Integration of Structure−Activity Relationship and Artificial Intelligence Systems To Improve in Silico Prediction of Ames Test Mutagenicity

Paolo Mazzatorta,*,† Liên-Anh Tran,† Benoît Schilter,† and Martin Grigorov‡

Nestlé Research Center, Quality and Safety Department and BioAnalytical Science Department, P.O. Box 44, Vers-chez-les-Blanc, CH-1000 Lausanne 26, Switzerland

Received September 21, 2006

The Ames mutagenicity test in *Salmonella typhimurium* is a bacterial short-term in vitro assay aimed at detecting the mutagenicity caused by chemicals. Mutagenicity is considered as an early alert for carcinogenicity. After a number of decades, several (Q)SAR studies on this endpoint yielded enough evidence to make feasible the construction of reliable computational models for prediction of mutagenicity from the molecular structure of chemicals. In this study, we propose a combination of a fragment-based SAR model and an inductive database. The hybrid system was developed using a collection of 4337 chemicals (2401 mutagens and 1936 nonmutagens) and tested using 753 independent compounds (437 mutagens and 316 nonmutagens). The overall error of this system on the external test set compounds is 15% (sensitivity = 15%, specificity = 15%), which is quantitatively similar to the experimental error of Ames test data (average interlaboratory reproducibility determined by the National Toxicology Program). Moreover, each single prediction is provided with a specific confidence level. The results obtained give confidence that this system can be applied to support early and rapid evaluation of the level of mutagenicity concern.

INTRODUCTION

The aim of this investigation was the improvement of the current in silico methodologies for the prediction of mutagenicity. The term mutation refers to the induction of permanent transmissible changes in the amount or structure of the genetic material of cells or organisms. Since exposure to mutagenic chemicals carries the risk of inducing germline mutations with the possibility of inherited disorders and the risk of somatic mutations including those leading to cancer, testing mutagenic potential is considered an essential step in the safety evaluation of chemicals. Currently, mutagenicity is assessed through the application of a battery of experimental in vitro and in vivo test systems. In the food context, this approach has proved suitable to support the safety of newly developed chemicals such as additives or pesticides. However, incidental food contaminations requiring rapid decisions and management in the absence of hard toxicological data have highlighted the need to develop alternative strategies, such as based on structure−activity relationship (SAR), capable of providing early and reliable information on potential mutagenic concern.

The Ames test[1,2] is a short-term bacterial in vitro assay detecting chemical induced point and frameshift mutations. Most of the carcinogens acting by induction of genetic damage are mutagenic in the Ames test. Therefore, it has been extensively applied as a screening tool for establishing an initial level of mutagenicity and carcinogenicity concern. The National Toxicology Program (NTP) determined the average interlaboratory reproducibility of a series of Ames test data to be 85%.[3] Since a large database including a wide array of chemical structures and experimental conditions (various bacterial strains used in absence or presence of activating enzymes) is publicly available, the Ames test appears as a promising candidate for developing the predictive model of mutagenicity.

In the 1970s James and Elizabeth Millet introduced the electrophilic theory[4,5] and paved the way for the use of structure−activity relationships (SAR) in the prediction of mutagenicity and/or carcinogenicity. Later, Ashby and Tennant[6−9] catalogued mutagens on the basis of their chemical structures, proposing a first list of 21 moieties highly correlated with mutagenicity. Following these pioneer works, several mutagenicity programs were developed. White et al.[10] challenged the most commonly used programs (CASE/MULTICASE,[11,12] DEREK,[13,14] TOPKAT[15,16]) with a test set of over 500 proprietary pharmaceuticals and reported a concordance (accuracy) between 72% and 81%. Two recent reviews[17,18] examining numerous (Q)SAR models for noncongeneric chemicals reported similar results and stressed the lack of sensitivity (the measure of the ability to correctly identify true positives) as the main limitation of these systems.

Recently, Benigni and Bossa[19] summarized the efforts aimed at expanding and/or refining the knowledge on the structural alerts (SAs) of mutagenicity and acknowledged their fundamental role in risk assessment. To date, the best set of SAs of mutagenicity known was published by Kazius et al.:[20] starting from the original Ashby's list, they assembled a set of 29 toxicophores, together with associated detoxifying substructures, able to classify the mutagenicity of the training set with an accuracy of 82%.

