Bioorganic & Medicinal Chemistry 19 (2011) 3173–3182

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Avoidance of the Ames test liability for aryl–amines via computation

Patrick McCarren [a], Gregory R. Bebernitz [a], Peter Gedeck [b], Susanne Glowienke [c], Melissa S. Grondine [a], Louise C. Kirman [a], Jacob Klickstein [a], Herbert F. Schuster [a], Lewis Whitehead [a,*]

[a] Novartis Institutes for Biomedical Research, 100 Technology Square, Cambridge, MA 02139, USA
[b] Novartis Pharma AG, Forum 1, Basel CH-4056, Switzerland
[c] Novartis Pharma AG, Auhafenstrasse, Muttentz CH-4132, Switzerland

ARTICLE INFO

Article history:
Received 13 February 2011
Revised 25 March 2011
Accepted 30 March 2011
Available online 3 April 2011

Keywords:
Aryl–amines
Toxicity prediction
Nitrenium
Ames
Density functional theory

ABSTRACT

Aryl–amines are commonly used synthons in modern drug discovery, however a minority of these chemical templates have the potential to cause toxicity through mutagenicity. The toxicity mostly arises through a series of metabolic steps leading to a reactive electrophilic nitrenium cation intermediate that reacts with DNA nucleotides causing mutation. Highly detailed in silico calculations of the energetics of chemical reactions involved in the metabolic formation of nitrenium cations have been performed. This allowed a critical assessment of the accuracy and reliability of using a theoretical formation energy of the DNA-reactive nitrenium intermediate to correlate with the Ames test response. This study contains the largest data set reported to date, and presents the in silico calculations versus the in vitro Ames response data in the form of beanplots commonly used in statistical analysis. A comparison of this quantum mechanical approach to QSAR and knowledge-based methods is also reported, as well as the calculated formation energies of nitrenium ions for thousands of commercially available aryl–amines generated as a watch-list for medicinal chemists in their synthetic optimization strategies.

1. Introduction

The aryl–amine chemical substructure is often without toxicity and appears in numerous successful medicines and therapeutics.[1] However, numerous anecdotal examples of faltered chemical series and terminated drug discovery projects have made the substructure anathema to some chemists. Aryl–amines present in concentrations of parts per million as impurity, metabolite, or degradation product can trigger a genotoxicity alert. This subsequently introduces increased developmental scrutiny with loss of valuable time, higher economic costs and competitive disadvantage to a clinical candidate.[2] Therefore, increased vigilance at the time of reagent selection is essential.

Because of the long history of study of aryl–amines as environmental carcinogens,[3,4] the mechanisms of aryl–amine genotoxicity are generally well established. Aryl–amines most often require metabolic activation to yield a more electrophilic compound. One of the more famous examples is the rat carcinogenicity of 2-acetylaminofluorene.[5,6] This was shown by Miller and Miller to occur through metabolism to an *N*-hydroxyester–amine. When an acyl anion is lost, a reactive, electrophilic intermediate is generated. This intermediate, a nitrenium ion, is a di-coordinate nitrogen intermediate with a lone pair and a positive charge on nitrogen. These intermediates are short-lived and can generally be studied only using flash photolysis and ultra-fast spectroscopy.[7,8] However, their products with DNA and proteins are quantifiable in human and animal subjects using mass spectrometry.[3] An unusual feature of the aryl nitrenium ion is its ability to reach higher spin states.[9]

The most common in vitro experimental genotoxicity test is the Ames test.[10,11] It is a highly sensitive surrogate test for carcinogenicity[12] and is generally required by regulatory authorities.[13] The *Salmonella typhimurium* bacterial strains used in the test have mutations in the His operon which prevents the biosynthesis of histidine. When cultured on histidine-deficient medium, few colonies will be formed. However, when mutagenic chemicals are tested, usually both with and without metabolic enzymes, they cause mutations in this operon region leading to a dose-dependent increase in colony formation. The strains have also been carefully selected to have a high mutation frequency, relatively low DNA repair capacity, and high cell membrane permeability.[14] Unfortunately, the test is quite expensive and slow on the scale of modern drug discovery. Although high-throughput genotoxicity tests exist, their correlation to the regulatory test has been fairly low for the considerable added expense.[15,16]

Computational screening can often provide inexpensive, new solutions to medicinal chemistry problems. Recent studies have provided an extremely good correlation between accurate calculations of nitrenium ion stability and the Ames test result.[17–20] Though the calculations are conceptually straightforward, they

* Corresponding author. Tel.: +1 617 871 7115.
E-mail addresses: lewis.whitehead@novartis.com, lewlou1@verizon.net (L. Whitehead).

0968-0896/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2011.03.066

are technically challenging and time-consuming. However, advances in computing capabilities have made this approach feasible.

This study critically assesses the accuracy of the approach across four sets of experimental results comprising the largest collection of aryl–amines considered in any study to date (846 unique). This work also considered the electronic spin states of the nitrenium ions, the relationship to molecular weight and applied the prediction method to reagents from available chemical catalogs. The four data sets studied are from different sources as detailed in Table 1: (1) a small collection of heterocyclic aryl–amine ligands used as synthons for an internal drug discovery program[21–26] (shown in Fig. 1); (2) a larger collection of in house data generated for multiple initiatives also of concern to discovery operations; (3) an external evaluation set constructed from primary literature[27,29,30] and the EPA Genetox database;[28] and (4) a large external evaluation set compiled by other groups.[31,32] An additional set of molecules includes over 14,000 commercially available aryl–amine synthons that have yet to be experimentally tested for mutagenicity. These chemical species were investigated to serve as an internal guideline useful for scientists in reagent selection for pharmaceutical small-molecule optimization strategies.

