Review

Quantitative Structure-Activity Relationship (QSAR) Studies on the Toxic Effects of Nitroaromatic Compounds (NACs): A Systematic Review

Tao Huang [1], Guohui Sun [1,*] , Lijiao Zhao [1] , Na Zhang [1], Rugang Zhong [1] and Yongzhen Peng [2]

[1] Key Laboratory of Environmental and Viral Oncology, College of Life Science and Chemistry, Faculty of Environment and Life, Beijing University of Technology, Beijing 100124, China; huangtaovincent@gmail.com (T.H.); zhaolijiao@bjut.edu.cn (L.Z.); nanatonglei@bjut.edu.cn (N.Z.); lifesci@bjut.edu.cn (R.Z.)

[2] National Engineering Laboratory for Advanced Municipal Wastewater Treatment and Reuse Technology, College of Environmental and Chemical Engineering, Faculty of Environment and Life, Beijing University of Technology, Beijing 100124, China; pyz@bjut.edu.cn

*   Correspondence: sunguohui@bjut.edu.cn; Tel.: +86-10-6739-1917

Abstract: Nitroaromatic compounds (NACs) are ubiquitous in the environment due to their extensive industrial applications. The recalcitrance of NACs causes their arduous degradation, subsequently bringing about potential threats to human health and environmental safety. The problem of how to effectively predict the toxicity of NACs has drawn public concern over time. Quantitative structure–activity relationship (QSAR) is introduced as a cost-effective tool to quantitatively predict the toxicity of toxicants. Both OECD (Organization for Economic Co-operation and Development) and REACH (Registration, Evaluation and Authorization of Chemicals) legislation have promoted the use of QSAR as it can significantly reduce living animal testing. Although numerous QSAR studies have been conducted to evaluate the toxicity of NACs, systematic reviews related to the QSAR modeling of NACs toxicity are less reported. The purpose of this review is to provide a thorough summary of recent QSAR studies on the toxic effects of NACs according to the corresponding classes of toxic response endpoints.

Keywords: nitroaromatic compounds; in silico modeling; QSAR; animal testing; toxicity

1. Introduction

Nitroaromatic compounds (NACs) are a type of aromatic compound with at least one nitro group (-NO$_2$) located at the benzene ring. They comprehensively exist in the atmospheric, aquatic, and terrestrial environment, as well as human foods. The nitro group on benzene ring delocalizes π-electrons of the ring, satisfying the deficiency of its own charge [1]. This unique structure enables NACs to exhibit various applications. For example, most explosive materials manufactured in the last century were NACs, such as 2,4,6-trinitrotoluene (TNT), 1,3,5-trinitrobezene (TNB), tetryl nitramine, trinitrophenols, hexanitrobezene, etc. [2,3]. In the dye industry and pharmaceutical factories, 2-nitrophenol derivatives are widely used. 2,4-Dinitrotoluenes (DNTs) are used in organic synthesis industries such as polyurethane, dyes, medicine, and rubber. Dinitrophenols (DNPs) have been used in herbicides and pesticides [4]. Table 1 lists the names, sources, and applications of NACs, and their toxic effects observed in different organisms. The main source of NACs in the environment is artificial manufacture (industrial emission, automobile exhaust and domestic sewage) due to the comprehensive application in industry. Apart from anthropogenic activities, some NACs can be produced by certain microbes, parasites or plants as secondary metabolites or pheromones [5]. The psychrotolerant, gram-negative bacterium *Salegentibacter* sp. strain T436, found in the Arctic, can produce 21 monocyclic NACs as secondary metabolites. The fermentation can be optimized to improve their

antimicrobial and cytotoxic ability, of which 2-nitro-4-(2′-nitroethenyl)phenol is the most potent one [6,7]. The electron-withdrawing nitro groups cause NACs to be recalcitrant for degradation [8]. Furthermore, the degradation of NACs is not sustainable and efficacious enough, directly leading to their persistent accumulation in the environment. Hence, the contamination of NACs poses a serious threat to the ecological environment and human health.

Table 1. Names, sources, applications, toxicity, and research objects of several representative nitroaromatic compounds.

Name	Source or Application	Toxicity	Organism	References
TNT	explosive	inhibition of cell growth and cell viability, mutagenic, liver damage, cataract	mammal	[9,10]
DNT	dye, medicine, rubber	central nervous system and respiratory system depression, ataxia, reproductive toxicity	rat	[11]
DNP	dye, pesticide, herbicide	weight loss, lethal, cataract, tumorigenicity	human	[12]
Nitrofuran	antibiotics	tumorigenicity, mutagenicity, carcinogenicity	human	[13]
Mononitrophenol	explosive, dye, pharmaceutical, rubber,	carcinogenicity	human	[14]
1-Nitropyrene	diesel	mutagenicity, carcinogenicity	human, bacteria	[15]
2-Nitronaphthalene	vehicle fuel	DNA damage	human, bacteria	[16]

NACs exhibit various toxic effects towards living organisms, such as mutagenicity, carcinogenicity, liver damage, jaundice, skin irritation, etc. It is reported that nitro polycyclic aromatic hydrocarbons (NPAHs) can be assimilated by plants and accumulated in the food chain, posing a potential danger to human [17]. Certain NACs like arylhydroxylamines, arylamines, azo, and azoxy compounds even display higher toxicity than their parent NACs [18]. The exhaust of diesel includes NPAHs such as 6-nitrochrysene, 1-nitropyrene, 1,6-dinitropyrene, etc., which are tested to be carcinogenic and mutagenic to mammals and human beings [19]. US Environmental Protection Agency (EPA) already listed them as priority pollutants [19,20]. In 2007, Germany identified more than 10 kinds of nitrophenols and nitrobenzoic acids in ground water and leachate; adjacent residents were found to show different symptoms like nausea, vomiting, headache, etc. [21]. Interestingly, a recent study concluded that 3-trifluoromethyl-4-nitrophenol (TFM) destroyed the balance of ATP supply and demand in trout [22]. TNT (including its metabolites) and DNT are manifested to impair the reproduction of crickets and salamander as well [23,24].

In general, there are three paths (physical, chemical and biological methods) to remove NACs in the environment. Physical methods involve absorption, adsorption, extraction, ultrafiltration, photo-oxidation, and volatilization. Among them, adsorption is employed most extensively [5]. Chemical methods include hydrolysis and advanced oxidation processes (AOPs). AOPs utilize oxidants like Fenton reagent, H_2O_2, and metal as photocatalysts to oxidize NACs. In this case, the stability of the aromatic ring is destroyed by radicals generated from oxidants, thus leading NACs to be decomposed into smaller molecules [25]. Biodegradation refers to using microorganisms to decompose NACs, typically including aerobic or anaerobic approach. Even if particular NACs are inclined to aerobic degradation, a number of them are recalcitrant to be oxidized. NACs with high electronegativity are amenable to be reduced first. Besides, it is difficult to decompose NACs completely solely by bacteria, and microbes differ in their ability to degrade NACs. Therefore, the selection of appropriate microorganisms is of vital importance [26]. The major defect of physical methods is that adsorbents usually adsorb organic compounds indiscriminately and such pollutants require further treatment to detoxify. AOPs are costly and microorganisms are inefficient in high concentration scale [5]. Consequently, physico-chemical and biological

methods are integrated to remove NACs in modern industry. In this case, NACs may be further detoxified to meet the environmental standard. However, the economic costs are very high and non-negligible. To avoid these NACs with high toxicity entering the environment, assessing various toxic effects of NACs in advance is necessary. As there is a massive number of NACs, it is impossible to evaluate all the toxic response endpoints through experiments in vitro or in vivo. Consequently, in silico predictive modeling techniques are introduced to supplant these experiments since they can reduce the number of animal testing, satisfy 3R rules (reduction, replacement, refinement), and rapidly predict the toxicity [27].

