Derivation and Validation of Toxicophores for Mutagenicity Prediction

Jeroen Kazius,[†,‡] Ross McGuire,[‡] and Roberta Bursi*[,‡,§]

Molecular Design & Informatics Department, and Department of Clinical Pharmacology & Kinetics, N.V. Organon, P.O. Box 20, 5340 BH Oss, The Netherlands

Received May 4, 2004

Mutagenicity is one of the numerous adverse properties of a compound that hampers its potential to become a marketable drug. Toxic properties can often be related to chemical structure, more specifically, to particular substructures, which are generally identified as toxicophores. A number of toxicophores have already been identified in the literature. This study aims at increasing the current degree of reliability and accuracy of mutagenicity predictions by identifying novel toxicophores from the application of new criteria for toxicophore rule derivation and validation to a considerably sized mutagenicity dataset. For this purpose, a dataset of 4337 molecular structures with corresponding Ames test data (2401 mutagens and 1936 nonmutagens) was constructed. An initial substructure-search of this dataset showed that most mutagens were detected by applying only eight general toxicophores. From these eight, more specific toxicophores were derived and approved by employing chemical and mechanistic knowledge in combination with statistical criteria. A final set of 29 toxicophores containing new substructures was assembled that could classify the mutagenicity of the investigated dataset with a total classification error of 18%. Furthermore, mutagenicity predictions of an independent validation set of 535 compounds were performed with an error percentage of 15%. Since these error percentages approach the average interlaboratory reproducibility error of Ames tests, which is 15%, it was concluded that these toxicophores can be applied to risk assessment processes and can guide the design of chemical libraries for hit and lead optimization.

Introduction

Screening of drug candidates for mutagenicity is a regulatory requirement for drug approval[1] since mutagenic compounds pose a toxic risk to humans. Mutagenicity is the ability of a compound to cause mutations into DNA. This effect can take place via several and different mechanisms. A compound's reactivity toward DNA can result in the creation of DNA adducts or base deletions,[2] which distort the DNA structure. Nonreactive compounds can be converted into DNA-reactive metabolites through enzyme-catalyzed metabolic activation. DNA distortion can also be caused by intercalation, a process of reversible, noncovalent fixation into the DNA. For instance, compounds with an aromatic polycyclic backbone can intercalate, i.e., insert themselves between and parallel to base pairs of the DNA double helix,[2] thus forming stabilizing π stacking interactions. The distortion of the DNA structure through DNA reactivity and/or intercalation can disrupt enzymatic DNA repair and replication, which increases the chance of erroneous base replacements or deletions or insertions of base pairs,[2] i.e., mutations.

The Ames test is a short-term in vitro assay designed to detect genetic damage caused by chemicals[3−5] and has become the standard test for mutagenicity determinations because it is relatively simple, fast, and inexpensive. Ames tests use a histidine-free medium with an engineered strain of bacteria that can only proliferate into colonies after certain mutations restore their ability to synthesize histidine. A chemical is considered Ames test positive when its addition to the assay causes a significant increase in the number of grown bacterial colonies with respect to a control experiment. In addition, a metabolic activation mixture, which contains liver microsomes, can be added to this test to mimic in vivo metabolism. The term Ames test, however, does not refer to a unique assay, as evidenced by the different standardized experimental methods, bacterial strains, and metabolic activation mixtures that are currently available.[5]

The reproducibility of Ames tests is limited by the purity of the tested chemical, inconsistencies in the interpretation of dose−response curves, interference of further toxic side effects (such as cytotoxicity), variations in the methodology employed, and variations in the materials used (bacterial strains and metabolic activation mixtures). Nevertheless, the average interlaboratory reproducibility of a series of Ames test data from the National Toxicology Program (NTP) was determined to be 85%.[6]

Although the ability of in vitro genotoxicity and mutagenicity tests to predict in vivo toxicity has limits, Ames test results have been applied as predictors for rodent carcinogenicity. It has been established that the predictive power of positive Ames test results for rodent carcinogenicity is high, ranging from 77% to 90%.[5] No other in vitro assay has been reported that better predicts carcinogenicity.[5,7−9]

* Corresponding author: Phone: +31 (0)412 662321. Fax: +31 (0)-412 662506. E-mail: roberta.bursi@organon.com.

† Current Address: Department of Medicinal Chemistry, Leiden-Amsterdam Center for Drug Research, Universiteit Leiden, P.O. Box 9502, 2300 RA Leiden, The Netherlands.

‡ Molecular Design & Informatics Department, N.V. Organon.

§ Current Address: Department of Clinical Pharmacology & Kinetics, N.V. Organon.

10.1021/jm040835a CCC: $30.25
Published on Web 12/17/2004

Following earlier efforts that catalogued mutagens and carcinogens on the basis of their chemical structure,[10,11] Ashby and Tennant[12−15] have derived and categorized a set of structural alerts for DNA reactivity that can identify potentially mutagenic compounds. These knowledge-based studies together with Klopman and Rosenkranz's computational studies[16,17] have established many qualitative structure−toxicity relationships for mutagenicity prediction. More recent literature methods[18−20] and currently available commercial software packages[21−27] also apply substructural descriptors for mutagenicity predictions. The differences between these knowledge-based or statistics-driven in silico tools as well as their individual limitations and capabilities have been discussed in numerous reviews.[28−34]

In general, the current methods have shown acceptable capabilities of predicting mutagenicity depending upon the approach used and particularly the type and the size of the dataset under investigation. While in the initial validation studies error percentages ranging from 15% to 35% were achieved[15−17] for the investigated datasets, in subsequent studies no predictive method was reported to correctly predict mutagenicity with an error percentage smaller than 24%[29,30,35−37] with the sole exception of one dataset for which an error percentage of 19% was obtained.[30] On the other hand, a recent analysis[38] has demonstrated that such methods identified mutagenic pharmaceuticals with a maximal sensitivity of only 52% and pointed to the need of developing additional toxicophores.

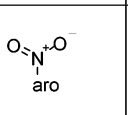
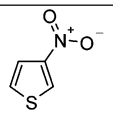
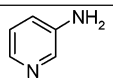
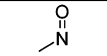
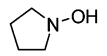
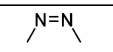
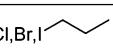
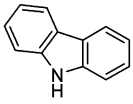
For these reasons, the primary aim of our study was the identification of new structural moieties as toxicophores and—if necessary—the (re)definition of known toxicophores in order to increase the reliability and accuracy of mutagenicity predictions by this approach. A dataset of 4337 compounds with available Ames data was therefore assembled and subsequently analyzed in order to derive new criteria for toxicophore selection and validation. *Here, toxicophores are substructures that indicate an increased potential for mutagenicity, whether this is caused by DNA reactivity or not*. In other words, a toxicophore can represent a reactive substructure or a substructure that is prone to either metabolic activation or intercalation. The novel selection/validation criteria and the available knowledge on the chemistry and the metabolism of specific substructures combined with a statistical analysis based on sensitivity- and *p*-values led to the derivation of 29 toxicophores. No attempt has been made in this work to review or assess the mechanisms of mutagenicity of all individual toxicophores. Rather, several important mechanisms considered representatives for toxicophores by established scientific literature are highlighted.

