Theoret. Chim. Acta (Berl.) 56, 133–148 (1980)

THEORETICA CHIMICA ACTA

Theoretical Aspects of Some Physiological Correlations

Waldemar Ulmer

Theoretische und Physikalische Chemie der Universität Frankfurt, Robert-Mayer-Str. 11, D-6000 Frankfurt (Main), Federal Republic of Germany

In the case of the non-substituted aromatic hydrocarbons there exist many different correlations between certain molecular properties of these compounds and their carcinogenic activity. Some correlations can be successfully extended to other carcinogenic compounds (aromatic amines, dyes, and molecules with mustard groups). With the help of perturbation theory methods we are able to reveal some interrelationships between the different correlations and to show that the nucleic acids are the specific bioreceptors of carcinogenic molecules.

Key words: Statistical multiregression analysis – Interaction with nucleic acids.

1. Introduction

The characterization of a molecule (or a class of molecules) according to certain chemical and physical properties, representing important features for a specific biological action (e.g. cancer induction), is an interesting problem in biology, biochemistry, and also toxicology. Although this problem is still unresolved – or only partially solved – in many domains, its practical interest is rather evident. Thus a very interesting question is, whether a (synthesized) molecule should be classified as carcinogenic or not. This problem has been most widely studied in the case of the aromatic polycyclic hydrocarbons (PHC) and their derivatives (e.g. aromatic amines, dyes, 3-methylcholanthrene), because 3,4-benzopyrene is highly carcinogenic, whereas 1,2-benzopyrene is not (however, recently this compound is considered as very weakly carcinogenic!) [1]. A central aspect in cancer research incorporates the finding of characteristic bioreceptors, of which the interaction with a cancerogene should represent an essential step in cancer induction. At first, proteins have been considered as the most relevant bioreceptors, but this assumption turned out to be insufficient, because many noncarcinogens showed also comparable interactions (e.g. covalent binding) with proteins [2]. So the interaction of carcinogens with nucleic acids (RNA, DNA) has become a main object of experimental research, and the correspondence between cancer induction and interaction of carcinogens with nucleic acids has revealed to be much more significant. However,

0040-5744/80/0056/0133/$03.20

this interaction did not turn out to be unique [3–6], since besides the covalent binding other forms of interactions have to be taken into consideration (intercalation, charge transfer etc.). But there are many carcinogens not belonging to the PHC and their derivatives, where the covalent binding is guaranteed (e.g. the alkylation of N_7 of guanine by the mustard or nor-N-mustard).

In the case of the PHC the methods of quantum chemistry have played an important role with respect to a relation between molecular properties of the PHC and their classification as cancerogens. The first correlation has been offered by A. and B. Pullman [7] with the famous K–L region hypothesis. This hypothesis, based on a computational analysis of 37 PHC, states that a PHC should exhibit K and L region reactivity indices within a certain threshold to be classified as carcinogenic, and it was claimed by the authors [8] that these properties represent necessary and sufficient conditions for a correct classification. It is striking that the conception of A. and B. Pullman is basing on a consideration of chemical reactivity indices, obtained by a restriction to ground state properties (free valences, charge densities). However, such a conception has some disadvantages:

1. Numerous exceptions exist, when the K-L region hypothesis is extended to further molecules of this class (see e.g. Schribner [1], Clar [9], and Sung [10]).

2. The K-L region hypothesis does not provide a tool to test other classes of carcinogenic compounds with respect to this property and already the extension to the aromatic amines cannot be achieved!

Point 1 was the main motivation of many authors to look for further molecular properties of the PHC quite different from the reactivity indices of the K-L-region hypothesis, in order to find new correlations. So a connection between cancer induction of the PHC and their charge transfer formation ability has been studied by many authors [11]. Further correlations have been offered by Birks [12] and Mason [13] by considering the excited states of the PHC. Birks [12] established a correlation by comparing the low excited states of tryptophan with those of the PHC and postulated a specific energy transfer via dipole–dipole interaction between the amino acid tryptophan and the carcinogenic PHC. This correlation would imply to provide an information on the relevant bioreceptor of the carcinogens, whereas we cannot make the same conclusion from the K–L-region hypothesis. A rather interesting correlation was found by Mason [13], who could classify many PHC by the assumption of a critical energy interval $3.24\,\text{eV} \pm 0.11\,\text{eV}$, where a carcinogen should have an excited singlet state. However, the authors [8] were able to show that the correlations [11–13] also invoke exceptions. Such a situation indicates a shortcoming of methodology [10], as it would be surprising to find a single molecular property to be the only cause for cancer induction! Mason's observa-

tion can be ameliorated [14], if two excited states in the interval 3–3.5 eV are taken into account, satisfying the condition.

$$P_0 = |p - \alpha - 0.15\ \text{eV}| > 0. \tag{1}$$

The term scheme corresponding to (1) can be verified in Fig. 1a. A statistical multidimensional regression analysis revealed many interesting interrelationships [10, 14] between the above mentioned correlations:

$$\text{Carcinogenic Index (0 to 1)} = -0.33\text{K} + 0.37\text{L}$$
$$-0.46E_2 + 0.28E_1 - 0.34P_0 + 0.36\beta - 0.26I_\beta. \tag{2}$$

K, L mean K–L-region reactivity indices, E_1 and E_2 electronic excitation energies, β and I_β β-absorption band energies and their intensities, resply. For 59 molecules of PHC the significance level of (2) is better than 0.001! We shall return to this analysis (2) after having considered the theoretical background of the above correlations and further classes of carcinogenic molecules.

