John D. Scribner
Fred Hutchinson Cancer Research Center
1102 Columbia Street
Seattle, Washington 98104

Molecular Orbital Theory in Carcinogenesis Research

Molecular orbital calculations are a language used to explain a body of chemical data and to predict the ease and course of reactions of chemical compounds. Unfortunately, the use of this language in carcinogenesis research has suffered some abuse by its practitioners and more by those who have recognized that often too much has been claimed for it in attempts to explain the mechanisms of chemical carcinogenesis.

With the development of quantum mechanics, it was immediately apparent that the exact calculations which so accurately explained the properties of a simple species such as H_2^+ could only be approximated in handling larger molecules (*1*). The coming of high-speed computers permitted more sophisticated approximations, but approximations they remained. The most useful approximation for practicing organic chemists was the linear combination of atomic orbitals into molecular orbitals method (LCAO–MO), for this procedure permitted quantitative predictions of chemical behavior for many organic molecules, with a requirement only for desk calculators. The simplest (Hückel) form of the LCAO–MO method could be expanded to cover many other molecules with the use of computers requiring only the simplest type of programming. The ability of the LCAO–MO method to predict chemistry of polycyclic aromatic hydrocarbons is truly remarkable, considering the many simplifying assumptions made for this procedure.

Many of the approximations have been improved by semiempirical methods, i.e., those in which knowledge of the thermochemical or electrochemical properties of compounds is used to modify the calculations for better agreement with experimental observations. Although this may be an unsatisfying approach for those concerned with the niceties of quantum theory, it is a most practical method for constructing a language that will be generally useful for the practicing chemist. Others who preferred to improve on the theory itself chose to observe the extent and direction of deviation between theory and fact and to delete approximations previously made for the sake of convenience. The result has usually been in the same direction as the semiempirical methods.

With the blossoming of systematic carcinogenesis research, particularly with polycyclic hydrocarbons, at the time of development of the MO theory and of a romantic belief among physical scientists that the riddles of biology would yield to the new knowledge of chemistry and physics (*2*), it was only natural that attempts would be made to explain the biological activity of carcinogenic compounds on the basis of predicted chemical behavior in the cell. At this point this desirable merging of chemistry and biology went awry because of two assumptions underlying most attempts to explain carcinogenic activity directly with MO methods: 1) Reactions characteristic of the carcinogen in the artificial systems of the organic chemist will be the controlling reactions biologically; 2) the carcinogen reacts directly with the sensitive target in the cell, and this reaction (repeated, if necessary) is sufficient to produce cancer. It is probably a worthwhile project for a historian of science to determine whether such assumptions were justifiable at the time. With the benefit of 20 years of hindsight, however, we can say that such an approach cannot be justified now. However, chemical carcinogens do undergo chemical processes in the cell and they do react in some form with cell macromolecules. An assumption, which still seems quite justifiable and accepted by many, is that one or more of these reactions is necessary for the carcinogenic process, even if the reaction is not sufficient cause in itself. The events leading to this reaction and other parallel events can be described by MO calculations, at whatever level of sophistication is necessary and available. Let us consider some of these events.

TRANSPORT OF THE CARCINOGEN

Hansch developed the concept that the action of all drugs in a particular class involves interaction with the receptor at roughly the same rate, and that the efficacy of the drug depends on the rate at which it reaches the receptor from the site of administration. This rate, in turn, is proposed to depend on an optimum partition coefficient for the particular system. This basic hypothesis has worked remarkably well and even produced a fairly good correlation between optimum partition and carcinogenic activity for an extended series of amino azo dyes (*3*). Herndon (*4*) later combined a molecular size factor with resonance theory calculations in a regression analysis that appeared to lead to a reasonable ability to predict carcinogenicity from these properties alone. However, a study of his data shows no improvement in reliability over previous approaches that neglect molecular size. Franke and Büchner (*5*) and Franke (*6*) have also suggested that the partition coefficient of a carcinogen may play a major role in its activity. If this property is indeed critical for carcinogenic activity, the ability to predict it reliably by theoretical methods would be desirable. Since partition coefficient is intimately related to the electron distribution within the molecule, such a prediction