The results obtained in this study give confidence that these toxicophores can be applied to different phases of the drug optimization process, from supporting early risk and hazard assessments to guiding the design and synthesis of chemical libraries as well as the ranking of compounds classes.

Results and Discussion

General Toxicophores. Eight different substructure representations detected over 70 mutagens with an accuracy of at least 70%. Together, these eight toxicophores detected 75% of all mutagens in the dataset, which was considered satisfactory. Table 1 shows the substructure representation of each general toxicophore as well as an example compound that contains this general toxicophore. Table 2 shows the accuracy and *p*-values of each individual general toxicophore in the presence (columns A) and in the absence (columns B) of compounds that contain different general toxicophores, respectively. The latter values better reflect the actual predictivity of the individual toxicophores for this dataset.

Table 1. Substructure Representations and Example Compounds of General Toxicophores[a]

Toxicophore name	Substructure representation	Example compound
aromatic nitro		
aromatic amine		
three-membered heterocycle		
nitroso		
unsubstituted heteroatom-bonded heteroatom		
azo-type		
aliphatic halide		
polycyclic aromatic system		

[a] "aro" indicates an aromatic atom. "arom. rings" indicates an atom that is part of multiple aromatic rings.

The aromatic nitro and amine groups are well-recognized toxicophores for mutagenicity.[11,17] The nitroso and azo-type groups and the three-member heterocycles are moieties that are similar to existing toxicophores.[11,17] Two other simple substructure representations that detected over 70 mutagens with an accuracy of about 70% are the aliphatic halide group (excluding the fluorine atom)[12−17] and the unsubstituted heteroatom-bonded heteroatom group[11] (a substructure that contains an unsubstituted heteroatom that is attached with a single bond to another heteroatom). Finally, one general toxicophore was represented by large polycyclic aromatic systems, i.e., systems of three

Table 2. Statistics of General Toxicophores

toxicophore	compounds		mutagens		nonmutagens		% accuracy		*p*-value	
	A[a]	B[b]	A	B	A	B	A	B	A	B
aromatic nitro	644	366	562	301	82	65	87	82	≪0.05	≪0.05
aromatic amine	508	288	401	192	107	96	79	67	≪0.05	≪0.05
three-membered heterocycle	233	187	194	152	39	35	83	81	≪0.05	≪0.05
nitroso	122	94	116	88	6	6	95	94	≪0.05	≪0.05
unsubstituted heteroatom-bonded heteroatom	128	101	97	73	31	28	76	72	≪0.05	≪0.05
azo-type	158	95	120	65	38	30	76	68	≪0.05	≪0.05
aliphatic halide	416	330	297	217	119	113	71	66	≪0.05	≪0.05
polycyclic aromatic system	660	321	614	285	46	36	93	89	≪0.05	≪0.05

[a] A: including compounds containing different toxicophores. [b] B: excluding compounds containing different toxicophores).

Table 3. Statistics of Approved and Unapproved Substructures as Specific Toxicophores

toxicophore	compds	mutagens	nonmutagens	% accuracy	*p*-value	specific toxicophore
specific arom nitro	632	561	71	89	≪0.05	approved
specific arom amine	441	380	61	86	≪0.05	approved
aromatic nitroso	32	30	2	94	≪0.05	approved
alkyl nitrite	6	6	0	100	<0.05	approved
nitrosamine	80	77	3	96	≪0.05	approved
epoxide	196	159	37	81	≪0.05	approved
aziridine	33	33	0	100	≪0.05	approved
azide	14	14	0	100	≪0.05	approved
diazo	7	7	0	100	<0.05	approved
triazene	20	19	1	95	≪0.05	approved
aromatic azo	88	67	21	77	≪0.05	approved
aromatic azoxy	9	3	6	33	0.95	
unsubstituted heteroatom-bonded heteroatom	128	97	31	76	≪0.05	approved
hydroperoxide	8	7	1	88	0.06	
oxime	18	10	8	56	0.59	
1,2-disubstituted peroxide	7	1	6	14	1	
1,2-disubstituted aliphatic hydrazine	12	2	10	17	1	
aromatic hydroxylamine	53	45	8	85	≪0.05	approved
aliphatic hydroxylamine	17	10	7	59	0.46	
aromatic hydrazine	10	8	2	80	0.10	
aliphatic hydrazine	16	12	4	75	0.09	
diazohydroxyl	4	4	0	100	0.09	
aliphatic halide	416	297	119	71	≪0.05	approved
carboxylic acid halide	26	23	3	88	≪0.05	approved
nitrogen or sulfur mustard	67	64	3	96	≪0.05	approved
aliphatic monohalide	254	192	62	76	≪0.05	
α-chlorothioalkane	18	15	3	83	<0.05	
β-haloethoxy group	32	29	3	91	≪0.05	
chloroalkene	26	22	4	85	≪0.05	
1-chloroethyl	19	18	1	95	≪0.05	
polyhaloalkene	10	10	0	100	≪0.05	
polyhalocarbonyl	4	4	0	100	0.09	
bay-region in polycyclic aromatic hydrocarbons	125	117	8	94	≪0.05	approved
K-region in polycyclic aromatic hydrocarbons	128	122	6	95	≪0.05	approved
polycyclic aromatic system	660	614	46	93	≪0.05	approved

or more fused aromatic rings, whose corresponding substructure representation consists of one aromatic atom that is connected to at least two atoms belonging to multiple aromatic rings.

Considering the size of the dataset and the chemical diversity herein, the number of general toxicophores that ultimately was identified was small. While for several toxicophores experimental data were abundant, for many other potentially mutagenic substructures data were often scarce (less than 10 compounds).

As Table 2 shows, a large spread was observed in the calculated accuracy of the individual toxicophores, either in the presence (71−95%) or in the absence (66−94%) of compounds with different substructures/toxicophores. In particular, Table 2 (columns B) shows that in this dataset, each general toxicophore is nonredundant because it detects over 60 mutagens undetected by other general toxicophores with an accuracy of at least 66%. Finally, it must be mentioned that the accuracy and *p*-values calculated in this study depend on the dataset composition, which, in turn, depends on the availability of Ames test data. We trust, however, that the large size of this dataset can guarantee the generality of the results here discussed.

Specific Toxicophores. This section provides an overview of the more complex toxicophores that were derived from the general toxicophores and discusses their structural, mechanistic, and statistical aspects. Table 3 shows the accuracy- and *p*-values of the 19 approved as well as of the 16 unapproved toxicophores, while in Figure 2 the substructures of the 19 approved specific toxicophores are displayed.

