# QUANTITATIVE PATTERN RECOGNITION FOR STRUCTURE-CARCINOGENIC ACTIVITY RELATIONSHIP OF N-NITROSO COMPOUNDS BASED UPON DI-REGION THEORY*

DAI QIAN-HUAN (戴乾圜) AND ZHONG RU-GANG (钟儒刚)
*(Center for Chemistry & Bioengineering of Cancer, Beijing Polytechnic University)*

Received November 14, 1986; revised March 20, 1988.

ABSTRACT

In this paper, it is evidenced by the quantitative structure-carcinogenic activity relationship (QSCAR) and the pattern recognition treatment of N-nitroso compounds (NNC) that the key step of carcinogenesis induced by NNC is the cross-linking on the complementary base pair of DNA, through the bifunctional alkylation between $\alpha$-carbon and another carbon within the same chain. The alkylation by the $\alpha$-carbon atom is through the diazonium salt, but that by the atom other than the $\alpha$-position is through the active ester formed from the hydroxylated metabolite of the chain. Therefore, the alkylation by the $\beta$-position of NNC, or by its $\gamma$-position under suitable conditions, of which the distances from the $\alpha$-position both approach 2.80—3.00Å, would be the most favourable positions along with the $\alpha$-position for the cross-linking to occur between the complementary base pairs of DNA, which will yield the carcinogenic activity of NNC. The above conception of bifunctional alkylation can reduce the QSCAR of NNC to a reasonable structure-chemical reactivity relationship under the complex biological conditions, and is the successful extension of the Di-region theory to the carcinogenesis mechanism of the important NNC series. In the light of the above viewpoint, for 153 NNCs including the nitrosamines and nitrosamides which have been tested reliably with animals, the correct discrimination ratio by quantitative pattern recognition according to carcinogenic activity indexes divided into 5 degrees comes up to as high as 97%.

**Key words: N-nitroso compounds, Di-region theory, pattern recognition, chemical carcinogenesis, structure-carcinogenic activity relationship.**

Since the discovery of the carcinogenic effect of nitrosodimethylamine by Magee and Barnes[1], three decades have passed, and the importance of NNC (including the nitrosamines and nitrosamides) for environmental cancer has been widely recognized. In the theoretical aspect of NNC research, there are two most interesting and yet pending problems: the structure-carcinogenic activity relationship and the structure-organotropy specificity relationship of NNC. In this study, it is discovered by quantitative pattern recognition method that the Di-region theory proposed by one of the present authors[2,3] can extend successfully to the QSCAR of NNC, thus, the

* Project supported by the National Natural Science Foundation of China.

puzzling QSCAR of NNC in the past can be explained by the chemical reactivity and physical property relationship. As space forbids, this paper will discuss only the QSCAR of NNC, although encouraging results related to the latter problem have been obtained in our laboratory based on the Di-region theory, which will be discussed later. Although the enzymological systems in different animals or in the various organs of animals of the same genus have their own specificities, yet they have much in common. In line with this aspect, we can discuss the above two problems separately.

## I. Di-region Theory and Its Extension to N-nitroso Compounds

The goal of structure-biological activity relationship research (SBAR) is to reduce it to a structure-chemical reactivity or a structure-physical property relationship, or the combination of both. Once this is achieved, we can get a deep insight into the mechanism of the related phenomenon. Conversely, a correct knowledge of mechanism will help to establish the related SBAR. At present the presumption that the covalent combination of the carcinogen metabolite with biomacromolecules induces chemical carcinogenesis gets much more support than that of physical combination with them. Most of the carcinogens during their metabolized course *in vivo* often give electrophilic intermediates in a considerable degree, which produce alkylation of biomacromolecules. By summarizing these facts, a hypothesis called electrophilic theory for the carcinogenic mechanism, which proposed that the monofunctional alkylation agents formed by metabolism were the so-called "ultimate carcinogens", was made by J. A. Miller and E. C. Miller[4]. The popular view, even that of Millers themselves, does not agree with the notion that the electrophilic theory can establish the structure-carcinogenic activity relatioship (SCAR) of any series of chemical carcinogens. Serious research on SCAR has been carried out only in a limited series of parent benzenoid polycyclic aromatic hydrocarbons (PAH), in which the K-region theory put forward by A. Pullman and B. Pullman guided the SCAR research over almost 30 years[5], and the Bay-region theory proposed by D. M. Jerina attempted to apply the Miller's theory to the experimental facts of PAH[6]. However, it is a pity that all the theories could not establish the SCAR of PAH[7]. All the above-mentioned theories lay stress on monofunctional covalent combination of carcinogens with biomolecules. As regards the common supposition that DNA molecules are the target molecules of carcinogen attackings[8], decisive proof is difficult to find. It is evidenced by animal tests and the epidemiological investigation that the action of chemical carcinogens has a very long incubation period, and in the case of human cancer, it can be as long as 20 years. The covalent combination of chemical carcinogens with DNA *in vitro* is as high as 10%, and even *in vivo* their combination ratio with DNA in individual organs comes up also to several thousandths[9]. It is hardly conceivable that this sort of covalent combination with so high a ratio would be the mechanism of chemical carcinogenesis. Owing to the tremendous difficulty in the research of SCAR and carcinogenic mechanism, it is generally considered that the sensitivity and precision or current experimental measures are insufficient to reveal the crucial steps of the mechanism of chemical carcinogenesis[10]. Nevertheless, through rigorous theoretical treatment of a vast amount

of experimental facts, it would be a possible means to uncover the essence of the mystery which cannot be recognized solely by means of experiments.