Editor's note: Periodically the Journal publishes solicited guest editorials as a means of transmitting to investigators in cancer research the essence of current work in a special field of study. The Board of Editors welcomes suggestions for future editorials that succinctly summarize current work toward a clearly defined hypothesis regarding the causes or cure of cancer.

should be possible. Rogers and Cammarata (*7*) have shown, with several classes of compounds, that this can indeed be done. Although the procedure requires computer aid, here too the initial program is not excessively difficult. Once developed, the method is capable of generating theoretical parameters for large numbers of compounds with far less effort than would be required to actually make the experimental determination.

METABOLISM OF THE CARCINOGEN

Although it is widely accepted that nonreactive carcinogens must be metabolized to exert their effect, studies of carcinogen metabolism are at a relatively primitive state. Extensive product studies on several aromatic hydrocarbons and a few aromatic amines have been undertaken with the major conclusion being that both classes of compounds are predominantly converted to hydroxylated derivatives by a system requiring NADPH and molecular oxygen. The fact of an intermediate epoxide in some hydrocarbon oxidation has been established, and the oxidation activity has been localized in certain protein fractions. Except for occasional efforts to establish nonenzymatic model oxidizing systems, however, there has been virtually no effort devoted to answering several long-standing questions. Why is a particular set of metabolites produced from a given compound? Why is an alteration in this pattern produced by certain inducing agents? Why do different inducing agents have different effects on metabolism of a given reference substrate? An additional assumption implicit in many attempts to correlate MO parameters with carcinogenicity is that electronic properties of the carcinogen are solely or primarily responsible for the action of the compound. The Hansch analysis does not detract from this position, but rather reinforces it, as demonstrated by the study of Rogers and Cammarata (*7*). Based on predictions of reactivity, however, such a position in untenable.

Any theoretical prediction of reactivity describing the electronic structure of the substrate only is delineating that structure in a random environment and its response to random perturbations of the environment. An article of faith among biochemists and biologists, however, is that metabolic processes are not random. Enzymes that are critical for the correct functioning of a biological system are highly specific for their correct substrates. A general assumption, however, is that the microsomal mixed function oxidases are nonspecific, both with respect to variety of substrate and to site of attack. MO calculations (and model in vitro systems, when available) should enable us to test this assumption of the random nature of the microsomal oxidases. Deviations from the predicted reactivity of a substrate toward such oxidation indicate a biological order imposed on the normal chemistry of the compound involved. Consider two carcinogens of considerable significance in current research, 2-acetamidofluorene (AAF) and benzo[*a*]pyrene (BP). Frontier orbital coefficients for AAF predict that the carbon atoms in AAF which should be most susceptible to electrophilic substitution are (in order) 7, 5, 4, and 3. Yet 4-hydroxy(OH)-AAF is not found in the urine of rats treated with AAF (*8*) or in the mixture of products resulting from incubation of AAF with rat liver microsomes in vitro (*9*). A similar order is predicted for free radical attack. For BP, the most reactive position toward electrophilic, radical, or nucleophilic substitution is the 6-carbon, a prediction confirmed by extensive laboratory experience. Yet the predominant metabolite of BP in microsomal incubations is 3-OH-BP (*10, 11*). Even the sum of all metabolites involving oxidation at the 6-position may be less than the amount of either 3-OH-BP or the 9,10-diol (*10, 11*). Unfortunately, even current analytical methods cannot yet resolve completely all possible metabolites of BP. Using the same principles, we can make similar predictions about numerous other compounds; the results should prod investigators into a critical evaluation of metabolic mechanisms. The ability of even the simplest MO techniques to predict relative reactivities in this manner was demonstrated by Dewar and Warford (*12*), who calculated the relative reactivities of positions in phenanthrene toward electrophilic substitution with a paper-and-pencil method. Their predictions were confirmed in a detailed and tedious experiment of nitrating phenanthrene, separating the products, and comparing them with the various nitrophenanthrenes prepared by unambiguous routes. Presumably, the development and use of higher-order methods should permit accurate predictions for many compounds not in the theoretically simple class of polycyclic aromatic hydrocarbons.