QSAR means quantitative structure–activity relationship, which is a process that quantitatively bridges molecular descriptors and biological activities. Molecular descriptor is the "fingerprint" of one molecule on the micro level, including theoretical and experimental parameters such as atom numbers; chemical bonds numbers; MLI (molecular connectivity index); ionization constant, pKa; electric dipole moment μ; MR (molecular molar refractive index), logP (lipophilic parameter), etc. With the aid of integrated modeling softwares and well-built mathematical models, the biological properties of specified functional groups can be analyzed and further utilized for chemical design to enhance or decrease the corresponding response endpoints. Common mathematical modeling algorithms include: Multiple Linear Regression (MLR), Genetic Algorithm (GA), Principal Component Analysis (PCA), Partial Least-square method (PLS), Artificial Neural Network (ANN), and Support Vector Machine (SVM). MLR, PCA, and PLS are favorable to process linear relationships, while GA, ANN, and SVM deal with non-linear relationships better [28,29]. Typically, the selection of molecular descriptors determines the quality of models. More descriptors are not meant to be more precise, on the contrary, it may result in overfitting. Accordingly, the selected molecular descriptors must be mutually independent or less correlated.

The results from QSAR are necessary to be verified and evaluated through statistical approaches, which involve internal and external validation metrics. Internal validation commonly employs leave-one-out cross validation (LOO), leave-more-out (LMO) and Y-scrambling method (also called Y-randomization). R^2 is the determination coefficient of the model, which indicates its fitting ability. Q^2 is the cross-validation coefficient of the model, which indicates its robustness and stability. The classical formulations of R^2 and Q^2 are listed as Equations (1) and (2).

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \overline{y})^2} = 1 - \frac{RSS}{TSS} \tag{1}$$

Y_i is the observed values of the response; $\overline{y}$ is the average of all the observed values; $\hat{y}_i$ is the calculated value; *RSS* represents the residual sum of squares; and *TSS* represents the total sum of squares.

$$Q^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \overline{y})^2} = 1 - \frac{PRESS}{TSS} \tag{2}$$

Y_i is the observed values of the response; $\overline{y}$ is the average of all the observed values; $\hat{y}_i$ is the calculated value for each object apart from the training set; and *PRESS* means the predicted residual sum of squares.

The performance of one model is acceptable when it satisfies $R^2 > 0.6$ (for training and test sets) and $Q^2 > 0.5$ (for training set) [30]. A good outcome of internal validation is the prerequisite of external validation. In fact, the actual predictive capability of a QSAR model can only be assessed by datasets that are never used in model development [31–33]. Test set, collected outside modeling data, is applied to verify the predictive veracity of a QSAR model. Generally, the ratio of a training set to test set is approximately set as 3:1. Notably, in general, the number of descriptors in the model equation should be less than one fifth of the compounds in the training set [34]. Extra external datasets will further validate the predictive performance of a QSAR model. As of now, QSAR has been modified and

ameliorated relentlessly with the development of deep learning and artificial intelligence, which guarantees the predictive ability to keep pace with the modern scientific field. Therefore, the pharmaceutical industry commonly uses QSAR methodology to design new drugs, thereby avoiding the introduction of specified toxic substructures. According to the results obtained from QSAR, the toxicity of new chemicals can be predicted, laying a solid foundation on their toxicity classification and risk assessment. Further, compared with traditional living animal testing or in vitro experiments, QSAR tremendously saves cost, time, and human resources. According to Registration, Evaluation, Authorization and Restriction of Chemicals (REACH) regulation in EU, QSAR has been promoted to apply in multiple fields since 2006 [35]. In particular, the principles for QSAR validation recommended by OECD (Organization for Economic Co-operation and Development) for regulatory purpose include: (1) a defined endpoint; (2) an unambiguous algorithm; (3) a defined domain of applicability; (4) appropriate measures of goodness-of fit, robustness, and predictivity; and (5) a mechanistic interpretation, if possible [36]. The prevalence of QSAR minimizes the use of living animals in the experiments of acute systemic toxicity and toxicokinetics [37].

Till today, a number of studies related to the QSAR modeling or applications for NACs toxicity have been reported. However, the information of QSAR studies on the toxic effects of NACs have not been systematically documented, compared, and reviewed. It is of vital importance to elucidate the toxic mechanism of NACs and perform the rapid toxicity prediction for developing greener and safer chemicals. In this review, we aim to summarize recent advances in QSAR studies of NACs according to the corresponding classes of toxic response endpoints.

2. QSAR Studies on Toxic Effects of NACs

2.1. Aquatic Toxicity

2.1.1. QSAR Studies on Aquatic Crustaceans

QSAR studies have been performed to quantitatively evaluate the ecotoxicity of pesticides to aquatic crustaceans, such as *Daphnia magna*. According to the mode of action (MOA), most chemicals are classified as inert (baseline), polar, reactive, and specifically acting compounds [38–40]. It was concluded that the toxicity of inert chemicals exhibits a linear relationship with hydrophobicity, which usually measured by octanol/water partition coefficient ($LogK_{ow}$). In the study performed by Wang et al., a total of 24 molecular descriptors including both physico-chemical properties and quantum chemical parameters were used to develop QSAR models [41]. The modeling dataset comprised 57 pesticides (most of them are NACs, such as Fenitrothion, Parathion, etc.); the result rendered that seven descriptors exhibited prominent correlation to the toxicity, namely, polar surface area (PSA), fraction of ionization (F^+), heat of formation (H_f), polarity (S), hydrogen bonding basicity (B), molecular volume (V), and Cosmo volume (CV). The validation of all QSAR models presented decent results with $R^2 > 0.8$ and $Q^2 > 0.6$. The increase of hydrophobicity contributed to an increase of the toxicity to *D. magna*. Chemical ionization can enhance the toxicity to *D. magna* as well because ionized compounds generally exhibit strong capability of interacting with biological macromolecules. Other descriptors that represent chemical stability and the strength of interaction with biological macromolecules also affect the toxicity. Molecules with strong electronegative substructures like O or N atom may show less toxicity by enhancing their overall polarity [41].

2.1.2. QSAR Studies on Algaes

Among early QSAR studies, Zhao et al., constructed QSAR models to predict the NACs toxicity on *Scenedesmus obliguus* [42]. The toxicity of 26 NACs was demonstrated in modeling, which solely adopted internal validation (without external validation) and showed $R^2 = 0.86$. As a corollary, descriptors such as halfwave reduction potential (E_{12}), bioconcentration factor (BCF), the highest occupied molecular orbital energy (E_{HOMO}), the lowest unoccupied molecular orbital energy (E_{LUMO}), and octanol/water partition

coefficient (K_{OW}) were determined as algae toxicity indicators. In another study, QSAR models established by Schmitt et al., exhibited reliable results in predicting the *Scenedesmus vacuolatus* toxicity of NACs [43]. Sixteen quantum chemistry descriptors along with K_{ow} participated in modeling, yielding an excellent $R^2 = 0.9$. Importantly, the prediction can be further improved by adding another descriptor maximum net atomic charge at the nitro nitrogen ($q_{nitro-N}$). As a result, parameters such as E_{HUMO}, E_{LUMO}, K_{OW}, and distribution coefficient proved to be highly relevant to the algae toxicity of NACs.