2. Properties of Further Carcinogenic Compounds

Formulae (1) and (2) of the preceding section cannot be applied to other classes of carcinogenic molecules in a simple manner, although (2) appears to be rather encouraging in the case of the PHC. However, we should particularly note that K–L-region reactivity indices as a casual factor of (2) incorporate the main hindrance in order to test other classes with respect to the carcinogenic activity according to the carcinogenic index, given by (2). A further critical aspect of (2) is its only consideration of the PHC without regarding the properties of the corresponding metabolites. Thus the unsatisfactory situation is still present that the strength of the carcinogenicity (e.g. that of 3,4-benzopyrene) is tried to be derived from the properties of the original molecule alone. In the following we shall observe that such a restriction may lead to contradictory results. However, the classification according to the excited states and transition moments, as expressed by Eq. (2), remains a rather promising starting-point, and therefore Mason's observation may be of interest in the case of other carcinogenic compounds. In particular, it is possible to test other classes, whether they possess a similar term scheme, as it is required in the case of (1) and (2), but it is also necessary to take into account the metabolites and to compare the transition moments.

It is striking that the specific term scheme (Fig. 1a) can be replaced by the term scheme according to Fig. 1b, where an excitable singlet state (which may be equivalent to the p-state) is quasi-degenerate with a triplet state, whereas the direct excitation of the triplet state from the singlet ground state is negligible. Thus the triplet state in Fig. 1b may play the same role as the α-states of the cancerogenic PHC. A term scheme of this kind show β-naphtylamine [16, 17] and the cytotoxic drug cyclo-phosphamide (CP) [15]. CP acts as both carcinogenic [2] and cancerostatic compound [18–21] and this may seem to be a

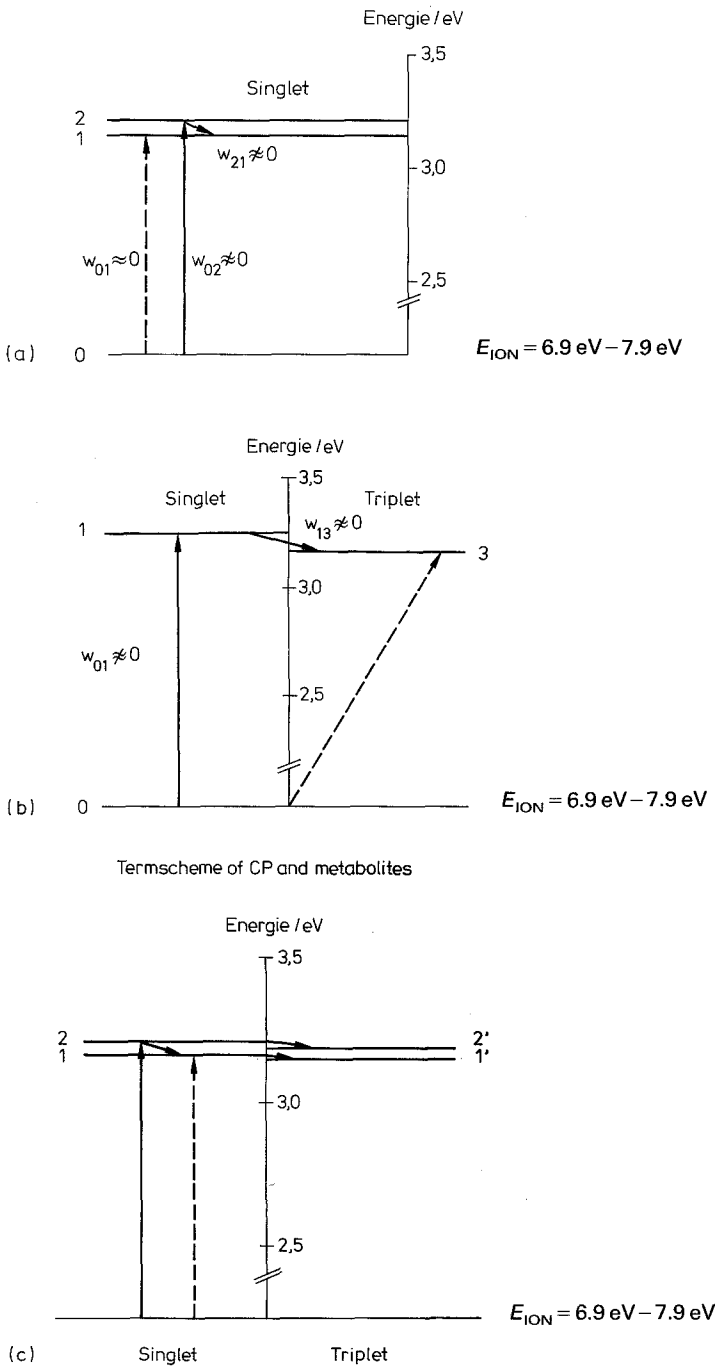

Energie / eV
3,5
Singlet
2
1
$w_{21} \not\approx 0$
3,0
$w_{01} \approx 0$
$w_{02} \not\approx 0$
2,5
(a) 0
$E_{ION} = 6.9\,eV - 7.9\,eV$
Energie / eV
3,5
Singlet
Triplet
1
$w_{13} \not\approx 0$
3
3,0
$w_{01} \not\approx 0$
2,5
(b) 0
$E_{ION} = 6.9\,eV - 7.9\,eV$
Termscheme of CP and metabolites
Energie / eV
3,5
2
1
2'
1'
3,0
2,5
$E_{ION} = 6.9\,eV - 7.9\,eV$
(c) Singlet Triplet