In order to determine the most relevant energetic steps for the nitrenium ion activation, the reaction energies representing a number of metabolically relevant pathways were calculated. Possible chemical reaction sequences are shown in Figure 2, while path E considered the theoretical one-step nitrenium formation through heterolytic cleavage of the amine N–H bond. Metabolism of aryl–amines can occur through a number of oxidation and conjugation reactions by cytochrome P450 enzymes,[6] prostaglandin H synthase (PHS),[33] *O*-acetyltransferases (OAT)[34], *N*-acetyltransferases (NAT),[35] sulfotransferases,[36] and UDP-glucuronosyl transferases.[37] In vitro experiments, in silico calculations, and discussions of aryl–amine toxicity often assume a hydroxylamine or hydroxylamine ester as the initial intermediate. However, the exact sequence or number of metabolic steps for even a modest set of aryl–amines is not known.[38] While it would be ideal to include metabolic enzyme selectivity in this study, it is a challenging area in itself.[39,40] The goal of this approach was to consider whether thermodynamics of some possible reaction steps would differentiate Ames+ compounds from Ames−.

2. Results and discussion

2.1. Reaction energies for steps leading to nitrenium intermediates and the beanplot

To investigate calculations and data in more detail, the use of the 'beanplot' was found to be a useful visual display element for analysis and presentation of this work.[41] The beanplot provides a quick side-by-side comparison of the calculated energies for Ames+ and Ames− compounds. Figure 3 provides a schematic of the components of a beanplot for path E with the Ames− distribution of calculated energies shown in the top left and the Ames+ distribution below. The calculated energies for each compound appear as short lines. In the Figure, the molecules with the maximum and minimum nitrenium formation energy of each group are shown. The average of the energies is depicted with a long, bolded line, and a kernel density estimate of the distribution of values is drawn. This latter element provides a histogram-like visualization of the concentration of points and normality of the distribution. It is filled green for the compounds measured to be Ames− and red for those measured to be Ames+. The two graphs are then stacked to compare the data directly and rotated to provide a compact visualization. In all graphs, the Ames+ distribution will be shown on the left and the Ames− will be shown on the right.

The beanplots for reaction path A (with 3 steps) and the single-step path E are displayed in Figure 4. The beanplots for steps 1 and 2 of path A are generally symmetric with little differentiation between Ames+ and Ames− compounds. The side-by-side comparisons in path A, step 3 and path E clearly show that the approximately 20 kcal/mol gap between the mean values appears to be a relevant delineation. These are the nitrenium forming steps of paths A and E. That the reaction energy is lower for Ames+ compounds is significant and reflects a lower energy cost for forming the reactive intermediate. Beanplots for all the steps in Figure 2 are provided in the Supplementary data.

The step in each pathway with the highest likelihood of differentiating Ames+ and Ames− compounds was the nitrenium-forming step. Using the Wilcoxon two-sample test,[42] the difference in means for the other steps was unlikely to be significant with calculated *p*-values in the range of 0.44–0.94. The nitrenium forming steps show a significant separation of 15–20 kcal/mol between the means and Wilcoxon *p*-values ranging from 0.002 to 0.006. The steps in Table 2 represent four different nitrenium formation reactions for set A with the nitrenium forming step of path C and B being identical. This finding was in agreement with what was found by other authors.[18]

Though attempts to use quantum mechanics to predict Ames toxicity have been made for many years, the quality of calculations possible for sets has been generally poor until recently. Semi-empirical calculations have been most represented[43,44] and typically used in conjunction with statistical models.[30,45] Semi-empirical quantum mechanics calculations have the benefit that they are fast and provide rough energetics for bond-breaking and bond-forming reactions. Recent calculations have focused on density functional theory (with the B3LYP hybrid functional)[18–20] or Hartree Fock reaction energies.[44] These are generally much more accurate for reaction energies than the parameterized semi-empirical methods. All studies looking at reaction energies of the nitrenium formation step have shown significant differences between Ames+ and Ames− compounds. The mean nitrenium formation energies for Ames+ and Ames− compounds from these studies are provided in Table S3 in the Supplementary data. These studies reported similar separations between the average formation energies of the Ames+ and Ames− compounds as reported here.

2.2. Larger basis sets, thermodynamic corrections and solvation

For set A, frequency calculations, single-point energies with larger basis sets (6-31+G**, 6-311G*), and solvation calculations were performed for each of the steps. Larger basis sets altered the relative rankings for sulfur-containing compounds slightly as might be expected. However, these larger basis sets shifted the distributions of Ames+ and Ames− compounds nearly equally resulting in nearly identical separation of the means. As shown in the beanplots (Figure S5) in the Supplementary data, thermodynamic corrections were not significant to the distributions and difference in the averages for path E. The lack of improvement using thermodynamic corrections combined with the large additional time required for these calculations led us to assume these would

Table 1
Aryl–amine data sets used in this investigation of nitrenium ion stability ('Ames+' denotes a compound found to be mutagenic by the Ames test, 'Ames−' is nonmutagenic.)

Set	Description	N=	Ames+	Ames−
A	Glucokinase activators[21–26]	32	6	26
B	Novartis	326	72	254
C	EPA-Toxnet[27–30]	189	136	53
D	Kazius[31] and Hansen[32]	459	326	133