The mechanism of mutagenicity of the compounds containing the aromatic amine, nitro, nitroso, or hydroxylamine moieties can be explained by partially overlapping metabolic activation pathways.[7,49−52] Although an aromatic nitro group requires enzymatic reduction (catalyzed by both cytosolic and microsomal

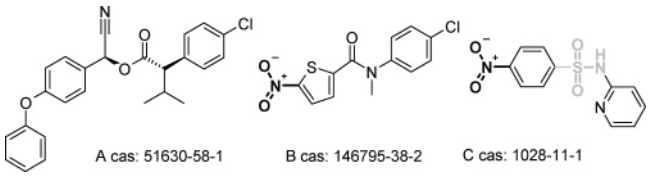


Figure 1. Example compounds for toxicophore-based prediction. Compounds **A** and **C** are nonmutagenic and compound **B** is mutagenic. The aromatic nitro toxicophore is shown in black and bold. The sulfonamide group is a detoxifying substructure for this particular toxicophore and is shown in bold and gray.

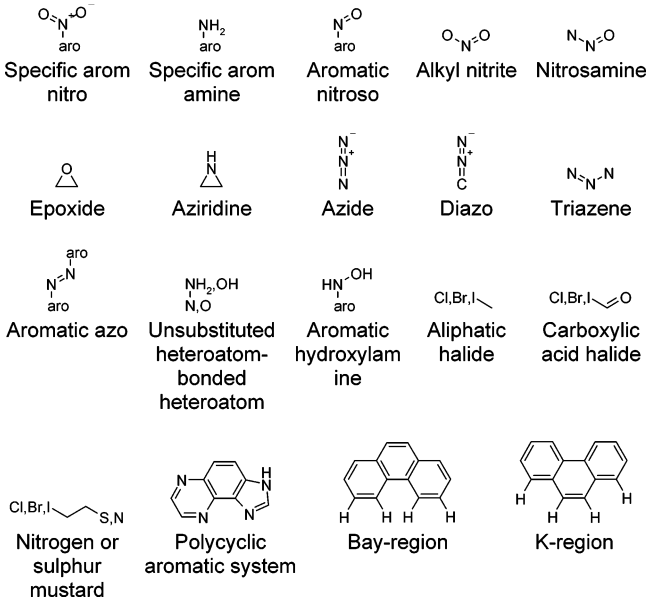


Figure 2. Example substructures of specific toxicophores.

enzymes) to form an aromatic hydroxylamine intermediate,[7,50] the analogous reduction of an aromatic nitroso group is probably nonenzymatic.[51] An aromatic amine moiety, on the other hand, requires enzymatic oxidation to form the same aromatic hydroxylamine intermediate.[52] Subsequent activation of aromatic hydroxylamine intermediates by O-acetylation, O-sulfatation, or O-protonation is suggested to form electrophilic intermediates that covalently bind to DNA.[7,51]

It has been discussed above that as a general toxicophore the aromatic nitro substructure[11−17] detected many dissimilar compounds while it maintained its high accuracy throughout different compound classes and therefore it has high predictive capability. However, the accuracy of this group could be increased by the identification of detoxifying substructures that were present in ortho, meta, and/or para position(s) with respect to this toxicophore. These substructures were the trifluoromethyl, the sulfonamide, the sulfonic acid, and the arylsulfonyl derivatives, and these have been incorporated into the specific aromatic nitro toxicophore. Although these substructures possess a strong electron-withdrawing character, a steric hindrance mechanism to mutagenicity cannot be excluded. For instance, steric hindrance of the aromatic nitro group has already been shown to prevent metabolic activation of this toxicophore. Compounds where a bulky *tert*-butyl substituent was added to the ortho position of the aromatic nitro group showed strongly diminished mutagenicity,[50] while the same addition to both the aromatic nitroso and the hydroxylamine analogues did not alter their mutagenicity.[51] At present, it is unclear whether these detoxifying substructures prevent mutagenicity by inhibiting metabolic activation of the aromatic nitro group through steric hindrance or by disrupting the electronic charge distribution near the toxicophore.

Like the aromatic nitro toxicophore, the general aromatic amine toxicophore[11,17] detected a large variety of compounds and its accuracy was comparable throughout different compound classes. Also in this case, detoxifying substructures located on the same ring as the aromatic amine were identified and incorporated into a specific toxicophore. These substructures were the trifluoromethyl, the sulfonamide, the sulfonic acid, the arylsulfonyl, and the carboxylic acid and its ester derivatives. As expected, many of these substructures were identical to those identified for the aromatic nitro toxicophore. Analogously to that toxicophore, in some cases the metabolic activation of the aromatic amine group has been shown to be primarily prevented by steric hindrance.[52] However, a mechanism of disruption of the charge distribution—required for either metabolic activation or DNA reactivity of the activated aryl hydroxylamine derivative—caused by these groups, the carboxylic acid in particular, cannot be excluded.

The general nitroso toxicophore consisted of several substructures with different reactivity, such as an aromatic nitroso,[51] an alkyl nitrite, and a nitrosamine group.[11−17] Each of these substructures showed high accuracy and was approved as a specific toxicophore and, together, they replaced the general toxicophores. Even though the general toxicophore also included a nitroso group bonded to an aliphatic carbon atom, the latter substructure was not approved, since its *p*-value was larger than 0.05, i.e., the occurrence of this particular group did not correlate significantly with the actual occurrence of mutagenicity. The mutagenic mechanism of action of aromatic nitroso substructures has been described above. The nitrite moiety probably induces chemical reactivity by increasing the electrophilicity of its nearest carbon atom together with its ability to act as a leaving group after protonation of its nitrogen. Nitrosamine groups require enzymatic hydroxylation of their neighboring α-carbon prior to the formation and cleavage of the carbon−diazonium bond, resulting in an electrophilic carbocation.[11,49] No structural properties of nonmutagens containing a nitroso group were found that could explain the absence of mutagenicity.

Epoxides[11−17] and aziridines[11−17] are electrophilic, alkylating substructures that possess significant intrinsic reactivity, and they were the only three-membered heterocyclic substructures that were present in various compounds of the constructed dataset. Epoxides showed an accuracy of 81% for a large number of compounds and aziridines detected mutagens with 100% accuracy. The original general toxicophore was replaced by these two specific toxicophores. No detoxifying structural properties could be identified for either of them.

The general toxicophore representing the azo-type group contained several groups with different reactivity, e.g. an azide, a diazo, and a triazene group as well as an azo[11−17] or an azoxy[11−17] group that are located between two aromatic rings. The azide, diazo, and

triazene groups were approved as specific toxicophores. Their high degree of reactivity is their most plausible cause for mutagenicity. The azoxy group located between two aromatic rings showed low accuracy and was consequently not approved. The azo moiety of diaryl diazenes can be cleaved via enzymatic reduction steps into two corresponding arylamines.[11,49] However, as observed before,[13] those compounds that contained a sulfonic acid group at both ring systems connected to the aromatic azo group were nonmutagenic. Once again, these sulfonic acid groups probably hinder the oxidation of the aromatic amines, as discussed above. After incorporation of this detoxifying substructure, the aromatic azo toxicophore was also approved. This combination of approved specific toxicophores replaced the original general toxicophore.