Fig. 1

contradiction, but indeed it is not so, since the latter property is related to a specific cytotoxicity leading to a destroying of tumor cells (but also of normal tissue), whereas the first one is related to a possible action after a latency period (the same fact is also true for the derivatives of nucleic acids, e.g. 5-fluorouracil, and the X- and γ-rays). In the case of CP we have also investigated its metabolites and precursors (mustard and nor-N-mustard: NNM), and because such a class of carcinogenic compounds does not show any topological similarity with the PHC it is justified to analyse the results of our investigations [15]: NNM and other single mustard compounds exhibit a very considerable singlet–triplet-splitting of the two lowest excited states due to the presence of the C–Cl bonds; so the singlet states lie in the domain 7–7.4 eV and only the corresponding triplet states being in the interval 3.1–3.5 eV possess an importance according to Eq. (2). CP and its metabolites result from a binding of the mustard group to a phosphate ester and such a binding has the remarkable consequence that the two lowest excited triplet states, which are related to the excitation of C–Cl bonds, are quasi-degenerate with two singlet states in the same interval (see term scheme 1c) and now the C–Cl bonds may be excited via intramolecular charge transfer of the 3d electron of the P=O double bond. CP and most of its metabolites [15] are characterized by the property that these compounds exhibit almost vanishing dipole transition moments for the two lowest excited singlet and triplet states (this fact is only true for the isolated molecules or environmental conditions being realistic in physiological systems), of which the energy levels in the domain 3.1 eV–3.5 eV are nearly unaffected by the progression of metabolization. CP and its metabolites show the main difference in the term scheme in the domain 4.5 eV–6.5 eV, being related to chemical reactions at the phosphate ester. The two lowest excited states, which we have observed to be of great interest with reference to the carcinogenic activity (Mason's observation), indicate a specific chemical reaction at the C–Cl bonds of the mustard group via splitting off chloride ions Cl^-, and this is known as an alkylation. But CP itself and the metabolites (excepted N–mustard–diamido–phosphoric acid: Friedman acid) can only be considered as *virtually* alkylating agents due to the vanishing dipole transition moments. Only the metabolite Friedman acid [19, 15] exhibits the ability to act as an alkylating agent. If we had only analysed CP, which is itself pharmacologically inefficacious [18–19], we would not have taken account of the fact that the dipole transition moments may be changed by the metabolization and thus an incorrect classification may be obtained!

With respect to Pullmans analysis of the PHC by the means of charge densities and free valences we should note that in the case of CP and its metabolites [15] we have also considered these ground state properties. They indicate for CP that besides the two C–Cl bonds the C4 carbon of the phosphate ester is chemically interesting, but it is impossible to decide whether a chemical reaction takes place at the C–Cl bonds (splitting off Cl^- ions) or at the C4 carbon of the oxazaphosphorine ring (electrophilic addition reaction). Only by taking account of the excited states and the transition moments we are able to

exclude the first possibility. However, it appears that the problem of an assessment of the chemical reactivity of the *nonsubstituted* PHC by charge densities and free valences is solved more reliably, since we have only to consider an unperturbed π-electron system. A further interesting problem which should be pointed out is the increasing of the static dipole moment (SDM) parallel to the metabolization of CP. So the metabolite Friedman acid [15, 18, 19] is the only alkylating agent (under physiological conditions), but a direct in vivo application of this substance turned out to be pharmacologically inefficacious [18], because the ability for passive permeation through membranes is drastically reduced due to the very high SDM. Such an example may illustrate the dependence of the pharmacokinetics of a drug on the SDM, and the biological efficacy of any substance is quite different, whether an active drug is immediately produced in a cell or not. A comparable situation exists with respect to the carcinogenic PHC. This class of molecules is characterized by the metabolic pathway:

Formation of the corresponding phenols (hydroxylation) via an intermediate formation of epoxydes (K-region epoxydes).

However, the experimental facts seem also to be somewhat ambiguous. Many carcinogenic PHC interact with DNA by intercalation, but there are also cases, where highly carcinogenic compounds do not exhibit this ability [2, 27].

On the other hand, some epoxydes did not prove to be so carcinogenic after an in vivo application as the PHC themselves [29] and the ability to intercalate DNA could not be observed [28]. In complete analogy with the results on CP and the metabolite Friedman acid we should note that an epoxydation and hydroxylation yield an increasing of the corresponding SDM and thus the direct use of epoxydes or phenols has not to lead to a clarification of the problem of the ultimate cancerogen. Although it *might* be possible that some PHC represent themselves the ultimate cancerogens, the corresponding metabolites may also lead to a contribution with reference to the carcinogenic activity, because according to formula (2) it is to anticipate that the energy levels of the lowest excited states of the PHC are only weakly affected by the formation of epoxydes and phenols (the same fact is also true for all metabolites of CP [15]) and preferably the transition moments may be changed more significantly by the metabolization. It is also apparent that the higher excited states of the PHC metabolites may differ drastically from those of the PHC themselves, since e.g. a Hydroxylation leads to a remarkable modification of an original σ-bond. With the help of some recent investigations [22–26] we are able to note similar properties in the case of some carcinogenic compounds of rather practical interest: the dyes. The spectroscopic properties can be characterized by the term scheme in Fig. 1a (Ref. [22]) and the interaction with DNA and tRNA has been studied by Ref. [22–26]. In the case of DNA a so-called non-competitive bond (equivalent to intercalation) has been observed [22–24], whereas the interaction of the dyes with tRNA can be described by a competitive bond [22, 23, 25, 26] (this may be a covalent binding). The interaction of this class of carcinogenic molecules with nucleic acids might be considered as