Although SAs have always been the favorable method for mutagenicity prediction, several other approaches, spanning

* Corresponding author phone: +41 (0) 21 7858252; fax: +41 (0) 21 7859486; e-mail: paolo-francesco.mazzatorta@rdls.nestle.com.
† Quality and Safety Department.
‡ BioAnalytical Science Department.

10.1021/ci600411v CCC: $37.00
Published on Web 12/23/2006

the whole universe of artificial intelligence (AI) and QSAR, have been successfully applied.[18,21−27] Because of an improved capacity to cover the space of training set, the combination of different models into a hybrid system has been known to perform at least as well as the best available model.[10,28,29] This triggered the idea of maximizing the current capability of available system to predict Ames test mutagenicity by integrating expert system SAR-based approaches and AI-based systems into a robust hybrid classifier (RHC), with the ultimate goal to get as close as possible of the reliability of experimental tests. In addition, the possibility to attach a level of confidence to each prediction was studied.

MATERIALS AND METHODS

Mutagenicity Data Sets. The data set used for the development of the system was assembled by Kazius et al.[20] from the Chemical Carcinogenicity Research Information System (CCRIS) database (http://toxnet.nlm.nih.gov), adopting a protocol, for ensuring a well-defined endpoint, and quality criteria. Data were restricted to standard Ames tests of *Salmonella Typhimurium* strains TA98, TA100, TA1535, and either TA1537 or TA97 performed with the standard plate method or the preincubation method,[1] either with or without a metabolic activation mixture. A compound was categorized as mutagenic if at least one Ames test result was positive. The final training set was formed by 4337 chemicals (2401 mutagens and 1936 nonmutagens). The test set was collected, harmonized, and cleaned by Young et al.[30] using several public sources including the U.S. EPA, NIH, and the open literature. Chemicals that were present also in the training set were deleted from the test set. In conclusion, a test set of 753 compounds with corresponding molecular structure and toxicity categorizations (437 mutagens and 316 nonmutagens) was constructed. Molecular structures were represented by SMILES strings.[31] We acknowledge that the results here discussed, of course, reflect the data sets used, which, in turn, depend on the availability and quality of Ames test data. Nevertheless, we believe that the dimension of these data sets can corroborate the reliability of the findings.

Structural Alerts. The list of SAs used in this study was derived and validated by Kazius et al.[20] Their work started spotting eight SAs responsible for detecting 75% of all mutagens in the training set. Then the structural complexity of these general toxicophores was improved by combining mechanistic knowledge gained from the literature, statistical tests, and data mining. Eventually, they ended up with 29 toxicophores and some detoxifying substructures (i.e., substructures inhibiting toxicophore action by, for example, steric hindrance or by a distruption of the required electronic charge distribution near the toxicophore), which were encoded in SMARTS strings (http://www.daylight.com/dayhtml/doc/theory/theory.smarts.html). In the following, we will refer to such a model as an SA model (SAm).

Artificial Intelligence. The AI model (AIm) is based on the LAZAR system (http://www.predictive-toxicology.org/lazar/index.html) developed by C. Helma[26,27] and uses the training set as an inductive database that derives its predictions from the experimental measurements of the neighbors for a query structure with a modified k-nearest-neighbor algorithm. The whole process can be summarized as follows: (i) linear fragments are generated automatically from the data set using the molecular feature miner (MOLFEA[32]); (ii) fragments relevant for the toxicity activity are selected; (iii) redundant fragments are removed; and (iv) prediction is obtained using a majority weighted vote from all neighbors of the query structure with a similarity (Tanimoto similarity score) above a predefined threshold.

