Research for partial support of this work. Illuminating discussions and critical remarks made by Prof. R.B. King are appreciated.

Table 1

Electron absorption spectra of the β bands, λ_β, of the homologous series of benzenoid hydrocarbons shown in Fig. 3. $^1\chi(T)$'s are molecular connectivity indices of the equivalent pseudocaterpillar trees.

Hydrocarbon*	$^1\chi(T)$	λ_β (Å)	λ_p(Å)
A	3.914	2570	
B	4.684	2870	
C	5.293	3130	
D	5.812	3430	
1	1.000		2068
2	1.414		2850
3	1.732		3745
4	2.000		4710
5	2.236		5755
6	2.449		6930

*c.f. Fig. 3

Table 2

Percent retrieval of the total number of self-avoiding paths for the series of nonbranched benzenoid hydrocarbons shown in Fig. 2. The regression line is Y = 4.374 + 0.484X with a correlation coefficient = 1.0000.

Hydrocarbon*	Calcd. $\Sigma p_i(B)$*	Actual $\Sigma p_i(B)$*	Difference	% Retrieval[a)]
E	3254	3277	23	99.298
F	7391	7413	22	99.704
G	16055	16055	0	100.000
H	33827	33879	52	99.846
I	69675	69857	182	99.740
J	141060	142645	1585	98.889

*c.f. Fig. 2

a) = 100 - 100|Calcd-Actual|/Actual

Fig. Legends

Fig. 1

Polyhex graphs of benzenoid hydrocarbons studied in this work together with the molecular graphs of their corresponding caterpillar trees.

Fig. 2

Combinatorial properties of benzenoid hydrocarbons studied: K(B), $\Sigma p(B)$ and $\Sigma R_{tot}(B)$ respectively indicate number of Kekulé structures, total number of self-avoiding paths and total number of conjugated circuits of a benzenoid system, B, plotted versus the connectivity index of the corresponding caterpillar tree, $\chi(T)$ raised to various powers. (The hydrocarbons and their caterpillars are shown in Fig. 1).

Fig. 3

Electronic absorption spectra of benzenoid hydrocarbons: Points 1-6 indicate benzene, naphthalene, anthracene, tetracene, pentacene and hexacene; the line passing through these points is a plot of their absorption para band (in ln units) vs. the connectivity index of the corresponding caterpillar trees, $\chi(T)$. The line passing through points A-D is a plot of the beta band vs. $\chi(T)$ where the hydrocarbons and their trees are shown in Fig. 1.

Fig. 4

A plot of the natural logarithms of the heats of atomizations of some benzenoid hydrocarbons, ln Ht At(B), against the connectivity indices, $\chi(T)$'s, of the corresponding caterpillar trees shown in Fig. 1.

References

(1) The statement is quoted from: M. Planck, Survey of Physical Theory, Dover, New York (1960).

(2) E.A. Smolenskii, Russ. J. Phys. Chem. 38 (1964) 700; 38 (1964) 1204.

(3) M. Gordon, and J.W. Kennedy, J. Chem. Soc. Faraday Trans 2, 69 (1973) 484.

(4) A.T. Balaban, "Chemical Applications of Graph Theory" (Academic, London 1977).

(5) S. El-Basil, J. Chem. Soc. Faraday Trans II, 82 (1986), 299.

(6) H. Hosoya, Bull. Chem. Soc. Japan, 44 (1971), 2332.

(7) H. Hosoya and T. Yamaguchi, Tetrahedron Letters, (1975), 4659.

(8) N. Ohkami, A. Motoyama, T. Yamaguchi, H. Hosoya and I. Gutman, Tetrahedron, 37 (1981) 1113.

(9) I. Gutman, Z. Naturforsch 37a (1982) 69; I. Gutman and S. El-Basil Z. Naturforsch 39a (1984) 276.