In a QSAR study performed by Tugcu and co-workers, the toxicity of nitrophenols on *Chlorella vulgaris* was predicted and strictly evaluated according to OECD principles for QSAR validation [44]. The QSAR models were constructed by MLR and Ordinary Least Square (OLS) method using QSARINS software developed by Gramatica et al., [34]. A series of descriptors such as constitutional index, topological index, connectivity index, ETA index, functional group counts, etc. were totally involved in the model calculation. Individual chemical toxicity to algae is characterized by the average growth rate, which can be expressed as the half maximal inhibitory concentration (IC_{50}). The results showed a sound validation both internally and externally. For training sets, $R^2 > 0.6$, $Q^2_{LOO} > 0.6$, $RMSE < 0.45$; and for test sets, Q^2_{F1}, Q^2_{F2} and $Q^2_{F3} > 0.70$, $r^2_m > 0.65$ and concordance correlation coefficient (CCC) > 0.85. Further, the predictive quality was assessed by mean absolute error (MAE)-based criteria proposed by Roy et al. [45]; model quality derived from Equation (3) was "good". Considering the comprehensive performance, Equation (3) was selected to be further validated using two true external sets. As a result, the observed and predicted values of the true external chemicals were close to each other. Moreover, more than 94% chemicals in these two true external sets fell within the applicability domain (AD) of the model defined by Equation (3).

$$pToxicity = 0.20(\pm 0.17)LogP - 2.75(\pm 1.33)Hardness + 12.70(\pm 6.18) \tag{3}$$

According to previous studies, the octanol-water partition coefficient ($LogP/K_{OW}/logK_{OW}$) is an indispensable factor in chemical toxicity [1,4,24,31,41,44,46–48]. In some cases, though, the determinant factors of toxicity need to be replenished to account for specific phenomena like ionization [49]. In sum, descriptors like hydrophobicity, harness, and electrophilicity as key factors to algae toxicity are commonly identified from the mechanistic analysis of QSAR models.

2.1.3. QSAR Studies on Aquatic Interspecies Toxicity

Interestingly, Tugcu et al., also developed Quantitatively Toxicity-Toxicity Relationship (QTTR) models to predict interspecies toxicity [44,50]. The toxicity order for dinitrophenols to *C. Vulgaris* is *para*-dinitrophenols > *meta*-dinitrophenols > *ortho*-dinitrophenols, which is identical in *T. pyriformis*. Also, the relationship among fish, bacteria, algae, and *Daphinia* toxicities had been proven in previous studies [51,52]. Ciliate–algae and algae–algae interspecies toxicity correlation were calculated through QTTR modeling. The results showed $R^2_{train} = 0.75$, $Q^2_{LOO} = 0.72$, $RMSE_{train} = 0.32$, $R^2_{test} = 0.82$, $RMSE_{test} = 0.28$; $R^2_{train} = 0.93$, $Q^2_{LOO} = 0.91$, $RMSE_{tr} = 0.18$, $R^2_{test} = 0.83$, $RMSE_{test} = 0.27$, respectively. The high values of R^2_{test} indicated a strong toxicity interrelationship within aquatic species. These models may give us insight to predict aquatic toxicity to various species through limited data [44,49].

2.2. Acute Toxicity

2.2.1. QSAR Studies on Fish and Algae

QSAR studies on Carp acute toxicity (LC_{50}) were established using six descriptors, namely, the first order valence molecular connectivity index (X_v), the molecular shape index (K_α), the sum of substituent constant ($\Sigma\sigma-$), E_{LUMO}, $logP$, and indicator variable (I) [53]. Nineteen NACs were involved in modeling and the best prediction was obtained

according to Equation (4), inferring a close relationship in electrophilicity and toxicity of dinitro compounds as *I* (indicator variable) were considered as electron descriptors.

$$-LogLC_{50} = 3.522 + 0.926\Sigma\sigma - +0.173I \quad (N = 19, R = 0.907, S = 0.271) \quad (4)$$

In Yan and co-workers' study, three descriptors K_{OW}, E_{LUMO}, and Q_{NO_2} were employed to establish the QSAR models of acute toxicity (EC_{50}) of algae (*Scenedesmus obliguus*) [54]. Twenty five NACs were divided into two groups: mononitro and dinitro aromatic compounds. For mononitro aromatic compounds, the authors presented a good prediction generated from K_{OW}, which showed $R = 0.904$, indicating a strong connection between hydrophobicity and toxicity. For dinitro aromatic compounds, however, the conclusion was different due to their high electrophilic nature. Thus, the authors selected E_{LUMO} and Q_{NO_2} to predict dinitro aromatic compounds toxicity and showed a reasonable result: $R = 0.926$, $SE = 0.206$. Different toxic mechanisms result in different key parameters in QSAR models, thus presenting inevitable challenges to researchers. The toxicity of mononitro aromatic compounds commonly can be expressed as the narcotic toxicity, which concentrates on the penetration of the cell membrane.

2.2.2. QSAR Studies on Rodents

For rodent acute toxicity, lots of QSAR models have been constructed by using rat or mouse as experimental objects [47,48,55–58]. QSAR models based on the simple representation of molecular structure (SiRMS) method were constructed by Kuz'min et al., to predict the rat oral acute toxicity (LD_{50}) of 28 NACs [59]. In this study, both 1D and 2D descriptors-based models generated decent prediction: $R^2 = 0.96\sim0.98$; $Q^2 = 0.84\sim0.93$; $R^2_{test} = 0.89\sim0.92$. The results showed that hydrophobicity, electrostatic, and Van der Waals interactions contributed to the toxicity. What is more, the addition of fluorine and hydroxyl groups in NACs increases the toxicity, while the insertion of methyl groups decreases the toxicity in most cases. Importantly, the toxicity of substituents is non-additive. The insertion of chlorine in *ortho*-position to the nitro group apparently enhances the toxicity while the second chlorine insertion in *para*-position significantly reduces the toxicity. Furthermore, recently, Hao et al., performed the QSTR modeling of the rat acute oral toxicity of NACs and constructed an interspecies QTTR models between rat and mouse [55]. Seven simple 2D descriptors were utilized to achieve good prediction ($Q^2_{LOO} = 0.7003$, $R^2_{adj} = 0.7292$, $RMSE_{train} = 0.5628$; $Q^2_{F1} = 0.7458$, $Q^2_{F2} = 0.7340$, $Q^2_{F3} = 0.7993$, $R^2_{test} = 0.7593$, $RMSE_{test} = 0.5024$). Specifically, the van der Waals surface area, the presence of C-F at topological distance 6, and high frequency of C-N at topological distance 9 directly affect the acute oral toxicity of NACs to rat. The rat–mouse and mouse–rat interspecies QTTR models also had excellent internal and external predictive performance and thus can be used for interspecies data gap filling [55]. More importantly, the models were used for a true external set containing hundreds of NACs without experimental values; the AD analysis showed that more than 90% compounds fell within the AD of the models. To better comprehend the mechanism of NACs acute toxicity in mammals, the authors also developed QSAR models by using lipophilic descriptor *logP* as the only descriptor, showing a poor correlation between the rat acute toxicity of NACs and *logP* [55]. Unlike aquatic toxicity, *logP* ($logK_{OW}$) has little impact on the rat acute toxicity of NACs. One possible explanation is that the toxic mechanisms of terrestrial mammals are far more intricate than aquatic toxicity since NACs' reactions with organisms in vivo may play more important roles than molecular diffusion.

Using physico-chemical and quantum chemistry descriptors sometimes are not enough to obtain a satisfactory toxicity prediction, thus Mondal et al., established QSAR models by using specific substructures generated by Monte Carlo optimization as descriptors to predict the rat acute oral toxicity of NACs [58]. The best regression-based QSAR model showed a satisfactory validation: $R^2_{train} = 0.719$, $Q^2_{train} = 0.695$; $R^2_{test} = 0.739$, $Q^2_{test} = 0.631$. Some of the substructures are shown in Figure 1. As a result, substructures such as the presence of double-bonded oxygen, sp² carbon with double bond, any hetero atom with

double-bonded oxygen etc. can enhance the acute oral toxicity. Others substructures like the presence of an NH_2 group and an sp^3 carbon attached to the aromatic ring, the presence of sp^3 carbon with branching, the presence of a hetero aromatic ring containing nitrogen, or the presence of oxygen and carbon etc. contribute to the decrease of NACs toxicity. More detailed substructures are listed in the original published paper [58]. Using QSAR incorporated with Monte Carlo approach significantly improved the specificity of toxicological prediction. Meanwhile, the analysis of toxicity–structure relationship towards specific substructures will further facilitate deeper understanding of the toxic mechanism.
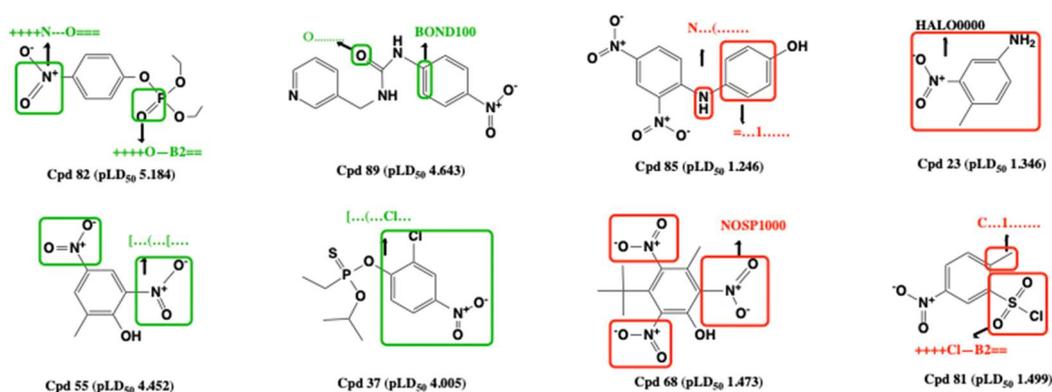