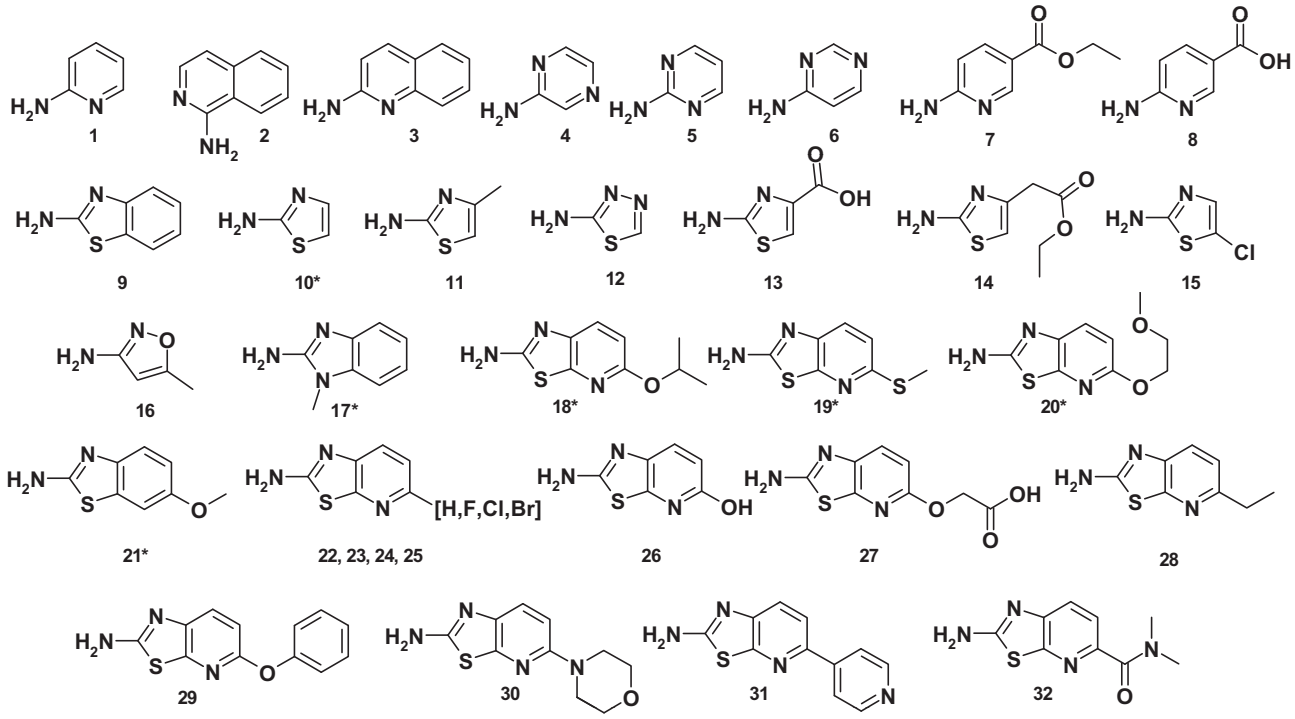

Figure 1. Molecules included in set A. Those highlighted with an asterisk show an Ames+ response.

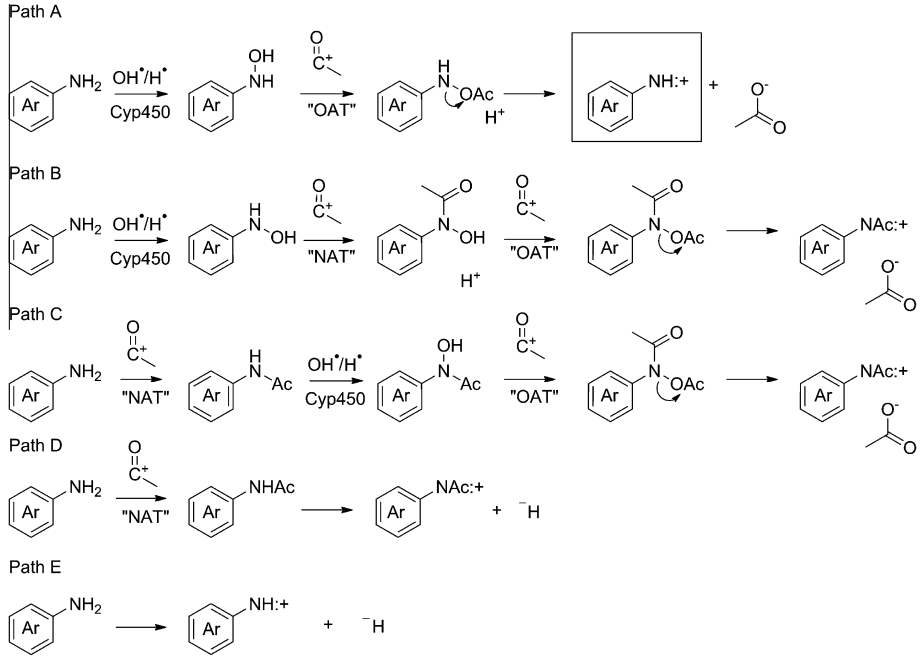

Figure 2. Nitrenium ion generation pathways for aryl–amines represented schematically as aniline labeled Ar. Acetylation by *N*-acetyltransferase (NAT) or *O*-acetyltransferase (OAT) is calculated thermodynamically as acetyl cation with loss of H+. The N-oxidation by cytochrome P450 is calculated as hydroxyl radical addition.

be negligible to the results in the larger set. Solvation energies had an observable effect (Figure S6) but overall did not improve the separation.

2.3. Singlet–triplet gaps in nitrenium ions

Previous calculations on nitrenium ions have shown that spin multiplicity can be important in these species, particularly for electron-donating groups at the *meta*-position to the nitrenium ion.[9] Even in this small set, some triplet nitrenium ions were preferred over the singlet at the UB3LYP level. This was more significant in pathways B–D. For path D, the N-acetylated nitrenium species have an even lower singlet–triplet gap. Aryl–amines **5** and **6** have lower energy triplet spin states for their nitrenium species by a relatively small amount- by 1–2 kcal/mol (Table S1). After acetylation, this gap increased to 4–7 kcal/mol and **1**, **4**, **7**, and **8** have triplet

