

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



---

# Assessment of Potential Carcinogenicity by Quantitative Structure-Activity Relationship (QSAR)

---

Davor Zeljezic

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.75420>

---

## Abstract

Already in 1978, Elisabeth C. Miller and James A. Miller came with a presumption that electrophilic molecules are predicted to be carcinogens. It is because DNA molecule is reached in nucleophilic centres that may covalently bind to such substances. Rules deduced by Millers are even nowadays irrefutable, and they are used as the basis of testing of the substance for its carcinogenicity potential. Toxicological discipline that emerged from Millers' research is based on dependence of chemical structure of the substance and their biological activity. Even further, there are strict regularities between molecular structures and activities. The tool used in assessment of biological activity of a substance is known as SAR, an abbreviation from structure–activity relationship. Besides electrophilic centres, in assessment of carcinogenic potential of a substance, the SAR also encounters chemical surrounding (neighbouring functional groups), size of the substance, its lipophilicity, number and position of aryl rings, substitutions of hydrogens, epoxides in aliphatic moieties or rings, resonance stabilisation, etc. To these days, SAR has been upgraded to quantitative SAR (QSAR) which applies multivariate statistical methods quantitatively comparing detected characteristics of “alerts” with biological activity of known carcinogens. Nowadays, chemical industry developing novel active substances is unthinkable without application of QSAR.

**Keywords:** structure–activity relationship, mutagenicity, carcinogenicity, QSAR, molecule descriptors

---

## 1. Introduction

Already in the 1940s, for the first time, Auerbach and Robson have reported that a chemical agent may induce mutations which are the main driving event in the process of carcinogenesis [1].

---

They studied the effect of exposure to nervous poison gas yperite (mustard gas (bis [2-chloroethyl] sulphide)) first used in World War I by the German army. They came to this conclusion by exposing the vine flies to components of mustard gas.

Several years later, in 1947, Beerenblum and Schubick [2] published the results of the study confirming that chemicals may induce mice skin carcinoma, thus being able to act as carcinogens.

Following these pioneer reports, even pore evidences have been gathered indicating that chemical agents, as it had been first acknowledged for ionising radiation, may interact with human genome by changing nitrogenous bases and inducing mutations. It was considered that mechanism by which DNA bases are changed is covalent binding of small functional group from the chemical substance or by binding the entire substance to the base. In that way position of polar functional groups (e.g. hydrogens bound to highly electronegative atom) in nitrogenous bases is changed. Such covalently modified bases have changed the ability of forming hydrogen bonds. Thus, in the process of DNA replication, instead of binding complementary nitrogenous base, they will form hydrogen bonds in a way that change the genetic code and produce mutations. Besides changing the way of base hydrogen binding, chemicals may affect DNA replication fork in a way which may lead to insertion or deletion of bases due to structural changes of DNA caused by covalent binding of bulky substances recognised as forming of bulky DNA adducts [3].

Based on all gathered knowledge regarding the interaction of chemical substances and DNA, changing in the ability of hydrogen binding of nitrogenous bases, in 1978 Miller concluded that majority of electrophilic molecules are predicted to be initiators of carcinogenicity due to their affinity to covalently bind to nucleophilic centres in DNA [4]. The conclusion was based on the researches that Miller spouses have been conducted from 1951. Finally, in 1983 Millers and the associates have concluded that there is a strong correlation and regularity between chemical structure of the substance and its biological activity (potentially direct carcinogenicity) and pathway of its metabolic transformation (potentially activation and indirect carcinogenicity) [5].

Based on their findings, a novel chapter in predicting biological activity of chemical substances, besides, in regard to their carcinogenic activity, has been initiated. The principle in assessing the mutagenic and carcinogenic potential is based on comparing the structure, with special concern regarding functional groups of the chemical, with already evaluated substances with known outcome. It is entirely a theoretically approach-based principle that relies on the database for chemicals that have been previously tested in vitro or in vivo or were proven to be biologically (in)active in other ways. Since it is based on relation of the molecular structure of substance of interest and reordered activities in read-across approach considering relevant molecules containing corresponding functional groups, the assessment approach was named as the structure–activity relationship approach or, as abbreviated, SAR [6]. The SAR identifies potential electrophilic centres in the substance of the interest by comparing them to those that have the potential to attack and bind nucleophilic centres in DNA. It also identifies structural moieties and fragments which may contribute to DNA covalent binding. These electrophilic centres are assigned as alerts.

## 2. Electrophilic centres

### 2.1. Types of electrophilic centres

There are two types of electrophilic centres that may interact with nitrogenous bases in DNA and change their ability to form hydrogen bonds with other, not any more, complementary bases. First of them are natively present in the chemical's structure and do not require any metabolic change. The second group of potential electrophilic centres is functional groups that require metabolic activation to be transformed into electrophilic centres [7].

#### 2.1.1. Direct-acting electrophilic species

Direct-acting electrophilic species in terms of the ability to attack nucleophiles in DNA in minority is concerning yet identified carcinogens. Molecules bearing these functional groups do not require metabolic activation to be able to interact with nitrogenous bases and induce mutations. There are four major classes of electrophilic species that may directly bind to DNA (**Figure 1**).

Concerning the formaldehyde ion mode of action, its DNA-binding activity is more complex than just transferring the functional group or entirely binding nitrogenous bases (**Figure 2**). In **Figure 2**,  $R^1$  stands for primary nitrogenous base attacked by the carbon atom as electrophilic centre of formaldehyde ion. In this way formaldehyde binds DNA by a peptide bound. However, in the second step, the carbon atom attacks exocyclic amino group of the complementary nitrogenous base and forms a covalent bond. By acting as such, two complementary DNA strands become covalently bound, instead by hydrogen bonds, which hinder the gene transcription and DNA replication resulting in base insertions or deletions. Such agents that form covalent bonds between complementary DNA strands are considered as cross-linking agents.

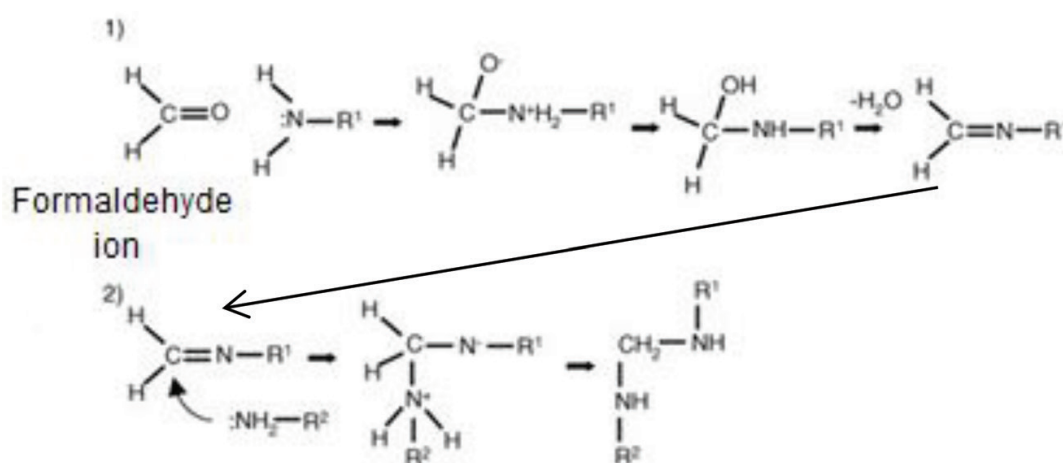
#### 2.1.2. Indirect-acting electrophilic species

Indirect-acting electrophilic species are those potentially electrophilic groups that require a metabolic transformation to be activated and able to interact with a DNA [7].

Carbonyls and carboxylates, in the course of metabolic activation, expose carbon atoms, like in case of formaldehyde ion, which becomes electrophilic and is able to attack nucleophilic centres in nitrogenous bases (**Figure 3**). Further, cyclophosphamide, an antineoplastic drug, by metabolic activation and P450 oxidase activity dissociates to phosphoramidate mustard and acrolein that belongs to carboxylates. Although majority of acrolein is excreted in urine as mercapturic acid following its conjugation with glutathione, small ratio of acrolein will form epoxide glycine aldehyde and reacts with guanine in DNA, which results in the changed base. Such formed chimeric structure hinders its ability to form hydrogen bonds with cytosine, as the complementary base, in the course of DNA replication, thus inducing changes in newly synthesised DNA strand (**Figure 3**).