an indication that metabolization is not absolutely a request for cancer induction. But we are also able to present a striking example, where metabolization is required: β-naphtylamine (β-NA). When this molecule is locally applied in an organism, experiments show only a very weak tumor yield in contrast to the local application of the metabolite 1-hydroxy-2-naphtylamine [2] (1-OH-2-NA). The term scheme of β-NA corresponds to Fig. 1b, but the oscillator strength is only small. This fact is changed remarkably, when the above mentioned metabolite 1-OH-2-NA is taken into account. At first, we observe that the position of the energy level of the excited singlet state has slightly decreased (β-NA: 3.56 eV [17, 16] 1-OH-2-NA: 3.46 eV, see appendix). The position of the triplet state of relevance is rather unaffected, and because in Ref. [16] the triplet states have been calculated with great approximations, we have recalculated the lowest triplets, using the CNDO/S-CI method [32]. The dipole transition moment for the first excited singlet state of 1-OH-2-NA has significantly increased, and by that the metabolite exhibits a tendency for singlet–triplet transitions being 35 times greater than that of β-NA. As the resonance condition (1) now is realized by a singlet–triplet transition and the intensity of a singlet excitation appears also in formula (2) to be an essential factor for carcinogenicity, we have qualitatively given a support of this formula, but 1-OH-2-NA is a metabolite (further informations; see appendix A).

3. Theoretical Conceptions of Chemical Reactivity and the Connection with the Chemical Process

It has been pointed out by Sung [10] that the only consideration of the chemical reactivity of the aromatic PHC (as done by A. and B. Pullman) does not tell the whole story with respect to the carcinogenesis of this class of molecules, because other authors [11–13] (see also Ref. [10]) have offered correlations and theories using other molecular properties of the PHC than reactivity indices (e.g. certain excited states). In order to achieve some progress in this situation (each correlation contains some exceptions) Sung has performed a statistical analysis and formula (2) is the corresponding result yielding the interrelationships between the correlations. In the following we shall analyse these interrelationships and we shall observe that the so-called shortcoming of methodology [10] in chemical cancerization is indeed the consequence of a restriction of the chemical reactivity to ground state properties (charge densities and free valences) without taking account of the energy. Such an approach [7,33] has been called a static (or electrostatic) approach by Fukui *et al.* [34], and it is conceivable that this approach may provide some insights, in particular if molecules of a certain class shall be compared. On the other hand, such an approach does not provide an access to thermodynamical properties (e.g. activation energy, Arrhenius equation, etc.). Since the velocity of a chemical reaction depends on the affinity of the corresponding reactants, which is usually expressed by the activation energy of each reactant, it becomes evident that a static approach for chemical reactivity may only work, when the difference between the activation energies of the reactants remains nearly

constant. Therefore more general approaches to the chemical reactivity have been proposed by Fukui [34] and Hartmann [35] by taking account of the HOMO and LUMO of each reactant, and such an approach has been called frontier-electron method.

However, such an approach to reactivity incorporates actually the consideration of the lowest excited state of each reactant, and it is quite conceivable that the subsequent excited states have also to be considered, if the transition moment to the lowest excited state is approximately 0. Within the present methodology of quantum chemistry (CNDO, MINDO, *ab initio*, etc.) there are two main approaches to describe the interaction between any two molecules A and B.:

1. Assume the Hamiltonian of the total system of the form:

$$H = H_A + H_B + H_{AB}\,(\text{interaction}). \tag{3}$$

For the sake of simplicity we suppose that the total system as well as the subsystems A and B shall remain in the singlet ground state. Then we have to calculate (e.g. Hartree–Fock approximation) the ground state energies of the subsystems A and B and of the total system in dependence on all nuclear co-ordinates (in realistic calculations restricted). The Hamiltonian is assumed as usual in quantum chemistry (Coulomb interactions), and spin–orbit coupling is only treated as a perturbation. However, although the total system shall remain in the singlet ground state, the corrections of the Hartree–Fock value by the contributions of excited states obtained via CI methods may become very essential, when A and B undergo interactions, because now the molecules have to be stronger deformed and distorted. But this fact is already true for the separated molecules (without H_{AB}, and the validity of the non-crossing rule is also an indication for the relevance of excited configurations.

2. Assume the Schrödinger equation for H_A and H_B is solved (exactly or only approximately). For the following considerations the impossibility of the first case is of no relevance, as we can measure and classify the eigenstates and transition moments under various conditions. Now we expand the eigenfunctions of H according to (3) in terms of the eigenfunctions of H_A and H_B

$$\psi = \sum_{k=0}^{\infty} \sum_{A,B} C_k \psi_k^{A,B} \tag{4}$$

and make use of the perturbation theory in order to classify the degree of approximation (the applicability of the usual perturbation theory is presumed, but this problem shall be discussed separately). In a first order approximation we obtain the well-known relations for the coefficients:

$$\left.\begin{aligned} C_k^1 &= \sum_{A,B} \sum_{m \neq k} \frac{H_{AB,km}}{E_m^{A,B} - E_k^{A,B}} \\ H_{AB,km} &= \langle \psi_k^{A,B} | H_{AB} | \psi_m^{A,B} \rangle . \end{aligned}\right\} \tag{5}$$

The second-order approximation does not provide any new fundamental insights, since only quadratic terms additionally appear:

$$C_k^2 = \sum_{A,B} \sum_{m,n \neq k} \left\{ \frac{H_{AB,kn} \cdot H_{AB,nm}}{(E_m^{A,B} - E_k^{A,B})(E_m^A - E_n^B)} - \frac{H_{AB,km} \cdot H_{AB,nm}}{(E_m^{A,B} - E_k^{A,B})^2} \right\}$$
$$+ \sum_{A,B} \sum_{m \neq k} \frac{C_m^1 H_{AB,km}}{(E_m^{A,B} - E_k^{A,B})}. \tag{6}$$

We shall now discuss some specific differences between the two approaches and we shall verify that perturbation theory (Eqs. (5) and (6)) has many advantages in biochemical problems, as e.g. formula (2) can be made apparent in a simple manner.