Figure 1. Molecular fragments for toxicity promotors (green) and inhibitors (red).

Other than recent QSAR models based on complex quantum chemistry descriptors, Keshavarz et al., introduced models with several simple constitutional descriptors such as n_{NO_2}, n_S, n_P (number of NO2, sulfur, phosphorous) and two adjustable variables T_{ox}^+ and T_{ox}^- [60]. The two variables quantitatively represent a specific molecular moiety that affects the rat acute oral toxicity of NACs. Although their models were derived from simple parameters, a satisfying correlation was exhibited: $R_{adj}^2 = 0.850$, $Q_{F1}^2 = 0.867$, $Q_{F2}^2 = 0.864$, $Q_{F3}^2 = 0.896$. Compared with using complex quantum chemistry descriptors like other studies mentioned in this review [47,48,51,55,57,58], which usually are calculated from various softwares, the utilization of more interpretable descriptors inspires us with a new and simple approach to comprehend the structure–toxicity relationship of NACs.2.3. QSAR Studies on Mutagenicity and Carcinogenicity

2.3. QSAR Studies on Mutagenicity and Carcinogenicity

Mutation is a normal phenomenon in nature. From the point of view of evolution, as far as the entire biological group is concerned, the evolution of organisms and the emergence of new species are partly attributed to mutations, as well as the crossing-over process. In this case, mutations are advantageous. However, in most cases, mutations are neutral or even harmful and lethal since teratogenicity and carcinogenicity that result from mutation are detrimental. The mechanism of gene mutation is that the sequence of base pairs in DNA changes under the action of mutagen, leading to protein synthesis errors under gene control, which can be manifested as the destruction and loss of enzyme functions or structural functions. As a result, cell genetic characteristics change and subsequent mutations happen [61]. Carcinogenesis is a complex process comprised of two stages: The initiation stage is the cell gene mutation that is triggered by carcinogens; then the growth promotion stage happens, where the mutant cells change the expression of genetic information, causing the mutant cells and cancerous cells to proliferate and become tumors [62]. Generally, it is believed that carcinogenesis is closely related to mutagenesis, and most of the substances that can trigger mutations have carcinogenic effects. In this review, QSAR studies on carcinogenicity and mutagenicity of NACs are discussed in terms of different species.

2.3.1. QSAR Studies on Bacteria

Among most bacteria strains, TA98 and TA100 of *S. typhimurium* strains are normally considered as the most effective candidates for mutagenicity test since they are modified to sensitively detect base-pair mutations [63]. Wang et al., built QSAR models integrated with comparative molecular field analysis (CoMFA) to predict TA98 mutagenicity of NACs [64]. In this study, classical QSAR models based on quantum chemistry descriptors, 3D-QSAR model, and 2D/3D-QSAR joint model were developed, respectively. After validation and comparison towards their predictions, 2D/3D-QSAR joint model exhibited the best performance with R^2 = 0.835 and Q^2_{Loo} = 0.672. It is concluded that the mutagenicity of NACs incorporates molecular transportation, interaction of bio-macromolecules, and nitroreduction. Consequently, the hydrophobicity and E_{LUMO} were considered as the indicators of nitroaromatic mutagenicity. Zhang et al., constructed QSAR models based on PLS to predict the mutagenic activity (MA) of TA98 strain of 1- and 2-nitronaphthalenes (NNs) and methylnitronaphthalenes (MNNs) [65]. The best model validation rendered a plausible outcome: Q^2 = 0.711, R = 0.876, p = 1.863 × 10^{-5}, SE = 0.53. Among 15 quantum chemistry descriptors, ΔH_f, core-core repulsion energy (CCR), $E_{HOMO\text{-}1}$, and E_{LUMO} + E_{HOMO} were demonstrated to correlate mutagenicity significantly, as shown in Equation (5). To be specific, the increase of E_{LUMO} + E_{HOMO} means the increase of electrophilicity, leading to more reactions with cellular nucleic acids. As for the other three parameters, the stability of mutagens decreased with the increase of their values.

$$LogMA \ = \ -10.527 - 2.828E_{HUMO-1} - 0.050\Delta H_f - 0.021CCR + 2.335(E_{LUMO} + E_{HOMO}) \quad (5)$$

In 2019, Hao et al., reported the prediction on the mutagenicity of NACs towards *Salmonella typhimurium* TA100 strain using quantum chemistry descriptor-based QSAR model and machine learning-based classification approaches [47]. They identified one quantum chemistry descriptor (E_{LUMO}) along with four 2D descriptors (Infective-50, Hypnotic-80, CATS2D_04_LL and TIC2) that were associated to the mutagenicity of NACs, of which E_{LUMO} and TIC2 negatively correlated with the mutagenicity, whereas the other three descriptors had a positive correlation. The key statistical parameters were R^2 = 0.961, Q^2_{LOO} = 0.950, R^2_{ext} = 0.836, Q^2_{F1} = 0.808, Q^2_{F2} = 0.808, and Q^2_{F3} = 0.826. It is worthy to note that some specific molecular properties or privileged substructures responsible for the high mutagenicity of NACs were obtained through the classification methods (substructure frequency analysis). Further, novel QSAR models combined with hierarchical support vector regression (HSVR), and PLS methods were developed by Ding et al., to predict TA98 strain mutagenicity of NACs [66]. Various descriptors were selected in different models. PLS model showed R^2 = 0.72, Q^2 = 0.45, $RMSE_{train}$ = 1.02, and $RMSE_{test}$ = 0.97. Apparently, this model was not ideal because several great prediction faults were observed. For HSVR models, the validation showed a satisfactory result: R^2 = 0.9, Q^2 = 0.84, $RMSE_{train}$ = 0.63 and $RMSE_{test}$ = 0.32, exhibiting high robustness and external predictivity. According to the prediction of QSAR models, descriptors including electrophilic index (ω), hydrophobicity, E_{LUMO}, partial atomic charge on the carbon attached to the nitro group (q_{c2}), the sum of molar refractivity of substituents at the *ortho* positions (M_{Ro}), diNO$_2$(I), and dipole moment can account for the complex mutagenicity of NACs—from the preliminary permeation to final base-pair mutations. Compared with previous QSAR models based on PLS, HQSAR, CoMFA, and GFA, HSVR model still outcompetes them, giving a superior stability and prediction. All these descriptors exert various mutagenic actions in different stages; further studies on mutagenic mechanism of NACs and more comprehensive QSAR models are necessary to promote the study on nitroaromatic mutagenesis.