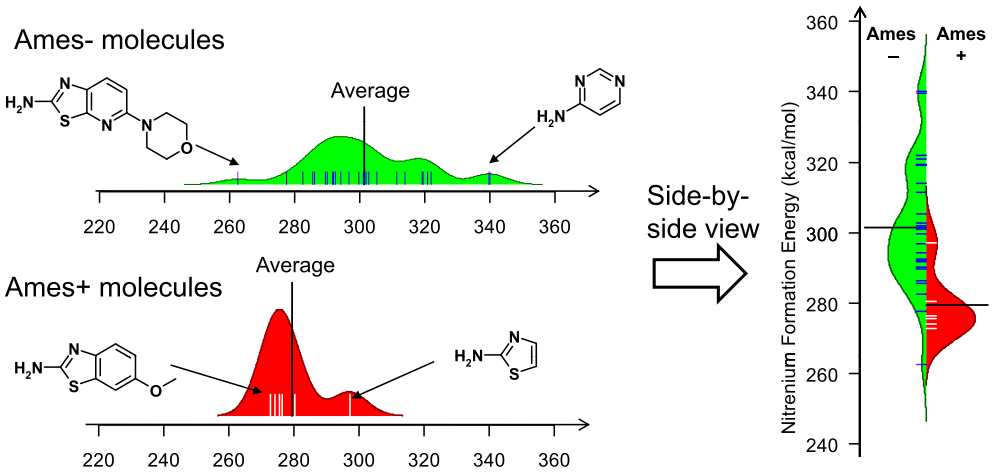

Figure 3. The beanplot depiction of data allows a visualization of individual data points and the statistics by showing each point as a short whisker line, an approximate distribution using a kernel density approximation, and the mean of the data shown as a long line. The plots in this paper use this depiction in a rotated, stacked form as shown on the right. Reaction energies of Ames+ molecules are always shown on the right with white lines inside a red distribution and those for Ames− molecules are always shown on the left of the center with dark individual points (blue) in a green filled distribution.

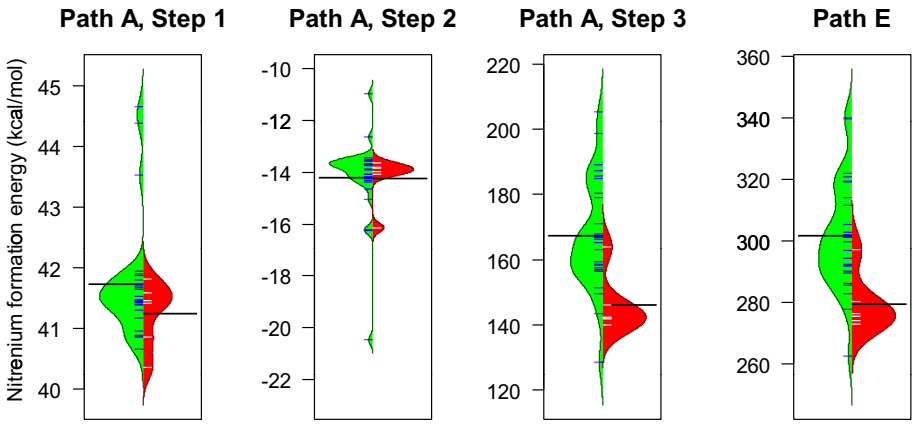

Figure 4. A beanplot for (a) reaction path A, and (b) path E. The *y*-axis is the reaction energy in kcal/mol. The distribution of energies on the left is for experimentally Ames− molecules (green) and that on the right is the distribution for Ames+ molecules (red). Small individual lines are data responses for 32 set A chemical species, two bold lines mark the average response for nitrenium ion creation in reaction path E.

Table 2
Statistics for the nitrenium forming steps of paths shown in Figure 4 for set A

Path (reaction step)	Ames+ mean (σ)	Ames− mean (σ)	Wilcoxon *p*
A (3)	146.0 (9.0)	167.4 (17.4)	0.002
B (4)/C (4)	149.9 (9.3)	164.8 (13.1)	0.006
D (2)	284.7 (8.4)	299.4 (12.8)	0.003
E (1)	279.4 (9.0)	301.5 (18.1)	0.002

energies favored by 1–2 kcal/mol (Table S2). This data led to the decision to calculate both singlet and triplet multiplicity states of the nitrenium ions in the larger sets. Other recent studies have ignored this consideration because the singlet is often preferred. Hartman and Schlegel[46] previously considered the nitrenium singlet–triplet gap as a single descriptor of aryl–amine mutagenicity, however.

2.4. Molecular weight dependence in the Novartis set

Reaction energies for the steps in path A (Figure S7) and path E for the compounds in set B were also calculated. Set B contains all aryl–amines that have been tested using the Ames screening test at Novartis and represents a fairly even distribution for molecular weights up to 500 g/mol. There are 115 compounds below 250 g/mol, 182 between 250 and 500 g/mol, and 29 above 500 g/mol. The data was binned into molecular weight categories to elucidate the relationship between molecular size of the aryl–amine chemical species and its relevance to the Ames response. Molecular weight is well-known to indirectly reflect information about other properties and complexity. In fact, it is an important component of the 'rule-of-five' chemical descriptors for small molecule drug-likeness of orally administered therapeutics.[47]

Consideration of molecular weight was an important choice as can be seen in the beanplots in Figure 5. Each of the beanplots is plotted on the same scale but represent data from different molecular weight ranges as noted. Larger aryl–amines generally have lower formation energies to the extent that the mean for the energies of the Ames+ compounds in one bin essentially becomes the Ames− mean for the next bin. The statistical data summarized in Table 3 provides a quantitative measure of the differences in the distributions. Compared to the overall average, the compounds in the <250 g/mol molecular weight bin that are Ames+ have an average that is 3 kcal/mol higher than that for all Ames+ compounds. The average of the Ames− compounds in this bin is 7 kcal/mol higher than the average of all Ames−. Between 250 and 500 g/mol, the Ames+ bin is 2 kcal/mol lower than the