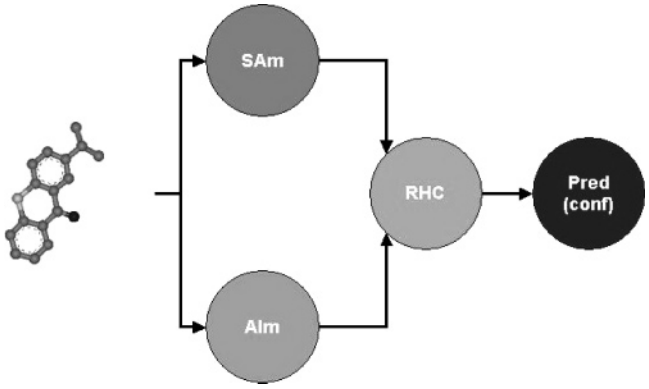

Figure 1. Scheme of the robust hybrid system (RHC).

Table 1. Overall Statistics of the Models[a]

	training (4337)			test (753)		
	accuracy (%)	sensitivity (%)	specificity (%)	accuracy (%)	sensitivity (%)	specificity (%)
SAm	81.7	83.6	77.7	80.5	82.2	78.2
AIm	81.8	83.6	79.5	78.5	77.2	80.4
RHC	88.7	91.1	85.6	85.3	85.4	85.1

[a] In brackets the number of compounds is reported.

Implementation. The two models (SAm and AIm) and the RHC were implemented in C++ using OpenBabel 1.100.2 libraries (http://openbabel.sourceforge.net/wiki/Main_Page). They were compiled with gcc on Windows 2000.

RESULTS AND DISCUSSION

The scheme of the hybrid system developed is shown in Figure 1: SAm and AIm independently predict the mutagenicity of a given chemical, then results are combined by RHC which returns the Ames prediction together with the confidence level.

The performances of independent models and hybrid system are reported in Table 1 and summarized by the receiver operating characteristic (ROC) curve[29] in Figure 2.

The performances are measured in terms of accuracy (the overall percentage of correct positive and negative calls), sensitivity (the measure of the ability to correctly identify true positives), and specificity (the ratio of true negatives to the sum of true negatives and false positives), where positive and negative refer respectively to mutagenic and nonmutagenic compounds.

The National Toxicological Program (NTP)[3] assessed the variability of different laboratories in reproducing measurements (ring test). It was found that the concordance among laboratories for a given chemical tested under code and under strictly controlled protocols was 85%. This creates a boundary condition for the internal performance and predictivity of (Q)SAR models. If the variation in measurement is 15%, it would be unrealistic to expect (Q)SAR models based on those measurement to have higher predictivity. Therefore this

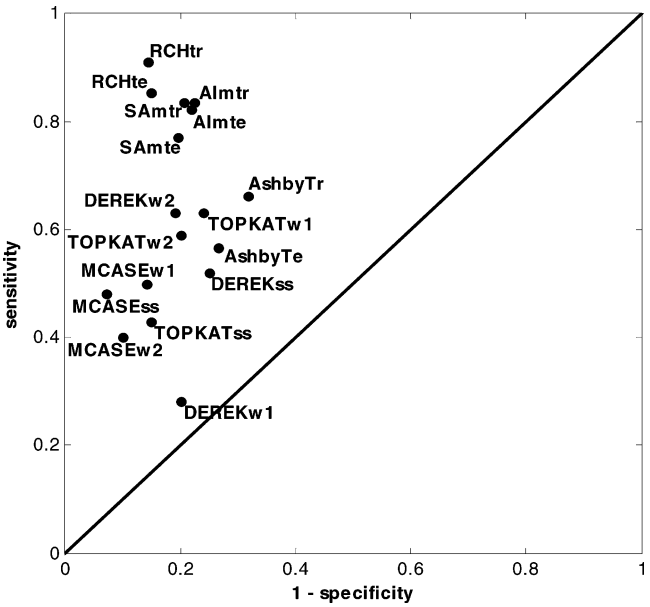

Figure 2. The ROC graph shows the performances of a number of models on different data set. In particular, SAm on training (SAmtr) and test (SAmte) sets, AIm on training (AImtr), and test (AImte) sets, RHC on training (RHCtr), test (RHCte) set; and Ashby/Tennant moieties[9] on training (AshbyTr) and test (AshbyTe) sets; and MULTICASE (MCASEss), DEREK (DEREKss), TOPKAT (TOPKATss) on a data set of over 500 proprietary pharmaceuticals (as reported by Snyder and Smith[17]); and TOPKAT (TOPKATw1, TOPKATw2), CASETOX (MCASEw1, MCASEw2), DEREK (DEREKw1, DEREKw2) on two different data sets of 520 and 94 compounds, respectively (as reported by White et al.[10]).