2.3.2. QSAR Studies on Mammals

Numerous QSAR studies on bacteria mutagenicity of NACs have been conducted, whereas little research on human cells has been done. Papa and co-workers constructed QSAR models to predict NPAHs mutagenicity towards human h1A1v2 cells (a line of B-lymphoblastoid cells), which were modified to express an enzyme that can metabolize

PAHs [67]. Initially, 364 descriptors were selected to participate in modeling of 11 NPAHs toxicity. Unlike pure quantitative prediction, Papa et al., adopted *k*-Nearest Neighbour (*k*NN) and Classification and Regression Tree (CART) to classify prediction into two succinct groups: mutagen and non-mutagen. In this case, the validation of their QSAR models subsequently changed somehow. They used sensitivity (SE, percentage of mutagens predicted correctly), accuracy (ACC, total percentage of chemicals predicted correctly), and specificity (SP, percentage of non-mutagens predicted correctly) to quantitatively measure the performance of models. The best *k*NN model showed ACC = 81.6%, SE = 87.1%, SP = 72.2%, while CART model displayed ACC = 77.5%, SE = 87.1%, and SP = 61.1%. The two best models were chosen since they displayed the highest sensitivity in their group respectively. As a result, two vital descriptors, the 2D-topological descriptor 1-path Kier alpha-modified shape index (S1K) and the number of aromatic nitro groups ($nArNO_2$) were proven to have high correlation with cellular mutagenicity.

Given the similarities of physiological structure among mammals, the toxicological investigations in rats usually promote precaution in human health. Morales et al., established QSAR models by multiple regression analysis (MRA) to predict rat carcinogenicity of nitrocompounds [68]. The carcinogenicity was quantitatively determined by TD_{50} values of female rats. After calculation of different QSAR models, the best model was generated after orthogonalization and standardization, as shown in Equation (6). These spectral moment parameters in Equation (6) refer to hydrophobicity, bond dipole moment, Gasteiger-Marsili charge, and molar refractivity. Importantly, the authors further utilized topological substructural molecular design (TOPS-MODE) to predict carcinogenicity based on the previous outcome of QSAR. Therefore, the carcinogenicity of 36 molecular fragments was described as seen in Figure 2. Table 2 quantitatively describes carcinogenicity of individual substructures. Substructures with negative values refer to the fact that they may promote carcinogenicity, while others with positive values mean the reduction of carcinogenicity.

$$\log TD50 = 1.01 + 0.29\Omega^1\mu_9^H - 0.23\Omega^2\mu_6^H - 0.38\Omega^3\mu_{15}^H - 0.22\Omega^4\mu_0\mu_{15}^{GM} + 0.32\Omega^5\mu_1 + 0.16\Omega^6\mu_0 - 0.33\Omega^7\mu_{11}^{MR}$$
$$R^2 = 0.791 \quad Q^2 = 0.666 \tag{6}$$

Table 2. Molecular fragments related to carcinogenic activity.

Studied Fragments	Fragment Contributions	Studied Fragments	Fragment Contributions
F_1	−0.036	F_{19}	+0.188
F_2	−0.420	F_{20}	+0.649
F_3	−0.718	F_{21}	−0.235
F_4	−0.134	F_{22}	+0.581
F_5	−0.260	F_{23}	−0.466
F_6	−0.058	F_{24}	+0.791
F_7	−0.040	F_{25}	−1.390
F_8	−0.243	F_{26}	−2.329
F_9	−0.233	F_{27}	−0.799
F_{10}	−0.321	F_{28}	−0.779
F_{11}	−0.186	F_{29}	−0.661
F_{12}	−0.175	F_{30}	−0.202
F_{13}	−1.112	F_{31}	−0.092
F_{14}	−0.638	F_{32}	−2.770
F_{15}	−0.912	F_{33}	+0.010
F_{16}	−0.666	F_{34}	−0.871
F_{17}	+0.180	F_{35}	−0.605
F_{18}	−1.206	F_{36}	+0.890

Overall, the toxicity of NACs may result from various mode of interaction. One path is attacking the electro-rich position in endogenous biomolecules to exert malfunctions of protein, the other is the further reduction of hydroxylamines with individual toxicity. In addition, single-electron reduction can produce free radical anions, leading to subsequent

oxidative stress and cytotoxicity after redox cycles [43]. Table 3 summarizes the recent QSAR studies on NACs toxicity in accordance with their toxic response endpoints. The mechanism of NACs toxicity in aquatic environment is not complicated compared to other toxic response endpoints. Hence, QSAR models in aquatic toxicity normally can be derived from relatively limited descriptors; even using one descriptor such as $LogK_{OW}/LogP/K_{OW}$ is enough to get a decent prediction in some cases. For mammals, though, it usually requires more descriptors to generate satisfactory results.
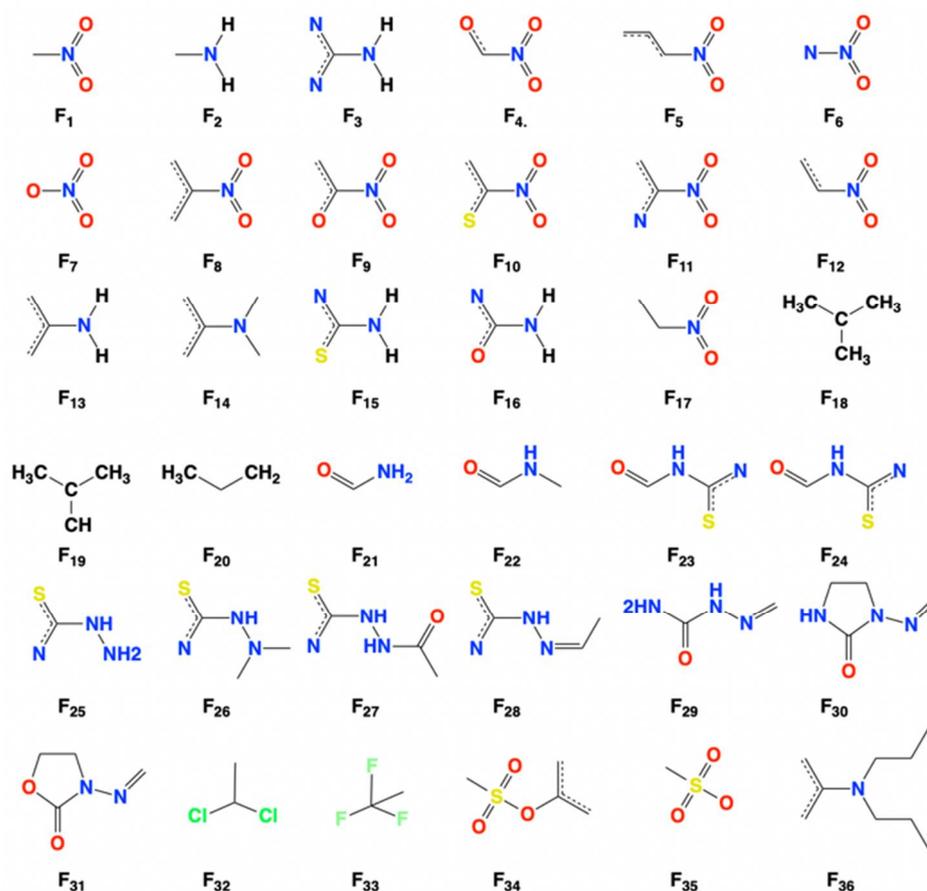

Figure 2. Molecular fragments contributing to carcinogenic activity of nitrocompounds.

Table 3. QSAR studies on the toxicity effects of NACs.