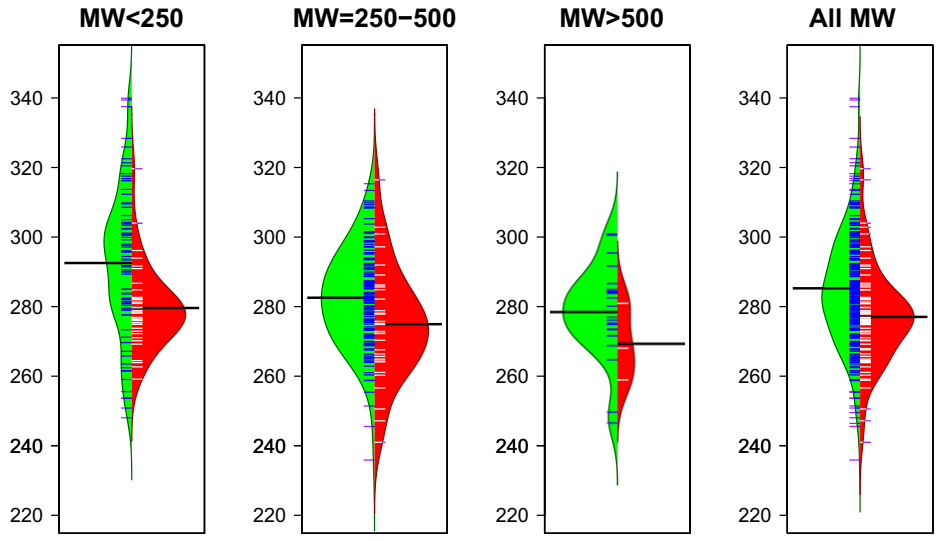

Figure 5. Beanplots showing the difference in distribution of nitrenium formation energies of set B for different molecular weight ranges. As molecular weight increased, the separation of Ames+ and Ames− energy distributions decreased.

Table 3
Analysis by binning the Ames test results and calculated energies shown in Figure 5 by molecular weight

MW	Ames+			Ames−			Significance
	Mean	SD	*N*	Mean	SD	*N*	*p*
<250	279.6	11.6	36	292.5	21.1	79	0.00019
250–500	275.0	16.2	33	282.6	13.9	150	0.00540
>500	269.2	11.1	3	278.8	12.8	26	0.14997
All	277.0	14.0	72	285.3	17.1	255	0.00006

average and the Ames− average is 3 kcal/mol lower. This is likely caused by multiple components arising both from the increased size and complexity. Shape, multiple sites of metabolism, intramolecular hydrogen bonding, larger ring systems, or alternate modes of DNA action are all reasonable explanations.

2.5. Determining the most predictive energetic threshold

For such a complicated process as DNA toxicity, it is not surprising that the Ames+ and Ames− compounds would be overlapping in almost any coordinate space of descriptors. However, excellent performance is achieved through this single parameter, in conjunction with a molecular weight binning. As noted in previous studies[18] and evidenced by the active area of developing quite complex statistical models for mutagenicity,[48–50] it is difficult to find such a simple classification from common chemical descriptors. To make this energy a useful parameter, it is necessary to determine an optimal translation of this number into an Ames test prediction. Previous studies seeking to use this approach as a tool, have generally used the nitrenium formation energy for aniline as a reference, which is known to be Ames−.[17,44]

In order to determine the optimal cutoff value for this single quantum mechanical quantity, the predictivity for sets A and B were divided into the performance for the Ames+ and Ames− compounds. For the nitrenium formation energy, the two performances are tied to the cutoff chosen. Using a threshold for nitrenium formation energy in this set of 340 kcal/mol (energies less than 340 being Ames+), the maximum found in this study, results in a 100% correct prediction of the Ames+ compounds. This also falsely predicts all Ames− compounds to be positive, therefore, a compromise value must be determined. In the discussion, it is useful to use terms from detection theory, which have also been applied to clinical tests. The fraction of Ames+ compounds correctly predicted is termed sensitivity. The more sensitive the test, the more Ames+ molecules that will be caught. The fraction of correctly predicted negative compounds is termed specificity or the false positive rate. The more specific a test is, the fewer negative compounds will be falsely classified as positive. Accuracy is simply the number of correctly predicted compounds divided by the total number of compounds.

The two graphs in Figure 6 demonstrate the effect of cutoff on the sensitivity (red), specificity (blue), and accuracy (green) for sets A and B from Novartis for the subset with molecular weight less than 250 g/mol. As the cutoff is increased from 274 to 294 kcal/mol, the sensitivity increased but specificity decreased. Cutoffs are labeled along the *x*-axis with dotted lines marking the 282–284 kcal/mol cutoffs in order. For set A, sensitivity stayed the same in this window while specificity and accuracy decreased minimally. For the larger Set B of Novartis aryl–amines with molecular weights less than 250 g/mol, sensitivity climbed steadily until 283 kcal/mol and plateaued around 80%. Accuracy and specificity was relatively stable in this window as well. Using a threshold of 283 kcal/mol, 28 of the 36 Ames+ compounds were correctly predicted to be positive giving a sensitivity of 78% (Table 4). This cutoff also resulted in a 66% specificity correctly predicting 52 of 79 Ames− compounds correctly.

Table 4 also shows the performance for aryl–amines ranging in weight from 250 to 500 g/mol. As was shown previously in Figure 5, the separation in means for the Ames+ and Ames− compounds was much lower in this range. Due to the low accuracy, an attempt to find a better cutoff for this weight range was not undertaken. Using the 283 kcal/mol cutoff yielded a sensitivity of 73% but a much lower specificity of 52%.

2.6. Predictivity in external data sets

In conjunction with calculations on the Novartis Ames test data, the approach was also applied to external datasets. One dataset was constructed from the EPA Genetox database and literature reviews.[27–30] This set, denoted set C, contains 189 aryl–amines with molecular weight less than 500 g/mol as detailed in Table 1. In addition, a number of authors have compiled large datasets of mutagenicity results with thousands of molecules including aryl–amine data as a sub-set, the most recent of which are those in Kazius et al.[31] and Hansen et al.[32] When combined, these two sets