The disadvantages of the approach (point 1) in biochemistry and pharmacology are easy to see: It is rather hopeless to compute the total system "drug-bioreceptor", where "drug" shall mean any molecule (e.g. CP) and "bioreceptor" shall be associated with a large biomolecule (DNA, RNA or a protein). This approach is already hopeless, if one wishes to define the Hamiltonian of such a bioreceptor (one may think of the very complicated geometry of a double-stranded DNA interacting with chromatin).

If a calculation could be achieved, only numerical results would be obtained. This is not quite satisfying, as numerical results are only related to a specific example, which has been calculated, and thus an enormous number of different examples would have to be considered in order to get an insight on a certain class of molecules (e.g. the carcinogenic activity of many PHC).

Therefore we have to restrict ourselves to theoretical means, which also permit to use some experimental properties (e.g. the measurement of excited states), and with respect to such a starting-point the methods of perturbation theory appear to be more appropriate. We have already mentioned the correlation of Birks (dipole–dipole interaction between tryptophan and any carcinogen) and the role of the excited states of the PHC in the domain 3.1–3.5 eV (see also Fig. 1a–1c). This may not be an incident, because the lowest triplet states of the nucleic acids lie also in this domain [31]: *guanine* (G: 3.3 eV), *adenine* (A: 3.35 eV), *thymine* (T: 3.25 eV), *uracil* (U: 3.17 eV), *and cytosine* (C: 3.45 eV)

It has been observed [30, 31] that the triplet states are rather unaffected by the keto-enol tautomery. If the above bases are built in DNA or RNA, we observe rather an energy band of triplet states in the above mentioned interval than separate energy levels (triplet conduction band [30]). The first excited singlet states are much more influenced by the tautomeric equilibrium induced by the H bonds between certain base pairs, and therefore these excited states lie in the large domain $\approx$3.8 eV–4.7 eV [30, 31]. It should be also noted that the amino acid tryptophan possesses two excited singlet states between 4 and 4.3 eV and the lowest triplet state lies in the same interval as the triplet states of the nucleic acids [31, 12]. The other amino acids (inclusive phenylalanine) only

possess singlet and triplet states beyond 4.7 eV and 3.6 eV, respectively. This fact is also true for biomolecules, consisting mainly of saturated bonds, and thus it appears that the energy domain 3.17 eV–3.45 eV is specific for nucleic acids and trypotphan. With the help of this experimental background we are able to analyse the perturbation theoretical approach given by (5) or (6). At first, we should note that we are not able to make any precise statement about the matrix elements $\langle \psi_k^A | H_{AB} | \psi_m^B \rangle$, as we do not know the interaction part H_{AB} of the Hamiltonian (3). In the same sense we have also to treat H_A (if related to DNA or RNA or any protein), but we can make use of the results of a triplet conduction band measurement! This situation looks much better with respect to the matrix elements $\langle \psi_k^B | H_{AB} | \psi_m^B \rangle$, because they may be represented by a multipole expansion and so the dipole transition moments of the separated molecule B (which shall interact with a biomolecule) represent a good approximation. In particular, the matrix element $\langle \psi_0^A | H_{AB} | \psi_0^B \rangle$ is related to ground state properties as considered by A. and B. Pullman. With respect to *all* matrix elements we should point out, that it is actually impossible to give a unique and characteristic information on them. This *might* be possible only for a very narrow class of molecules with a lot of many similar properties. However, the ability to make reliable statements on the basis of incomplete informations increases considerably, if we consider the contributions from the energy levels $(E_k^A - E_m^B)^{-1}$ and $(E_k^B - E_m^B)^{-1}$. As the triplet states under consideration are quasi-degenerate (else the above approach would not be applicable if they would be completely degenerate), the contribution of the terms $(E_k^A - E_m^B)^{-1}$ is then of *great relevance*, if many excited states of the external molecule are themselves quasi-degenerate in the same energy interval. Hereby we have only considered the lowest excited states, because the higher excited states should only represent a higher order perturbation (slight corrections). But a further aspect is more interesting: In the energy domain beyond 4 eV (for triplet states) and 4.7 eV (for singlet states) we should have to take notice of an increasing number of biomolecules, which have lowest excited states in this domain, where the nucleic acids possess only high excited states. Thus the class of bioreceptors would become more and more unspecific.

The importance of the additional resonance condition $(E_k^B - E_m^B)^{-1}$ becomes immediately clear, if we regard the term schemes in Fig. 1a–1c, since this condition expresses a significant contribution for two excited states being quasi-degenerate (besides the quasi-degeneracy with the triplet levels of the bioreceptors), and the consequences of this contribution may be seen in formulae (1) or (2), and in the already mentioned term schemes, resply. With the help of these results we are now able to make apparent the statistical multiregression analysis of Sung [10] (formula (2)), performed in the case of the non-substituted aromatic PHC (see also Ref. [38]).

Besides the already discussed contribution of the energy levels E_α and E_ρ (in the interval 3.1 eV – 3.5 eV) and the resonance condition for $P_0 = |p - \alpha - 0.15\,\text{eV}|$ we mention the contributions of the K–L region indices and I_β. The band intensity I_β follows from the matrix element $\langle \psi_0^B | H_{AB} | \psi_\beta^B \rangle$, and