**Figure 1.** Direct-acting electrophilic species known to be able to act as carcinogens.



**Figure 2.** Formaldehyde ion mode of mutagenic action.

As far as acyl halides are concerned as potential carcinogens, it has to be notified that an electrophilic centre to be formed, halide atom, should dissociate in the process of metabolic activation. In this reaction cytochrome P450 is involved, and in the course of this reaction, the carbon atom becomes available to attack nucleophilic centre in DNA.

As an example of carbenes registered as potential carcinogens, carbon tetrachloride has been present. As shown in **Figure 4**, it can form tricyclic structure with nitrogenous base. It is highly reactive and may open attacking nitrogenous bases in DNA. Its use has been associated to refrigerant agents.

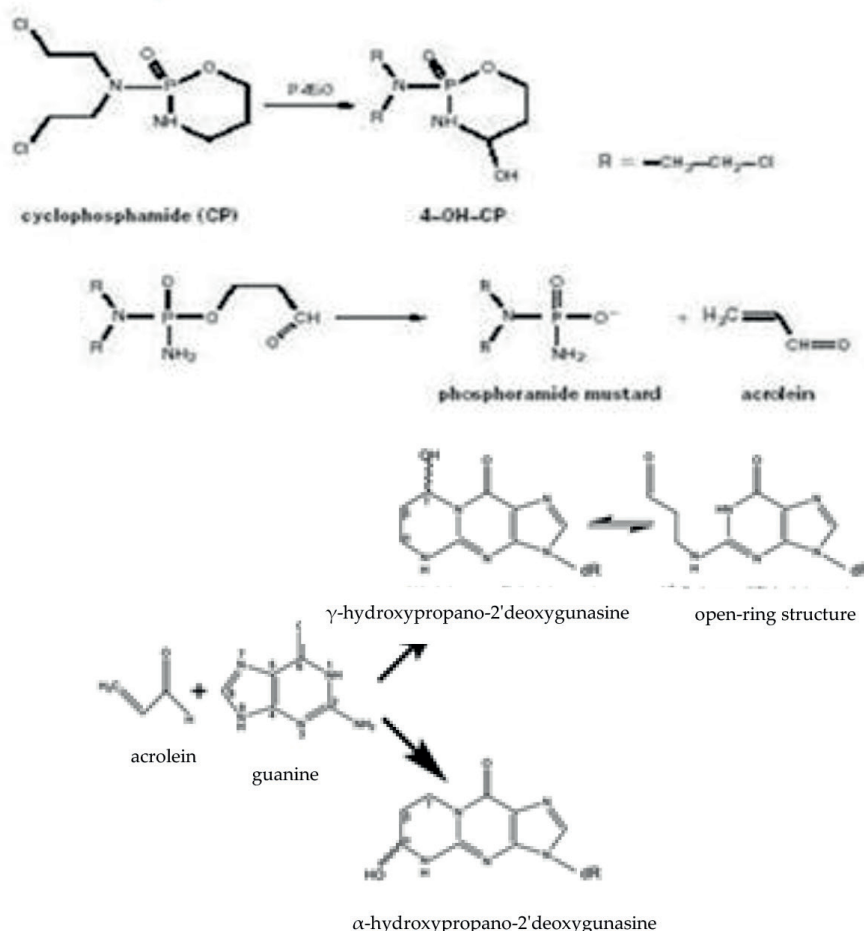
Considering the nitrogen groups as potential electrophilic centres, there are several possible examples (**Figure 5**).

Again, cyclophosphamide as antineoplastic drug by metabolic activation forms aziridinium which is energetically instable, and the ring opens easily. Opened ring interacts

## $\alpha$ , $\beta$ -unsaturated carbonyls / carboxylates

Carbonyls (aldehydes  $\text{R}-\text{C}(=\text{O})\text{H}$  ketones  $\text{R}-\text{C}(=\text{O})\text{R}'$ )

Carboxylates



**Figure 3.** Interactions of carbonyls, carboxylates and carboxylates with DNA.

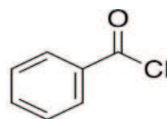
with nitrogenous bases in DNA modifying their ability to pair complementary bases (Figure 5). In more details reaction of aziridinium ion and nitrogenous base is presented in Figure 6.

The consequence of such interaction of aziridinium ion to DNA cross-linking is formed. After opening the aziridinium ring, carbon atom reacts with intra-ring nitrogen atom. Nevertheless, chlorine atom from the second moiety dissociates, and formed ion interacts with intra-base nitrogen of complementary nitrogenous base forming cross-links. The same interaction is presented in Figure 7.

Nitrogen radicals are formed by metabolic activation either. They may interact with DNA in several ways. One of them is presented in Figure 8, concerning mutagenic activity of benzidine.



**Acyl halides** (Benzoyl chloride)  
(activation by CYP450)



**Carbenes** (Carbon tetrachloride)

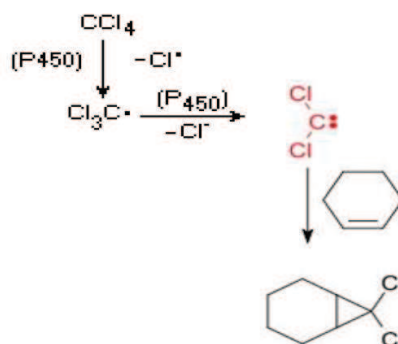


Figure 4. Acyl halides and carbenes as metabolically induced electrophilic species.

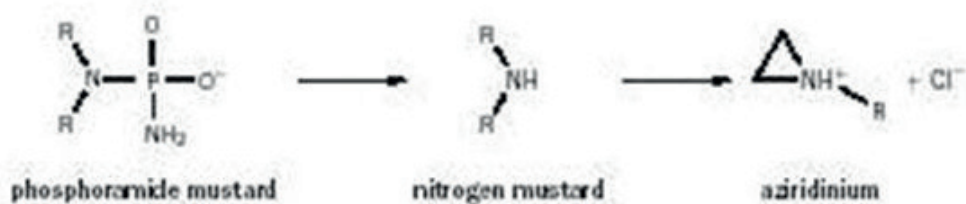
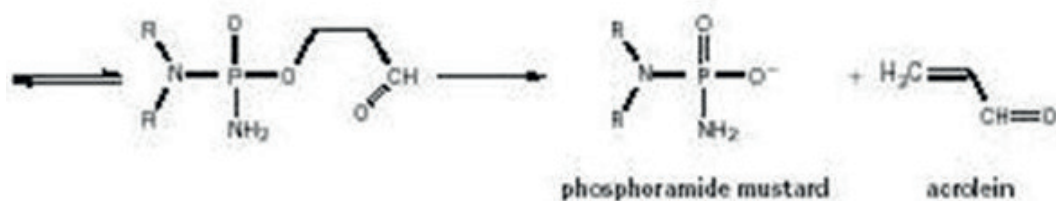
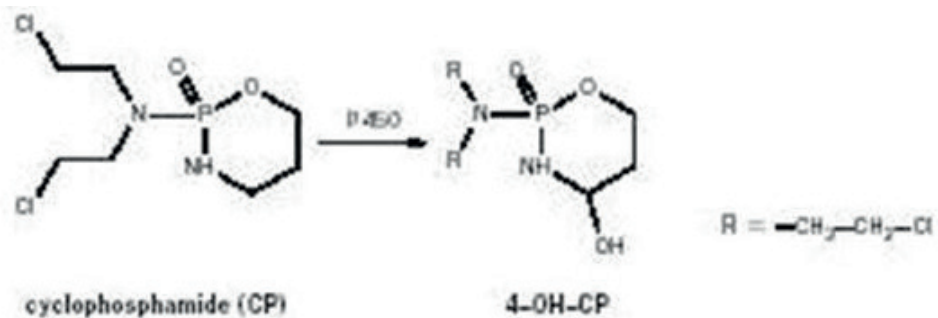


Figure 5. Aziridinium as metabolically induced electrophilic site.