The general unsubstituted heteroatom-bonded heteroatom[11] toxicophore included aliphatic groups such as the hydroperoxide,[11] the oxime, and the diazohydroxyl groups as well as the hydrazine[11] and the hydroxylamine[11] groups connected to either aromatic ring(s) and/or aliphatic atom(s). Several unsubstituted heteroatom-bonded heteroatom moieties are metabolized via many one-electron oxidation and/or reduction steps into various unstable radicals and other, more stable[53] reactive substructures. Monoalkyl aliphatic hydrazine moieties can be metabolized into unstable, electrophilic diazene, diazohydroxyl, and carbon-diazonium intermediates.[11] An aryl hydrazine group requires one-electron oxidation step to form its arene diazonium derivative,[53] which can react with DNA, as mentioned above.

Of all these groups, only the aromatic hydroxylamine substructure could separately be approved as a specific toxicophore. Other more specific substructures were not approved because of too high *p*-values, which either resulted from relatively low accuracy values (oximes and aliphatic hydroxylamines) or from the detection, although accurate, of only few mutagens (hydrazine derivatives). The unsubstituted heteroatom-bonded heteroatoms toxicophore was clearly a case where the available data did not allow the identification of multiple, specific toxicophores representing distinct mechanisms of mutagenicity. Since this general toxicophore satisfied all four criteria and no detoxifying substructures could be identified, it alone was approved as a specific toxicophore. This toxicophore did not include any unsubstituted heteroatom-bonded heteroatom moiety that contained a sulfur atom. Finally, the 1,2-disubstituted peroxide and the 1,2-disubstituted aliphatic groups, such as the hydrazine groups, which were related to (but not part of) the general toxicophore, could not be approved because of their low accuracy values.

The general aliphatic halide toxicophore[12−17] detected a variety of mutagens containing aliphatic chloride, bromide, and iodide substructures. The carboxylic acid halide group and nitrogen and sulfur mustard groups possessed significantly more intrinsic reactivity than other aliphatic halides, and they were approved as specific toxicophores. Seven halide derivatives with remarkably high accuracy, i.e., the aliphatic monohalide, the α-chlorothioalkane, the *β*-halo ethoxy, the chloroalkene, the 1-chloroethyl, the polyhaloalkene, and the polyhalocarbonyl group, were not approved as specific toxicophores because of lack of scientific literature describing distinct mutagenic mechanisms related to these structures. No detoxifying substructures were identified for aliphatic halide toxicophores. The general toxicophore fulfilled the four criteria and was approved without modifications as specific toxicophore.

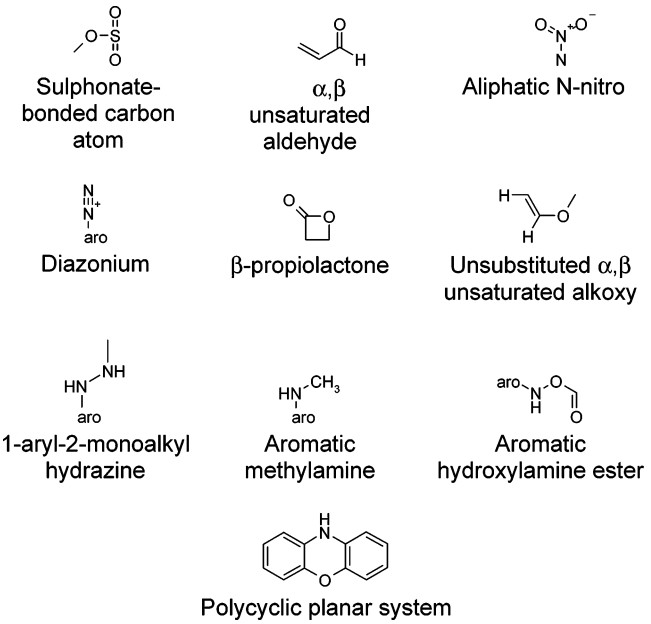


Figure 3. Example substructures of additional toxicophores.

The polycyclic aromatic toxicophore consisted of a system of three or more fused aromatic rings. This toxicophore included polycyclic aromatic hydrocarbons with and without bay- or K-regions,[49] which are displayed in Figure 2, as well as polycyclic heteroaromatic compounds. Many compounds with a polycyclic aromatic system have been reported to intercalate into DNA[2]. In addition, polycyclic aromatic hydrocarbons can become DNA reactive after enzymatic epoxidation at relatively large unsubstituted aromatic regions, such as bay- or K-regions, since diol−epoxide derivatives of polycyclic aromatic hydrocarbons have been shown to form DNA adducts.[49,54] Because the general toxicophore showed an accuracy that was high and comparable to that of the individual specific bay- or K-region toxicophores and because it detected considerably more mutagens, it was approved as a specific toxicophore. As a restriction, only tricyclic aromatic systems that were present in various mutagens of the dataset were included in this specific toxicophore. No further detoxifying substructures or other structural factors were identified, probably because the dataset contained few polycyclic aromatic compounds with large substituents.

Additional Toxicophores. The substructures of the approved additional toxicophores are shown in Figure 3, while their corresponding statistics are shown in Table 4.

The sulfonate-bonded carbon atom,[11−15] the α,*β* unsaturated aldehyde[14] (including the α-carbonyl aldehyde), and the aliphatic *N*-nitro group satisfied all criteria for approval. The sulfonate-bonded carbon atom and the α,*β*-unsaturated aldehyde contain an electrophilic carbon atom, which can explain the DNA reactivity and mutagenicity of these substructures.

Despite factors such as the elevated electrophilicity of its adjacent carbon atom and its capability to act as

Table 4. Statistics of Approved Additional Toxicophores

toxicophore	compds	mutagens	nonmutagens	% accuracy	*p*-value	additional toxicophore
sulfonate-bonded carbon (alkyl alkane sulfonate or dialkyl sulfate)	15	15	0	100	≪0.05	approved
aliphatic *N*-nitro	8	8	0	100	≪0.05	approved
α,β-unsaturated aldehyde (including α-carbonyl aldehyde)	37	30	7	81	≪0.05	approved
diazonium	3	3	0	100	0.17	approved
β-propiolactone	3	3	0	100	0.17	approved
α,β-unsaturated alkoxy group	25	24	1	96	≪0.05	approved
1-aryl-2-monoalkyl hydrazine	6	6	0	100	<0.05	approved
aromatic methylamine	17	15	2	88	≪0.05	approved
ester derivative of aromatic hydroxylamine	4	4	0	100	0.09	approved
polycyclic planar system	839	762	77	91	≪0.05	approved

a precursor to a nitrosamine metabolite, the exact mutagenic mechanism of action of the *N*-nitro group is still unclear.