value can be regarded as the intrinsic experimental error associated to this toxicological endpoint. Keeping this value in mind, it can be noticed that both SAm and AIm are exceptionally promising and not too far from the variability of the experimental test. Nevertheless, none of the models alone can reach the reliability of laboratory measurement of mutagenicity. Only when the two models are combined into the RHC can the accuracy in prediction be considered quantitatively equivalent to experimental tests. Yet more, the fact that both sensitivity and specificity equalize performances of experimental determination of mutagenicity has to be regarded as a proof of the exceptional reliability, stability, and robustness of the proposed system.

The above appears clearer in the ROC graph (Figure 2), where the diagonal line represents random models, whereas the best models should cluster in the upper left corner (sensitivity = 1, specificity = 1). It can be easily noticed that RHC is superior to others models and moreover overcomes the characteristic limitations of previous in silico programs,[17] i.e. poor sensitivity. It must be mentioned that the results here shown (Figure 2) depend on the composition of the data sets, and the benchmark of programs' performances on different data sets is not good practice. More yet this figure may not take into account the recent development of commercial models. The general trends, however, clearly emerge.

The second step of our study was to assess the accuracy of the individual toxicophores and therefore the confidence of every prediction (between zero and one). In order not to bias RHC, this was carried out considering only the compounds included in the test set, and the values were calculated in the absence of compounds that contain multiple SAs. Doing so, the actual predictivity of the individual toxicophores is better reflected. The estimation of the confidence of individual toxicophores and the associated *p*-value are reported in Table 2.

The level of confidence represents the ratio of the number of mutagens in the test set that contain only a given toxicophore to the total number of all compounds in the test set having only that particular moiety. In Table 2, the *p*-value refers to the hypothesis that a random selection of an equal number of compounds from the test set will have an accuracy that equals or exceeds the accuracy of a given SA, so it can be regarded as a test for assessing the reliability of the accuracy of a given toxicophore: a small value (<0.05) proves that the information is statistically relevant. Only six moieties (specific aromatic nitro, specific aromatic amine, nitrosamine, epoxide, aliphatic halide, polycyclic aromatic system) were enough represented in the test set to allow a reliable assessment of their confidence (bold in Table 2). These toxicophores are in fact the most discriminative moieties for mutagenicity and are comparable to the six substructures recently selected using an elaborated graph-based technique.[33] The remaining SAs all together have a confidence of 0.81 and a *p*-value of 0.00098.

A further analysis was performed on nonconsensus predictions (18.4% in the training set, 17.1% in the test set), i.e. when SAm and AIm are in disagreement. Table 3 summarizes the performances of the individual models in predicting the subsets of compounds with disagreeing results.

The dramatic worsening of the performances reflects the uncertainty in predicting these compounds. We strove to identify possible patterns among these chemical structures. Particularly, we focused on the nonconsensus subsets of compounds and investigated the possibility to find the reasons for conflicting prediction and/or expert rules for the choice of the best predictor. For these reasons, more than 60 000 fragments/substructures were generated using MOLFEA,[32] the molecular fragment miner (MoFa),[34] and Omega[35,36] programs from the training set and analyzed by statistical means. Unfortunately, apart from a few weak relationships (*p*-value $\gg 0.05$) that could not be considered statistically relevant, no clear patterns emerged.

Therefore, for the assessment of the confidence of the prediction we decided to use the scheme summarized in Table 4.

The confidence level and error associated with individual toxicophores are summarized in Table 5, where e_i is the error associated with individual toxicophores.

In practice, according to the results of SAm and of AIm four cases are possible:

(i) if both models predict the compound as nonmutagenic, RHC considers it negative attaching a confidence equal to the overall specificity of the system (0.85);

(ii) in case there is a consensus regarding the mutagenicity of the given compound, RHC consider it as mutagen and the confidence is equal to the sensitivity of RHC weighted by the product of the individual error associated to the SAs present in the compound.