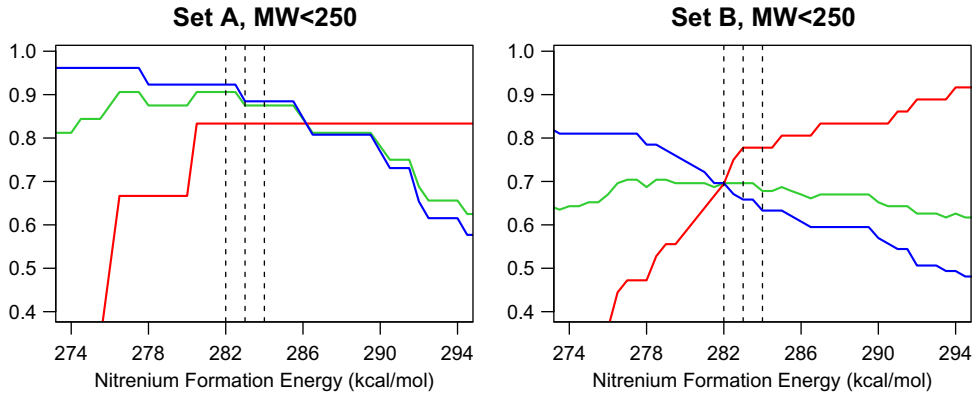

Figure 6. Prediction accuracy (green), sensitivity, or fraction of correctly predicted Ames+ compounds (red), and specificity, or fraction of correctly predicted Ames− compounds (blue) graphed by nitrenium formation cutoff for Novartis sets A and B with molecular weight below 250 g/mol. The three lines mark cutoffs of 282–284 kcal/mol.

Table 4
Predictivity of the nitrenium formation energy using the 283 kcal/mol cutoff (lower predicted to be Ames+)

Set	Sensitivity	Specificity	Overall accuracy	Sensitivity	Specificity	Overall accuracy
	MW <250			250 < MW <500		
A	83% (5/6)	88% (23/26)	88% (23 + 5)/32	—	—	—
B	78% (28/36)	66% (52/79)	70% (52 + 28)/115	73% (24/33)	52% (77/149)	55% (77 + 24)/182
C	84% (108/129)	52% (23/44)	76% (24 + 108)/173	100% (7/7)	33% (3/9)	60% (2 + 7)/16
D	79% (226/285)	74% (83/112)	78% (83 + 226)/397	90% (37/41)	76% (16/21)	85% (16 + 37)/62

of aryl–amine Ames data points resulted in 461 aryl–amines of molecular weight less than 500 g/mol, which are denoted as set D. There was one discrepancy discovered in set D, namely 3-chloro-4-methylaniline, which was marked positive in Hansen et al. but negative in Kazius et al. This value was kept as negative based on a literature report on this compound which found that it was Ames- in multiple *S. typhimurium* strains, though chromosome aberration tests were positive.[51] There were only 11 compounds in the Novartis set, set B, present in either of these test sets, which yielded a total of 846 unique compounds with MW <500 g/mol. Together these sets represent the largest set of collected aryl–amine mutagenicity data reported to date. Sets C and D were not completely independent: there was an overlap between sets C and D of 121 compounds, which left 68/189 (36%) unique compounds in C and 329/459 (72%) unique aryl–amines in D. A summary of the set membership of unique compounds is given in the Supplementary data (Figure S8).

As shown in the beanplots of Figure 7, if considering the aryl–amines with molecular weights less than 250 g/mol, there is also a clear differentiation of Ames+ and Ames− compounds in sets

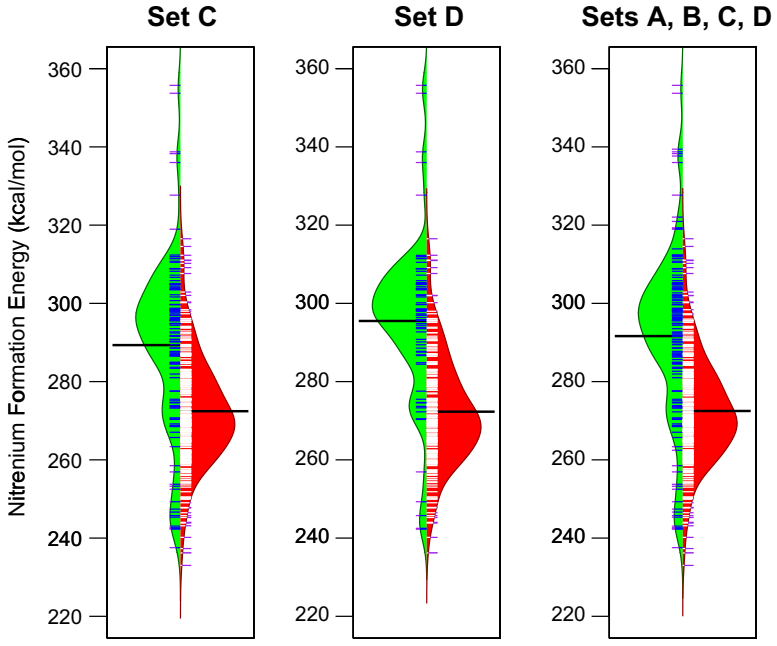

Figure 7. Beanplots for the aryl–amine datasets with molecular weight less than 250 g/mol. The lines and distribution (green) pointing left are for the experimentally Ames− compounds and Ames+ are shown to the right (distribution in red).

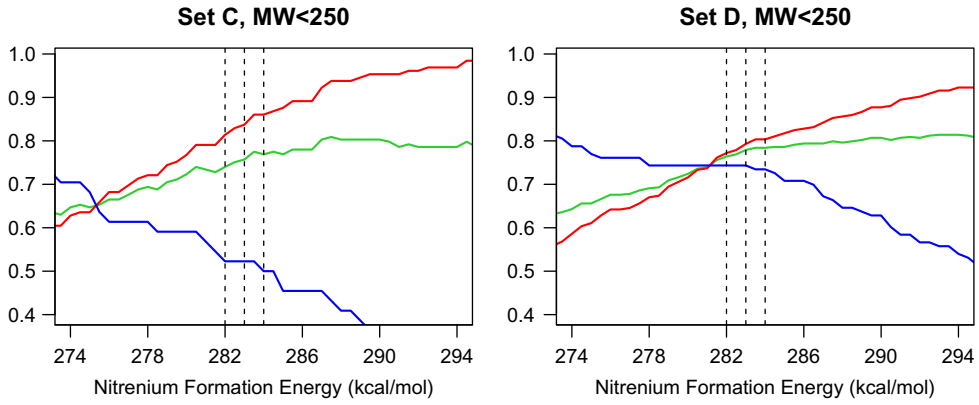

Figure 8. Prediction accuracy (green), sensitivity, or fraction of correctly predicted Ames+ compounds (red), and specificity, or fraction of correctly predicted Ames− compounds (blue) graphed by nitrenium formation cutoff for Novartis sets C and D. The three lines mark cutoffs of 282–284 kcal/mol.