According to the SCAR of PAH and the experimental facts of biological metabolism, through the systematic quantum chemical calculation, one of the present authors discovered that the necessary condition for the carcinogenic activity of a PAH is the formation of two electrophilic reactive centers in its metabolism course, and that the most favorable distance between these two centers is 2.80—3.00 Å, which is decisive to yielding the carcinogenic activity of the corresponding compound, and just matches the distance between the paired negative atoms in the interstrand complementary base pair of DNA. In the light of these facts concluded from the above calculation, obviously, the key step in chemical carcinogenesis should be the cross-linking between the DNA complementary base pairs, and the mechanism of chemical carcinogenesis should be the frameshift mutation which is very difficult to repair. The above rules discovered by the quantum chemical calculation have been named the Di-region theory. In order to explain these phenomena reasonably, a hypothesis had been put forward in this laboratory too in 1979[10], which holds that the effective attacking site of chemical carcinogens is very possibly the transcriptase gene of DNA, which may be transferred into the reverse transcriptase gene through the complementary frameshift. Thus these wrong DNA produced even at an extremely small probability beneath one millionth will occur potentially in the nucleus of cells, and integrate gradually their own mistaken information into the surrounding normal DNA molecules through reverse transcription mechanism, leading to oncological mutation through the incubation period. Our hypothesis, different from the prevalent "oncogene hypothesis" proposed by Huebner and Todaro in 1969, denies the existence in the normal cell of a specific mysterious gene for carcinogenesis, and based upon the theoretical conclusion of experimental facts, predicts that the carcinogenesis is incited by the variation of a normal gene, the transcriptase gene. It is very possible that the final target of physical or virulent carcinogensis is originated through this mechanism predicted by the Di-region theory[3]. Now, the Di-region theory is not only applicable to PAH but can also extend smoothly to non-alternant PAH, alkyl-substituted PAH, aromatic amines as well as azo dyes and so on[3]. At the same time, the corresponding quantitative relations for all the above series have been established based upon the above bifunctional alkylation conception. Therefore the mysterious structure-carcinogenic activity relationships of a series of compound classes have now become explainable structure-chemical reactivity relationships by theoretical organic chemistry.

Obviously, the above calculation results support the viewpoint that the final target molecules of carcinogenesis are DNA, but the conclusion of the Di-region theory is different from the traditional conception, but implicates that the effective site of carcinogenic alkylation on the base is possibly the site represented by $O^6$-alkylguanine, $O^2$-alkycytosine, $N^6$-alkyladenine, $O^4$-alkylthymine and so on, which are situated between the complementary base pairs, the site where the cross-linking occurs, all these alkylation products appear only with minor contents but can be separated from the culture. We have indicated that some experimental facts independent of the Di-region theory prove consistent with the conclusion of the Di-region theory[3]. We have predicted that

the metabolic products of the indeno[1, 2, 3-cd]pyrene other than those on its angular ring should occur at the 1,2-K-region instead of the 11, 12-position concluded from the tradition conception, the later experiment turned out to be completely consistent with our prediction[7]. Therefore, the Di-region theory is being evidenced both in theoretical and experimental aspects.

The idea about the $\alpha$-position monofunctional alkylation of NNC through diazonium salt proposed by Magee and Farber cannot explain the SCAR of NNC[6]. In this study, we used the experimental data of the carcinogenic behavior of NNC as starting information. It is evidenced by pattern recognition method that the key step of NNC carcinogenesis is also the bi-functional alkylation, which induces the cross-linking between the complementary base pairs of DNA. Thus the SCAR of NNC has been reduced to the reasonable structure-chemical reactivity relationship.

## II. Quantitative Pattern Recognition of SCAR of NNC

The experimental data set of SCAR of NNC, the method and result of corresponding pattern recognition have been published in another paper which was sent to publish at the same time with this one[9].

The original goal of tradition computer pattern recognition based upon the mathematical statistics is to simulate the intelligence of human eyes to recognize the figure pattern. The chemical application of pattern recognition has been influenced deeply by these circumstances. In the pattern recognition of SCAR of NNC made by Jurs and others, the majority of the descriptors (as many as 22 terms) are figured parameters, except the individual descriptor of the so-called mechanism conception of monoalkylation. Obviously, its result is beyond explanation and can hardly predict the unknown facts[10,11]. Our work is somewhat different from the previous. To start with the potential of pattern recognition in analysing abstract rules without any subjective desire should be fully utilized. And if we use the terms representing the regularity of biological metabolism as descriptors, the inherent rules which cannot be seen by the intelligence of human eyes can be obtained. Furthermore, the traditional pattern recognition treatment is only classification according to two degrees (i. e. "yes" or "no"), which is not in tune with the real biological grades possibly achievable. This work aims at harmonizing the theoretical method with the five-degree real carcinogenic indexes, i. e. non (—), slight (+), certain (++), marked (+++) and potent carcinogenic activity (++++), based on our previous work[12], which may bring the strong capacity of pattern recognition for revealing inherent law into full play. Although the mathematical principle of the pattern recognition method in our work is known, the new artificial intelligent strategy applied in our papers for SBAR research is without precedent. This new pattern recognition is named now quantitative analysing pattern recognition (QAPR).

Of the 20 terms selected to simulate metabolic rules as descriptors, through optimization by multivariate regression analysis, we screened out the terms with smaller variances, and got 10 descriptors with unequivocal biochemical meaning. Using them as basis, after binary classification and iteration, the correct discriminating ratio for the carcinogenic activity of 153 NNC came to 97.4%. The 10 descriptors and their

biological contents are listed in Table 1.

From the regression coefficients and variance ratios, it can be concluded that the bi-functional alkylation between $\alpha$- and $\beta$-positions or between $\alpha$- and $\gamma$-positions (if $\gamma$-position just corresponds to $\omega$-site and has neighboring assistant effect) will give the most important contribution to the carcinogenesis of NNC. In terms of the metabolism experiments, the alkylation routes yielded by NNC——just like the metabolism of aliphatic acids——are through $\beta$-oxidation[13] and $\omega$-oxidation[14] as the major trend, alongside the alkylation by $\alpha$-position through diazonium salt. Thus, the intermediate alcohols are alkylated through the active esters formed by the catalysis of sulpho- or phospho-transferase. But the detoxication of NNC can be realized through further oxidation of these alcohols to form acids. The dual routes of NNC for metabolic activation and detoxication are summarized in Fig. 1, where, if the $\gamma$-position