(iii) when AIm detects a given compound as mutagenic, but SAm does not, RHC considers it as mutagenic, but the confidence is equal to the difference between SAm specificity and AIm sensitivity on the nonconsensus subset.

Table 2. Example Substructures of Toxicophores and Associated Accuracy[a]

Toxicophore	Example substructure	Confidence	P-value
specific aromatic nitro		**0.81**	**<< 0.05**
specific aromatic amine		**0.79**	**< 0.05**
aromatic nitroso		0.5	0.80
alkyl nitrite		1.00	0.55
nitrosamine		**0.90**	**<< 0.05**
epoxide		**0.85**	**< 0.05**
aziridine		1.00	0.55
azide		NA	NA
diazo		NA	NA
triazene		0.00	1
aromatic azo		1.00	0.09
unsubstituted heteroatom-bonded heteroatom		0.80	0.10
aromatic hydroxylamine		NA	NA
aliphatic halide		**0.79**	**<< 0.05**
carboxylic acid halide		NA	NA
nitrogen or sulphur mustard		NA	NA
polycyclic aromatic system		**0.90**	**<< 0.05**
bay-region in polycyclic aromatic hydrocarbons		NA	NA
K-region in polycyclic aromatic hydrocarbons		NA	NA
sulphonate-bonded carbon (alkyl alkane sulfonate or dialkyl sulfonate)		1.00	0.55
α,β-unsaturated aldehyde (including R-carbonyl aldehyde)		0.81	0.17
aliphatic N-nitro		NA	NA
diazonium		0.00	1
β-propiolactone		NA	NA
α,β-unsaturated alkoxy group		1.00	0.09
1-aryl-2-monoalkyl hydrazine		NA	NA
aromatic methylamine		NA	NA
aromatic hydroxylamine ester		NA	NA
polycyclic planar system		0.78	0.15

[a] Toxicophores that can be considered significant are depicted in bold. If there were no chemicals with the given toxicophore, the estimation of the confidence and the calculation of *p*-value is not applicable (NA).

Table 3. Statistics of the Models on the Subsets of Compounds with Nonconsensus Prediction[a]

	training (796)			test (129)		
	accuracy (%)	sensitivity (%)	specificity (%)	accuracy (%)	sensitivity (%)	specificity (%)
SAm	50.6	54.5	46.7	57.4	66.2	45.5
AIm	48.6	45.8	54.6	42.6	34.3	56.6

[a] In brackets the number of compounds is reported.

Table 4. Prediction Scheme[a]

SAm	AIm	RHC	confidence
0	0	0	0.85
1	1	1	$0.85\ (1 - \Pi\ e_i)$
0	1	0	0.11
1	0	1	$(1 - \Pi\ e_i) - 0.57$

[a] 0 = nonmutagen; 1 = mutagen.

Table 5. Confidence and Error Associated with Individual Toxicophores

toxicophore	confidence	error
specific aromatic nitro	0.81	0.19
specific aromatic amine	0.79	0.21
nitrosamine	0.90	0.10
epoxide	0.85	0.15
aliphatic halide	0.79	0.21
polycyclic aromatic system	0.90	0.10
other SAs	0.81	0.19

(iv) otherwise (SAm positive and AIm negative), RHC considers the given compounds as mutagenic, but the product of confidences of each individual toxicophore present is adjusted by the specificity of AIm on the nonconsensus subset.

In other words, in case of nonconsensus prediction, SAm is preferred, because rather than being a merely statistical model it is based on well-documented experimental evidence and its accuracy is generally superior, but the confidence of the prediction is accordingly lowered.

CONCLUSIONS

The system proposed in this study is an improved fast, reliable, and easy method for the prediction of the Ames test results. Starting from the molecular structures encoded in SMILES notation, the results of each model are combined and interpreted. Each single prediction is provided with a confidence level that gives its statistical reliability rather than quantifying actual errors. The overall error of the proposed system, 15% (sensitivity = 85%, specificity = 85%), is lower than any other predictive system (to author's knowledge) and is equal to the interlaboratory reproducibility of experimental test, which can be regarded as the intrinsic error of the test. These characteristics make the system suitable for early determination of levels of genotoxicity concern. Moreover, the use of two independent predictive models and a confidence level allow risk assessors to depict worst and best case scenarios, and the user can apply consensus, worst case, or best case strategies, according to the issue faced. The system is easily scalable allowing the integration with other predictive models.