Toxicity Endpoints	Chemicals	Molecular Descriptors	Organisms	References
Aquatic Toxicity	Pesticides	PSA, F^+, H_f, S, B, V, CV	aquatic crustaceans	[41]
	Nitrophenols	Hydrophobicity, harness, electrophilicity	*Chlorella vulgaris*	[44,49]
	26 NACs	E_{12}, BCF, E_{HOMO}, E_{LUMO}, K_{OW}	*Scenedesmus obliguus*	[42]
	19 NACs	E_{HOMO}, E_{LUMO}, K_{OW}, distribution coefficient	*Scenedesmus vacuolatus*	[43]
Acute Toxicity	19 NACs	Xv, $K\alpha$, $\Sigma\sigma-$, E_{LUMO}, $logP$ and I	carp	[53]
	25 NACs	K_{OW}, E_{LUMO} and Q_{NO_2}	*Scenedesmus obliguus*	[54]
	28 NACs	Hydrophobicity, electrostatic, Van der Waals interactions, -F and -OH	rat	[59]
	128 NACs	The van der Waals surface area, the presence of C-F at topological distance 6, and high frequency of C-N at topological distance 9	rat	[55]
	90 NACs	Presence of specific substructures generated from Monte Carlo	rat	[58]

Table 3. *Cont.*

Toxicity Endpoints	Chemicals	Molecular Descriptors	Organisms	References
Mutagenicity & Carcinogenicity	219 NACs	Hydrophobicity and E_{LUMO}	TA98/100	[64]
	48 NACs	E_{HOMO}, Hypnotic-80, Infective-50, TIC2, and CATS2D_04_LL	TA100	[47]
	16 NNs and MNNs	H_f, CCR (core-core repulsion energy), E_{HOMO-1}, $E_{LUMO} + E_{HOMO}$	TA98	[65]
	282 NACs	ω, hydrophobicity, E_{LUMO}, q_{c2}, M_{Ro}, $I(\text{diNO}_2)$ and dipole moment	TA98	[66]
	11 NPAHs	S1K and nArNO$_2$	human h1A1v2 cells	[67]
	48 NACs	Hydrophobicity, bond dipole moment, Gasteiger–Marsili charge, molar refractivity and specific molecular fragments	female rat	[68]

3. Conclusions and Future Scope

In this review, QSAR studies on toxicities of NACs are classified into different genres. For aquatic toxicity, it is concluded that hydrophobicity is one of the pivotal parameters to affect the spread of NACs. This may be because the narcotic toxic effect plays a predominant role in the aquatic environment, thereby hydrophobicity determines the penetration and stimulation of toxicants towards cell membranes. Further, the toxic response endpoints of most NACs are generally close to electron-rich sites due to their electrophilic nature. Thus, E_{LUMO} is always an indispensable indicator to correlate toxicity of NACs; the lower the E_{LUMO} value is, the more electrophilic they are. The toxicity of NACs with more nitro groups generally depends on their electronic reactivity. That is to say, electron-withdrawing substructures like halogen will enhance toxicity. In the reductive environment, NACs are easily reduced to hydroxyl amines or other derivatives and produce free radicals, interfering with normal metabolic activities and syntheses of genetic materials in the cell. As for acute toxicity, numerous studies are conducted based on data of rodents, which further help us understand the corresponding toxicity of NACs towards human. In this case, the complexity of toxicology in mammals results in the various descriptors utilized in QSAR studies. Hydrophobicity is no longer regarded as a key factor to toxicity, instead, distinct quantum chemistry descriptors characterize toxicity to different extents. Besides, specific molecular fragments, generated from Monte Carlo or Gauss approach, are quantitatively correlated to NACs toxicity through QSAR studies, aiding us considerably in the prediction of new chemicals. Researches on the mutagenicity of NACs are imperative since genetic damages are usually irreversible. Therefore, precaution of the toxicants using QSAR models can assure environmental safety and human health. Due to the high sensitivity of TA98/100 strain to genetic mutation, they are selected as the most suitable objects to detect mutagenicity. QSAR studies reveal that dipole moment, hydrophobicity, electrophilicity, etc. contribute to mutations triggered by NACs. The QSAR studies presented above exhibit good performance, though they show some limitations in the types of test objects or applicability. Further studies in QSAR are still encouraged to establish more comprehensive and accurate models to satisfy toxicological prediction of new compounds under the regulatory framework.

Author Contributions: Conceptualization, G.S. and T.H.; resources, L.Z. and N.Z.; writing—original draft preparation, T.H.; writing—review and editing, G.S.; supervision, R.Z. and Y.P.; funding acquisition, G.S. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ju, K.S.; Parales, R.E. Nitroaromatic Compounds, from Synthesis to Biodegradation. *Microbiol. Mol. Biol. Rev.* **2010**, *74*, 250–272. [CrossRef]
2. Akhavan, J. *The Chemistry of Explosives*; The Royal Society of Chemistry: London, UK, 2011.
3. Lee, P.R. Explosives development and fundamentals of explosives technology. In *Explosive Effects and Applications*; Springer: New York, NY, USA, 2002; pp. 23–45.
4. Mulla, S.I.; Bharagava, R.N.; Belhaj, D.; Saratale, G.D.; Bagewadi, Z.K.; Saxena, G.; Kumar, A.; Mohan, H.; Yu, C.-P.; Ninnekar, H.Z. An overview of nitro group-containing compounds and herbicides degradation in microorganisms. In *Microbial Metabolism of Xenobiotic Compounds*; Springer: Singapore, 2019; pp. 319–335.
5. Tiwari, J.; Tarale, P.; Sivanesan, S.; Bafana, A. Environmental persistence, hazard, and mitigation challenges of nitroaromatic compounds. *Environ. Sci. Pollut. Res. Int.* **2019**, *26*, 28650–28667. [CrossRef]
6. Parry, R.; Nishino, S.; Spain, J. Naturally-occurring nitro compounds. *Nat. Prod. Rep.* **2011**, *28*, 152–167. [CrossRef] [PubMed]
7. Schuhmann, I.; Yao, C.B.F.-F.; Al-Zereini, W.; Anke, H.; Helmke, E.; Laatsch, H. Nitro derivatives from the Arctic ice bacterium *Salegentibacter* sp. isolate T436. *J. Antibiot.* **2009**, *62*, 453–460. [CrossRef] [PubMed]
8. Zhang, C.-L.; Yu, Y.-Y.; Fang, Z.; Naraginti, S.; Zhang, Y.; Yong, Y.-C. Recent advances in nitroaromatic pollutants bioreduction by electroactive bacteria. *Process Biochem.* **2018**, *70*, 129–135. [CrossRef]
9. Berthe-Corti, L.; Jacobi, H.; Kleihauer, S.; Witte, I. Cytotoxicity and mutagenicity of a 2,4,6-trinitrotoluene (TNT) and hexogen contaminated soil in S. typhimurium and mammalian cells. *Chemosphere* **1998**, *37*, 209–218. [CrossRef]
10. Bilal, M.; Bagheri, A.R.; Bhatt, P.; Chen, S. Environmental occurrence, toxicity concerns, and remediation of recalcitrant nitroaromatic compounds. *J. Environ. Manag.* **2021**, *291*, 112685. [CrossRef]
11. Rickert, D.E.; Butterworth, B.E.; Popp, J.A.; Krahn, D.F. Dinitrotoluene: Acute toxicity, oncogenicity, genotoxicity, and metabolism. *Crit. Rev. Toxicol.* **1984**, *13*, 217–234. [CrossRef]
12. Grundlingh, J.; Dargan, P.I.; El-Zanfaly, M.; Wood, D.M. 2,4-Dinitrophenol (DNP): A Weight Loss Agent with Significant Acute Toxicity and Risk of Death. *J. Med. Toxicol.* **2011**, *7*, 205. [CrossRef]
13. Vass, M.; Hruska, K.; Franek, M. Nitrofuran antibiotics: A review on the application, prohibition and residual analysis. *Vet. Med.* **2008**, *53*, 469. [CrossRef]
14. Karim, K.; Gupta, S. Biotransformation of nitrophenols in upflow anaerobic sludge blanket reactors. *Bioresour. Technol.* **2001**, *80*, 179–186. [CrossRef]
15. Marshall, T.C.; Royer, R.E.; Li, A.P.; Kusewitt, D.F.; Brooks, A.L. Acute and genetic toxicity of 1-nitropyrene and its fate after single oral doses to rats. *J. Toxicol. Environ. Health Part A* **1982**, *10*, 373–384. [CrossRef]
16. Wang, W.; Jariyasopit, N.; Schrlau, J.; Jia, Y.; Tao, S.; Yu, T.-W.; Dashwood, R.H.; Zhang, W.; Wang, X.; Simonich, S.L.M. Concentration and photochemistry of PAHs, NPAHs, and OPAHs and toxicity of PM2. 5 during the Beijing Olympic Games. *Environ. Sci. Technol.* **2011**, *45*, 6887–6895. [CrossRef]
17. Deng, K.; Wong, T.-Y.; Wang, Y.; Leung, E.M.K.; Chan, W. Combination of Precolumn Nitro-reduction and Ultraperformance Liquid Chromatography with Fluorescence Detection for the Sensitive Quantification of 1-Nitronaphthalene, 2-Nitrofluorene, and 1-Nitropyrene in Meat Products. *J. Agric. Food Chem.* **2015**, *63*, 3161–3167. [CrossRef]
18. Kurian, J.R.; Chin, N.A.; Longlais, B.J.; Hayes, K.L.; Trepanier, L.A. Reductive Detoxification of Arylhydroxylamine Carcinogens by Human NADH Cytochrome b5 Reductase and Cytochrome b5. *Chem. Res. Toxicol* **2006**, *19*, 1366–1373. [CrossRef] [PubMed]
19. Benbrahim-Tallaa, L.; Baan, R.A.; Grosse, Y.; Lauby-Secretan, B.; El Ghissassi, F.; Bouvard, V.; Guha, N.; Loomis, D.; Straif, K. Carcinogenicity of diesel-engine and gasoline-engine exhausts and some nitroarenes. *Lancet Oncol.* **2012**, *13*, 663–664. [CrossRef]
20. Keith, L.; Telliard, W. ES&T Special Report: Priority pollutants: I-a perspective view. *Environ. Sci. Technol.* **1979**, *13*, 416–423.
21. Preiss, A.; Elend, M.; Gerling, S.; Berger-Preiss, E.; Steinbach, K. Identification of highly polar nitroaromatic compounds in leachate and ground water samples from a TNT-contaminated waste site by LC-MS, LC-NMR, and off-line NMR and MS investigations. *Anal. Bioanal. Chem.* **2007**, *389*, 1979–1988. [CrossRef] [PubMed]
22. Birceanu, O.; Sorensen, L.A.; Henry, M.; McClelland, G.B.; Wang, Y.S.; Wilkie, M.P. The effects of the lampricide 3-trifluoromethyl-4-nitrophenol (TFM) on fuel stores and ion balance in a non-target fish, the rainbow trout (Oncorhynchus mykiss). *Comp. Biochem. Physiol. Part C Toxicol. Pharmacol.* **2014**, *160*, 30–41. [CrossRef]
23. Brüning, T.; Thier, R.; Bolt, H.M. Nephrotoxicity and Nephrocarcinogenicity of Dinitrotoluene: New Aspects to be Considered. *Rev. Environ. Health* **2002**, *17*, 163–172. [CrossRef] [PubMed]
24. Karnjanapiboonwong, A.; Zhang, B.; Freitag, C.M.; Dobrovolny, M.; Salice, C.J.; Smith, P.N.; Kendall, R.J.; Anderson, T.A. Reproductive toxicity of nitroaromatics to the cricket, Acheta domesticus. *Sci. Total Environ.* **2009**, *407*, 5046–5049. [CrossRef]
25. Deng, Y.; Zhao, R. Advanced Oxidation Processes (AOPs) in Wastewater Treatment. *Curr. Pollut. Rep.* **2015**, *1*, 167–176. [CrossRef]
26. Min, J.; Wang, J.; Chen, W.; Hu, X. Biodegradation of 2-chloro-4-nitrophenol via a hydroxyquinol pathway by a Gram-negative bacterium, *Cupriavidus* sp. strain CNP-8. *AMB Express* **2018**, *8*, 43. [CrossRef]
27. Liedtke, C.; Luedde, T.; Sauerbruch, T.; Scholten, D.; Streetz, K.; Tacke, F.; Tolba, R.; Trautwein, C.; Trebicka, J.; Weiskirchen, R. Experimental liver fibrosis research: Update on animal models, legal issues and translational aspects. *Fibrog. Tissue Repair* **2013**, *6*, 1–25. [CrossRef]