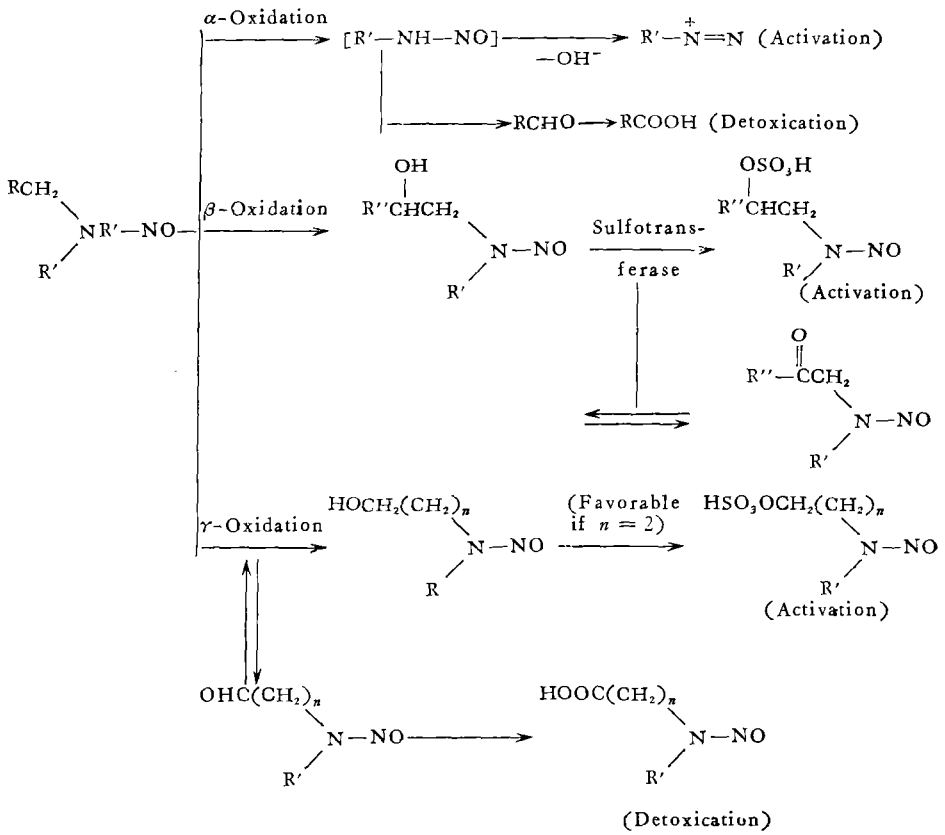



Fig. 1. The major routes of NNC for metabolic activation and detoxication.

just corresponds to the $\omega$-position, it will contribute to the carcinogenic activation. Illustrated in Fig. 2 are the distances both between $\alpha$- and $\beta$-positions and between $\alpha$- and $\gamma$-positions approaching 2.80—3.00 Å, which is just the favourable distance for carcinogenic activity indicated by the Di-region theory, i. e. the distance for the cross-linking between complementary base pairs.

**Table 1**

Descriptors for Pattern Recognition of SCAR of NNC

| No. | Name | Definition | Content | Coef.[a] | Variance |
|---|---|---|---|---|---|
| 1 | $\alpha$-H product | Product of H-atom number on two $\alpha$-carbons | Reflecting the probability and stereohindrance of diazonium produced on $\alpha$-carbons | 0.548 | 1.139 |
| 2 | $\beta$-H sum | Sum of H-atom numbers on $\beta$-carbons with the largest number of H | Reflecting the probability and stereohindrance of active ester produced on $\beta$-carbons | 0.438 | 0.452 |
| 3 | $\beta$-position activating | 3, if Cl, Br or $OSO_3H$ on $\beta$-position; 1, if OH, C=C on $\beta$-position; 0, other conditions | Groups contribute to activation and cross-linking. | 0.610 | 0.178 |
| 4 | $\gamma$-position assistance | 1 with O, N, S, assistance through 3- or 5-member ring, 0 without | $\gamma$-carbon becomes the second effective activating. | 1.340 | 0.158 |
| 5 | Correction on $\varepsilon$-, $\beta$-positions | Sum of H on another chain, if one $\alpha$-carbon joins the N or O atom, $-4$ where both $\alpha$-carbons join N or O atoms, $-4$ where F,N on $\beta$-carbons | Possible through non-enzymatic activation, lose activity by instability of diazonium salt, blocking $\beta$-site activation | 0.569 | 0.920 |
| 6 | Corrected C number | Value $= 2 + \lvert N_C - 4 \rvert$; $N_C$ = sum of carbons | Simulating influence of hole of active center in enzyme | $-0.119$ | 0.082 |
| 7 | Detoxication group | 1 for $CH_2OH$, $-C{=}C$, $-ArCH_3$, 3 for $-CHO$, $-COOH$ | Reflecting detoxication trend | $-0.458$ | 0.093 |
| 8 | Ring effect | $-1$ when <6-member ring, 0 for 6-member ring, 1 when >6-member ring, 0.1 added to every $CH_3-$ on ring | Simulating effects of Baeyer & Pitzer stress on stability of $\beta$-carbonium ion | 0.705 | 0.065 |
| 9 | Corrected N number | Unbound N number with nitroso group | Strengthening activation by kin to special organ | 0.820 | 0.119 |
| 10 | Analogous aromatic ring effect | 0 for both chains with block ring like benzene in volume; number of H on other chains, such ring only on one chain | Simulating key-lock effect of reaction between enzyme & substitute | 0.445 | 0.102 |

a) Regression coefficients.


Fig. 2. Favorable distance between $\alpha$- and $\beta$-positions or between $\alpha$- and $\gamma$-positions.

Because of certain flexibility between the DNA double strands, the distance between both reactive centers is slightly apart from this characteristic distance, such as in dimethylnitrosamine, and cross-linking will occur too, but the activity will be decreased if their difference becomes much larger. When the $\omega$-position is far apart from the $\alpha$-position, the possibility of the reaction on this second center through the bending of chain will be decreased. Additionally, the epoxidation on aromatic ring can produce the second active center, then the cross-linking occurs.