ACKNOWLEDGMENT

The authors wish to express their gratitude to Dr. Christoph Helma (In Silico Toxicology, Germany) for providing the

LAZAR source code and for his kind support in the compilation of the program. Further, the authors wish to thank Dr. Roberta Bursi (NV Organon, The Netherlands) and Dr. Christophe Lambert (Golden Helix Inc., U.S.A.) for providing their data sets as well as Dr. Romualdo Benigni (Istituto Superiore di Sanità, Italy) for his precious suggestions. We thank Dr. Manuel Dominguez and Dr. Gabriele Scholz (Nestlé Research Center, Switzerland) for their helpful comments.

REFERENCES AND NOTES

(1) Ames, B. N.; McCann, H.; Yamasaki, E. Methods for detecting carcinogens and mutagens with the Salmonella/mammalianmicrosome mutagenicity test. *Mutat. Res.* **1975**, *31*, 347−364.
(2) Josephy, P. D.; Cruz, P.; Nohmi, T. Recent advances in the construction of bacterial genotoxicity assays. *Mutat. Res.* **1997**, *386*, 1−23.
(3) 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.
(4) Miller, J. A.; Miller, E. C. *Origins of human cancer*; Cold Spring Harbor Laboratory: Cold Spring Harbor, 1977.
(5) Miller, E. C.; Miller, J. A. Searches for ultimate chemical carcinogens and their reactions with cellular macromolecules. *Cancer* **1981**, *47*, 2327−2345.
(6) Ashby, J. Fundamental structural alerts to potential carcinogenicity or non carcinogenicity. *Environ. Mutagen.* **1985**, *7*, 919−921.
(7) Ashby, J.; Tennant, R. W. Chemical structure, Salmonella mutagenicity and extent of carcinogenicty as indicators of genotoxicity carcinogenesis among 222 chemicals tested in rodens by the U.S. NCI/NTP. *Mutat. Res.* **1988**, *204*, 17−115.
(8) Tennant, R. W.; Ashby, J. Classification according to chemical structure, mutagenicity to Salmonella and level of carcinogenicty of a further 39 chemicals tested for carcinogenicity by the U.S. National Toxicology Program. *Mutat. Res.* **1991**, *257*, 209−227.
(9) 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.
(10) White, A. C.; Mueller, A. R.; Gallavan, R. H.; Aaron, S.; Wilson, A. G. E. A multiple in silico program approach for the prediction of mutagenicity from chemical structure. *Mutat. Res.* **2003**, *539*, 77−89.
(11) Klopman, G. Artificial intelligence approach to structure-activity studies. Computer automated structure evaluation of biological activity of organic molecules. *J. Am. Chem. Soc.* **1984**, *106*, 7315−7321.
(12) Cunningham, A. R.; Rosenkranz, H. S. Estimating the extent of the health hazard posed by high-production volume chemicals. *Environ. Health Perspect.* **2001**, *109*, 953−956.
(13) 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.
(14) Ridings, J. E.; Barratt, M. D.; Cary, R.; Earnshaw, C. G.; Eggington, C. E.; Ellis, M. K.; Judson, P. N.; Langowski, J. J.; Marchant, C. A.; Payne, M. P.; Watson, W. P.; Yih, T. D. Computer prediction of possible toxic action from chemical structure: an update on the DEREK system. *Toxicol.* **1996**, *106*, 267−279.
(15) Enslein, K.; Blake, B. W.; Borgstedt, H. H. Prediction of probability of carcinogenicity for a set of ongoing NTP bioassays. *Mutagenesis* **1990**, *5*, 305−306.
(16) 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.
(17) Snyder, R. D.; Smith, M. D. Computational prediction of genotoxicity: room for improvement. *Drug Discovery Today* **2005**, *10*, 1119−1124.
