RELATIONSHIP BETWEEN THE QUANTUM CHEMICAL INDICES OF REACTIVITY AND THE CARCINOGENIC ACTIVITY OF CHEMICAL COMPOUNDS. AROMATIC AMINES

V. M. Promyslov, A. S. Kabankin,
S. P. Kozlova, and G. M. Barenboim

UDC 614.73.074+615.277.4.074

Pollution of the enviornment with chemical compounds (CC) as a result of active industrial activity of man has raised an acute problem of evaluating the biological hazard of CC for living system. The CC that affect the human genetic apparatus, including mutagens, carcinogens, gonadotoxins, and embryotoxins, are recognized as especially hazardous chemical environmental factors. According to the data of the International Cancer Research Agency, 80-90% of all the cases of carcinogenesis are induced by envirnmental factors. Among the CC hazardous for man are benzidine, 4-aminobiphenyl, 2-naphthylamine, benzene, and vinyl chloride; the drugs diethylstilbestrol, azathioprin, chlorbutin (Chlorambucil), and cyclophosphane (cyclophosphamide) are attracting special attention [13]. However, the possibilities of testing CC for carcinogenic activity by traditional methods are extremely limited, since it requires considerable time and extensive facilities. According to the data of the US National Cancer Institute, the investigation of one substance for carcinogenic activity on animals takes about three years and requires 300,000-500,000 dollars. In this case the results obtained are frequently insufficiently convincing to evaluate the hazard of CC for man. These circumstances were responsible for the necessity of creating a rapid test for predicting the toxicologic properties of CC. In connection with this, calculation methods of classifying CC with the aid of computers, based on the use of the interrelationship of structure and activity, the detection of correlations between the biological activity of CC and their physicochemical or quantum chemical parameters, are of substantial interest.

The interrelationship between the quantum chemical reactivity indicates (RI) and the carcinogenic activity of CC has been attracting the attention of researchers for many years, beginning with the well known studies of [20, 21], devoted to aromatic hydrocarbons. Recently the study of the electronic structure of aromatic hydrocarbons has been conducted in especially great detail by methods of quantum chemistry [2, 3, 6, 8, 11, 16, 17, 19, 24, 25]. Quantum chemical calculations for aromatic amines are not so numerous [4, 15].

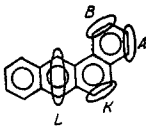
Studies of this kind are based on the fact that the carcinogenicity of CC is related to their metabolic conversions in the mammalian organism. Therefore, we can assume that in the determination of a number of related compounds, the manifestation of carcinogenic properties will be determined by the ability of these CC to enter into chemical reactions of a definite type. A well known method based on a comparison of the RI of CC molecules, i.e., the values obtained by quantum chemical calculations and characterizing the chemical activity both of individual positions in the molecule (atoms or chemical bonds) and of the molecule as a whole, is frequently used to study the reactivity in series of related compounds. Such quantities include the charges on the atoms, the populations on the bonds, the ionization potential, etc. If the mechanism of action is the same for CC of a given class, we might expect correlations between the manifestation of carcinogenic properties and a definite assortment of RI of the molecules of these CC. In the case when a quantitative measure of hazardous properties is known for CC, correlation equations relating the quantum chemical characteristics to the toxicological properties of CC can be constructed. If, however, the experiment characterizes the carcinogenic properties of CC only quantitatively (hazardous—nonhazardous), then methods of the type of image recognition can be used. They permit the deciding rules to be obtained, i.e., conditions that should be satisfied by a given assortment of RI for assigning CC to the class of carcinogenic or noncarcinogenic compounds.

Scientific-Research Institute for the Biological Testing of Chemical Compounds, Moscow Province. Translated from Khimiko-Farmatsevticheskii Zhurnal, Vol. 19, No. 12, pp. 1456-1462, December, 1986. Original article submitted November 29, 1984.

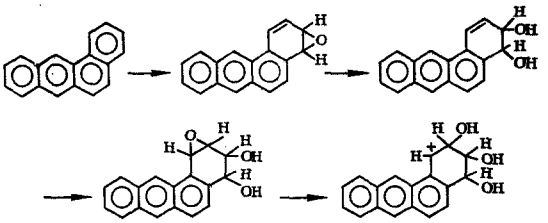
0091-150X/85/1912-0857$09.50

The correlation equations or deciding rules obtained can be used for an initial estimation of the carcinogenicity of CC. Such an approach, of course, cannot replace the traditional methods of biological testing, but as a simple and rapid method of preliminary screening they undoubtedly merit attention. Moreover, an anlysis of the results of such calculations may be of substantial interest from the standpoint of detection of the bulk of the metabolic activity of CC of the given class, establishment of the nature of the limiting step of the process of carcinogensis, and the more profound understanding of the mechanism of the carcinogenic action of CC.