The diazonium[53] and β-propiolactone[11,13,14,47] substructures detected only few mutagenic compounds and, hence, did not satisfy all criteria. They were nevertheless approved as toxicophores because of their evident intrinsic reactivity.

Most mutagenic aflatoxins contain an unsubstituted α,β-unsaturated alkoxy group in their bisfuranoid substructure, which can be metabolized into an electrophilic epoxide derivative.[55] Since this unsubstituted α,β-unsaturated alkoxy substructure also detected mutagens across different compound classes with high accuracy, it was approved as an additional toxicophore.

The 1-aryl-2-monoalkyl hydrazine,[11] the aromatic methylamine group, as well as the ester derivatives of an aromatic hydroxylamine moiety were also approved. Since the mutagenicity of 1-aryl-2-monoalkyl hydrazine possibly results from metabolic activation into an activated aromatic hydroxylamine derivative, they are all structural derivatives of the previously approved aromatic amine and nitro toxicophores, whose metabolic pathways have been discussed previously.

The search for additional toxicophores showed that the percentage of mutagens among compounds with a polycyclic planar system is only slightly lower than the percentage of mutagens among compounds with a polycyclic aromatic system. This polycyclic planar system toxicophore was defined by three fused five- or six-membered rings that consist of atoms with a connectivity of two or three, i.e., the atoms included have either one π electron or one or more free electron pairs. Compounds containing polycyclic planar systems can act as intercalating agents,[2] while mechanisms of metabolic activation cannot be excluded. Since the polycyclic planar system toxicophore satisfied all criteria, it was also approved as specific toxicophore.

New Toxicophores and Detoxifying Substructures. The primary aim of our analysis was the identification of new structural features that could be linked to mutagenicity. For this purpose, new criteria for toxicophore derivation and validation were developed and applied to a large collection of selected Ames test data. The resulting set of 29 toxicophores was a combination of (i) substructures that are known and well-defined in the literature, (ii) substructures that are common knowledge but which are not defined or are ill-defined in the literature, and (iii) substructures that are new. For instance, the aromatic nitro, aromatic amine, nitrosamine, epoxide, aziridine, aromatic azo, nitrogen or sulfur mustard, α,β-unsatured aldehyde, β-propiolactone, and several aliphatic halide and heteroatom-bonded heteroatom derivatives are examples of well-defined known toxicophores, which in our study were identified and validated in agreement with expert analysis[11−15] and statistics-driven methods.[16,17] On the other hand, the α,β-unsaturated alkoxy toxicophore, the detoxifying substructures for the aromatic nitro and amine groups, and, in particular, the polycyclic planar system toxicophore belong to a novel set of additional toxicophores, which fulfills the aim of our study.

Table 5. Overall Statistics of the Complete Set of Approved Toxicophores

dataset	% error	correct negative	correct positive	false positive	false negative
constructed training set (4337 compds)	18	1539	2019	397	382
external test set (535 compds)	15	163	290	30	52

Mutagenicity Classification and Toxicophore Validation. According to the mutagenicity categorization method that was applied in this study, 54% of the investigated dataset of 4337 compounds was mutagenic (2401 mutagens and 1936 nonmutagens). The mutagenicity classification of these compounds by means of the 29 approved toxicophores resulted in a total classification error of 18%, as shown in Table 5 and Figure 4.

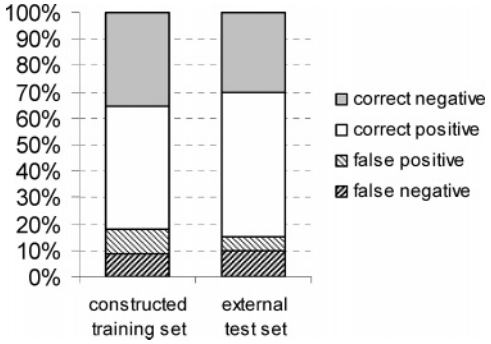


Figure 4. Overall statistics of the complete set of approved toxicophores.

Actual validation was performed by collecting[19] a second, independent dataset of Ames test data generated with standardized protocols of either the National Toxicology Program (NTP)[41] or the Environmental Protection Agency (EPA).[43]

Compounds already in the classification set, inorganic compounds, organometallic compounds, and additional occurrences of enantiomers and diastereoisomers were then removed from the validation set, leading to a final external set of 535 compounds, including 342 mutagens (64%) and 193 nonmutagens (36%). The derived set of 29 toxicophores was subsequently applied to the valida-

tion set, resulting into an error percentage in prediction of 15%, as indicated in Table 5 and Figure 4.

These encouraging results clearly confirmed the usefulness of the toxicophore approach for mutagenicity predictions. The question remained, however, as to whether a better performance of such an in silico tool could be expected. The answer is most likely negative, because of the intrinsic limitations present in both the experimental data and the (sub)structure−toxicity relationship approach. Experimentally, it has been evaluated that because of interlaboratory differences and reproducibility limitations, Ames data are generally affected by a 15% error. Predictions, therefore, with an average accuracy superior to 85% cannot be expected. Methodologically, the toxicophore approach is clearly incapable of including consecutive metabolic activation steps or of taking into account the diversity and the polymorphism of the enzymes responsible for metabolic activation and detoxification. Furthermore, toxicophores fail to detect, for instance, selective noncovalent binding to cellular components that induce error-prone DNA repair, as suggested by Singer.[54] This hypothesis could be supported by the observation that about 200 mutagenic compounds that contained no toxicophore (60% of the false negatives) showed mutagenicity in absence of a metabolic activation mixture.

Despite these limitations, the results obtained in this study support the use of the toxicophore approach as a valuable tool for mutagenicity predictions and highlight the necessity of improving its current degree of reliability and accuracy by identifying additional structural features of the chemical-toxicity space.

Conclusions

In this study, a final set of 29 approved toxicophores was developed from a constructed mutagenicity dataset of 2401 mutagens and 1936 nonmutagens by applying new toxicophore selection/validation criteria and statistics in combination with mechanistic and chemical knowledge. This novel set of approved toxicophores could classify and predict mutagenicity for different datasets with error percentages as low as 18% and 15%, respectively. Since these error percentages approach the average intrinsic error of the assembled dataset (11%) and the experimental error of Ames tests in general (15%), the ability of this set of toxicophores to accurately classify and predict Ames test mutagenicity was confirmed. It was concluded that these approved toxicophores can aid the prediction of mutagenicity in early risk assessment as well as in the design of chemical libraries for hit and lead optimization.