The alkylation by any alkyl nitrosamide, including nitrosoureas and alkyl nitrosourethanes, can take place through the diazonium salt formed from non-enzymatic decomposition without metabolic activation. The decomposition of the former can happen during the reaction with hard base such as $OH^-$, and that of the latter can be realized under the catalysis of soft base such as $^-SR$. However, the formation of the next active center is analogous to the above-mentioned routes.

## III. Explanation by Di-region Theory for SCAR of NNC

The quantitative analysing pattern recognition (QAPR) is different from the routine one. It can explain the SCAR of NNC based upon the descriptors and the inferred Di-region theory conclusion. Thus the SCAR of NNC can now be reduced to the reasonable structure-chemical reactivity relationship under the complex biological condition. The data set for the carcinogenic activity of NNC is compiled in terms of the reports or reviews of animal tests by Magee and Barbes[15], Druckrey and Preussman et al.[16], as well as Lijinsky's recent summary[6], which is based upon the oral administration results as the major basis. In light of the theoretical conception of the Di-region theory for NNC, the metabolic intermediates of NNC can be classified into three types, i. e. the *non-active species*, which includes the metabolites such as alcohols and acids, among them the soluble acids as the detoxified products are easy to eliminate, but the alcohols have the duality of elimination and further activation through the active esters; the *awaiting activation species* which includes the mono-alkylation species on the $\alpha$-position or on the $\beta$- or $\gamma$-position; the *effective activation species* which is the bi-functional alkylation agent formed through mechanism of further activation. However, the classification of the last species does not mean the denial of the possibility that the bi-functional alkylation may take place after the first monofunctional alkylation on DNA. For reference, the relationship between the carbon atom number and the carcinogenic activity of normal NNC is illustrated in Fig. 3 before our discussion. The general trend of the carcinogenic activity of various series NNC

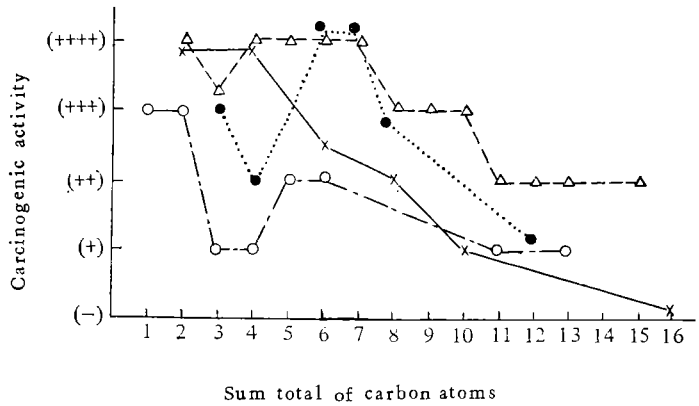



Fig. 3. Relationship between carbon atom number and carcinogenic activity of NNC. △——△, Methyl normal alkyl nitrosamines; ×——×, symmetric di-*n*-alkyl nitrosamines; ●···●, N-nitroso aza-carbocyclic compounds; ○—·—○, N-nitroso *n*-alkyl ureas (skin).

is decreasing during the increasing of the carbon number, which is related to the increase of non-active species producing possibility. However, along the first several members of every series, the possibility of hydroxydation on the chain will increase rapidly, and because of the high polarity of the further oxidation products, the detoxication through draining will be easier than the former members; so there appear the turns in most lines. Now, in terms of the Di-region theory, the regularity of the SCAR of NNC can be explained in the following paragraphs.

1. The hydrophilic group will be of benefit to the elimination of compound from the body, and lower its permeability through the cell membrane, thus decreasing the carcinogenic activity. The influence of —COOH group appears more striking, while the —COOH group is held on the activation-awaiting species, its activity will be a complete loss as shown in (III). But while it cannot be held on the activation-awaiting species, the influence will be relatively small, such as in compounds (I) and (II). In these cases, both activation-awaiting species are the methyl diazonium salts. Their activities can be held even two grades lower than the parent compounds.

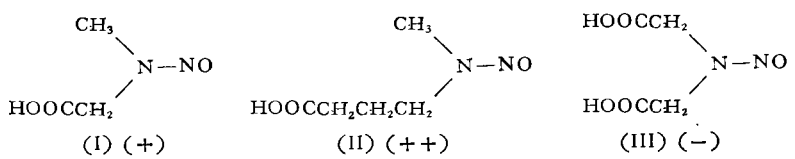
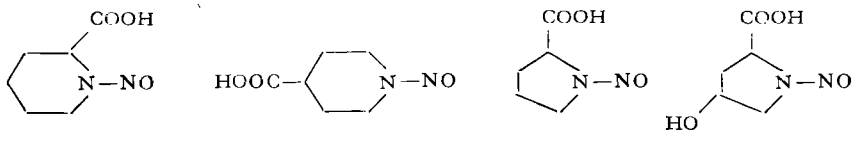



Fig. 4. Remaining or losing of activity based on whether —COOH is on activation-awaiting species.

It is difficult to explain by the previous mono-alkylation conception why once a cyclic nitrosamine has a —COOH group on it, its activity will be a complete loss. But now it is easy to understand because the —COOH group is always on the activation-awaiting species for the cyclic compound.

The esters are very easy to hydrolyze by esterase catalysis during oral administration, so the law is the same with the acid series, i. e. they will lose activity if the ester group is on activation-awaiting species as shown in (IX) and (X), but hold the

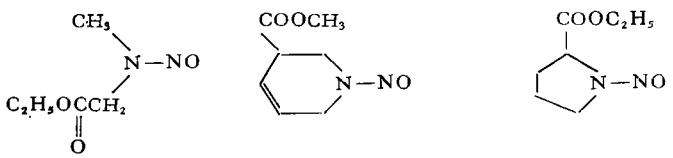


(IV) (−) (V) (−) (VI) (−) (VII) (−)

Fig. 5. Complete loss of activities because of COOH on the activation-awaiting group.