the K–L region indices from ground state properties $\langle \psi_0^A | H_{AB} | \psi_0^B \rangle$, but these two informations are actually not independent, as treated by Sung, since the dipole transition moments for the transition $\langle \psi_0^B | H_{AB} | \psi_\beta^B \rangle$ depend on the ground state wave function ψ_p^B, and so it is possible to express many ground state properties by the transition moments. Furthermore, although $I_\alpha < I_\beta$, it seems not justified to neglect the contribution of I_α as done in Ref. [10]. As the K–L region index of a PHC follows from $\langle \psi_0^A | H_{AB} | \psi_0^B \rangle$, the only question arises, whether the matrix elements $\langle \psi_k^A | H_{AB} | \psi_m^B \rangle$ can also be interpreted in formula (2) since ψ_k^A is related to the triplet conduction band. Matrix-elements of such a form can also be represented by a multipole expansion and thus we see that a first order approximation yields a dipole–dipole interaction. But contrary to the dipole transition moment $\langle \psi_0^B | X | \psi_\beta^B \rangle$ being related to the external molecule alone, we now obtain a dipole–dipole interaction, which is proportional to the dipole transition of the biomolecule A as well as to the dipole transition of the molecule B. Therefore it depends on the specific form of the wavefunctions ψ_0^A, ψ_k^A and ψ_k^B, whether the matrix elements $\langle \psi_m^A | H_{AB} | \psi_k^B \rangle$ can be related to a charge transfer from B to A and (or) conversely or not.

Finally we should make some remarks on the applicability of the perturbation expansion under consideration. This approach works only, if all excited states are non-degenerate or quasi-degenerate, as it is usually true. However, it follows directly from the expressions $(E_k^A - E_1^B)^{-1}$ that this kind of perturbation theory does not work, if E_k^A and E_1^B identically agree to become a singularity. With respect to such a situation it is necessary to make some distinctions. If it is possible to characterize the system by a finite number degrees of freedom, the usual Schrödinger perturbation theory with degeneracy can be applied without additional restrictions and leads to the same conclusions as in the case of non-degeneracy with reference to the classification of the energy levels. This perturbation theory is described in an excellent way in the textbook of Ludwig [39]. However, in particular in solid state physics there are phenomena, where the perturbation theory [39] is not applicable, e.g. in superconductivity (here we meet a situation, where an *infinite* number of atoms are in the same state and thus the usual perturbation theory fails). If a similar case *should* happen in biochemical problems, we have not to work in an "empty space", because the propagator perturbation theory still works and leads to the same consequences with respect to a classification of molecules than the usual perturbation theory:

Let K_0 be the propagator of the separated molecules A and B (but without H_{AB}), then the Dyson expansion for $H = H_A + H_B + H_{AB}$ (interaction between A and B) reads

$$K_H = K_0 - i \int K_0(x, y) H_{AB}(y) K_0(y, x')\, dy + \cdots + \tag{7}$$

where K_0 is given by the expansion:

$$K_0 = \sum_{k=0}^{\infty} \psi_k^\dagger(x) \psi_k(y) \exp\{-iE_k t/\hbar\}. \tag{8}$$

It is assumed that the set of eigenfunctions $\{\psi_k\}$ belongs as well to H_A and to H_B, and E_k are the corresponding energy eigenvalues. If we neglect higher order contributions in (7), then the spectral representation of the Greens function (in first-order approximation) $K^1_{H,k}$ assumes the shape:

$$K^1_{H,k} \sim \sum_m \frac{1}{(E(k) - E_m + i\tau_m \hbar)} \cdot C_{m,k}. \tag{9}$$

This spectral representation contains a sum over all discrete states (bound states) and continuous scattering states, but we are not interested in the latter case. In principle, we again obtain the well-known resonance conditions, and the possible singularities are removed by an integration over the complex plane. The imaginary part of (9) τ is related to the lifetime of the excited state ψ_m, and thus we are able to get the informations of eq. (5) at once. If only quasi-degeneracy is present (as assumed in our considerations), the Greens function approach does not yield any new information with respect to the classification of interacting molecules. However, it appears that the use of perturbation theory may be useful in problems of biochemistry, biophysics, pharmacology, etc., because this approach can be used to classify interactions of complete classes of molecules with certain biomolecules, of which we know only some experimental properties (e.g. the lowest excited states of nucleic acids) and already the determination of the Hamiltonian of such systems cannot be achieved.

4. Conclusions

In the preceding sections we have observed that the only unique feature of cancerogens, which may belong to rather different classes of molecules, may be characterized by the their tendency to interact with nuclear acids, and it *may* depend on many factors of a cancerogen, in which way such an interaction occurs. We have verified that the term scheme 1a may be replaced by a term scheme 1b (realized by β-NA and 1-OH-2-NA) or 1c (realized by CP and its metabolites), but such a replacement may change an interaction mechanism with nucleic acids in a remarkable way. Many compounds, containing C–Cl bonds, exhibit in particular triplet states being quasi-degenerate with the lowest triplet states of nucleic acids [15] and so intersystem-crossing is involved in a chemical reaction mechanism. The alkylation of nucleic acids by these compounds via splitting off Cl atoms takes place by an intermediate formation of *radicals*, when the interacting compounds exhibit a term scheme according to Fig. 1b and 1c. Because 1-OH-2-NA possesses a term scheme according to Fig. 1b, we assume that the same reaction mechanism plays also an essential role, when this compound interacts with nucleic acids. In the case of carcinogenic PHC and dyes, which have usually a term scheme according to Fig. 1a, biradical reactions are rather unlikely, and so the non-competitive bond (intercalation) may become more interesting, but it is impossible to exclude other interaction mechanisms with nucleic acids, because the specific form of

such an interaction may be influenced by parameters quite different from the term scheme 1a (e.g. the geometry, higher excited states etc.).