(10) A caterpillar tree P_n $(m_1, m_2, \ldots, m_n)$ is formed by the addition of m_j monovalent vertices to vertex v_j of path P_n, $1 \leq j \leq n$; where n is the number of vertices in P_n. c.f., F. Harary and A.J. Schwenk, Discrete Math. 6 (1973) 359.

(11) Polyhex graphs considered here are assumed to have at least one perfect match (i.e. one Kekulé structure) c.f., H. Hosoya, Croat Chem. Acta, 59 (1986) 583. See also ref. 8. The treatment of perfect matchings in hexagonal systems is given by H. Sachs, Combinatorica, 4 (1984) 89.

(12) Two vertices in a Clar graph[9] are adjacent if the corresponding hexagons in the polyhex graph are nonresonant and if the corresponding two edges in the caterpillar tree[10] are adjacent. Hence the Clar graph of a benzenoid

hydrocarbon is the line graph of its caterpillar tree, c.f., F. Harary, Graph Theory, Addison Wesley, Reading, Mass. (1972).

(13) I. Gutman, Bull. Soc. Chim. Beograd, 47 (1982) 453.

(14) C.L. Lui, Elements of Discrete Mathematics, McGraw-Hill Bock Company (1977).

(15) I. Gutman, Theor. Chim. Acta, 45 309 (1977).

(16) For definitions see, S. El-Basil, Bull. Chem. Soc. Japan, 56 (1983) 3152.

(17) S. El-Basil, Unpublished Results (available on request).

(18) N. Ohkami and H. Hosoya, Theor. Chim. Acta, 64, (1983), 153.

(19) A self-avoiding path in a molecular graph is a sequence of adjacent bonds with no atom (vertex) appearing more than once in the sequence, c.f., M. Randić and C.L. Wilkins, J. Phys. Chem. 83 (1979) 1525.

(20) M. Randić, Chem. Phys. Letters 38 (1976) 68; J. Am. Chem. Soc. 99 (1977) 444; Tetrahedron, 33 (1977) 1906; J.A.F.N. Gomes, Theor. Chim. Acta, 59 (1980) 333.

(21) The molecular connectivity index of a graph G, $\chi(G)$ is defined by

$$\chi(G) = \Sigma (d_i d_j)^{\frac{1}{2}}$$

where d_i = the number of nearest neighbors to the ith' vertex and the summation is taken over all edges in G, c.f., M. Randić, J. Am. Chem. Soc., 97 6609 (1975).

(22) Data are collected from E. Clar, The Aromatic Sextet, Wiley New York (1972).

(23) c.f., M.J.S. Dewar, J.A. Hasmall and N. Trinajstić, J. Am. Chem. Soc., 92 (1970) 5555.

(24) D.H. Rouvray and S. El-Basil, Submitted for publication.

(25) It can be shown that there is a linear correlation between simple functions of the molecular connectivity indices of a series of Clar graphs and those

of the equivalent caterpillar trees.[17]

(26) A.T. Balaban and F. Harary, Tetrahedron 24 (1968) 2505.

(27) M. Randić, G.M. Brissey, R.B. Spencer and C.L. Wilkins, Computers and Chem. 3 (1979) 5.

(28) M. Randić and C.L. Wilkins, Chem. Phys. Letters 63 (1979) 332.

(29) J.R. Dias, J. Chem. Inf. Comput. Sci. 24 (1984) 124.