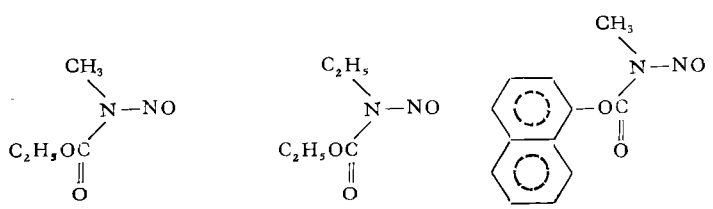
lowered activity if —COOH is not on it as in (VIII).

The condition in the urethane series is more complex, while by oral administration the hydrolysis of the ester will happen first, so the activation-awaiting species should be the alkyl diazonium salt; thus their activities will be analogous to their parents. However, in the case of skin painting, the oxidation decomposition of the alkyl group will become the major trend, so the —COOH group holds on the activation-awaiting species, then the activity approaches loss. Among them, because of the stereohindrance of the naphthyl group, the hydrolysis velocity may be slightly quicker than the oxidation, thus contrary to the general case.



(VIII) (++) (IX) (−) (X) (−)

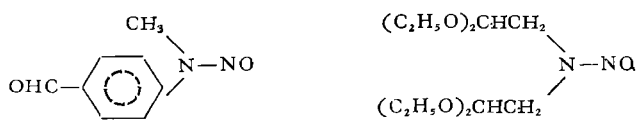
Fig. 6. Hydrolyzable esters show analogous behaviour to acids.



Oral: (XI) (++++) (XII) (++++) (XIII) (+++)
Painting: (−) (−) (+)

Fig. 7. The influence of different decomposition routes on activity.

In general, the activity of oxidase in animal organs is higher than that of the reductase, so the aldehyde group and the acetal group which is easily hydrolyzed to aldehyde *in vivo* can be oxidated smoothly to acid. And their behaviour shows an analogous trend to the acid series. Their activities will be lost if the aldehyde group is held on the activation-awaiting species.



(XIV) (−) (XV) (−)

Fig. 8. Analogous behavior of aldehyde group to carboxyl group.

The ketone group is different from the aldehyde group, which can hardly form

acid by oxidation, so the NNC with the ketone group often holds the activity of its parent. The hydroxyl group, on the one hand, weakens the molecular activity due to its hydrophilic nature; on the other hand, it can be transformed into the active ester, and is of advantage to the activation of cross-linking. Their behavior will be further discussed in the successive paragraphs.

2. While the leaving group such as $Cl^-$, $Br^-$, which makes the nucleophilic substitution proceed easily, has been introduced in the molecule; especially when it is placed at the $\beta$- or $\gamma$-position, the cross-linking will be more advantageous even when their —I effect will lower the metabolic activation. These compounds show comparable or stronger activity of the parent compounds. In Fig. 9, the carcinogenicities of compounds (XVI) and (XVII) approach, those of (XVIII) and (XIX) equal, and those of (XX), and (XXI) are one grade higher than the activities of their parent compounds. In compound (XXII), since the Cl atom has high reactivity due to Pitzer stress, it is a potent carcinogen even when it is two grades stronger in activity than that of its parent compound.

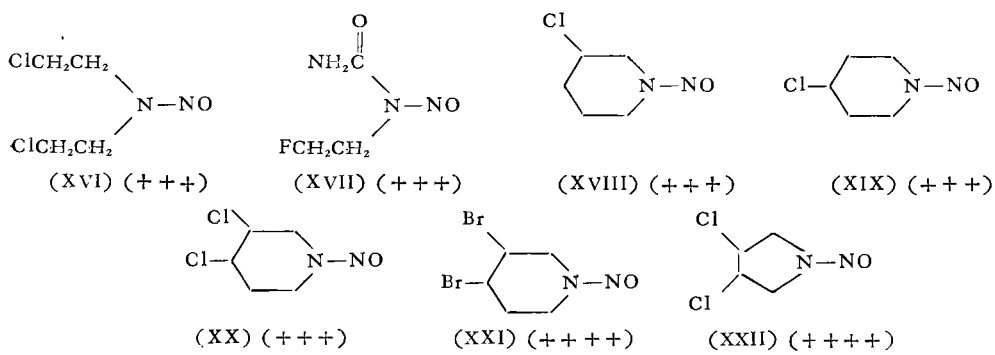
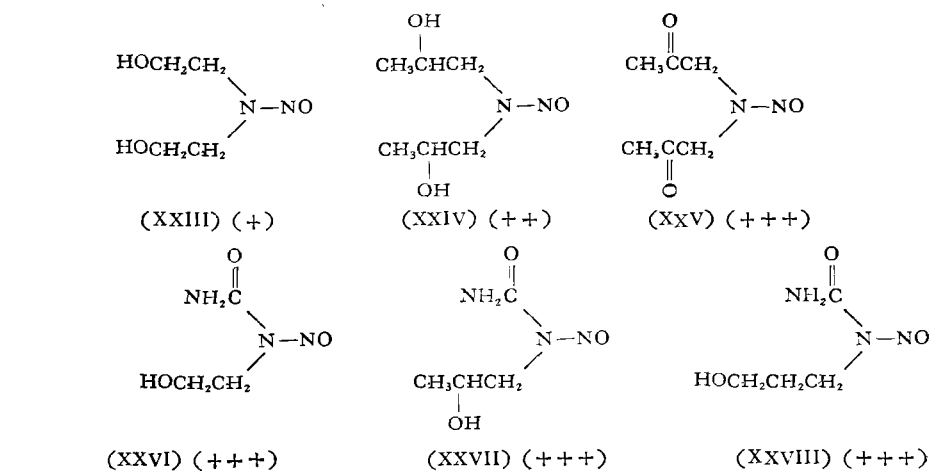



Fig. 9. $\beta$- or $\gamma$-position leaving group holds or strengthens activity of NNC.