28. Verma, J.; Khedkar, V.M.; Coutinho, E.C. 3D-QSAR in drug design-a review. *Curr. Top. Med. Chem.* **2010**, *10*, 95–115. [CrossRef]

29. Dudek, A.Z.; Arodz, T.; Gálvez, J. Computational methods in developing quantitative structure-activity relationships (QSAR): A review. *Comb. Chem. High Throughput Screen.* **2006**, *9*, 213–228. [CrossRef]

30. Golbraikh, A.; Tropsha, A. Beware of q2! *J. Mol. Graph. Modell.* **2002**, *20*, 269–276. [CrossRef]

31. Vehtari, A.; Gelman, A.; Gabry, J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat. Comput.* **2017**, *27*, 1413–1432. [CrossRef]

32. Gramatica, P.; Sangion, A. A Historical Excursus on the Statistical Validation Parameters for QSAR Models: A Clarification Concerning Metrics and Terminology. *J. Chem. Inf. Model.* **2016**, *56*, 1127–1131. [CrossRef]

33. Gramatica, P. Principles of QSAR models validation: Internal and external. *QSAR Comb. Sci.* **2007**, *26*, 694–701. [CrossRef]

34. Gramatica, P.; Chirico, N.; Papa, E.; Cassani, S.; Kovarich, S. QSARINS: A new software for the development, analysis, and validation of QSAR MLR models. *J. Comput. Chem.* **2013**, *34*, 2121–2132. [CrossRef]

35. Worth, A.P.; Bassan, A.; De Bruijn, J.; Gallegos Saliner, A.; Netzeva, T.; Pavan, M.; Patlewicz, G.; Tsakovska, I.; Eisenreich, S. The role of the European Chemicals Bureau in promoting the regulatory use of (Q)SAR methods. *SAR QSAR Environ. Res.* **2007**, *18*, 111–125. [CrossRef]

36. OECD. *Environment Health and Safety Publications Series on Testing and Assessment No. 69, Guidance Document on the Validation of (Quantitative) Structure-Activity Relationships [(Q) SAR] Models*; OECD: Paris, France, 2007.

37. Combes, R.; Grindon, C.; Cronin, M.T.D.; Roberts, D.W.; Garrod, J.F. Integrated Decision-tree Testing Strategies for Acute Systemic Toxicity and Toxicokinetics with Respect to the Requirements of the EU REACH Legislation. *Altern. Lab. Anim.* **2008**, *36*, 45–63. [CrossRef]

38. Schultz, T.W.; Cronin, M.T.D.; Walker, J.D.; Aptula, A.O. Quantitative structure–activity relationships (QSARs) in toxicology: A historical perspective. *J. Mol. Struct. THEOCHEM* **2003**, *622*, 1–22. [CrossRef]

39. Verhaar, H.J.M.; van Leeuwen, C.J.; Hermens, J.L.M. Classifying environmental pollutants. *Chemosphere* **1992**, *25*, 471–491. [CrossRef]

40. Wen, Y.; Su, L.; Qin, W.; Zhao, Y.; Madden, J.C.; Steinmetz, F.P.; Cronin, M.T.D. Investigation of Critical Body Residues and Modes of Toxic Action Based on Injection and Aquatic Exposure in Fish. *Water Air Soil Pollut.* **2015**, *226*, 1–11. [CrossRef]

41. Wang, J.; Yang, Y.; Huang, Y.; Zhang, X.; Huang, Y.; Qin, W.C.; Wen, Y.; Zhao, Y.H. Evaluation of modes of action of pesticides to Daphnia magna based on QSAR, excess toxicity and critical body residues. *Ecotoxicol. Environ. Saf.* **2020**, *203*, 111046. [CrossRef]

42. Zhao, Y.H.; Yuan, X.; Ji, G.D.; Sheng, L.X.; Wang, L.S. Quantitative structure-activity relationships of nitroaromatic compounds to four aquatic organisms. *Chemosphere* **1997**, *34*, 1837–1844. [CrossRef]