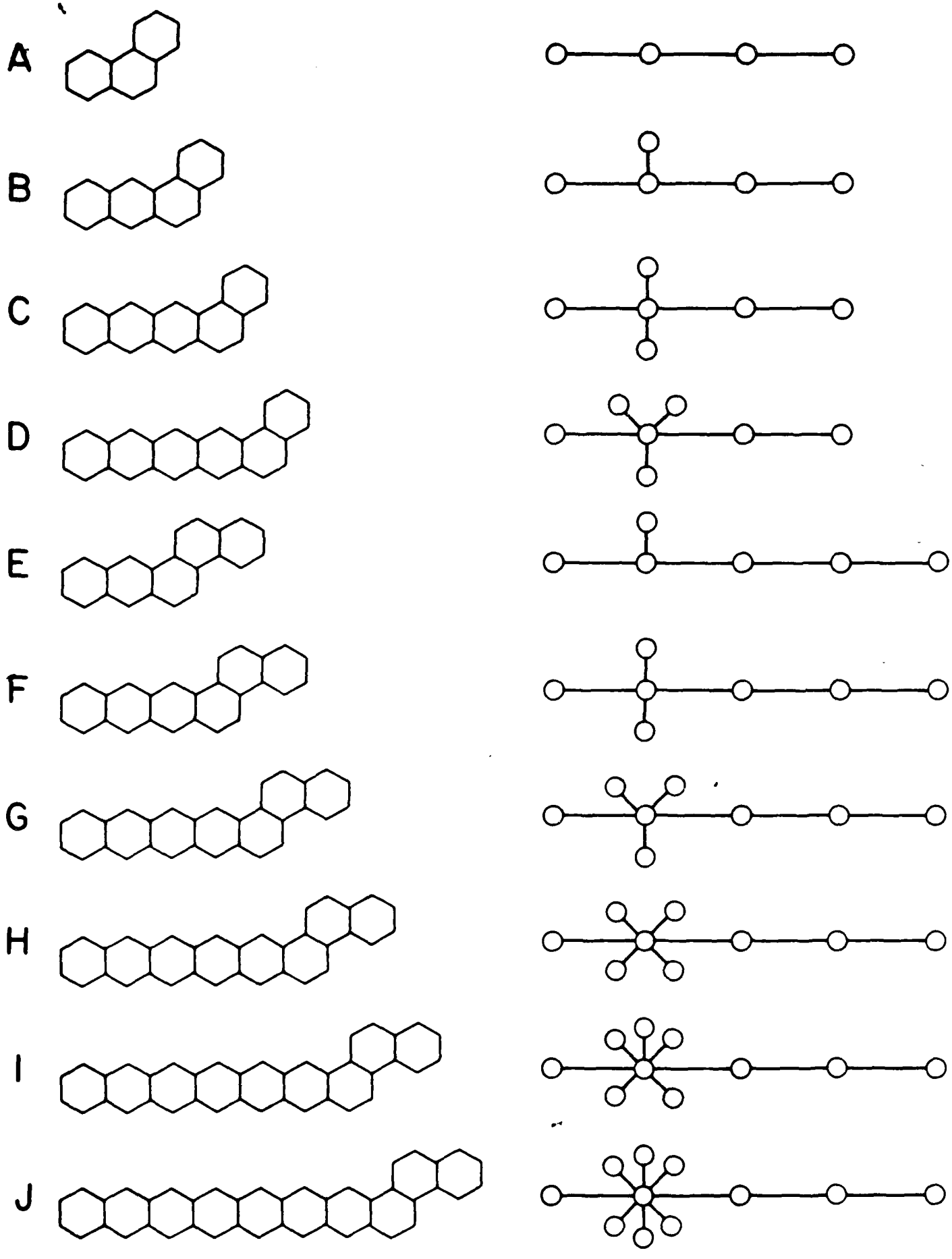


Fig.