As is discussed in the above, the hydroxyl group has the duality of strengthening activity and detoxication. In the case of oral administration, where the OH is an end group, the activity will be lowered markedly, because there is the possibility of its oxidation to form acid. When hydroxyl is placed in the middle of the chain, it can be oxidated to the ketone group only, the decreasing of activity is often negligible. And the hydrophilic property of ketone is relatively weak, but the ketone group has the possibility of activation through reduction. These differences are shown by the following compounds: (XXIII) is three grades lower than the parent, (XXIV) is only one degree lower but (XXV) is with analogous activity. In the case of skin painting, the xenobiotics cannot be drained away easily by the hydrophilic property of hydroxyl, the activity will be held regardless of the hydroxyl group at any place, as in the case of (XXVI), and even the activity is one grade higher than the parent as shown in the examples of (XXVII) and (XXVIII).

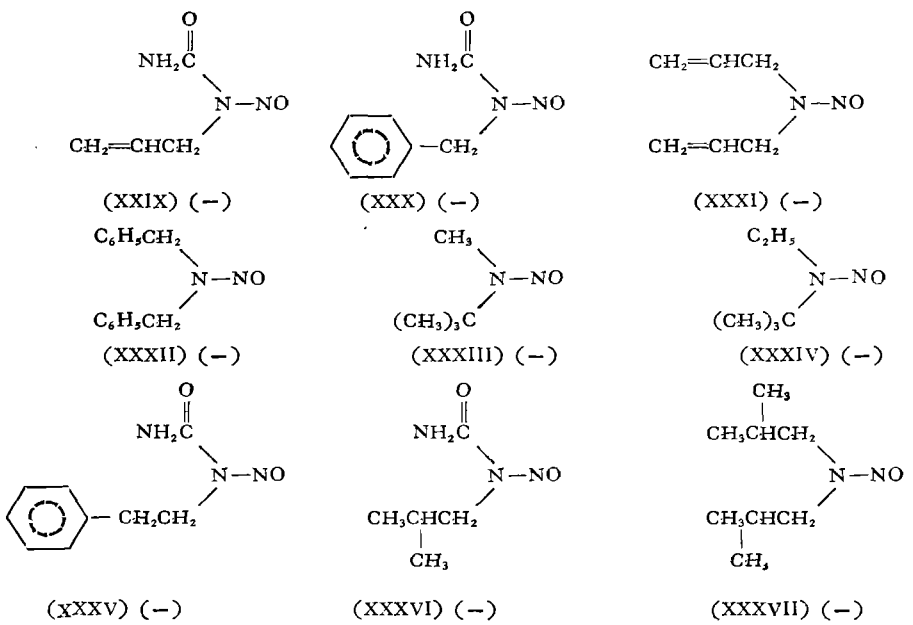
3. For the carbonium ions with high stability or those that are quick to decompose and form the olefine, the corresponding NNC are short of activity, because the activation awaiting species cannot enter the nucleus but react immediately in the cytoplasm.



(XXIII) (+) (XXIV) (++) (XXV) (+++)

(XXVI) (+++) (XXVII) (+++) (XXVIII) (+++)

Fig. 10. Influence of the hydroxyl and ketone groups on activity of NNC.



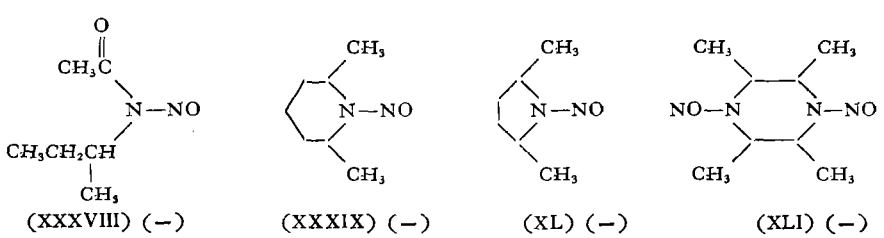
(XXIX) (−) (XXX) (−) (XXXI) (−)

(XXXII) (−) (XXXIII) (−) (XXXIV) (−)

(XXXV) (−) (XXXVI) (−) (XXXVII) (−)

Fig. 11. NNC with unstable activation-awaiting species are short of activity.

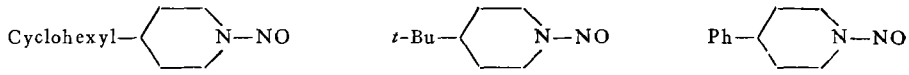
4. The activity of NNC will be lowered, if there is stereohindrance on the $\alpha$-position or on the $\beta$- or $\gamma$-position. There are many examples of stereohindrance on the $\alpha$-position which lowers the activity of NNC, such as



(XXXVIII) (−) (XXXIX) (−) (XL) (−) (XLI) (−)

Fig. 12. Activity loss due to stereohindrance on $\alpha$-position of NNC.

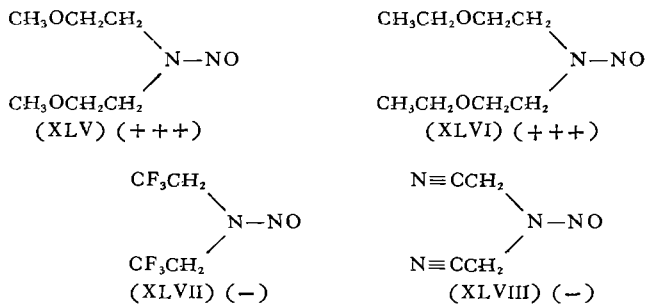
The activity of NNC will be lowered greatly too, if there is stereohindrance on the $\beta$- or $\gamma$-position. As examples show, (XLII), (XLIII) and (XLIV) are three, two and one degrees lower than the parent, respectively. This is difficult to understand by the previous monofunctional conception.




(XLII) (−) (XLIII) (+) (XLIV) (++)

Fig. 13. Activity decreased by stereohindrance on the $\beta$-$\gamma$-position.