Experimental Section

Mutagenicity Dataset. The Chemical Carcinogenicity Research Information System (CCRIS) database[39] contains scientifically evaluated Ames test data for approximately 7000 compounds and mixtures, which are identified with a CAS registry number and/or chemical name(s). Additional mutagenicity data, although in lower quantities, are available from other public toxicity databases.[40−43] It is noted that these databases also contain data from Ames tests that were performed before strict regulatory requirements were imposed for the authorization of new chemicals. The molecular structures of these compounds were either retrieved from the National Cancer Institute's Developmental Therapeutics Program database[44] and via Beilstein[45] by means of their CAS registry number or constructed from their chemical name(s). Inorganic compounds, organometallic compounds, and additional occurrences of enantiomers and diastereoisomers were then removed from this dataset.

To construct a consistent mutagenicity dataset from the available Ames test data, the following criteria were applied. First, to diminish data heterogeneity and avoid data pollution by nonstandard Ames tests, our analysis was restricted to standard Ames test data of *Salmonella Typhimurium* strains TA98, TA100, TA1535 and either TA1537 or TA97, which are required for regulatory evaluation of drug approval.[1] In addition, strains TA102 and TA1538 were also selected, since they are applied in cases where results of other strains are equivocal or difficult to interpret. Further, Ames tests were only considered if they were performed with the standard plate method or the preincubation method,[3−5] either with or without a metabolic activation mixture. Second, this study required the categorization of each compound as either a mutagen or a nonmutagen, which was based on the available, occasionally conflicting, Ames test results determined in different laboratories. In this study, *a compound was categorized as a mutagen if at least one Ames test result was positive*. Consequently, a false positive Ames test result will erroneously rendering a compound mutagenic, irrespective of the number of negative results. In general, the categorization of a compound as nonmutagenic is sufficiently reliable if at least four Ames tests, performed with different strains, give reproducible negative results.[4] In this study, to assemble a large dataset with maximal compound diversity, *a compound was categorized as a nonmutagen if exclusively negative Ames test results—one or more—were reported*. Further, the robustness of the above mutagenicity categorization of the CCRIS database was tested by applying the same categorization criteria to another set of Ames test results collected from the NTP.[41] The results obtained for approximately 1500 compounds present in both the NTP and the CCRIS databases showed contradicting categorizations in 11% of the cases. Because this error was smaller than 15%, which is the average interlaboratory reproducibility error of Ames tests,[6] the categorization applied in this study was considered satisfactory. To further increase the consistency of the dataset, compounds whose CCRIS data showed contradicting categorizations with the NTP data were removed from the dataset. In conclusion, a dataset of 4337 compounds with corresponding molecular structures and toxicity categorizations (2401 mutagens and 1936 nonmutagens) was constructed.

Derivation of General, Specific, and Additional Toxicophores. This section discusses the definitions of accuracy and *p*-value and the procedures that were applied during the identification of toxicophores for mutagenicity. The equations of accuracy and *p*-value, the SMARTS string[46] representations corresponding to the specific and additional toxicophores, as well as the dataset of 4337 compounds are provided in the Supporting Information.

Definition of Accuracy. The accuracy of a substructure (or toxicophore) is the percentage of experimentally determined mutagens in the subset of compounds containing this substructure.

Definition of *p*-Value. Given a subset of compounds containing a substructure (or toxicophore), the *p*-value is the chance that a random selection of an equal number of compounds from the assembled dataset will have an accuracy that equals or exceeds the accuracy of this substructure.

General Toxicophores. A first step toward the identification of toxicophores was the investigation of whether simple substructures were capable of detecting the majority of the mutagens in the dataset of 4337 compounds. For this purpose, a program was developed that describes every compound from the dataset as a comprehensive series of small substructure representations, each consisting of one atom and its neighboring atoms. The substructure representations accounted for the elemental type of each atom, as well as the aliphatic/aromatic character, the aromatic ring membership count, the larger halide character (for Cl, Br, and I), and the heteroatom

character (small for O and N, large for P and S). Then, for each available substructure representation, the accuracy and the *p*-value were computed from the corresponding mutagenicity data. Subsequently, only substructure representations that detected 70 mutagens with an accuracy of at least 70% were selected. Where substructures of similar overall character and performance were obtained, the definitions using the smallest number of atoms were chosen; where identical numbers of atoms resulted, the substructure with highest accuracy was used. The substructures fulfilling these criteria were named *general toxicophores*.

Specific Toxicophores. The second step consisted of improving the specificity of these simple, general toxicophores by increasing their structural complexity. General toxicophores were used to organize the data set into different subsets. Each of these subsets was then separately analyzed to derive specific toxicophores. A substructure was *approved as a specific toxicophore*, i.e., it was considered capable of predicting mutagenicity from molecular structure, if and only if it simultaneously satisfied the four criteria given below:

(1) The substructure must have a sufficient degree of either intrinsic reactivity or chemical similarity with an existing, knowledge-based toxicophore, or the substructure must be reported for several compounds as a critical component of a mechanism of action that leads to mutagenicity. In cases where new structural features (either more general or more specific) were identified for a given toxicophore, mechanisms of toxicity or detoxification were discussed.

(2) The substructure must be a toxicophore in at least three chemically different compound classes. This requirement is based on the assumption fundamental to this approach that the mutagenic character of a given substructure is generally conserved throughout chemically diverse classes.

(3) The accuracy of the substructure must be at least 70%. If compounds contain different substructures/toxicophores, then these substructures will identify the same compounds. These cases are examples of substructure intercorrelation. When this occurs, the computed accuracy of each individual substructure in a given compound does not reflect its unique ability to predict mutagenicity. To avoid this, the accuracy of the toxicophore candidates was also calculated by eliminating the compounds containing different substructures/toxicophores. In these cases, an accuracy of at least 60% was required.

(4) The substructure's *p*-value must be smaller than 0.05. This means that the chance that a random selection of an equal number of compounds will contain at least an equal number of mutagens needs to be smaller than 5%.

Once these criteria were defined and specific toxicophores were approved, each general toxicophore was reconsidered. The original general toxicophore was replaced by its derived, approved specific toxicophores if they (i) were considered to better distinguish between different mechanisms of mutagenicity, (ii) showed a higher accuracy, and (iii) detected most of the mutagens that were identified by the general toxicophore. If the original general toxicophore could not be replaced by specific toxicophores and it satisfied the four criteria itself, it was approved as a specific toxicophore, even if it covered different mechanisms of mutagenicity.

While searching for toxicophores, *detoxifying substructures* were also identified. Early examples of such substructures have been reported by Ashby and Tennant.[13] Detoxifying substructures can make toxicophore-containing compounds nonmutagenic because of their inhibiting action upon mechanisms such as metabolic activation, DNA reactivity, or intercalation. This effect may be caused by, for example, steric hindrance or by a disruption of the required electronic charge distribution near the toxicophore.

If one or more detoxifying substructure(s) was/were identified in a given general toxicophore, this general toxicophore was replaced by a specific toxicophore that also incorporated the detoxifying substructure(s). This new "toxicophore" needed to satisfy all four criteria in order to be approved like any other specific toxicophore. The aromatic nitro and the aromatic amine toxicophores are specific examples of how toxicophore accuracy could be improved by the introduction of detoxifying substructures.