Thus, for example, calculations of condensed aromatic hydrocarbons performedin [2, 3, 6, 8, 11, 16, 17, 19, 24, 25] have made it possible to construct the following model of the mechanism of action for these compounds: four characteristic portions are distinguished in the molecule: the K-, L-, A-, and B-regions.



The active K-region ensures binding of a substance to the cell and penetration into it; the active L-region leads to deactivation of the molecule by oxidation to an inactive hydroxy-derivative, while oxidation in the A- and B-regions leads to the formation of an epoxide, dihydrodiol, and finally, a carbonium ion, which interact with the macromolecules of the cell — DNA.

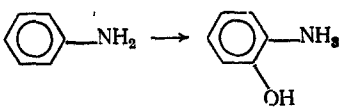


In this work we consider the class of aromatic amines. Interest in these compounds is due to their use in the chemical industry, in particular, the aniline dye and chemico-pharmaceutical industries, where aromatic amines are the starting materials or intermediate products in the technological process. In connection with this, the estimation of the biological hazard of aromatic amines is exceptionally important for these branches of industry.

Compounds of this class are being widely investigated for carcinogenic activity, and there are already definite ideas of the mechanism of their action. It is believed that the process of metabolic activation of aromatic amines includes as its most important step the N-hydroxylation of CC, follwed by esterification and the formation of the nitronium ion, which interacts with the DNA molecule and evidently is a direct carcinogen [5].

$$ArNH_2 \longrightarrow ArNH(OH) \longrightarrow ArNH(OX) \longrightarrow Ar\overset{+}{N}H$$

Another mechanism of metabolic activation of aromatic amines, associated with ortho-hydroxylation:



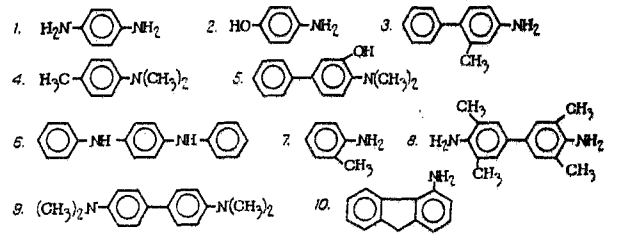
is also known.

These mechanisms of metabolic activation of aromatic amines also do not exclude other pathways of conversion of CC, such as acetylation, intramolecular arrangement, etc., leading to activation of CC.

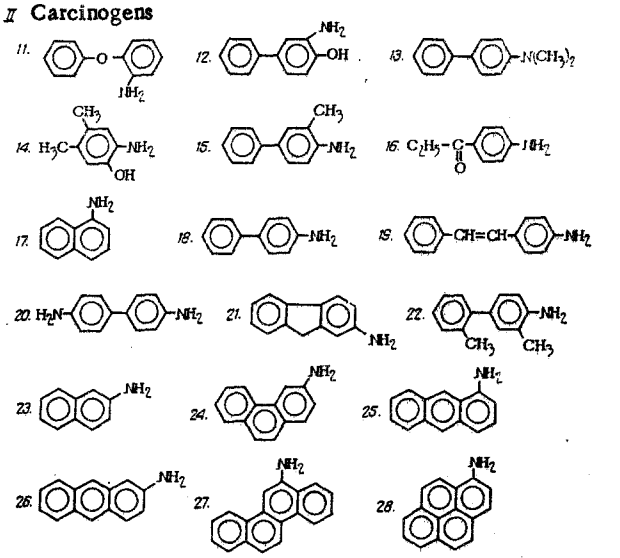
In this work we discuss 28 molecules of aromatic amines, for which the results of quali-

tative tests for carcinogenic activity are known [12, 18, 23].