A term scheme according to Fig. 1a exhibits also 3-methylcholanthrene [37] (see also the book of Birks [12], where further references may be found), and so it appears interesting to test all strongly carcinogenic substances under the aspects we have founded in the previous section with the help of perturbation theory. Besides the already mentioned compounds, containing C–Cl bonds, we have also studied some other compounds of this class (see appendix A), e.g. chloroform, and it appears that the carcinogenic activity of these compounds can be understood with the same principles. Therefore we arrive at the following conclusions:

Any molecule should exhibit a carcinogenic activity, if it possesses a term scheme according to Fig. 1a–1c with nonvanishing dipole transition moments $\gg 0$ and the static dipole moment must be of such a kind that passive permeation through membranes is favourable. The same properties should also be satisfied by the metabolites (or many metabolites), because only the common activity of the original molecule and the corresponding metabolites leads to a biological action, we finally observe. With respect to the results, we have obtained by perturbation theory in the previous section, it is possible to make two statements:

1. All matrix elements of the form $\langle \psi_k^{A,B} | H_{AB} | \psi_m^{A,B} \rangle$ may be related to the specific interaction mechanism of a molecule (with nucleic acids). This interaction may consist of a chemical reaction (alkylation binding via epoxydes, etc.) or of an intercalation.

2. The division of the matrix elements by the energy differences $(E_k^{A,B} - E_m^{A,B})^{-1}$ may lead to a specific chemical affinity between the interacting molecules, and the activation energy in the Arrhenius equation may be considered as an example for the role of the energy terms in chemical kinetics. In the same way, as we have founded the interaction of some classes of molecules with nucleic acids, we may also give an explanation for the correlation of Birks [12], since we have already stated the energy levels of tryptophan. This amino acid may indeed represent important bioreceptor, as it is an essential constituent of non-histone proteins, and biochemical research [36] shows that the interactions of chromatin with DNA represent a key for understanding of many phenomena (e.g. cell differentiation). Thus we cannot exclude that the interaction of some carcinogens with tryptophan may represent a relevant contribution in cancer induction.

If we give the interaction of carcinogens with nucleic acids a preference, and we have verified that there exist theoretical means to give a general foundation for this interaction, then we have *only* found a relationship between molecular properties and their mutagenic activity. Although there exist profound correlations between mutagenic and carcinogenic activity [2] these correlations *remain* correlations, and it appears impossible to give a profound explanation for such

a relationship, because a rather difficult process like cancer induction involves the consideration of further factors (e.g. repair, immunology, etc.), which lie beyond the scope of theoretical means.

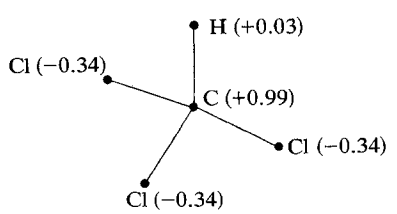
A. Appendix

We have calculated β-NA, 1-OH-2-NA, chloroform, CCl_4, and vinylchloride with the CNDO/S-CI method (Ref. [32]), and the parameters have already been prescribed in Ref. [15]. With respect to β-NA we adopt the enumerations of the carbons by Nishimoto [16], who has performed a calculation with the PPP-method, and we present only some results with respect to the term schemes 1a – 1c. The Cl carbon is the position, where β-NA exhibits the highest charge density [16], and thus this position is preferred for electrophilic addition reactions. β-NA is actually instabolized at this position by a hydroxylation to give 1-OH-2-NA. The subsequent results for β-NA and its metabolite 1-OH-2-NA show that the main differences between the two compounds are the transition probabilities:

	excited singlet states (in eV)		oscillator strength		triplet states (eV)
B-NA:	3.59	(3.56)[a]	0.031	(0.033)[a]	2.97
	4.40		0.12	(0.10)[a]	3.32
1-OH-2-NA:	3.46		0.25		2.95
	4.28		0.37		3.31

We have denoted by ( )[a] the experimental result of Ref [17]. It is immediately seen from the oscillator strength of the transition: ground state → excited singlet state that in the case of β-NA the transition is very weak, whereas for 1-OH-2-NA it has remarkably increased. If we norm the probability of a singlet-triplet transition for β-NA to 1 by taking account of the spin-orbit coupling, we obtain a drastical increasing of this transition probability for the metabolite 1-OH-2-NA, which is now 35! So we can see that this metabolite satisfies in an excellent way the conditions of term scheme 1b.

We have also performed calculations on CCl_3H, CCl_4, and CH_2CHCl (vinylchloride), but for brevity we only discuss chloroform, because the other compounds do not provide any new information with respect to the C-Cl bonds (triplet states). The charge densities of the corresponding atoms can be verified in the subsequent figure, exhibiting the C_{3v} symmetry of the molecule:



However, properties like the charge densities do not provide any indication, why this molecule should be highly carcinogenic. An analogous situation is present with respect to the excited singlet states, which again make apparent the C_{3v} symmetry:

singlet states of CCl_3H (eV)	**oscillator strength**
7.03	0.02
7.09	0.66
7.18	0.55
7.44	0.09
7.49	0.48
7.53	0.37

Much more interesting informations are provided by the lowest triplet states (in eV):

triplet excitation energies	**oscillator strength**
3.18	0.09
3.25	0.78
3.31	0.48

The oscillator strength of a triplet state is treated, as if it would be related to a singlet state. These triplet states are of interest with reference to the already discussed term schemes. Although the direct excitation of two triplets from the singlet ground state is not negligible, because the spin-orbit coupling of the Cl atoms is a significant factor for intersystem crossing, we do not assume that the carcinogenic activity of this compound (and other compounds, containing chlorocarbon bonds) is induced solely by the triplets in the interval 3.1 eV – 3.4 eV. However, we should refer to our results on cyclophosphamide [15] and its metabolites, where we have verified that the lone pairs at the Cl atoms interact with the 3d electron of the phosphate ester (range of the interaction: 3 – 5 Å). Such an interaction, which could be confirmed experimentally [15] yields an additional excited singlet state at 3.3 eV for each C–Cl bond. In the case of chloroform (and similar compounds) it is possible that the interaction of the lone pair electrons with environmental phosphate esters (e.g. those of DNA) may play an interesting role, when a compound of this class of molecules interacts with nucleic acids, and this interaction would mean that the term scheme in Fig. 1c has to be taken into account.