Despite the current acceptance of the proposition that nonreactive carcinogens must be metabolized to active forms, the oldest assumption is that unmetabolized carcinogen, particularly polycyclic hydrocarbon, reacts directly with the cellular target controlling neoplasia. This assumption demands that carcinogenic activity be directly correlated with a particular type of reactivity of the carcinogen. Fundamentally, a polycyclic hydrocarbon may undergo two types of reaction: addition or substitution. Either the original complement of carbon and hydrogen atoms is added to, or a substitution is made for one of the hydrogen atoms. A substitution reaction involves only one carbon atom of the hydrocarbon, whereas an addition must involve two carbon atoms. Generally, theorists of chemical carcinogenesis have not recognized that the same intermediate stage of an addition to one carbon atom to produce an unstable species is probably necessary for both types of metabolic events (*13, 14*). Since this is not necessarily the case for laboratory reactions, some confusion and ill will have been created in this area. The Pullmans (*15*) originally considered the possibility that a substitution reaction might be the critical one. However, this hypothesis provided no correlation with carcinogenicity and was temporarily abandoned. The assumption of addition to the most reactive double bond in the hydrocarbon (the "K-region") proved more rewarding, especially when modified by the further assumption that a deactivating addition of the Diels-Alder type could take place (at the "L-region"). The final modification satisfying the Pullmans was the assumption that each addition *has as its initial step a substitution*. The Pullmans (*15*) then combined the predictive indices for the substitution and the addition to obtain complex indices yielding a reasonable correlation with carcinogenicity. The basic assumption of addition as a necessary reaction was not directly challenged for 14 years, at which time two separate papers appeared (*14, 16*). Hoffman (*16*) applied semiempirical calculations to the compounds discussed by the Pullmans and concluded that there was no theoretical basis for choosing addition as the critical reaction. Scribner (*14*) showed that, once the assumption of the importance of a K-region was made, prediction of substitution at the K-region as an activating reaction plus substitution at the L-region as a deactivating reaction proved sufficient to give a correlation as good as that found by the Pullmans, and possible better. Recently, Herndon (*4*)

revived the original addition hypothesis, claiming that the use of a novel resonance theory method resulted in a better correlation between predicted and observed levels of carcinogenic activity. He was apparently unaware of Hoffman's work, which rested on calculations of the same degree of accuracy as those of his own. In addition, Herndon's presentation suffers from two critical difficulties. First, he based his correlation on the Iball index. This crude attempt at quantitating carcinogenesis suffers in itself from two defects: 1) it is not based on dose–response studies but rather on the tumor production from an arbitrary uniform dose for all carcinogens; and 2) the relative order of carcinogenicities of compounds often varies with the route of administration and the target organ. The second failing in Herndon's paper is more critical: When based on the simple question of predicting whether a compound will be carcinogenic, his correlation is less reliable than that obtained by Scribner. Thus Herndon's claim that his "results support the idea that the hydrocarbon itself is the ultimate carcinogen" is difficult to accept.