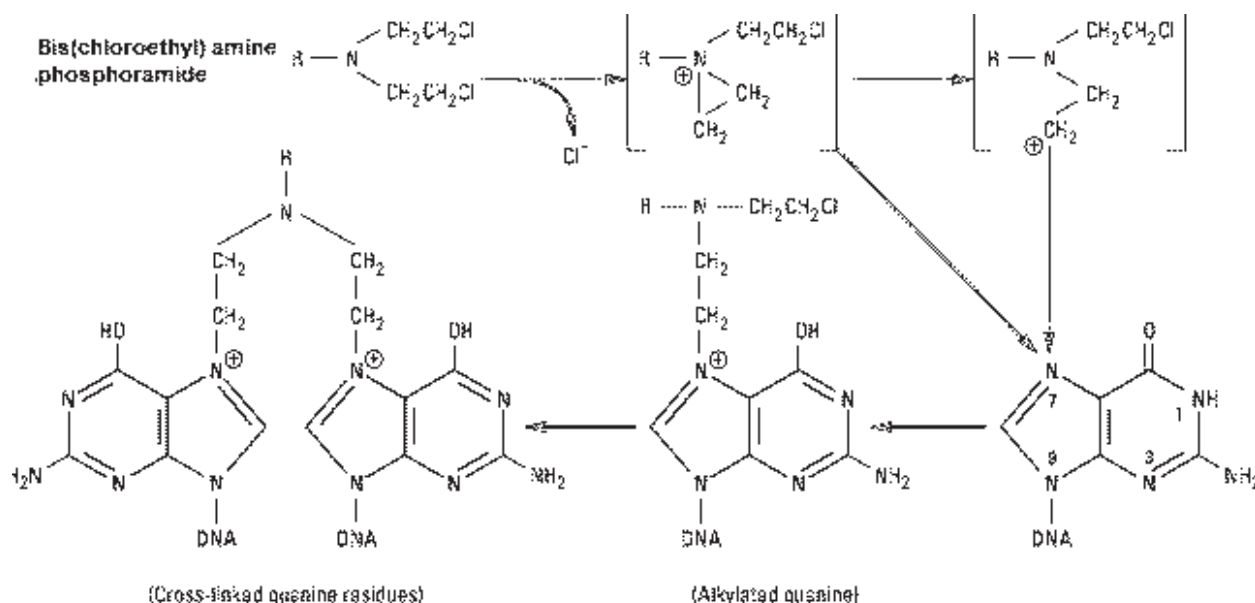


Figure 6. Yperite interaction with nitrogenous bases in DNA.

Benzidine is used as the solvent in dye production. By activity of N-acetyltransferase in the presence of acetyl coenzyme A, it is transduced to the form containing two acetyl functional groups. Acetyl groups easily dissociate leaving two nitrogen radicals which may interact with N1 or C2 of guanine in DNA and inducing irregular base pairing.

Phenylamine is another example of the substance that exhibits nitrogen radical several rounds of metabolic changes as shown in Figure 9, which interacts with the genome.

Metabolic transformation and toxicokinetics of phenylamine (arylamine) are rather complex processes as it can be seen in Figure 9. Phenylamine is used as the manufacture of precursors to polyurethane and other industrial chemicals. In the liver phenylamine is oxidised to hydroxylamine which enters the bloodstream to be excreted by the urinary tract. In urinary bladder epithelium, it may follow several pathways. In the first one, it is N-acetylated to form N-arylamide. This product is further activated by N-acetyltransferase 1 or 2 resulting in formation of acetylated derivate. Acetyl group easily dissociates under conditions present in bladder epithelium, leaving nitrenium ion being able to covalently bind DNA and induce mutation centres. Other scenarios foresee formation of sulfonyl ester or nitrenium ion, both of which are chemically instable and bind covalently to nitrogenous bases in DNA resulting in DNA mutations [7].

Peroxy radicals ( $\text{R-O-O} \cdot$ ) belong to reactive oxygen species (ROS) and are characterised with extremely long half-life (in order of seconds) compared to other ROS [8]. They are spontaneously formed by the process of autoxidation mostly of unsaturated fatty acids in the food, when hydrogen atom is removed and the rest of the molecule interacts with molecular oxygen producing the peroxy radical. It predominantly interacts with thymine forming highly mutagenic 5-(hydroperoxymethyl)-2'-deoxyuridine, 5-formyl-2'-deoxyuridine and 5-(hydroxymethyl)-2'-deoxyuridine [8, 9].



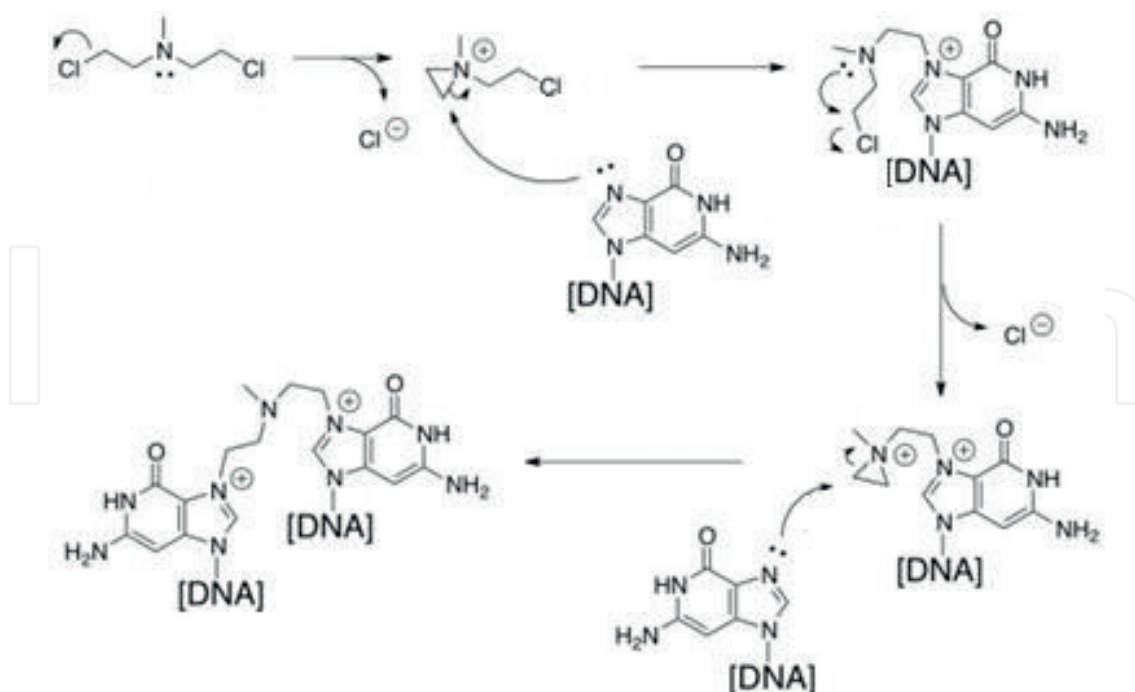


Figure 7. Yperite interaction and genotoxic mode of action.

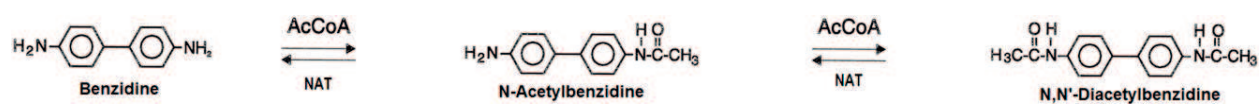


Figure 8. Activation of benzidine into DNA-binding substance.

Epoxides present the next electrophilic oxygen-containing group that requires metabolic processes to be formed. They are cyclic ethers in the form of equilateral triangles with one of the atoms being oxygen. Their high ring strain makes them highly reactive. In attacking nitrogenous bases in DNA, the ring opens, and carbon atom reacts with nucleophilic centre in the base.

Examples of epoxides together with their reaction with DNA are shown in **Figure 10**.