I Noncarcinogens



II Carcinogens



For each of the molecules a complete quantum chemical calculation was performed, and 11 parameters characterizing each of them were selected: 1) the electrophilic hyperdelocalizability (HDL) on the nitrogen atom, S_N; 2) the charge on the nitrogen atom, q_N; 3) the maximum HDL on the aromatic C—C—S_{CC} bond; 4) the maximum population (according to Mulliken) on the aromatic C—C—Q_{CC} bond; 5) the maximum interatomic HDL for the carbon atoms in the para-position to one another, S_{CC}^{para}; 6) the maximum interatomic population for carbon atoms in the para-position to one another, Q_{CC}^{para}; 7) the maximum HDL for a carbon atom in the ortho-position to the nitrogen atom, S_C^{ortho}; 8) the maximum charge for a carbon atom in the ortho-position to the nitrogen atom, q_C^{ortho}; 9) the energy gap between the upper occupied and lower free molecular orbital, ΔE; 10) the energy of the lower free molecular orbital, E_v; 11) the number of substituents at a nitrogen atom without one, N_R.

The values of S_N and q_N characterize the activity of the nitrogen atom in reactions with electrophilic reagents, in particular, in the reaction of N-hydroxylation, which is an important step in the metabolic activation of amines. The parameters S_{CC}, Q_{CC}, S_{CC}^{para}, and Q_{CC}^{para} characterize the activity of the aromatic portion of the molecule in oxidation reactions; the values of S_C^{ortho} and q_C^{ortho} reflect the possibility of orthohydroxylation of aromatic amines; ΔE is related to the ability of a given CC to enter reactions with radicals; E_v characterizes the reactivity of the molecule for interaction with nucleophilic reagents; the parameter N_R at the level of chemical structure considers differences between primary, secondary, and tertiary amines. The lengths of the chemical bonds and bond angles were assumed equal to the standard values, only in certain cases the experimental values were used. In general, as was verified, the indices calculated in this work are rather insensitive to variation of the geometry of the molecule within rather wide limits. An expanded Huckel method in the standard parameterization was used for calculation in this work [9, 10]. The values of S_A, q_A, S_{AB}, and Q_{AB} were calculated according to the formulas:

$$q_A = Z_A^- - 2\sum_{i=1}^{N}\left(\sum_{\mu\in A} C_{i\mu}^2 + \sum_{\mu\in A}\sum_{\nu\in B} C_{i\mu}C_{i\nu}S_{\mu\nu}\right);$$

$$Q_{AB} = 2\sum_{i=1}^{N}\sum_{\mu\in A}\sum_{\nu\in B} C_{i\mu}C_{i\nu}S_{\mu\nu};$$

$$S_A = 2\sum_{i=1}^{N}\left(\sum_{\nu\in A} C_{i\mu}^2/\varepsilon_i + \sum_{\mu\in A}\sum_{\nu\in B} C_{i\mu}C_{i\nu}S_{\mu\nu}/\varepsilon_i\right);$$

$$S_{AB} = 2\sum_{i=1}^{N}\sum_{\mu\in A}\sum_{\nu\in B} C_{i\mu}C_{i\nu}S_{\mu\nu}/\varepsilon_i.$$

Here the subscripts A and B number the atoms; μ and ν are the atomic orbitals; i is the molecular orbital; $C_{i\mu}$, $C_{i\nu}$ are the coefficients of the MO; ε_i is the energy of the i-th MO; Z_A is the charge of the nuclear core of the atom A; and $S_{\mu\nu}$ is the overlapping integral. The numerical indices obtained were treated on a computer using a method of step-by-step linear discriminant analysis. This method permits the construction of linear discriminant functions from a set number of variables, providing for the best (within the limits of the given approach) separation of the set of CC under consideration into classes of carcinogens and noncarcinogens, by successive selection of the most informative parameters. A knowledge of the discriminant functions is equivalent to a knowledge of the condition that should be satisfied by linear combination of the parameters selected in the course of discriminant analysis for the classification of CC as active or inactive. As a result of such treatment, the following condition of carcinogenic activity of aromatic amines was obtained:

$$-21.97\cdot S_N + 40.274\cdot Q_{CC} - 20.034\cdot q_C^{ortho} + 1.063\cdot N_R + 153.92 > 0 \qquad (1)$$

The difference of the group averages for such a set of parameters is at the 0.01 level of significance. The use of this inequality for classification of the CC under consideration gives seven errors: molecules 3, 7, 10, 13, 15, 23, and 25 are classified incorrectly.