Fukui and co-workers (*17*) recognized one fact that had not impressed the Pullmans: Carcinogenic alkylating agents reacted with nucleophilic centers in the cell. This observation suggested to them that hydrocarbons and aromatic amines also ought to react with nucleophilic centers. This group, however, also proceeded from the assumption that unmetabolized carcinogen reacted directly with the critical site. It is perhaps appropriate to use this example to comment on another attitude that has plagued theorists of chemical carcinogenesis, i.e., a tendency to disregard the realities of chemistry, both metabolic and laboratory. Nucleophilic substitutions on unsubstituted or alkylated polycyclic aromatic hydrocarbons are thermodynamically unfavorable. Under the most extreme laboratory conditions designed to produce such a substitution, it is more likely that an elimination would occur, leading to an unstable benzyne-type intermediate. Even such reactions as these can normally be performed only on hydrocarbons already bearing an appropriate leaving group. A nucleophilic substitution should proceed with even more difficulty (if that is possible) in the ring of an aromatic amine, since the electrons donated by the nitrogen must decrease reactivity toward an electron-bearing reagent (nucleophile). Since such reactions do not occur in vitro under the most vigorous conditions, it is illogical to imagine that they may take place under the mild ones of a biological system. Of course, enzymes do facilitate reactions difficult for a chemist, but even enzymes cannot violate thermodynamics. Nor are enzymes allowed under the original assumption.

REACTIONS OF THE ULTIMATE METABOLITE

If one assumes that, indeed, metabolic activation is necessary for a carcinogen to produce its effect, the rate of production of a specific metabolite might control the level of carcinogenicity. This assumption was the basis of Scribner's approach to hydrocarbon carcinogenicity (*14*), and not that of direct reaction between carcinogen and critical target. Dipple et al. (*13*), however, were the first to attempt to deal theoretically with the known facts concerning metabolism and reactivity of carcinogens. They assumed that the first reaction of a hydrocarbon was an electrophilic oxidation to an intermediate oxygen-bearing sigma complex with a positive charge delocalized throughout the aromatic system. They then postulated that this species could continue to a relatively stable epoxide, or that it could react directly with a critical target and initiate carcinogenesis. The stability of this intermediate would then determine its carcinogenic activity. Unfortunately, this postulate did not result in any useful correlation of theory with carcinogenicity for unsubstituted hydrocarbons. It did produce a remarkably good correlation for methylated benz[*a*]anthracenes and dibenz[*a,h*]anthracenes, based on the assumption that a carbonium ion is generated at the methyl group. This study is a good example of a valid combination of theoretical approach and experimental testing (which followed). Much fruitful research has followed this particular study, with a new class of carcinogen, the bromomethylbenz[*a*]anthracenes (*18–22*). Unfortunately, the original postulate that a methylated hydrocarbon is activated by generating a positive charge on the methyl group seems to have been refuted by recent work on the binding of hydrocarbons to DNA (*23, 24*). On the other hand, it suggests that, within a closely related series of carcinogens all having the same type of reactivity and similar physical properties, the ability to undergo a particular type of reaction controls the tumor-initiating potency of the carcinogen. The entire program has been an impressive example of a continuing dialogue between fact and theory, leading to significant conclusions concerning the nature of carcinogen chemistry.

Application of MO theory to aromatic amine carcinogenesis is in the fortunate position of having available accurate models for the ultimate reactive metabolites of aromatic amines and/or their *N*-acetyl derivatives. Esters of the *N*-hydroxy derivatives of AAF and 4-acetamidobiphenyl (AABP) have been found to yield the same adducts with DNA and RNA in vitro as those isolated from livers of rats fed *N*-OH-AAF or *N*-OH-AABP (*25, 26*). Kinetics experiments and product studies indicated that the reactive intermediate in these reactions is an *N*-aryl-*N*-acetylnitrenium ion (*27*). MO calculations were successfully used to explain the relative decomposition rates of esters of a series of different *N*-arylacetohydroxamic acids (*27*). Reaction rates did not explain the yield or substrate selectivity from these reactions, however. Recently, we have been able to show that an MO description of the charge distribution in the nitrenium ion may accomplish the latter goal (*28*). Further experiments are required to substantiate this possibility.