Aflatoxin B<sub>1</sub> is a mycotoxin predominantly produced by the fungus *Aspergillus flavus*. It is widely present in food and feed, especially in areas with warm and humid climate. It can be found in grain, various nuts and wine. It is suspected to be a potent hepatocarcinogen, although lately there is a strong indication that its carcinogenic effect is potentiated by coinfection with the virus of hepatitis B. Nevertheless, aflatoxin forms epoxide, and after the opening of ring, it forms DNA adducts by binding to N7 of guanine. N7 of guanine is a preferential site for adduct formation, especially those which formation is mediated by epoxide ring openings. Another such examples are vinyl chloride which is raw a material in production of plastic polymer polyvinyl chloride. Benzo(a)pyrene belongs to a wide group for chemical names like polycyclic aromatic hydrocarbons (PAHs). Benzo(a)pyrene is a class 1 carcinogen to humans according to the International Agency for Research on Cancer (IARC). It is a constituent of chimney soot, coal tar and exhaustion gases especially those from diesel engines, cigarette smoke and every

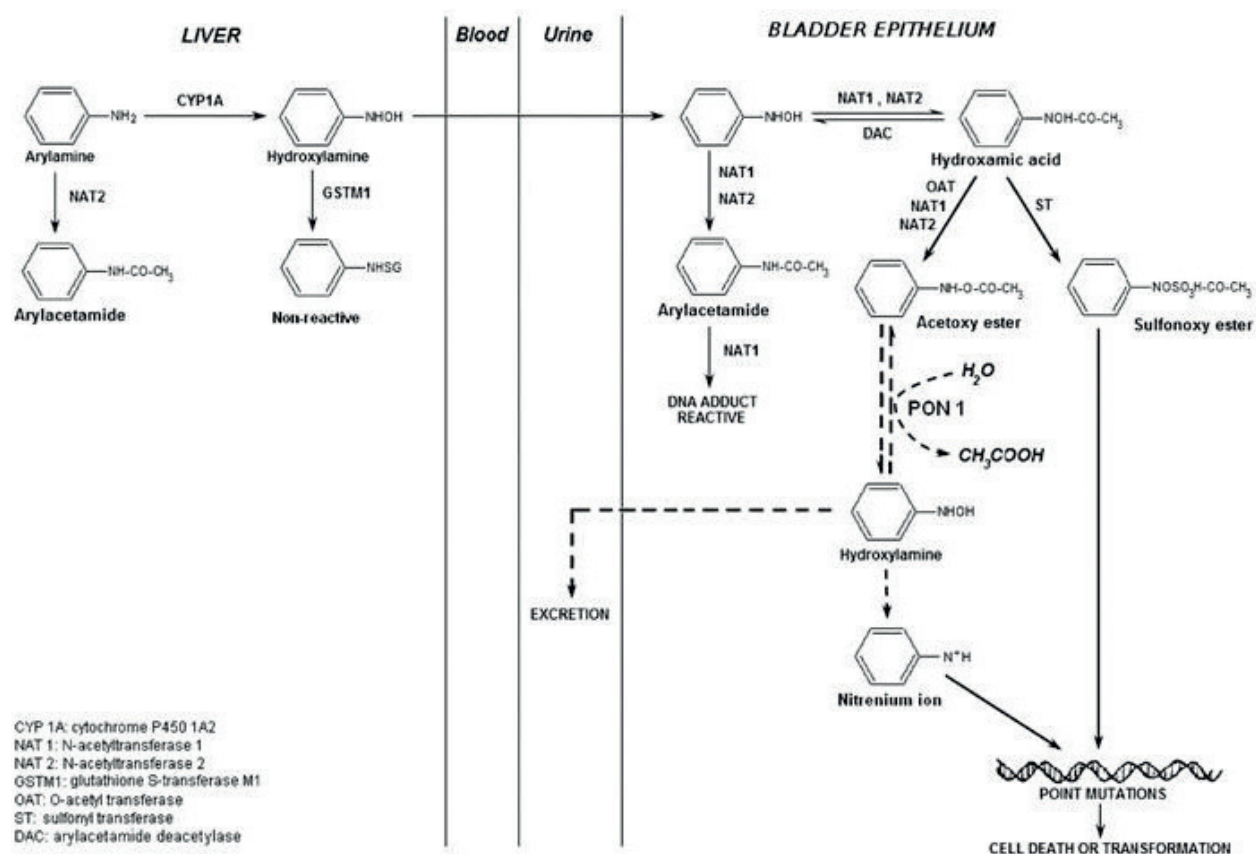


Figure 9. Mutagenic mode of action of phenylamine (arylamine).

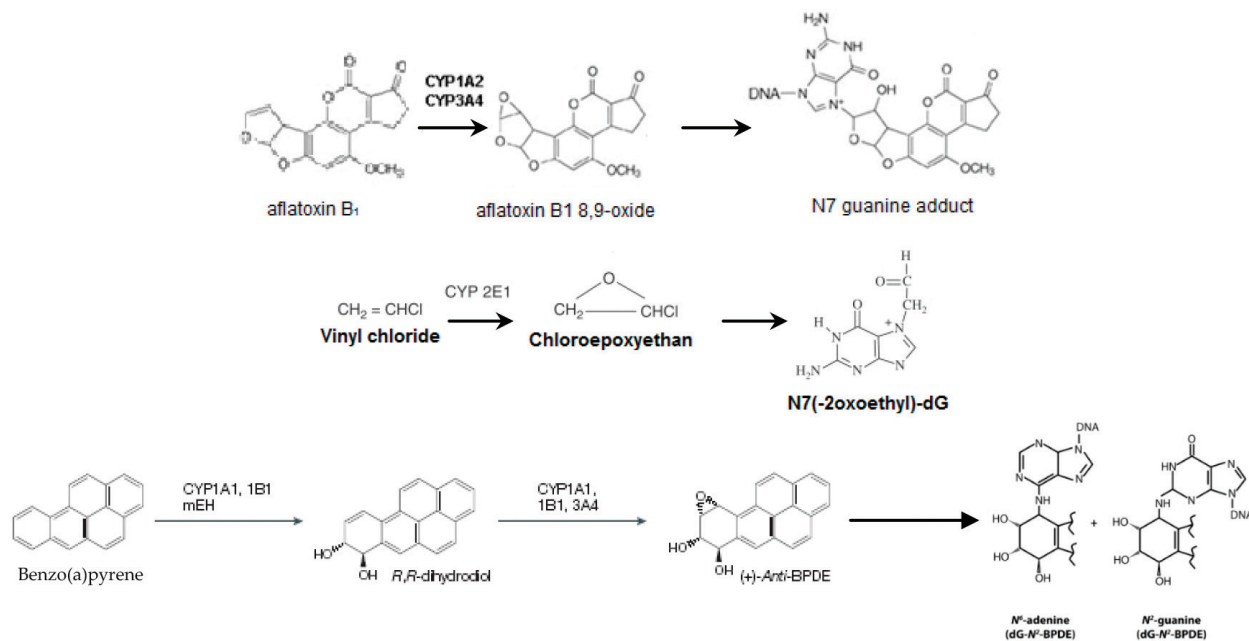
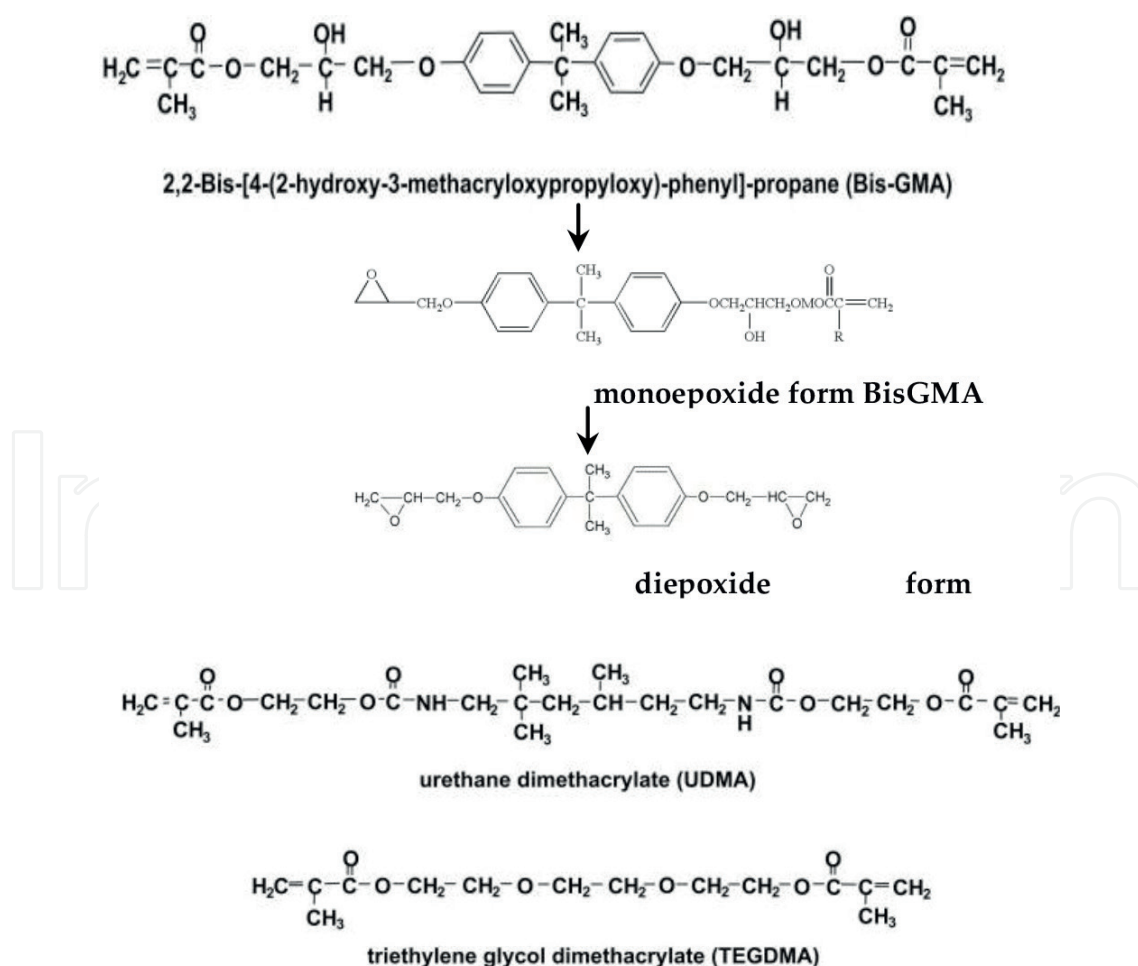