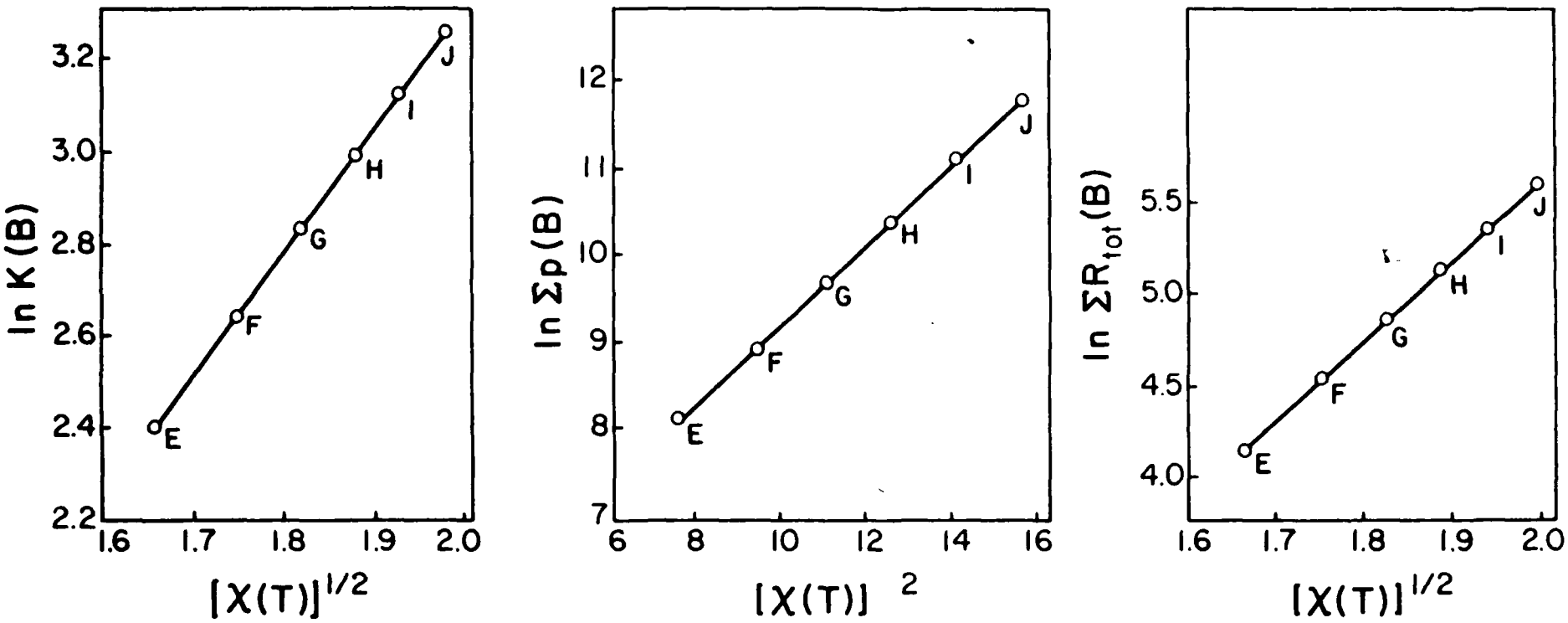