Acknowledgement. The author is deeply indebted to Prof. H. Hartmann, whose profound knowledge of many scientific disciplines was the most stimulating influence on this paper. He wishes also to thank Prof. A. Wacker and Prof. J. Petzold for critical advices.

References

1. Scribner J. D.: Journ. Nat. Cancer Inst. **50,** 1717 (1973)
2. Suess, R., Kinzel, V., Scribner J. D.: Krebs, Experimente und Denkmodelle. Heidelberg: Springer Press 1970
3. Handbuch für experimentelle Pharmakologie. Vol. **XVI,** 12,13 Berlin: 1966
4. Schmähl D.: Entstehung, Wachstum und Chemotherapie maligner Tumoren. Aulendorf: 1970 Ambrose, E. J. Roe, F. J.: The Biology of Cancer. London: 1966
5. Kuroki, T. Heidelberger, C.: Canc. Res. **31,** 2168 (1971)
6. Marquardt, H., Bendich, A., Philips F. S., Hoffman, D.: Chem. Biol. Interactions **3,** 1 (1971)
7. Pullman, A. B.: Cancer Res. **3,** 117 (1955)
8. Pullman, A. B.: Nature **196,** 228 (1962) Pullman, A. B.: Bipolymers. Symp. **1,** 41, 141 (1964)
9. Clar E.: Polycyclic Hydrocarbons. New York: Academic Press 1964
10. Sung, S. S.: Intern. J. Quantum Chem. **XII, Suppl.** I, 387 (1977)
11. Szent-Györgyi, A. Jsenberg, I., Baird, S. L.: Proc. Nat. Acad. Sci. U.S. **46,** 1444 (1960) Hoffman, T. A. Ladik, J.: Cancer Res. **21,** 474 (1962) Allison, A. C. Nash, T. Nature **197,** 758 (1963)
12. Birks, J. B.: Nature **190,** 232 (1961); Birks, J. B.: Photophysics of Aromatic Molecules. London: Wiley-Interscience 1970
13. Mason, R.: Nature **181,** 820 (1958) Mason, R.: Brit. Journ. Cancer **12,** 469 (1958)
14. Sung, S. S.: Arch. Geschwulstforsch. **47,** 109 (1977) Sung, S. S. Damiani, P.: Journ. Soc. Statist. Paris **117,** 14 (1976) Sung, S. S.: Compt. Rend. **257,** 1425 (1963)
15. Ulmer, W. Z.: Naturforsch. **34c,** 658 (1979)
16. Nishimoto, K.: Theoret. Chim. Acta (Bell.) **10,** 65 (1968)
17. Baba, H. Suzuki, S.: Bull. Chem. Soc. Japan **34,** 82 (1961)
18. Brock, N., Hohorst, H. J.: Zeits. Krebsforsch. **88,** 185 (1977)
19. Friedman, O. M. Boger, E.: Anal. Chem. **33,** 906 (1961)
20. Voelcker, G., Giera, H. P. Jäger, L., Hohorst, H. J.: Zeits. Krebsforsch. **91,** 127 (1978)
21. Peter, G. Wagner, T. Hohorst, H. J.: Canc. Treatm. Rep. **60,** 429 (1976)
22. Pauluhn, J. Zimmermann, H. W.: Ber. Bunsenges. Phys. Chem. **82,** 1265 (1978)
23. Pauluhn, J., Zimmermann, H. W.: Ber. Bunsenges. Phys. Chem. **83,** 76 (1979)
24. Sobell, H. M., Tsai, C. C., Jain, S. C., Gilbert, S. G.: J. Mol. Biol. **114,** 833 (1977)
25. Liebman, M., Rubin, J., Sundaralingan, M.: Proc. Nat. Acad. Sci. USA **74,** 4821 (1977)
26. Sakai, T. T. Cohen, S.: Prog. Nucleic Acid Res. Mol. Biol. **17,** 15 (1976)
27. Nagata, C., Kodama, M., Tagashira, Y., Imamura, A.: Gann **57,** 75 (1966)
28. Grover, P. L.: New Scient. **1,** 685 (1973)
29. Boyland, E., in: The Jerusalem symp. on quantum chemistry and biochemistry, Vol. I. New York: Academic Press 1969
30. Eisinger, J. Shulman, R. G.: Science **161,** 1311 (1968)
31. Barenboim, G. M., Domanskii, A. N., Turoverov, K. K.: Luminescence of biopolymers and cells. New York: Plenum Press 1969
32. Del Bene, J., Jaffé, H.: Journ. Chem. Phys. **48,** 1807 (1968)
33. Pullman, A. Pullman, B.: Experientia **2,** 364 (1946)
34. Fukui, K. Nagata, C., Shingu, H.: Journ. Chem. Phys. **22,** 1433 (1954); Fukui, K. Fujimoto, H.: Bull. Chem. Soc. Japan **42,** 3399 (1969)
35. Hartmann, H.: Theorie der chemischen Bindung 2. Aufl. Heidelberg: Springer 1964
36. Sonnenbichler, J.: Naturwissenschaften **5,** 244 (1979)
37. Stevens, B.: Spectrochim. Acta, **18,** 439 (1962)
38. Herndon, W. C.: Intern. J. Quant. Chem. **51,** 123 (1974)
39. Ludwig, G.: Theoretische Physik III (Quantenmechanik). Braunschweig: Vieweg 1976

Received July 18, 1979/October 10, 1979