Figure 10. Examples of molecules that by metabolic oxygenation form epoxides. In further steps epoxide rings open and DNA adducts are formed.

smoke originating from combustion of the organic matter. As with previous two chemicals, it is capable to form bulky DNA adducts by binding N7 of guanine. There are many other chemicals being able to via epoxide intermediate bind to DNA by forming bulky adducts. To those substances belongs bisphenol A, also identified as endocrine disruptor exhibiting its hormone poisonous activity by DNA binding. But also to via epoxide formation large group of resin monomers used in endodontic materials such as BisGMA, TEGDMA and UDMA (**Figure 11**) are activated [7].

As it was shown in the examples given in **Figure 10**, single epoxide may lead to single DNA adduct formation. Speaking of resin monomers, they are capable of forming two distant epoxides in the same molecule, thus being able to covalently bind two guanines. Thus, they may form cross-links by covalently linking guanines in complementary DNA strands or form dimers by binding two guanines in the same DNA strand.

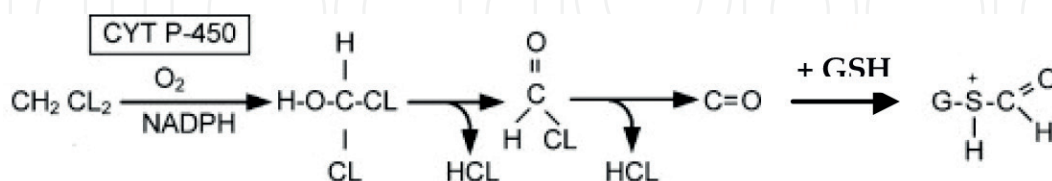
A last group of electrophilic centres that require metabolic transformation of precursor molecule to be formed are sulfonium ions. It is a species containing sulphur atom that has an octet of electrons but bears a formal charge of +1. It can be present in two different structures: open and ring (**Figure 12**).



**Figure 11.** Resin monomers used in endodontic materials and process of forming diepoxide of the molecules.

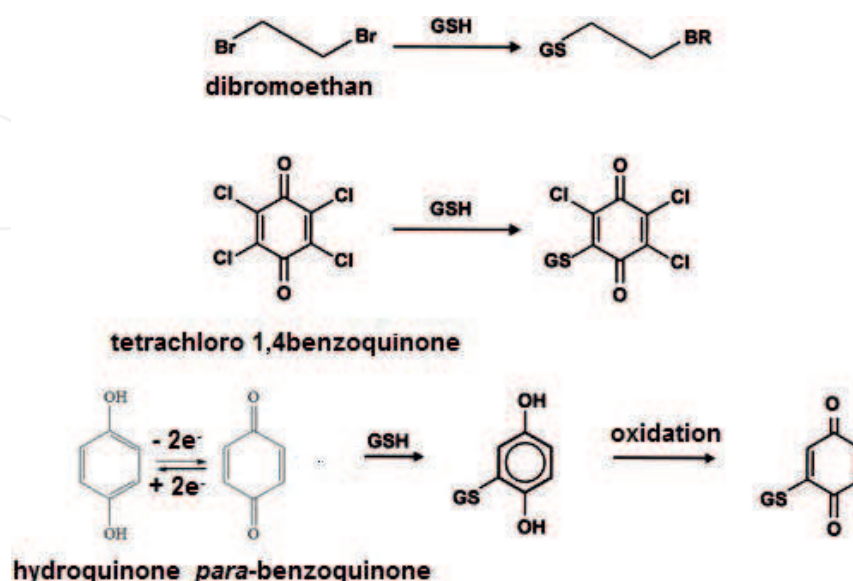


**Figure 12.** Forms of sulfonium ions. In the open form, H atoms and methyl group may be replaced by any organic moiety.

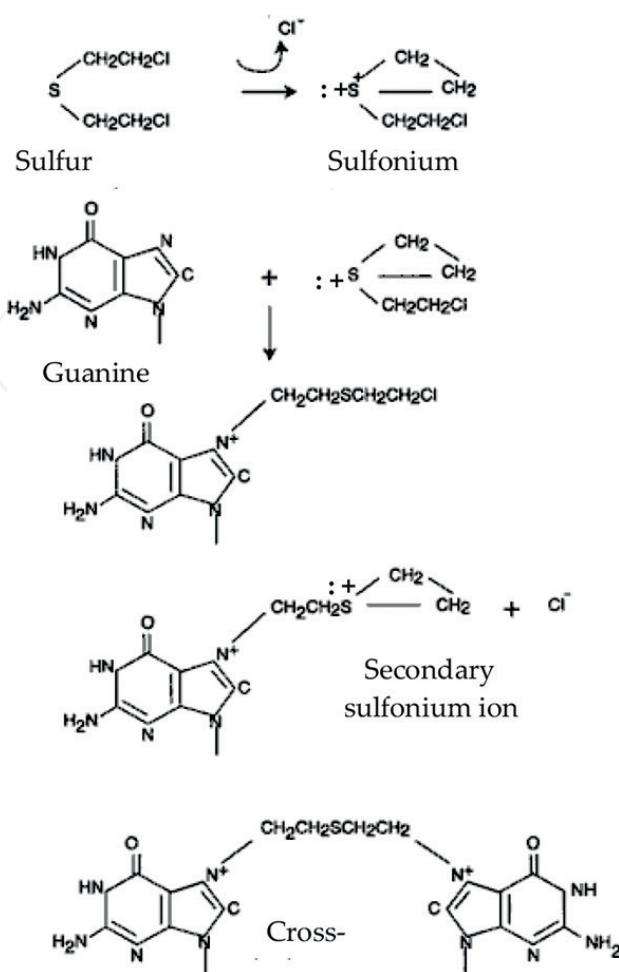


**Figure 13.** Formation of sulfonium ion by metabolic transformation of dichloromethane.

Formation of sulfonium ion is most frequently preceded by the second step of metabolic transformation, chelation with glutathione. Chelation produces sulfonium ion which, due to its electrophilic characteristics, may attack nucleophilic centres in DNA (**Figure 13**). One of examples for chemicals that form sulfonium ion is dichloromethane [10]. The substance is used as the leaching agent not only in the industry but also in production of decaffeinated coffee and tea. Except via sulfonium ion, it has been proved that in mice degradation of dichloromethane goes down to formaldehyde which induces mutations by forming protein-DNA cross-links [11]. Some other chemicals known to interact with nitrogenous bases through sulfonium forms are shown in **Figure 14**. Dibromoethane, for instance, is used as fungicide, insecticide and precursor in the production of insect repellents. Tetrachloro-1,4-benzoquinone is a fungicide, while hydroquinone is applied in the skin whitening cosmetics.