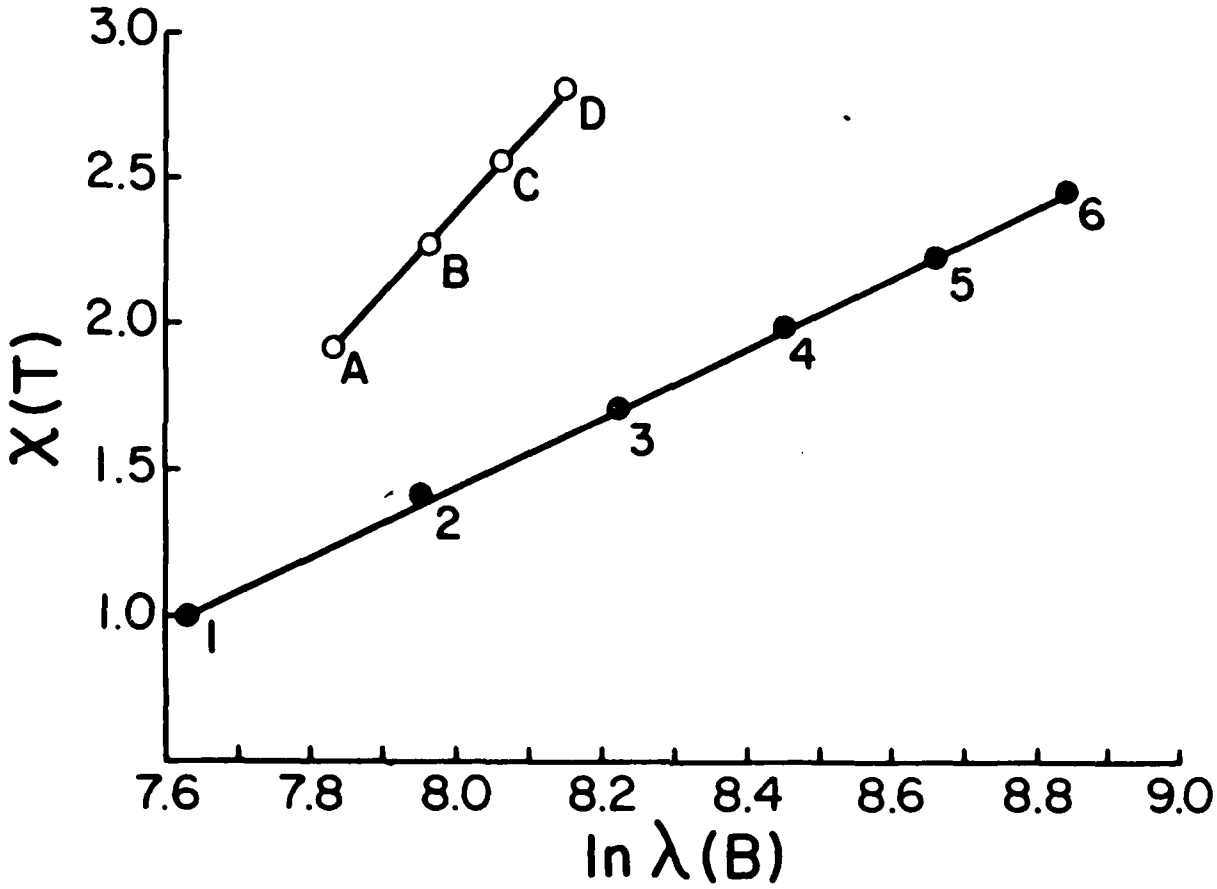