43. Schmitt, H.; Altenburger, R.; Jastorff, B.; Schüürmann, G. Quantitative Structure−Activity Analysis of the Algae Toxicity of Nitroaromatic Compounds. *Chem. Res. Toxicol.* **2000**, *13*, 441–450. [CrossRef]

44. Tugcu, G.; Ertürk, M.D.; Saçan, M.T. On the aquatic toxicity of substituted phenols to Chlorella vulgaris: QSTR with an extended novel data set and interspecies models. *J. Hazard. Mater.* **2017**, *339*, 122–130. [CrossRef]

45. Roy, K.; Das, R.N.; Ambure, P.; Aher, R.B. Be aware of error measures. Further studies on validation of predictive QSAR models. *Chemom. Intell. Lab. Syst.* **2016**, *152*, 18–33. [CrossRef]

46. Šarlauskas, J.; Polmickaitė-Smirnova, E.; Čėnas, N.; Krikštopaitis, K.; Anusevičius, Ž. The QSAR study for antibacterial activity of structurally diverse nitroaromatics. *Chemija* **2019**, *30*. [CrossRef]

47. Hao, Y.X.; Sun, G.H.; Fan, T.J.; Sun, X.D.; Liu, Y.D.; Zhang, N.; Zhao, L.J.; Zhong, R.G.; Peng, Y.Z. Prediction on the mutagenicity of nitroaromatic compounds using quantum chemistry descriptors based QSAR and machine learning derived classification methods. *Ecotoxicol. Environ. Saf.* **2019**, *186*, 109822. [CrossRef]

48. Fan, T.; Sun, G.; Zhao, L.; Cui, X.; Zhong, R. QSAR and Classification Study on Prediction of Acute Oral Toxicity of N-Nitroso Compounds. *Int. J. Mol. Sci.* **2018**, *19*, 3015. [CrossRef]

49. Ertürk, M.D.; Saçan, M.T. Assessment and modeling of the novel toxicity data set of phenols to Chlorella vulgaris. *Ecotoxicol. Environ. Saf.* **2013**, *90*, 61–68. [CrossRef]

50. Artemenko, A.G.; Muratov, E.N.; Kuz'min, V.E.; Muratov, N.N.; Varlamova, E.V.; Kuz'mina, A.V.; Gorb, L.G.; Golius, A.; Hill, F.C.; Leszczynski, J.; et al. QSAR analysis of the toxicity of nitroaromatics in Tetrahymena pyriformis: Structural factors and possible modes of action. *SAR QSAR Environ. Res.* **2011**, *22*, 575–601. [CrossRef]

51. Cassani, S.; Kovarich, S.; Papa, E.; Roy, P.P.; van der Wal, L.; Gramatica, P. Daphnia and fish toxicity of (benzo)triazoles: Validated QSAR models, and interspecies quantitative activity–activity modelling. *J. Hazard. Mater.* **2013**, *258–259*, 50–60. [CrossRef]

52. Aruoja, V.; Sihtmäe, M.; Dubourguier, H.C.; Kahru, A. Toxicity of 58 substituted anilines and phenols to algae Pseudokirchneriella subcapitata and bacteria Vibrio fischeri: Comparison with published data and QSARs. *Chemosphere* **2011**, *84*, 1310–1320. [CrossRef] [PubMed]

53. Lang, P.Z.; Ma, X.F.; Lu, G.H.; Wang, Y.; Bian, Y. QSAR for the acute toxicity of nitroaromatics to the carp (Cyprinus carpio). *Chemosphere* **1996**, *32*, 1547–1552. [CrossRef]

54. Yan, X.F.; Xiao, H.M.; Gong, X.D.; Ju, X.H. Quantitative structure–activity relationships of nitroaromatics toxicity to the algae (Scenedesmus obliguus). *Chemosphere* **2005**, *59*, 467–471. [CrossRef]

55. Hao, Y.X.; Sun, G.H.; Fan, T.J.; Tang, X.Y.; Zhang, J.; Liu, Y.D.; Zhang, N.; Zhao, L.J.; Zhong, R.G.; Peng, Y.Z. In vivo toxicity of nitroaromatic compounds to rats: QSTR modelling and interspecies toxicity relationship with mouse. *J. Hazard. Mater.* **2020**, *399*, 122981. [CrossRef]

56. Wang, L.L.; Ding, J.J.; Pan, L.; Fu, L.; Tian, J.H.; Cao, D.S.; Jiang, H.; Ding, X.Q. Quantitative structure-toxicity relationship model for acute toxicity of organophosphates via multiple administration routes in rats and mice. *J. Hazard. Mater.* **2021**, *401*, 123724. [CrossRef] [PubMed]

57. Sun, Y.Z.; Li, Z.J.; Yan, X.L.; Wang, L.; Meng, F.H. Study on the quantitative structure–toxicity relationships of benzoic acid derivatives in rats via oral LD 50. *Med. Chem. Res.* **2009**, *18*, 712–724. [CrossRef]

58. Mondal, D.; Ghosh, K.; Baidya, A.T.; Gantait, A.M.; Gayen, S. Identification of structural fingerprints for in vivo toxicity by using Monte Carlo based QSTR modeling of nitroaromatics. *Toxicol. Mech. Methods* **2020**, *30*, 257–265. [CrossRef] [PubMed]

59. Kuz'min, V.E.; Muratov, E.N.; Artemenko, A.G.; Gorb, L.; Qasim, M.; Leszczynski, J. The effects of characteristics of substituents on toxicity of the nitroaromatics: HiT QSAR study. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 747–759. [CrossRef]

60. Keshavarz, M.H.; Akbarzadeh, A.R. A simple approach for assessment of toxicity of nitroaromatic compounds without using complex descriptors and computer codes. *SAR QSAR Environ. Res.* **2019**, *30*, 347–361. [CrossRef] [PubMed]

61. Harvey, R.G. *Polycyclic Aromatic Hydrocarbons: Chemistry and Carcinogenicity*; CUP Archive: Cambridge, UK, 1991.

62. Cogliano, V.; Baan, R.; Straif, K.; Grosse, Y.; Secretan, B.; El Ghissassi, F. Carcinogenicity of human papillomaviruses. *Lancet Oncol.* **2005**, *6*, 204. [CrossRef]

63. Hornberg, J.J.; Laursen, M.; Brenden, N.; Persson, M.; Thougaard, A.V.; Toft, D.B.; Mow, T. Exploratory toxicology as an integrated part of drug discovery. Part II: Screening strategies. *Drug Discov. Today* **2014**, *19*, 1137–1144. [CrossRef]

64. Wang, X.; Lin, Z.; Yin, D.; Liu, S.; Wang, L. 2D/3D-QSAR comparative study on mutagenicity of nitroaromatics. *Sci. China Ser. B Chem.* **2005**, *48*, 246–252. [CrossRef]

65. Zhang, Z.; Niu, J.; Zhi, X. A QSAR model for predicting mutagenicity of nitronaphthalenes and methylnitronaphthalenes. *Bull. Environ. Contam. Toxicol.* **2008**, *81*, 498–502. [CrossRef] [PubMed]

66. Ding, Y.L.; Lyu, Y.C.; Leong, M.K. In silico prediction of the mutagenicity of nitroaromatic compounds using a novel two-QSAR approach. *Toxicol. In Vitro* **2017**, *40*, 102–114. [CrossRef] [PubMed]

67. Papa, E.; Pilutti, P.; Gramatica, P. Prediction of PAH mutagenicity in human cells by QSAR classification. *SAR QSAR Environ. Res.* **2008**, *19*, 115–127. [CrossRef] [PubMed]

68. Morales, A.H.; Pérez, M.Á.C.; Combes, R.D.; González, M.P. Quantitative structure activity relationship for the computational prediction of nitrocompounds carcinogenicity. *Toxicology* **2006**, *220*, 51–62. [CrossRef] [PubMed]