In summary, the prediction of mutagenicity was performed as follows:

(1) If a compound did not contain any toxicophore, like compound A in Figure 1, it was classified as a nonmutagen.

(2) If a compound contained a toxicophore, like the aromatic nitro group in compound B (see Figure 1), it was classified as a mutagen.

(3) If a compound contained not only a toxicophore but also one of its detoxifying substructures, like the aromatic nitro group and the sulfonamide group in compound C (see Figure 1), it was classified as a nonmutagen.

Additional Toxicophores. About 600 mutagenic compounds of the investigated dataset did not contain any general toxicophore. These compounds were further analyzed to determine whether additional, less common toxicophores could be identified and approved. To support this search, structural derivatives of specific toxicophores as well as potentially reactive functional groups that interfere with high-throughput screening results[47,48] were investigated. While specific toxicophores were approved if all four criteria were simultaneously met, in this analysis a few substructures were approved as additional toxicophores, even though the shortage of available data prohibited the *p*-value criterion and/or the compound class criterion from being fulfilled.

Acknowledgment. The authors wish to thank the reviewers for their helpful comments. Further, the authors wish to thank Dr. Stan S. Young (GlaxoSmithKline, USA) for providing his EPA and NTP dataset, Dr. Tjong Yih (NV Organon, The Netherlands) and Dr. Ad IJzerman (Leiden University, The Netherlands) for their contributions, as well as Dr. Jaap van der Louw (NV Organon, The Netherlands) for his assistance during the study.

Supporting Information Available: The definitions and the equations of the accuracy and *p*-value tests; SMARTS string representations of the specific toxicophores, the unapproved substructures, and the additional toxicophores, including the polycyclic aromatic and polycyclic planar system toxicophores. This material is available free of charge via the Internet at http://pubs.acs.org. The constructed dataset of 4337 molecular structures with mutagenicity categorizations is available from the author upon request.

References

(1) International conference on Harmonisation: Guidance on specific aspects of regulatory genotoxicity tests for pharmaceuticals. *Fed. Regist.* **1996**, *61* (*80*), 18197−18202.
(2) Garret, L. H.; Grisham, C. M. *Biochemistry;* Saunders College Publishing: Orlando, FL, 1995; pp 225, 929−932.
(3) Ames, B. N.; McCann, H.; Yamasaki, E. Methods for detecting carcinogens and mutagens with the Salmonella/mammalian-microsome mutagenicity test. *Mutat. Res*. **1975**, *31*, 347−364.
(4) Maron, D. M.; Ames, B. Revised methods for the Salmonella mutagenicity test. *Mutat. Res.* **1983**, *113*, 173-215.
(5) Mortelmans, K.; Zeiger, E. The Ames Salmonella/microsome mutagenicity assay. *Mutat. Res.* **2000**, *455*, 29−60.
(6) Piegorsch, W. W.; Zeiger, E. Measuring intra-assay agreement for the Ames Salmonella assay. Lecture notes in Medical Informatics; Springer-Verlag: Heidelberg, 1991; pp 35−41.
(7) Benigni, R.; Giuliani, A.; Franke, R.; Gruska, A. Quantitative structure−activity relationships of mutagenic and carcinogenic aromatic amines. *Chem. Rev.* **2000**, *100*, 3697−3714.
(8) Kim, B. S.; Margolin, B. H. Prediction of rodent carcinogenicity utilizing a battery of in vitro and in vivo genotoxicity tests. *Environ. Mol. Mutagen.* **1999**, *34*, 297−304.
(9) Zeiger, E.; Haseman, J. K.; Shelby, M. D.; Margolin, B. H.; Tennant, R. W. Evaluation of four in vitro genetic toxicity tests for predicting rodent carcinogenicity: Confirmation of earlier results with 41 additional chemicals. *Environ. Mol. Mutagen.* **1990**, *16*, 1−14.