**Figure 14.** Examples of substances that form sulfonium ion as the result of glutathione binding in process of metabolic transformation.



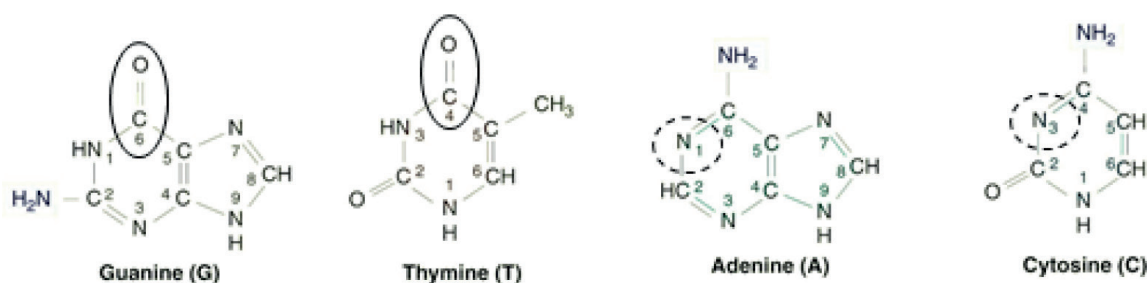
**Figure 15.** Formation of circular form of sulfonium ion and its interaction with DNA.

Earlier, we have talked by mustard gas, which in its initial form contains nitrogen atom that binds two chloroethyl moieties and attacks DNA via aziridinium ion formation. However, other form of mustard gas, sulphur mustard, contains sulphur instead nitrogen and forms circular form of sulfonium ion which, after opening, attaches to guanine. As in the case of mustard, sulphur mustard is capable of forming two rings in succession, thus covalently binding complementary DNA strands and acting as cross-linking agent (**Figure 15**; [12]).

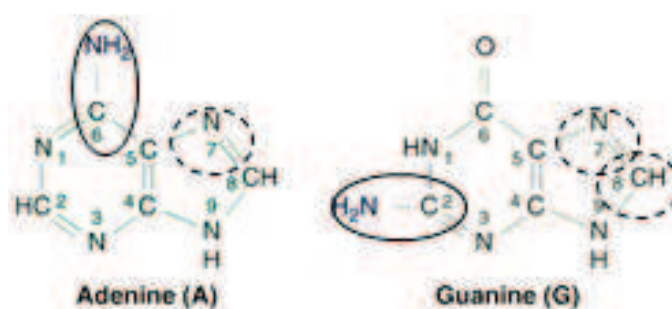
### 3. Nucleophilic centres in DNA

Each of four nitrogenous bases in DNA that form genetic code poses specific nitrogen and oxygen atoms that are capable of donating an electron pair to an electrophilic centre to form a covalent bond. In this way nitrogenous bases are structurally changed. Their ability to form hydrogen bonds with complementary bases is also altered, and, if not repaired, after the DNA replication, new synthesised DNA strand will contain a base with which damaged one can bind. In this way, a mutation is formed and fixed and will result in the change of the genetic code.





**Figure 16.** Nucleophilic sites in nitrogenous bases which are predominantly prone to alkylations. Solid-lined ellipses indicate the most frequent atoms being subjected to alkylation, while dash-lined circles roundless frequently alkylated atoms.



**Figure 17.** Nucleophilic sites in nitrogenous bases which are predominantly prone to formation of DNA adducts. Solid-lined ellipses indicate the most frequent atoms being subjected to binding of bulky molecules, while dash-lined circles round less frequently bound atoms.

Due to steric properties of DNA bases and DNA itself, there is a regularity which nucleophilic sites are available for binding of bulky molecules and adduct formation and which are predominantly alkylation sites (**Figures 16 and 17**).

Primary sites for transferring an alkyl functional group from mutagenic substances are exocyclic oxygen bound to C6 atom of guanine or C4 atom of thymine. Less frequently alkylation occurs at N1 atom adenine or N3 atom of cytosine.

Exocyclic nitrogen bound to C6 in adenine and C2 in guanine is a primary site of DNA adduct formation. Nevertheless, as shown in examples of electrophilic centres, many bulky adducts prefer N7 atom of guanine or adenine. Besides C8 atom of guanine is also prone to bounding of bulky molecules, but that site is primary nucleophile for oxidative changes of base by ROS.

#### 4. Characteristics that mediate electrophilic centre activity

Over the time even more substances have been analysed for SAR, and all the data helped the database which rely on the entire approach used in SAR to identify potential carcinogens. Broadening the database and expanding it with data on other molecular characteristics than



solely functional groups lead additional knowledge regarding dependence of chemical structure and biological activity to be acquired. It has been learned that other features such as the presence and distribution of other functional groups that are even not electrophilic centres, planarity of the molecule, distribution of aromatic rings, points in the molecules where epoxides are formed, and many other characteristics influence activation of electrophilic centre, its stability (half-life) and activity. Thus, all of them should be considered in prediction of possible mutagenic and carcinogenic potential of the substance of interest.

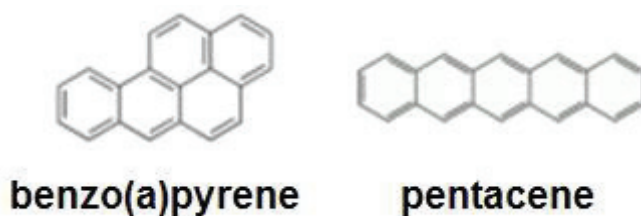
Size of the molecule is one of its characteristics that matters in prediction of its reactivity. Molecules with molecular weight beyond 1000 are not likely to be absorbed and enter the bloodstream. Even in the event of such scenario, it is highly unlikely that they will be able to enter the cell and be approached by the active sites of enzymes needed for metabolic transformation. Last but not least, such bulky molecules will not trespass the nuclear envelope in order to get into interaction with DNA.

Highly hydrophilic substances will be hardly absorbed in the organism, either. Even so, they are rapidly excreted which mitigates their DNA damaging activity. On the other site, highly lipophilic chemicals will not effectively dissolve in blood plasma or cytoplasm, which are aqueous media, which, again, hinders them to reach and damage genetic material. For an effective mutagen, a balance between its lipophilicity and hydrophilicity is needed.

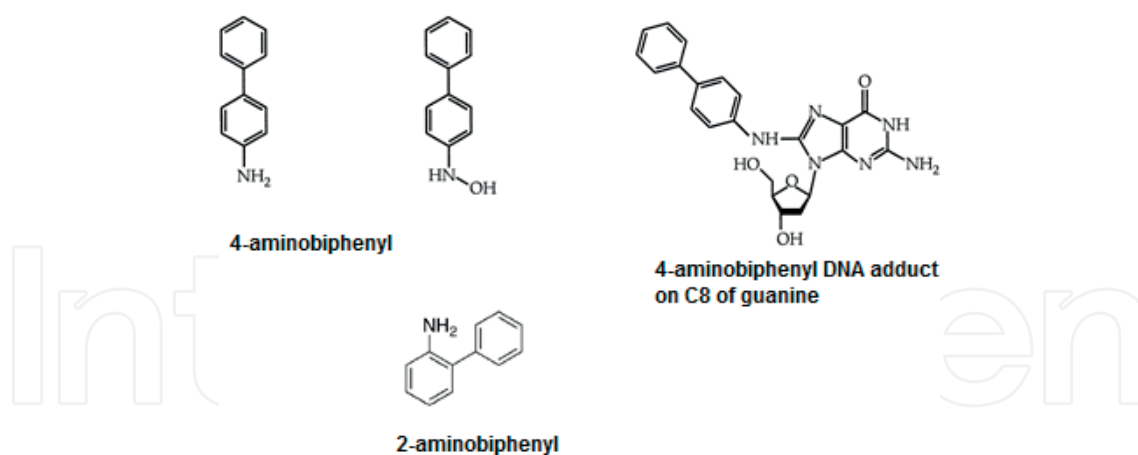
Regarding the polycyclic aromatic substances such as PAHs, dioxins (tetrachlorodibenzo-p-dioxin (TCDD)), aflatoxin B1, etc., planarity in structure is an important element which determines their mutagenic potential. Thus, molecules of planar in structure, with less than four aromatic rings connected in line and molecular size 100–150 Å, are potent mutagens (**Figure 18**).