In this work the number of variables included in the discriminant functions was limited to four. This is due to the small volume of the learning sample (28 compounds). The problem of the relationship between the size of the learning sample and the group of characteristics was discussed in detail in [1]. From the table cited in this work it follows that in the case of a learning sample consisting of 26 experiments, reliable recognition with the aid of linear discriminant functions of no more than five variables is possible, and this estimate is probably too high [1].

The predictive power of the discriminant functions obtained was verified by the sliding control method. The probability of correct classification of noncarcinogens proved equal to 0.80, that of carcinogens 0.72, and the probability of correct classification as a whole 0.75. The numberical values of the parameters entering into the inequality (1) are cited in Table 1 for all the molecules considered.

From the inequality (1) it is evident that the probability of assignment of CC to the class of carcinogens is greater the higher the reactivity of the aromatic portion of its molecule, i.e., the greater Q_{CC} and $|q_{CC}^{ortho}|$ ($q_{CC}^{ortho} < 0$ for a given class of CC, and the lower the reactivity of the nitrogen atom, i.e., the value of S_N (in both cases we have in mind reactions with electrophilic particles).

The results obtained in this work partially contradict the conclusion of [15] that the carcinogenic activity of aromatic amino acids increases with increasing reactivity both of the nitrogen atom and of the aromatic portion of the molecule. In [15] for eight aromatic primary amines, using an iteration variant of the expanded Huckel method, the contributions of the π-electrons to S_N, S_{CC}, and to the greatest HDL on the carbon atom were calculated. It is customary to consider that there must be at least five observations per independent parameter, i.e., in this case the set should contain no less than 15 molecules. Moreover, of the three molecules classified as noncarcinogens [15], there are data on carcinogenic activity, although weak, for two of them.

We should also note that the overwhelming majority of primary aromatic amines are oxidized at the amino group very readily under mild conditions; therefore, the hypothesis advanced in [4, 15] that carcinogenic amines differ from noncarcinogenic amines primarily by the greater case of oxidation of the amino group (N-hydroxylation), seems insufficiently substantiated from the chemical standpoint. On the other hand, the more readily the nitrogen atom of the amino group enters into a reaction with an electrophilic particle, evidently the higher the solubility of the corresponding amine in water will be and the more rapidly it is eliminated from the organism. At the same time, the more reactive the aromatic portion of an amine molecule, the more readily the process of its intercalation into DNA, i.e., penetration between

TABLE 1. Values of the Quantum Chemical Indices of Aromatic Amines Entering into Inequality (1)

n	S_N	Q_{CC}	q_C^{ortho}	N_R	n	S_N	Q_{CC}	q_C^{ortho}	N_R
1	9,104	1,093	—0,085	0	15	9,127	1,098	—0,090	0
2	9,103	1,098	—0,084	0	16	9,048	1,015	—0,091	0
3	9,096	1,095	—0,124	0	17	9,090	1,143	—0,146	0
4	9,321	1,102	—0,102	2	18	9,090	1,126	—0,132	0
5	9,336	1,105	—0,101	2	19	9,077	1,127	—0,147	0
6	9,231	1,100	—0,088	1	20	9,099	1,122	—0,121	0
7	9,102	1,100	—0,091	0	21	9,090	1,093	—0,091	0
8	9,126	1,100	—0,020	0	22	9,079	1,100	—0,091	0
9	9,317	1,101	—0,102	2	23	9,091	1,092	—0,091	0
10	9,083	1,103	—0,113	0	24	9,095	1,106	—0,129	0
11	9,095	1,098	—0,089	0	25	9,134	1,100	—0,089	0
12	9,095	1,103	—0,109	0	26	9,086	1,123	—0,135	0
13	9,316	1,101	—0,102	2	27	9,098	1,134	—0,164	0
14	9,104	1,103	—0,119	0	28	9,092	1,144	—0,103	0

Note. N_R is dimensionless; the remaining variables are in atomic units.

two neighboring complementary nitrogen base pairs and the fixation of such an associate through intermolecular forces, leading to a distortion of the form of the DNA molecule, can occur. The following step in the process is the formation of a covalent bond between the nitrogen atom of the amine and one of the nitrogen bases of DNA; moreover, this reaction occurs rapidly if there are no steric hindrances. The mechanism of interaction presented is confirmed by calculated and experimental data in [7, 22].