CONCLUSIONS

So far no technique for prediction of chemical reactivity from MO (or resonance theory) methods has explained successfully the relative carcinogenicity of an extended series of compounds, except for the bromomethylbenz[*a*]anthracenes. This fact in itself should make us acutely aware of the possibility that the factors governing carcinogenesis may include chemical reaction in only a minor, albeit necessary, role. On the other hand, because so few studies have been undertaken to understand transport, metabolism, and ultimate reactivity by theoretical methods, it is safe to say that this area has not really been explored. For all the benefits collaboration may offer, intimate contact is essential between investigators familiar with MO theory and those familiar with carcinogen chemistry and metabolism. Failure to maintain such contact could easily result in perpetuation of a situation in which theoreticians are not really aware of the complexities of the systems they are being asked to describe, and one in which experimentalists may be unaware of

significant studies which would be suggested by theoretical results.

The ultimate ideal in the goals of carcinogenesis research is the ability to predict carcinogenicity from knowledge of chemical structure alone. Yet no progress can be expected in this direction until there is a substantial improvement in the level of research into chemistry of carcinogens and the micromechanisms of metabolism. As I mentioned in the beginning, MO theory is a language. In terms of resources, it will require but a small expenditure to use this language, compared with the resources required to generate the data to be comprehended. In the evaluation of current and projected studies by investigators who wish to use MO theory, we must ask: Does this use of theory suggest specific experiments to be performed or specific classes of data to be collected, which will improve our understanding of carcinogen transport, metabolism, or reactivity? If so, then it appears wise to encourage such use and to improve its accessibility to the carcinogenesis research community.

REFERENCES

(1) ROBERTS JD: Notes on Molecular Orbital Calculations. New York, WA Benjamin, 1962

(2) STENT GS: The Coming of the Golden Age. Garden City, N.Y., Am Museum Natural History, 1969

(3) HANSCH C, FUJITA T: ρ-σ-π Analysis, a method for the correlation of biological activity and chemical structure. J Am Chem Soc 86:1616–1626, 1964

(4) HERNDON WC: Theory of carcinogenic activity of aromatic hydrocarbons. Trans NY Acad Sci 36:200–217, 1974

(5) FRANKE R, BÜCHNER M: Zusammenhänge zwischen Elektronenübergangsenergie, karzinogener Aktivität und hydrophober Eiweissbindung polyzyklischer aromatischer Kohlenwasserstoffe. Arch Geschwulstforsch 37:45–52, 1971

(6) FRANKE R: Structure–activity relationships in polycyclic aromatic hydrocarbons: Induction of microsomal aryl hydrocarbon hydroxylase and its possible importance in chemical carcinogenesis, Chem Biol Interact 6:1–17, 1973

(7) ROGERS KS, CAMMARATA A: A molecular orbital description of the partitioning of aromatic compounds between polar and nonpolar phases. Biochim Biophys Acta 193:22–29, 1969

(8) CRAMER JW, MILLER JA, MILLER EC. N-Hydroxylation: A new metabolic reaction observed in the rat with the carcinogen 2-acetylaminofluorene. J Biol Chem 235:885–888, 1960

(9) LOTLIKAR PD, ENOMOTO M, MILLER JA, et al: Species variations in the N- and ring-hydroxylation of 2-acetylaminofluorene and effects of 3-methylcholanthrene pretreatment. Proc Soc Exp Biol Med 125:341–346, 1967

(10) HOLDER G, YAGI H, DANSETTE P, et al: Effects of inducers and epoxide hydrase on the metabolism of benzo[a]pyrene by liver microsomes and a reconstituted system: Analysis by high pressure liquid chromatography. Proc Natl Acad Sci USA 71:4356–4360, 1974

(11) SELKIRK JK, CROY RG, ROLLER PP, et al: High-pressure liquid chromatographic analysis of benzo[a]pyrene metabolism and covalent binding and the mechanism of action of 7,8-benzoflavone and 1,2-epoxy-3,3,3-trichloropropane. Cancer Res 34:3474–3480, 1974