(10) Miller, J. A.; Miller, E. C. Ultimate chemical carcinogen as reactive mutagenic electrophiles. In *Origin of Human Cancers;* Hiatt, H. H., Watson, J. D., Winsten, J. A., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1977; pp 605−627.
(11) Fishbein, L. *Potential industrial carcinogens and mutagens;* Elsevier Scientific Publishing Company: Amsterdam, The Netherlands, 1979.
(12) Ashby, J. Fundamental structural alerts to potential carcinogenicity or noncarcinogenicity. *Environ. Mutagen.* **1985**, *7*, 919−921.
(13) Ashby, J.; Tennant, R. W. Chemical structure, Salmonella mutagenicity and extent of carcinogenicity as indicators of genotoxic carcinogenesis among 222 chemicals tested in rodents by the U.S. NCI/NTP. *Mutat. Res.* **1988**, *204*, 17−115.
(14) Tennant, R. W.; Ashby, J. Classification according to chemical structure, mutagenicity to Salmonella and level of carcinogenicity of a further 39 chemicals tested for carcinogenicity by the U.S. National Toxicology Program. *Mutat. Res.* **1991**, *257*, 209−227.
(15) Ashby, J.; Tennant, R. W. Definitive relationships among chemical structure, carcinogenicity and mutagenicity for 301 chemicals tested by the U.S. NTP. *Mutat. Res*. **1991**, *257*, 229−306.
(16) Klopman, G.; Frierson, M. R.; Rosenkranz, H. S. The structural basis of the mutagenicity of chemicals in *Salmonella typhimurium*: the Gene-Tox Data Base. *Mutat. Res.* **1990**, *228*, 1−50.
(17) Rosenkranz, H. S.; Klopman, G. The structural basis of the mutagenicity of chemicals in *Salmonella typhimurium*: The National Toxicology Program Data Base. *Mutat. Res.* **1990**, *228*, 51−80.
(18) Sello, G.; Sala, L.; Benfenati, E. Predicting toxicity: A mechanism of action model of chemical mutagenicity. *Mutat. Res*. **2001**, *479*, 141−171.
(19) Young, S. S.; Gombar, V. K.; Emptage, M. R.; Cariello, N. F.; Lambert, C. Mixture-deconvolution and analysis of Ames mutagenicity data. *Chem. Int. Lab. Sys.* **2002**, *60*, 5−11.
(20) Llorens, O.; Perez, J. J.; Villar, H. O. Toward the design of chemical libraries for mass screening biased against mutagenic compounds. *J. Med. Chem.* **2001**, *44*, 2793−2804.
(21) Sanderson, D. M.; Earnshaw, C. G. Computer prediction of possible toxic action from chemical structure; the DEREK system. *Hum. Exp. Toxicol.* **1991**, *10*, 261−273.
(22) Ridings, J. E.; Barratt, M. D.; Cary, R.; Earnshaw, C. G.; Eggington, C. E.; et al. Computer prediction of possible toxic action from chemical structure: an update on the DEREK system. *Toxicology* **1996**, *106*, 267−279.
(23) Smithing, M. P.; Darvas, F. Hazardexpert: An expert system for predicting chemical toxicity. *Food Safety Assessment;* American Chemical Society: Washington, DC, 1992; pp 192−200.
(24) Woo, Y.-T.; Lai, D.; Argus, M.; Arcos, J. Development of structure−activity relationship rules for predicting carcinogenic potential of chemicals. *Toxicol. Lett.* **1995**, *79*, 219−228.
(25) Klopman, G. Multi-CASE: A hierarchical computer automated structure evaluation program. *QSAR* **1992**, *11*, 172−184.
(26) Enslein, K.; Gombar, V. K.; Blake, B. W. International Commission for Protection Against Environmental Mutagens and Carcinogens. Use of SAR in computer-assisted prediction of carcinogenicity and mutagenicity of chemicals by the TOPKAT program. *Mutat. Res.* **1994**, *305*, 47−61.
(27) Bacha, P. A.; Gruver, H. S.; Den Hartog, B. K.; Tamura, S. Y.; Nutt, R. F. Rule extraction from a mutagenicity data set using adaptively grown phylogenetic-like trees. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1104−1111.
(28) Dearden, J. C.; Barratt, M. D.; Benigni, R.; Bristol, D. W.; Combes, R. D.; et al. The Development and Validation of Expert Systems for Predicting Toxicity; ATLA, 1997; pp 223−252. Available at: http://altweb.jhsph.edu/publications/ECVAM/ecvam24.htm.
(29) Greene, N. Computer systems for the prediction of toxicity: An update. *Adv. Drug Deliv. Rev.* **2002**, *54*, 417−431.
(30) Pearl, G. M.; Livingston-Carr, S.; Durham, S. K. Integration of computational analysis as a sentinel tool in toxicological assessments. *Curr. Top. Med. Chem.* **2001**, *1*, 247−255.
(31) Johnson, D. E.; Wolfgang, G. H. Predicting human safety: Screening and computational approaches. *Drug Discovery Today* **2000**, *5*, 445−454.
(32) Benfenati, E.; Gini, G. Computational predictive programs (expert systems) in toxicology. *Toxicology* **1997**, *119*, 213−225.
(33) Cronin, M. T.; Jaworska, J. S.; Walker, J. D.; Comber, M. H.; Watts, C. D.; et al. Use of QSARs in international decision-making frameworks to predict health effects of chemical substances. *Environ. Health Perspect.* **2003**, *111*, 1391−1401.
(34) Benigni, R.; Richard, A. M. Quantitative structure-based modeling applied to characterization and prediction of chemical toxicity. *Methods* **1998**, *14*, 264−276.
(35) Zeiger, E.; Ashby, J.; Bakale, G.; Enslein, K.; Klopman, G. et al. Prediction of Salmonella mutagenicity. *Mutagenesis* **1996**, *11*, 471−484.
(36) White, A. C.; Mueller, R. A.; Gallavan, R. H.; Aaron, S.; Wilson, A. G. A multiple in silico program approach for the prediction of mutagenicity from chemical structure. *Mutat. Res.* **2003**, *539*, 77−89.
(37) Cariello, N. F.; Wilson, J. D.; Britt, B. H.; Wedd, D. J.; Burlinson, B.; et al. Comparison of the computer programs DEREK and TOPKAT to predict bacterial mutagenicity. Deductive Estimate of Risk from Existing Knowledge. Toxicity Prediction by Computer Assisted Technology. *Mutagenesis* **2002**, *17*, 321−329.
(38) Snyder, R. D.; Pearl, G. S.; Mandakas, G.; Choy, W. N.; Goodsaid, F.; et al. Assessment of the sensitivity of the computational programs DEREK, TOPKAT, and MCASE in the prediction of the genotoxicity of pharmaceutical molecules. *Environ. Mol. Mutagen.* **2004**, *43*, 143−158.
(39) Chemical Carcinogenesis Research Information System is available through TOXNET at http://toxnet.nlm.nih.gov.
(40) Gene-Tox is available through TOXNET at http://toxnet.nlm.nih.gov.
(41) National Toxicology Program, at http://ntp-server.niehs.nih.gov.
(42) Carcinogenic Potency Database, at http://potency.berkeley.edu/CPDB-tab.html.
(43) EPA/IARC Genetic Activity Profile database, at http://www.epa.gov.
(44) Developmental Therapeutics Program, at dtp.nci.nih.gov/webdata.html.
(45) Beilstein CrossFire Commander, v 4.0, MDL Information Systems, Inc., San Leandro, CA.
(46) Daylight Chemical Information, Inc., Santa Fe, NM, at www.daylight.com/dayhtml/doc/theory/theory.smarts.html.
(47) Hann, M.; Hudson, B.; Lewell, X.; Lifely, R.; Miller, L.; et al. Strategic pooling of compounds for high-throughput screening. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 897−902.
(48) Rishton, G. M. Reactive compounds and in vitro false positives in HTS. *Drug Discovery Today* **1997**, *2*, 382−384.
(49) Testa, B. *Biochemistry of Redox Reactions (the Metabolism of Drugs and Other Xenobiotics);* Academic Press: New York, 1995.
(50) Glende, C.; Schmitt, H.; Erdinger, E.; Engelhardt, V.; Boche, G. Transformation of mutagenic aromatic amines in to nonmutagenic by alkyl substituents. Part I. Alkylation ortho to the amino function. *Mutat. Res.* **2001**, *498*, 19−37.
(51) Haack, T.; Erdinger, L.; Boche, G. Mutagenicity in *Salmonella typhimurium* TA98 and TA100 of nitroso and respective hydroxylamine compounds. *Mutat. Res.* **2001**, *491*, 183−193.
(52) Klein, M.; Voigtmann, U.; Haack, T.; Erdinger, L.; Boche, G. From mutagenic to nonmutagenic nitroarenes: effect of bulky alkyl substituents on the mutagenic activity of 4-nitrobiphenyl in *Salmonella typhimurium*. Part I. Substituents ortho to the nitro group and in 2′-position. *Mutat. Res.* **2000**, *467*, 55−68.
(53) Powell, J. H.; Gannett, P. M. Mechanisms of carcinogenicity of aryl hydrazines, aryl hydrazides, and arenediazonium ions. *J. Environ. Pathol. Toxicol. Oncol.* **2002**, *21*, 1−31.
(54) Singer, B.; Kusmierek, J. T. Chemical mutagenesis. *Annu. Rev. Biochem.* **1982**, *51*, 655−693.
(55) Guengerich, F. P.; Kim, M. S.; Muller, M.; Lowe, L. G. Chemical mechanisms of formation of DNA-carcinogen adducts, elucidation of potential of adducts for mutagenicity, and mechanisms of polymerase fidelity and mutation in the presence of adducts. *Recent Results Cancer Res.* **1997**, *143*, 449−63.

JM040835A