Two PAHs with six benzene rings may be demonstrated as an example to that rule. Benzo(a)pyrene has six benzene rings which are not all linearly bound. Pentacene also consists of six benzene rings, but they are all in a single line. Alignment of rings gives more planar structure to benzo(a)pyrene over pentacene, which results in benzo(a)pyrene being a potent carcinogen and pentacene an inert molecule in terms of mutagenicity (**Figure 19**).

The second example when difference in planarity significantly alters mutagenic potential of the molecules is 2-aminobiphenyl and 4-aminobiphenyl. They are both products contained in cigarette smoke. However, 4-aminobiphenyl is planar, and amino group can be N-hydroxylated forming the product that attacks nucleophilic centre on C8 of guanine and forms DNA adduct.



**Figure 18.** Example of two molecules with an equal number of benzene rings. However, opposite to pentacene, benzo(a)pyrene has less than four rings linearly aligned and exhibits potent mutagenicity.

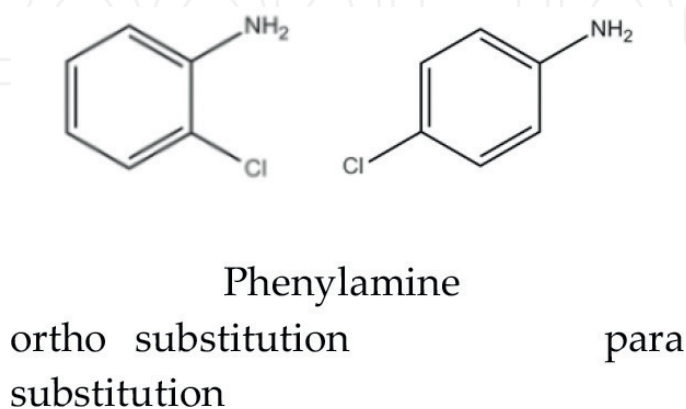


**Figure 19.** Significance of planarity in DNA binding on example of aminophenyls.

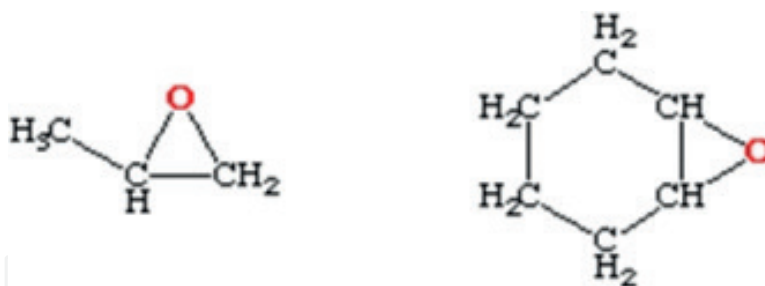
Opposite, 2-aminobiphenyl has lost its planarity, amino group is not achievable for metabolic transformation and the substance remains inactive in terms of carcinogenicity.

The effect of hydrogen substitution may also exert effects on mutagenicity of the substance of interest. It has been proven that substitution of a chloro, methoxy or methyl group in ortho-position to amino group of phenylamine enhances its mutagenic potency (**Figure 20**). It is because ortho-substituted phenylamine forms adducts with DNA that are more efficient in affecting hydrogen bonding with substitute instead of complementary base, thus having higher mutagenic potential.

However, in other cases when mutagenic potency is not determined at the level of already formed DNA adduct and its ability to bind non-complementary nitrogenous base, but earlier at the stage of metabolic transformation, the size of substituent is a critical factor affecting its later activity. For instance, bulky substituents sterically hinder N-hydroxylation or N-acetylation of neighbouring amino group. This may prevent its activation, for instance, later formation of nitrogen radical, as we talked earlier when discussing mechanisms of metabolic activation of electrophilic centres.



**Figure 20.** Ortho- and para-substitution of phenylamine as a cause of differences in mutagenic potency.



**Figure 21.** More reactive epoxide on aliphatic chain and less active one on cycloaliphatic ring.

Flexibility in the structure of molecule strongly affects several potential electrophilic centres that are formed by opening of the triangle structures (e.g. epoxides and sulfonium ions). Cycloaliphatic rings are rigid in comparison to aliphatic moieties. Thus, epoxide formed on the ring will be less active than the one formed on the chain structure which will more easily open (**Figure 21**).

Earlier, when discussing epoxides formed in resin monomers such as BisGMA, TEGDMA and UDMA, we already mentioned that as more potentially electrophilic centres are present in the molecule, the higher carcinogenic potential is (**Figure 11**). Distance between multiple electrophilic centres additionally contributes to reactivity of the substance.

Finally, resonance stabilisation of a metabolically formed electrophilic centre prolongs its half-life. Most electrophiles are readily hydrolysed or neutralised by antioxidant molecules in the cell (e.g. glutathione). Resonance stabilisation remains electrophile reactive providing them better chances to reach genetic material and interact with DNA. It is achieved by the presence of conjugated double bonds, aryl moiety, aromatic rings and structures that allow electrophilic centres to remain silent in cyclic form until they reach the target molecule [7].

## 5. QSAR as the method of choice

Predicting mutagenic and carcinogenic activity is a quite complex task. It demands not only the knowledge of chemical structure of substance of interest together with all functional groups being acknowledged but also its physical and steric properties, as well as metabolic pathways of the substance in organism. Thus, for efficient assessment, a vast database is crucial, which is obtained by assembling all available knowledge gathered by analysing and collecting data obtained for as large number of substances as possible experimentally, empirically, from results of epidemiological studies and case reports.

By using previously described SAR regularities in prediction of carcinogenicity of the substance of interest, quantitative structure–activity relationship (QSAR) was developed as a relevant non-experimental tool. It is considered that the first QSAR, at its simplified basic level, was Mills equation which predicted melting and boiling points of chemicals based on the number of carbon atoms in the chain [13]. Other pioneers in QSAR methodology are Overtone [14] and Mayer [15] who deduced that the toxicity of organic chemicals to

aquatic species is proportional with their partition coefficient. This research directly bound chemical structure and biological activity of chemicals in living organisms. However, Hansch [16] who is considered to be the father of QSAR publishes research showing that a range of biological activities could be modelled mathematically using simple physicochemical properties. Following the entrance of the substance into the organism, the chemical is subjected to the absorption, metabolism, excretion, metabolic changes and transport to nucleus where it can react with nitrogenous bases of DNA. As discussed earlier, predictive modelling studies aiming to exploring all such attributes that affect the activity, property and toxicity of chemicals are bases for developing the database that will be used in QSAR.

How does exactly the QSAR function? QSAR grew out of physical organic chemistry based on studies to show how differential reaction rates of chemical reactions depend on the differences in molecular structure. In classical means QSAR model was developed on the basis of comparison of so-called descriptors from the tested substance of interest with dataset obtained during the years of empirical testing of substances of known structural, steric and physical characteristics [17]. In the contest of QSAR, molecular descriptors are characteristics regarding specific information about a substance of interest [18]. These are its chemical structure and properties, steric properties, physical properties and, for substances already present in database, biological/toxicological activity. For the purpose of QSAR, descriptors from qualitative are transferred to the numerical or quantitative representations of chemical by using suitable algorithms. Thus, QSAR is a simple mathematical model that can correlate chemistry with the properties of substance of interest using various computationally or experimentally derived quantitative parameters known as descriptors [17]. They are used as independent variables for mutagenicity prediction model development [19]. The selection of relevant descriptors is a well-known issue in QSR, and its reliability depends on the quantity of substances that have been evaluated for all screened molecular descriptors and entered the referent database [17].