Thus, the results of the present work suggest that the carcinogenic activity of aromatic amines is determined by the ratio of the rates of two competing processes — binding of the aromatic portion of an amine molecule to DNA, which occurs more readily the more active the aromatic portion of the molecule in reactions with electrophilic reagents, and the elimination of CC from the organism, which occurs more readily, the more active the nitrogen atom of the amino group in reactions with electrophilic reagents. In this case the processes of metabolic activation of aromatic amines occur equally easily for carcinogens and noncarcinogens.

In conclusion, we should note that in certain cases the difference of carcinogens and noncarcinogens may not only be associated with the reactivity of CC but may also be determined by other factors. Thus, model calculations of the interaction of amino derivatives of fluorene [22] and naphthalene [7] (structures 10, 17, 21, 23) with DNA fragments showed that the differences in the carcinogenic activity for these molecules (molecule 10 is inactive, 17 has low activity, and 21 and 23 are active) are due to steric factors, whereas their electronic parameters differ negligibly. When a complex of these molecules with DNA is formed in the case of CC 10 and 17, the formation of a covalent bond between the nitrogen atom and the nitrogen base of DNA is hindered by steric factors, whereas for molecules 21 and 23 there is no such hindrance. In our opinion, to improve the separation of CC into classes of carcinogens and noncarcinogens, the integral parameters characterizing the chemical structure of the CC should be included in the analysis. We suggest that such topological characteristics of the molecules as the molecular bonding capacities, widely used for the study of structure-activity correlations of drug preparations [14], be used for this purpose.

LITERATURE CITED

1. G. S. Lbov, Methods of Processing Heterotypic Experimental Data [in Russian], Novosibirsk (1981).
2. G. D. Berger, I. A. Smith, P. G. Seybold, and M. P. Serve, Tetrahedron Lett., No. 3, 231-234 (1978).
3. G. D. Berger and P. G. Seybod, Int. J. Quant. Chem., Quant. Biol. Symp., No. 6, 305-311 (1979).
4. W. Brock Neely, Ibid., No. 2, 171-177 (1975).
5. D. B. Clayson and R. C. Garner, in: Chemical Carcinogens, ed. C. H. Searle, Washington (1976), pp. 366-461.
6. B. Y. Duke, D. R. Eilers, Y.S. Eilers, et al., Int. J. Quant. Chem., Quant. Biol. Symp., No. 2, 155-170 (1975).

7. N. Gresh, Int. J. Quant. Chem., 22, No. 4, 37-48 (1982).
8. W. C. Herndon, Int. J. Quant. Chem., Quant. Biol. Symp., No. 1, 123-134 (1974).
9. R. Hoffman, J. Chem. Phys., 39, 1397-1412 (1963).
10. R. Hoffman, Ibid., 40, 2474-2480 (1964).
11. A. Imamura, M. Koida, and S. Kato, Gann, 70, 291-296 (1979).
12. International Agency for Research on Cancer, Monograph. Suppl. 1, Lyon (1979).
13. Ibid., Suppl. 4, Lyon (1982).
14. L. B. Kier and L. H. Hall, Molecular Connectivity in Chemistry and Drug Research, New York (1976).
15. G. Loew, B. S. Sudhindra, and S. Burt, Int. J. Quant. Chem.: Quant. Biol. Symp., No. 6, 259-284 (1979).
16. G. H. Loew, B. S. Sudhindra, and Y. E. Ferrell, Chem. Biol. Interact., 26, 75-89 (1979).
17. M. A. Mainster and J. D. Memory, Biochim. Biophys. Acta, 148, 605-608 (1967).
18. J. McCann, E. Choi, E. Yamasaki, and B. N. Ames, Proc. Nat. Acad. Sci. USA, 72, 5135-5139 (1975).
19. J. D. Memory, Int. J. Quant. Chem., Quant. Biol. Symp., No. 2, 179-180 (1975).
20. A. Pullman, Bull. Soc. Chim. France, 21, 595-603 (1954).
21. A. Pullman and B. Pullman, Adv. Cancer Res., 3, 117-169 (1955).
22. A. Sawaryn and H. Bradaszek, J. Molec. Struct., 90, 271-278 (1982).
23. P. Shubik and J. L. Hartwell, Survey of Compounds Which Have Been Tested for Carcinogenic Activity, Suppl. 2, Washington (1969).
24. I. A. Smith, G. D. Berger, P. G. Seybold, and M. P. Serve, Cancer Res., 38, 2968-2977 (1978).
25. I. A. Smith and P. G. Seybold, Int. J. Quant. Chem., Quant. Biol. Symp., No. 5, 311-320 (1978).