5. The electronic effect has relatively small influence on the reactivity of diazonium salt on the $\alpha$-carbon atom because of its high reactivity. However, the electronic effect has important influence on the reactivity of carbonium ion on the $\beta$- or $\gamma$-position. Therefore, +M, +I and the neighboring group assistance will strengthen carcinogenicity, whereas −M and +I will lower the activity. When the $\beta$- or $\gamma$-position has been blocked, the potency of NNC will be lost. The potencies of the following compounds, (XLV) and (XLVI), are higher than the activity of di-*n*-butyl nitrosamine (++) with approximate molecular weight. But (XLVII) and (XLVIII) lose the carcinogenicity completely due to the blocking of the $\beta$-position.




(XLV) (+++) (XLVI) (+++)

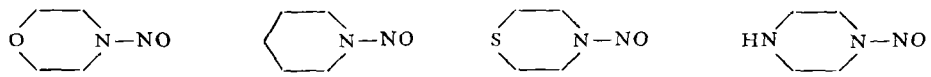
(XLVII) (−) (XLVIII) (−)

Fig. 14. Influence of electronic effect on carcinogenicity of NNC.

Equally clearly shown are the phenomena of the blocking of the formation of the second active center, and those of nitrosamines with aromatic ring. As shown in the example, in $ArN(CH_3)NO$, if Ar=Ph, the potency is (+++); but while its para-position is substituted by F or $NO_2$, the corresponding carcinogenicities are (++) and (−) because of the difficulty of epoxidation on the 3, 4-positions. When Ar=4-, 3- or 2-pyridinyl, the former two compounds will be non-carcinogens due to the inactive influence of the −M effect on epoxidation of the 3, 4-positions, and only the last one has slight activity due to the weakest influence.

It has been invariably considered that the SCAR of cyclic NNC is very hard to explain[6]. Once the importance of the hydroxylation on $\beta$- or $\gamma$-position has been recognized, however, all the problems can be readily solved. For instance, the following series used to leave the researchers at a loss what to do. But they have now become the chemical phenomena. The $\beta$-position carbonium ion of (XLIX) gets +M effect from the O atom, the potency is higher than its parent. It has been known that the sulfur atom in (LI) oxidates readily *in vivo*, and that when the sum trend decreases the electronic pushing effect to $\beta$-atoms, its activity shows slight lowering. In (LII), once the hydroxylation on its $\beta$-atom occurs, the hydrogen atom, due to the

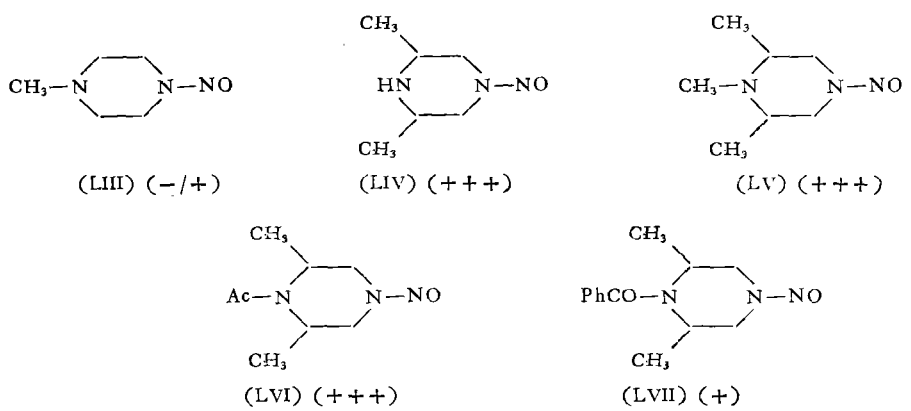
basicity of the nitrogen atom, shifts immediately to split the ring and forms the aldehyde and amine; then the aldehyde group oxidates to form the acid. Thus (LII) is non-carcinogen, because a carboxyl is always on its activation-awaiting species. The above shift is a well-known rule in foundamental organic chemistry.




(XLIX) (>+++) (L) (+++) (LI) (+++) (LII) (−)

Fig. 15. Influence of $\beta$-carbonium ion activity on carcinogenicity of cyclic NNC.

The N-methyl derivative of (LII) and (LIII) is almost non-carcinogen for the same reason. The high potencies of (LIV), (LV) and (LVI) in contrast with the inactive behavior of (LII) and (LIII) would seem an astonishing puzzlement in the previous view. Under this circumstance, however, not only is the hydroxylation on $\beta$-positions of these methylated compounds hindered by methyl group, but the ring split should stop at ketone and cannot be further oxidated to acid. In these cases, the hydroxylation on the $\gamma$-position, i.e. the $\omega$-methyl group would become the major tendency. Conversely, the active esters formed on $\gamma$-positions would get the neighboring group assistant effect from the nitrogen atoms, so they are all potent carcinogens. As for (LVII), it is reasonable that the lower potency is related to the stereohindrance of block group to hydroxylation of $\beta$- and $\gamma$-positions.




(LIII) (−/+) (LIV) (+++) (LV) (+++) (LVI) (+++) (LVII) (+)

Fig. 16. Importance of neighboring group participation on potency of NNC.

The importance of neighboring group assistance on the $\gamma$-position of NNC is also stressed by the fact that the following compounds show much higher carcinogenicity than the corresponding carbocyclic compounds. In Fig. 17, the dotted lines denote the possible direction of neighboring group assistant effect on $\gamma$-positions.