(18) Benigni, R. Structure-Activity Relationship Studies of Chemical Mutagens and Carcinogens: Mechanistic Investigations and Prediction Approaches. *Chem. Rev.* **2005**, *105*, 1767−1800.
(19) Benigni, R.; Bossa, C. Structural Alerts of Mutagens and Carcinogens. *Curr. Comput. Aid. Drug Des.* **2006**, *2*, 1−19.
(20) Kazius, J.; McGuire, R.; Bursi, R. Derivation and Validation of Toxicophores for Mutagenicity Prediction. *J. Med. Chem.* **2005**, *48*, 312−320.
(21) Benigni, R.; Giuliani, A. Quantitative structure−activity relationship (QSAR) studies of mutagens and carcinogens. *Med. Res. Rev.* **1996**, *16*, 267−284.
(22) Tuppurainen, K. Frontier orbital energies, hydrophobicity and steric factors as physical QSAR descriptors of molecular mutagenicity. A review with a case study: MX compounds. *Chemospshere* **1999**, *38*, 3015−3030.
(23) Sello, G.; Sala, L.; Benfenati, E. Predicting toxicity: a mechanism action model of chemical mutagenicity. *Mutat. Res.* **2001**, *479*, 141−171.
(24) Votano, J. R.; Hall, L. H.; Kier, L. B.; Oloff, S.; Tropsha, A.; Xie, Q.; Tong, W. Three New Consensus QSAR Models for the Prediction of Ames Genotoxicity. *Mutagenesis* **2004**, *19*, 365−377.
(25) Kho, R.; Hodges, J. A.; Hansen, J. A.; Villar, H. O. Ring Systems in Mutagenicity Databases. *J. Med. Chem.* **2005**, *48*, 6671−6678.
(26) Helma, C. Lazy Structure-Activity Relationships (lazar) for the Prediction of Rodent Carcinogenicity and Salmonella Mutagenicity. *Mol. Diversity* In press.
(27) Helma, C.; Cramer, T.; Kramer, S.; De Raedt, L. Data Mining and Machine Learning Techniques for the Identification of Mutagenicity Inducing Substructures and Structure Activity Relationships of Noncongeneric Compounds. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1402−1411.
(28) Benfenati, E.; Mazzatorta, P.; Neagu, C.-D.; Gini, G. Combining classifiers of pesticides toxicity through a neuro-fuzzy approach. In *Lecture Notes in Computer Science*; Roli, F., Kittler, J., Eds., Springer-Verlag: Berlin, Heidelberg, Cagliari, Italy, pp 293−303.
(29) Provost, F.; Fawcett, T. Robust Classification of Imprecise Environment. *Machine Learn.* **2001**, *42*, 203−231.
(30) 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. Syst.* **2002**, *60*, 5−11.
(31) Weininger, D. SMILES, a chemical language and information system 1. Introduction and encoding rules. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31−36.
(32) Kramer, S.; Frank, E.; Helma, C. Fragment generation and support vector machines for inducing SARs. *SAR QSAR Environ. Res.* **2002**, *13*, 509−523.
(33) Kazius, J.; Nijssen, S.; Kok, J.; Bäck, T.; Ijzerman, A. P. Substructure Mining using Elaborate Chemical Representation. *J. Chem. Inf. Model* **2006**, *46*, 597−605.
(34) Hofer, H.; Borgelt, C.; Berthold, M. Large Scale Mining of Molecular Fragments with Wildcards. In *Proceedings of the 5th International Symposium on Intelligent Data Analysis (IDA 2003, Berlin, Germany)*; Springer-Verlag: Heidelberg, Germany, 2003; pp 380−389.
(35) Boström, J. Reproducing the Conformations of Protein-Bound Ligands: A Critical Evaluation of Several Popular Conformational Searching Tools. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 1137−1152.
(36) Boström, J.; Greenwood, J. R.; Gottfries, J. Assessing the Performance of OMEGA with Respect to Retrieving Bioactive Conformations. *J. Mol. Graphics Modell.* **2003**, *21*, 449−462.

CI600411V