C and D using the nitrenium formation energy. The right-most beanplot in Figure 7, shows the combined results. As summarized in Table 4, using the same cutoff of 283 kcal/mol for the nitrenium formation energy led to the correct prediction of Ames+ compounds (sensitivity) of 84% for set C and 79% for compounds in set D. The specificity was much lower in set C but the high sensitivity and overall accuracy of 76% was an extremely helpful prediction especially considering the other prediction methods available which are discussed below. The cutoff plot is shown in Figure 8 and shows that while sensitivity increased steadily for sets C and D with increasing cutoff, the specificity began to decrease after 283 kcal/mol.

Because sets C and D contain significant overlap, a further measure of the inter-set performance is to evaluate the performance for all unique aryl–amines from the four sets with molecular weight less than 250 g/mol. This resulted in 560 unique molecules with 196 concurrent Ames− and 339 concurrent Ames+, 126 duplicates with identical results, and 25 duplicates that had different results. This 20% discrepancy rate was similar to reports in the experimental Ames test literature.[52] After the value for these duplicates was changed to Ames+, the Ames+ performance for the resulting unique set was also quite good: 291/364 were correctly predicted (80% sensitivity) and 58/196 Ames− molecules were correctly predicted (71% specificity). This yielded an overall accuracy of 429/560 correctly predicted results (77%).

2.7. Other available genotoxicity prediction methods

There are also external genotoxicity prediction methods available, which have been assessed in reviews for their performance on available sets of Ames data,[53–55] but not aryl–amines. Three such methods for sets A–D were compared to using a cutoff in nitrenium formation energy: TOPKAT,[56] the Benigni TA100 QSAR6 model in Toxtree,[29,57] and a modified version of DEREK[58,59] called ToxCheck.[60]

The most commonly used solution for predicting genotoxicity problems is a knowledge-based alert system. Just as expert medicinal chemists or toxicologists would know, these computer programs identify substructures in a molecule that are associated with toxicity and flag the molecule as potentially toxic. These substructure rules can be quite complicated and are routinely updated. The Deductive Estimation of Risk from Existing Knowledge[59] (DEREK) program is one of the most widely used programs in this arena. Novartis uses this program and has also uploaded some of its own alerts based on unique cases of toxicity. This modified rule set and internal interface to DEREK is called ToxCheck.[60] Because of the substructure approach, aryl–amines are quite likely to be flagged. As shown in Table 5, some molecules were correctly predicted to be of no concern, but overall, the specificity was low (between 13% and 35%). However, as an alert system, it is successful—the sensitivity for Ames+ compounds was greater than 96%. The much lower accuracy for the Novartis set reflects the larger prevalence of Ames− molecules than in sets C and D.

Another area of extensive work in the area of predicting mutagenicity is the use of statistical models. Most of these models have been built using data available in the public domain. The predictivity of an aryl–amine QSAR model developed by Benigni[29] implemented in Toxtree[57] and a commercially available global mutagenicity model called TOPKAT[56] were assessed for these sets as well. Benigni's mutagenicity QSAR model built for the mutagenicity of aryl–amines[29,49] was applicable to only 50 of the 115 Novartis aryl–amines in the MW <250 g/mol set. Table 5 summarizes the performance for those aryl–amines that were applicable (see the caption for the number of applicable molecules as determined by Toxtree). For this 50, the accuracy was only 50%, essentially random. The performance was significantly better for sets C and D which contain similar or the same molecules used to train the model. Similarly, TOPKAT provided a fairly low performance for the Novartis aryl–amines (39% accuracy) but extremely good performance for sets C and D (76% and 87% accuracy, respectively).

Table 5
Summary of the performance of available alert and QSAR models on the MW <250 subset of aryl–amine compounds

Set	Sensitivity	Specificity	Accuracy
Novartis Toxcheck			
A	100%	35%	47%
B	97%	25%	48%
C	98%	13%	76%
D	96%	29%	77%
TOPKAT[a]			
A	100%	23%	38%
B	67%	50%	39%
C	94%	24%	76%
D	98%	61%	87%
Toxtree Benigni TA100[b]			
B[c]	53%	48%	50%
C[d]	77%	71%	75%
D[e]	79%	53%	73%

[a] The default probability cutoff (0.722) was used.
[b] Set A analysis excluded due to only two applicable points.
[c] Set B, Toxtree applicable:(19 Ames+/31 Ames−).
[d] Set C, Toxtree applicable: (128 Ames+/45 Ames−).
[e] Set D, Toxtree applicable: (218 Ames+/74 Ames−).

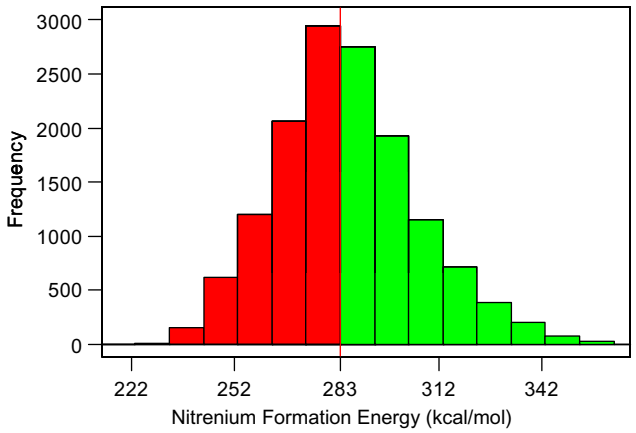

Figure 9. Distribution of nitrenium formation energies calculated for 14,000 commercially available aryl–amines. Bars are colored by predicted Ames test response; approximately 7200 of these aromatic amines are predicted to be 'safe'.