6. Based upon the experimental data of SCAR of asymmetric nitrosamine, it can be postulated that there is probably a hole which can include a block group over 5—6 Å, near the active center of the nonoxygenase of the rat liver and other organs, so the larger group other than the methyl group can be decomposed preferentially; and the asymmetric nitrosamines show analogous activity with dimethylnitrosamine. In the case of $CH_3N(R)$-NO, for instance, it will show a potent carcinogenicity (++++) like dimethylnitrosamine once R = $PhCH_2$-, cyclohexyl-, $PhCH_2CH_2$,

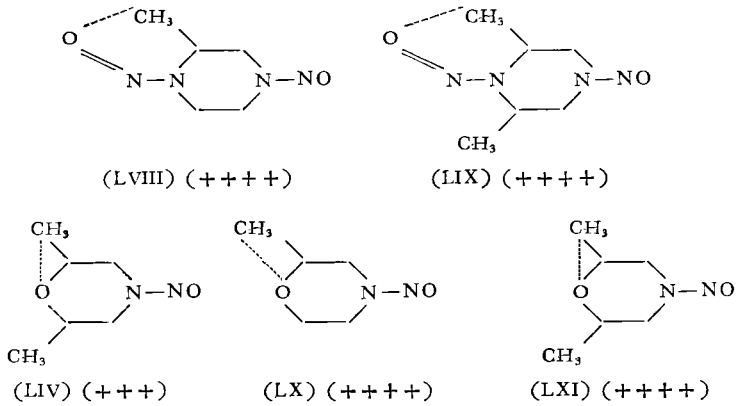



Fig. 17. Neighboring group assistant effect strengthens the carcinogenicity of NNC.

and $(CH_3)_3CCH_2$-. When R = *y*-pyridinyl-$CH_2$- and $CF_3CH_2$-, the activities will be lowered to (+++) because of the inactivation on $\alpha$-carbons by —M or —I effect. If the methyl group is substituted by a larger group, the "move" of R into the hole of the active center would be hindered, therefore, phenylbenzylnitrosamine has only slight activity (+), and both decyclohexylnitrosamine and dibenzylnitrosamine are non-carcinogens.

The SCAR of all 153 NNC we collected, including the above-mentioned ones, can be explained reasonably in terms of the Di-region theory conception. However, in order to consider comprehensively the various factors which might contradict each other, the quantitative treatment would be more convincing, (for which one can refer to Ref. [9]). Since the present pattern recognition method can reduce the SCAR of NNC to a reasonable structure-chemical reactivity relationship, it would be estimated that the QAPR in hand would have good predictive potential.

## IV. Conclusion

There has been ample experimental evidence besides the experimental facts of SCAR that support the above Di-region theory on the mechanism of NNC carcinogenesis. Culture experiment, for one, shows that the dimethylnitrosamine induces a wide methylation on DNA while diethylnitrosamine induces a negligible ratio of both methylation and ethylation. In practice, however, the latter has a higher potency than the former[16]. Obviously, this can be explained by the favorable potential of the ethyl group to cross-linking. Schoental indicated that in the culture of dimethylnitrosamine, the product of cross-linking between protein and DNA may be separated. It is a pity that Schoental proposed inappropriately that the key metabolite inducing the cross-linking was the formaldehyde to be oxidated very easily. And he mistook the cross-linking between the histone and DNA as the key step of the mechanism of NNC carcinogenesis, which, in fact, is the protective mechanism to prevent the cross-linking between DNA double strands. Since these conceptions proposed by Schoental have been quite controversial, his view did not attract NNC researchers' attention, though he guessed at the possible importance of cross-linking for the NNC carcinogensis.

It might not be an extravagant hope that the new conception for the mechanism of NNC carcinogenesis based upon the Di-region theory will further the research on the metabolism of $\beta$- and $\gamma$-positions of NNC, and the cross-linking induced thereom, which has been neglected in a certain degree for a long period of time. As NNC is an important domain for the epidemiology of human cancer, there should be a novel, significant breakthrough made in this field. Furthermore, it was discovered in this laboratory that the Di-region theory can be not only applied successfully to the sphere of SCAR of NNC, but also extended resultfully to other pending problems, i. e. the structure-organotropic specificity relationship (SOSR). Such are the advantages of the Di-region theory in dealing with these puzzling problems. It might implicate that the conscientious theoretical treatment of experimental facts can uncover some scientific mystery which would not be understood only by pure experiments.

## References


[1] Magee, P. N. & Barnes, J. M., *Brit, J. Cancer*, **10**(1965), 114.

[2] Dai Qianhuan, *Scietia Sinica*, **23**(1979), 10: 964.

[3] ————, *Mechanism, Methods and Metabolism in Polynuclear Aromatic Hydrocarbons* (Eds. Cooke, M. & Dennis, A.), Battelle Press, Columbus, 1984, 1045—1073.

[4] Temin, T. M., *Cancer Res.*, **34**(1974), 2835; Heidelberger C., *Ann. Rev. Biochem.*, **44**(1975), 79.

[5] Margison, G. P. & O'Connor, P. J., in *Chemical Carcinogens and DNA* (Ed. Grover, P. L.), CRC Press, 1979, 111.

[6] Lijinsky, W., in *Genetoxicology of Nitroso Compounds* (Eds. Rao, T. K. Lijinsky, W. & Epler, J. L.), Plenum Press, 1984, p. 189.

[7] Rice, J. E. et al., *Mechanism, Methods and Metabolism in Ploynuclear Aromatic Hydorcarbons* (Eds. Cooke, M. & Dennis, A.), Rattelle Press, Columbus, 1984, 1097.

[8] Magee, P. N. & Farber, E., *Biochem. J.*, **83**(1962), 114.

[9] Dai Qianhuan et al., *Environmental Chemistry*, **6**(1987), 1.

[10] Chou, J. T. & Jurs, P. C., *J. Med. Chem.*, **22**(1979), 792.

[11] Rose, S. L. & Jurs, P. C., *ibid.*, **25**(1982), 769.

[12] Dai Qianhuan & Shi Chengxi, *Kexue Tongbao*, **29**(1984), 1239.

[13] Micheida, C. J. et al., in *Nitroso Compounds* (Eds. Scanlan, R. A. & Tannenbaum, S. R.), ACS Symposium Series No. 174, 1981, 3.

[14] Okada, M. et al., *Gann.*, **67**(1976), 771.

[15] Magee, P. N. & Barnes, J. M., *Adv. Cancer Res.*, **10**(1976), 163.

[16] Druckrey, H. et al., *Z. Kreb.*, **69**(1967), 103.

[17] Burnotle, J. & Verly, W. G., *Biochem. Biophys., Acta.*, **262**(1972), 449.

[18] Schoental, R., *Br. J. Cancer*, **33**(1976), 668.