(12) DEWAR MJ, WARFORD EW: Electrophilic substitution. III. Nitration of phenanthrene. J Chem Soc 3570–3572, 1956

(13) DIPPLE A, LAWLEY PD, BROOKES P: Theory of tumor initiation by chemical carcinogens: Dependence of activity on structure of ultimate carcinogen. Eur J Cancer 4:493–506, 1968

(14) SCRIBNER JD: Formation of a sigma complex as a hypothetical rate-determining step in the carcinogenic action of unsubstituted polycyclic aromatic hydrocarbons. Cancer Res 29:2120–2126, 1969

(15) PULLMAN A, PULLMAN B: Electronic structure and carcinogenic activity of aromatic molecules. Adv Cancer Res 3:117–169, 1955

(16) HOFFMAN F: LCAO-MO-SCF indices of chemical reactivity and carcinogenic activity of polycyclic hydrocarbons. Theor Chim Acta 15:393–412, 1969

(17) FUKUI K, NAGATA C, YONEZAWA T, et al: Electronic structure and carcinogenic activity of conjugated compounds, substituted aromatic hydrocarbons, heteroaromatic compounds and azo compounds. Gann 51:67–81, 1960

(18) DIPPLE A, SLADE TA: Structure and activity in chemical carcinogenesis: Reactivity and carcinogenicity of 7-bromomethylbenz[a]anthracene and 7-bromomethyl-12-methylbenz[a]anthracene. Eur J Cancer 6:417–423, 1970

(19) ———: Structure and activity in chemical carcinogenesis: Studies of variously substituted 7-bromomethylbenz[a]anthracenes. Eur J Cancer 7:473–476, 1971

(20) DIPPLE A, BROOKES P, RAYMAN MP, et al: Reaction of 7-bromomethylbenz[a]anthracene with nucleic acids, polynucleotides and nucleosides. Biochemistry 10:4323–4330, 1971

(21) RAYMAN MP, DIPPLE A: Structure and activity in chemical carcinogenesis. Comparison of the reactions of 7-bromomethylbenz[a]anthracene and 7-bromomethyl-12-methylbenz[a]anthracene with deoxyribonucleic acid in vitro. Biochemistry 12:1202–1207, 1973

(22) ———: Structure and activity in chemical carcinogenesis. Comparison of the reactions of 7-bromomethylbenz[a]anthracene and 7-bromomethyl-12-methylbenz[a]anthracene with mouse skin deoxyribonucleic acid in vivo. Biochemistry 12:1538–1542, 1973

(23) BAIRD WM, BROOKES P: Isolation of the hydrocarbon–deoxyribonucleoside products from the DNA of mouse embryo cells treated in culture with 7-methylbenz[a]anthracene-^{3}H. Cancer Res 33:2378–2385, 1973

(24) BAIRD WM, DIPPLE A, GROVER PL, et al: Studies on the formation of hydrocarbon–deoxyribonucleoside products by the binding of derivatives of 7-methylbenz[a]anthracene to DNA in aqueous solution and in mouse embryo cells in culture. Cancer Res 33:2386–2392, 1973

(25) KRIEK E: On the mechanism of action of carcinogenic aromatic amines. II. Binding of *N*-hydroxy-*N*-acetyl-4-aminobiphenyl to rat liver nucleic acids in vivo. Chem Biol Interact 3:19–28, 1971

(26) ———: Persistent binding of a new reaction product of the carcinogen *N*-hydroxy-*N*-2-acetylaminofluorene with guanine in rat liver DNA in vivo. Cancer Res 32:2042–2048, 1972

(27) SCRIBNER JD, MILLER JA, MILLER EC: Nucleophilic substitution on carcinogenic *N*-acetoxy-*N*-arylacetamides. Cancer Res 30:1570–1579, 1970

(28) SCRIBNER JD, NAIMY NK: Adducts between the carcinogen 2-acetamidophenanthrene and adenine and guanine of DNA. Cancer Res 35:1416–1421, 1975