Regarding the database in relation to which descriptors are evaluated, it gathers the knowledge regarding physicochemical (hydrophobic, steric or electronic), structural (based on frequency of occurrence of a substructure) and molecular structure, relations between functional groups, their mutual influence and influence of entire chemical structure of the molecule, possible pathways and products of their metabolic activation and others (as spoken earlier) that may affect the biological properties of detected functional groups.

Based on above-discussed conditions for electrophilic centre to affect DNA, descriptors may be classified into several groups: substituent constants; whole molecular, topological and structural descriptors, indicator variables; thermodynamical descriptors; and electronic and spatial parameters.

Substituent constants are basically physicochemical descriptors mediated by differences in molecular structure of the molecule. Whole molecular descriptors represent expansions of the substituent constant approach and represent features other than functional groups. For instance, they are lipo-/hydrophilic ratio, dissociation constants, van der Waals volume, etc. Topological descriptors represent the position of the individual atoms and the bonded connections between them. Structural descriptors refer to content of functional groups [19]. Indicator variables are used for comparison of two molecules by utilising all other independent

variables. It can be employed only when the two sets of compounds are identical in every respect. Electronic parameters describe electronic aspects of both the entire molecule and its specific parts, such as atoms, bonds and molecular fragments. Spatial parameters reflect spatial arrangement of the molecules and the surface occupied by the molecules.

Several prerequisites are necessary for model development in classical QSAR: the compounds to be studied should be closely related congeners, thereby increasing the probability of having the same mechanism of action; the biological activity data to be used in modelling should be accurate and measured under uniform conditions; and the activity parameter must be intrinsically additive [20]. Thus, in certain restricted way, it may be indicated that QSAR relies on read-across approach, since it evaluates the substance of interest based accordingly on substances that have been evaluated and for which a structure-action relationship has been identified.

## 6. Conclusion

Several prerequisites are necessary for model development in classical QSAR: the compounds to be studied should be closely related congeners, thereby increasing the probability of having the same mechanism of action; the biological activity data to be used in modelling should be accurate and measured under uniform conditions; and the activity parameter must be intrinsically additive [21]. Thus, in certain restricted way, it may be indicated that QSAR relies on read-across approach, since it evaluates the substance of interest based on according substances that have been evaluated and for which a structure-action relationship has been identified.

## Conflict of interest

There is no conflict of interest.

## Author details

Davor Zeljezic

Address all correspondence to: [dzjeljezi@imi.hr](mailto:dzjeljezi@imi.hr)

Institute for Medical Research and Occupational Health, Zagreb, Croatia

## References

- [1] Frickel S. Working on mutations. In: Chemical Consequences. Environmental Mutagens, Scientist Activism, and The Rise of Genetic Toxicology. New Brunswick, New Jersey: Rutgers University Press; 2004. pp. 21-47. DOI: 10.1172/JCI24868



- [2] Beerenblum I, Schubick P. A new, quantitative, approach to the study of the stages of chemical carcinogenesis in the mouse's skin. *British Journal of Cancer*. 1947;383-391
- [3] Friedberg C, Walker GC, Siede W, Wood RD, Schultz RA, Ellenberger T. Introduction to mutagenesis. In: Errol CF, editor. *DNA Repair and Mutagenesis*. 2nd ed. ASM Press; 2005. pp. 71-74. DOI: 10.1128/9781555816704
- [4] Miller JA. Some current perspectives on chemical carcinogenesis in humans and experimental animals: Presidential address. *Cancer Research*. 1978;38:1479-1496
- [5] Miller EC, Swanson AB, Phillips DH, Fletcher TL, Liem A, Miller JA. Structure-activity studies of the carcinogenicities in the mouse and rat of some naturally occurring and synthetic alkenylbenzene derivatives related to safrole and estragole. *Cancer Research*. 1983;43(3):1124-1134. DOI: Published March 1983
- [6] Tong W, Welsh WJ, Shi L, Fang H, Perkins R. Structure-activity relationship approaches and applications. *Environmental Toxicology and Chemistry*. 2003;22(8):1680-1695. DOI: 10.1897/01-198
- [7] Lai DY, Woo Y. Structure-activity relationship (SAR) and its roles in cancer risk assessment. In: Choy WN, editor. *Genetic Toxicology and Cancer Risk Assessment*. New York, Basel: Marcel Dekker Inc; 2001. pp. 73-92. DOI: 10.1080/089583799196952
- [8] Ingold KU. Peroxy radicals. *Accounts of Chemical Research*. 1969;2(1):1-9. DOI: 10.1021/ar50013a001
- [9] Martini M, Termini J. Peroxy radical oxidation of thymidine. *Chemical Research in Toxicology*. 1997;10(2):234-241. DOI: 10.1021/tx960154l
- [10] Boerleider RZ, Olie JDN, van Eijkeren JCH, Bos PMJ, Hof BGH, de Vries I, Bessems JGM, Meulenbelt J, Hunault CC. Evaluation of three physiologically based pharmacokinetic (PBPK) modeling tools for emergency risk assessment after acute dichloromethane exposure. *Toxicology Letters*. 2015;232(1):1-27. DOI: 10.1016/j.toxlet.2014.10.010
- [11] Casanova M, Deyo DF, Heck H. Dichloromethane (methylene chloride): metabolism to formaldehyde and formation of DNA-protein cross-links in B6C3F1 mice and Syrian golden hamsters. *Toxicology and Applied Pharmacology*. 1992;114(1):162-165
- [12] Kehe K, Balszuweit F, Steinritz D, Thiermann H. Molecular toxicology of sulfur mustard-induced cutaneous inflammation and blistering. *Toxicology*. 2009;263(1):12-19. DOI: 10.1016/j.tox.2009.01.019
- [13] Mills EJ. On melting point and boiling point as related to composition. *Philosophical Magazine Series*. 1884;5(17):17387. DOI: 10.1080/14786448408627502
- [14] Overton E. Ueber die allgemeinen osmotischen Eigenschaften der Zelle, ihre vermutlichen Ursachen und ihre Bedeutung für die Physiologie. *Vierteljahrsschr Naturforsch Gesellschaft von Zurich*. 1899;44:88-135
- [15] Meyer H. On the theory of alcohol narcosis. Part 1. Which property of anaesthetics gives them their narcotic activity? *Naunyn-Schmiedeberg's Archives of Experimental Pathology and Pharmacology*. 1899;42:109-118



- [16] Hansch C, Maloney PP, Fujita T, Muir RM. Correlation of biological activity of phenoxycetic acids with Hammett substituent constants and partition coefficients. *Nature*. 1962;**194**:178-180. DOI: 10.1038/194178b0
- [17] van de Waterbeemd H, Carter RE, Grassy G, Kubinyi H, Martin YC, Tute MS, Willett P. Glossary of terms used in computational drug design (IUPAC recommendations 1997). *Annual Reports on Medical Chemistry*. 1998;**33**:397-409. DOI: 10.1351/pac199769051137
- [18] Roy K, Kar S, Das RN. Chemical information and descriptors. In: *Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment*. Amsterdam, Boston, Heidelberg, London, New York, Oxford, Paris, San Diego, San Francisco, Singapore, Sydney, Tokyo: Elsevier; 2015. pp. 58-78. DOI: 10.1016/B978-0-12-801505-6.00006-5
- [19] Todeschini R, Consonni V. *Handbook of Molecular Descriptors*. Weinheim, Germany: Wiley-VCH; 2000. DOI: 10.1002/9783527613106
- [20] Roy K, Kar S, Das RN. Introduction. In: *Understanding the Basic of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment*. Amsterdam, Boston, Heidelberg, London, New York, Oxford, Paris, San Diego, San Francisco, Singapore, Sydney, Tokyo: Elsevier; 2015. p. 81. DOI: 10.1016/B978-0-12-801505-6.00006-5
- [21] Roy K, Kar S, Das RN. *Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment*. Amsterdam, Boston, Heidelberg, London, New York, Oxford, Paris, San Diego, San Francisco, Singapore, Sydney, Tokyo: Elsevier; 2015. p. 456. DOI: 10.1016/B978-0-12-801505-6.00006-5

IntechOpen