ln K(B)
3.2
3.0
2.8
2.6
2.4
2.2
E
F
G
H
I
J
1.6
1.7
1.8
1.9
2.0
[X(T)]^1/2
ln Σp(B)
12
11
10
9
8
7
6
8
10
12
14
16
[X(T)]^2
ln ΣR_tot(B)
5.5
5.0
4.5
4.0
1.6
1.7
1.8
1.9
2.0
[X(T)]^1/2

Fig 2


3.0
2.5
2.0
1.5
1.0
χ(T)
A
B
C
D
1
2
3
4
5
6
7.6
7.8
8.0
8.2
8.4
8.6
8.8
9.0
ln λ(B)

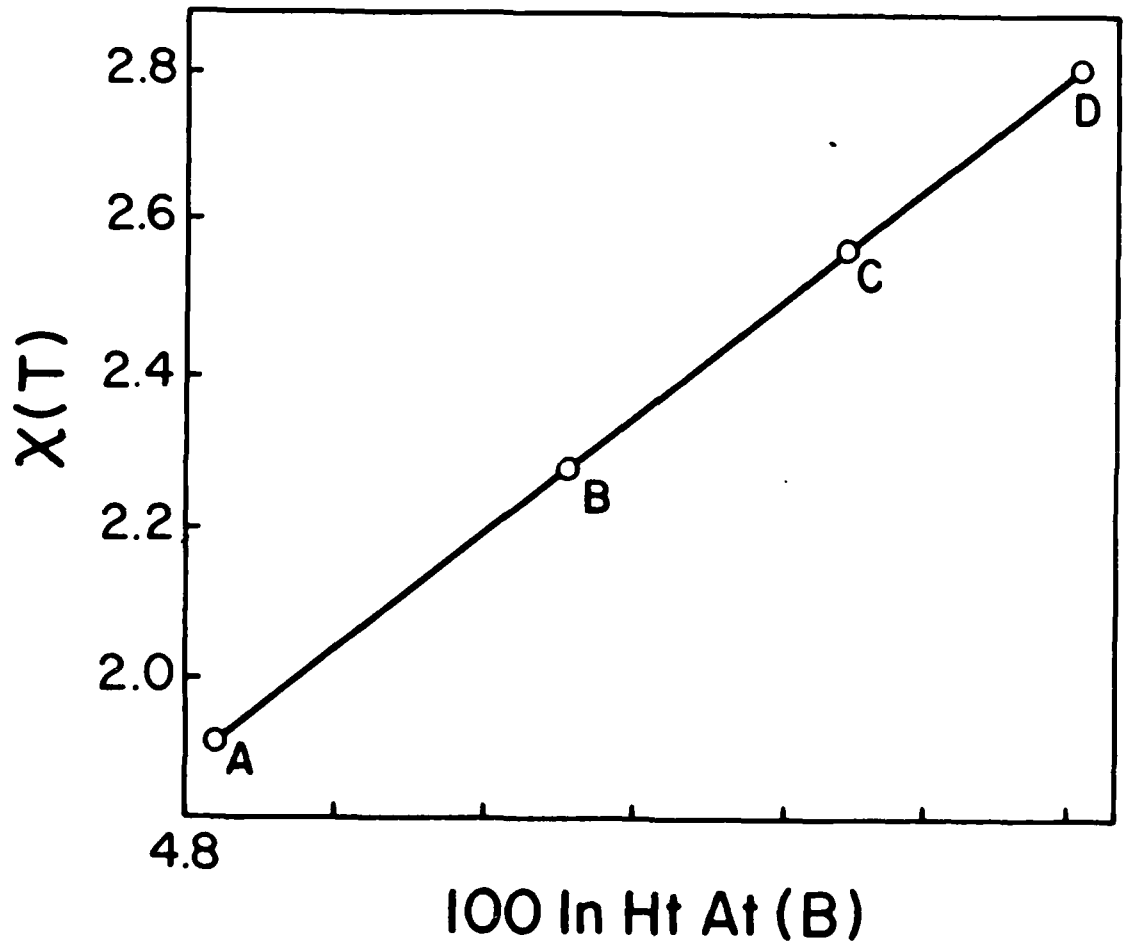


Fig 4

DL/1113/87/2

TECHNICAL REPORT DISTRIBUTION LIST, GEN

	No. Copies
Office of Naval Research Attn: Code 1113 800 N. Quincy Street Arlington, Virginia 22217-5000	2
Dr. Bernard Douda Naval Weapons Support Center Code 50C Crane, Indiana 47522-5050	1
Naval Civil Engineering Laboratory Attn: Dr. R. W. Drisko, Code L52 Port Hueneme, California 93401	1
Defense Technical Information Center Building 5, Cameron Station Alexandria, Virginia 22314	12 high quality
DTNSRDC Attn: Dr. H. Singerman Applied Chemistry Division Annapolis, Maryland 21401	1
Dr. William Tolles Superintendent Chemistry Division, Code 6100 Naval Research Laboratory Washington, D.C. 20375-5000	1
Dr. David Young Code 334 NORDA NSTL, Mississippi 39529	1
Naval Weapons Center Attn: Dr. Ron Atkins Chemistry Division China Lake, California 93555	1
Scientific Advisor Commandant of the Marine Corps Code RD-1 Washington, D.C. 20380	1
U.S. Army Research Office Attn: CRD-AA-IP P.O. Box 12211 Research Triangle Park, NC 27709	1
Mr. John Boyle Materials Branch Naval Ship Engineering Center Philadelphia, Pennsylvania 19112	1
Naval Ocean Systems Center Attn: Dr. S. Yamamoto Marine Sciences Division San Diego, California 91232	1

END

9-87

DTIC