TOPKAT was built with mutagenicity test results for more than just aryl–amines and contained set D.

Given the low overlap of the Novartis data set with previously published data, the poor results of these models are not completely surprising especially those of a global model like TOPKAT. Further efforts to build a more advanced model with better predictivity than this one-parameter quantum mechanical model were surprisingly challenging and will be reported in a future publication.

2.8. Commercially available aryl–amines

One desired outcome of this work was to provide a database of aryl–amine synthon information that would provide internal guidance and awareness of potential Ames liabilities. This data set also serves as a collection of prospective predictions that can be verified in the future as more Ames data becomes available. A list of aryl–amines in internal synthetic building block collections containing proprietary chemical scaffolds and the ASDI catalog[61] was compiled. This compilation yielded over 3000 compounds that have a molecular weight less than 250 g/mol. Additionally, use of the SciQuest catalog[62] to identify other aryl–amines that could be available commercially yielded approximately 50,000 aryl–amines with molecular weights less than 250 g/mol. Of this set, 12,000 compounds had molecular weights less than 200 g/mol and the remaining were in the range of 200–250 g/mol. Figure 9 shows the result of the nitrenium formation energy calculations for over 14,000 aryl–amines. Slightly more than half of the compounds (~7200) are predicted to be Ames− (shown in green) by applying a threshold of 283 kcal/mol. This provides a list of aryl–amine fragments to be used for designing new compounds and navigating away from genotoxic roadblocks.

3. Conclusions

Aryl–amines are important reagents in the construction of pharmaceutical compounds, however, they are all often labeled as increasing the liability for carcinogenicity of a potential clinical candidate. The results in this paper reinforce the rationale of considering the relative reaction energies forming the reactive nitrenium intermediate. This calculated value provides a robust, physically relevant predictor of a primary genotoxicity test, the Ames test. This has been shown by comparing the performance of predictions across different sets comprising the largest set of aryl–amines considered in the literature thus far. Using a cutoff value of 283 kcal/mol detected 78–84% of the reagent-sized Ames+ compounds in each of the sets while still keeping 53–88% of the Ames− reagents. Statistical models currently available for predicting mutagenicity do not provide this performance across sets. This method was used to compile a list of around 7200 aryl–amines to choose from when designing libraries or searching for safe replacements for problematic fragments.

4. Methods

4.1. Ames experiments

S. typhimurium strains TA98 and TA100 were used with a metabolic activation system of 5 μL per well of liver S9 mix from male rats pretreated with Aroclor 1254. All positive controls were dissolved in DMSO and used at one concentration level per strain only at 0.02 mL/plate. For positive controls, 2-aminoanthracene, benzo(a)pyrene, sodium azide were used at 0.6 micrograms/well and 2-nitrofluorene was used at 0.4 micrograms/well.

The mutagenic potential of the test item was determined using *S. typhimurium* in vitro, with and without the addition of a mammalian metabolizing system. The Miniscreen Ames test is a miniaturized version of the original test using smaller volumes and test item amounts, which was performed using 6-well plates instead of Petri dishes.[63,64] It has been modified and validated in house recently, showing an excellent correlation between the results in the Ames Miniscreen and in the Standard Ames test. The top concentration of 1000 micrograms/well used in the Ames Miniscreen approximately corresponds to the top concentration of 5000 micrograms/plate in the standard Ames test. The medium contains traces of histidine, which allow the bacteria to make a few cell divisions. After exhaustion of this histidine, only prototrophic, that is, mutant, bacteria are able to grow. Thus, only histidine prototrophic bacteria are able to form colonies during an incubation time of three days. The number of colonies on the tester wells is, therefore, directly proportional to the number of mutant bacteria. A clear increase in the colony numbers on the treated wells above the corresponding negative control values indicates a mutagenic potential of the test article.

4.2. Data set preparation

Molecules having confounding toxic substructures including azides, nitro, nitroxide, *N*-nitroso were removed as well as molecules with a fixed formal charge such as tetraalkylammonium and pyridinium compounds. Set C was constructed from a download of EPA GENETOX[28] data. Only entries with 'SAL+' or 'SAL−' values for mutagenicity were kept. Chemical structures were produced by converting names into chemical structures using NameToStructure from CambridgeSoft.[65] Ambiguous entries were checked using common name lookup or CAS numbers in SciFinder.[66] Literature entries were translated to chemical structures in similar fashion. In the event of duplicate compounds with different results, the Ames+ value was used. Set D was constructed using the Kazius et al.[31] results provided with Pipeline Pilot 7.5 as an SD file and the file published as Supplementary data in Hansen et al.[32]

4.3. Computational methods

The reaction energies were calculated using density functional calculations with Becke's 3-parameter hybrid functional[67,68] with the LYP correlation functional of Lee, Yang and Parr[69] (B3LYP) as implemented in Gaussian03.[70] A basis set scheme was adopted using 6-31G* for C, H, O, F, P, and S but the LANL2DZ basis and effective core potential[71–73] for Cl, I, and Br. When these halogen atoms were not present, the general basis set specification was still

used. In Gaussian03, the default *d*-orbitals in calculations with general basis set specifications are implemented using pure *d* orbitals, involving five basis functions, whereas Cartesian coordinate *d*-orbitals, using six basis functions, are used by default when 6-31G* is specified. Single-point energies using 6-311G* or geometry optimizations at 6-31+G** in place of 6-31G* in the scheme were investigated for set A.

For set A, frequency calculations were performed to confirm structures as minima and to evaluate the significance of thermal corrections. This was not performed for the larger set though for set A, only a couple structures had to be optimized further and changes in energy and geometry were negligible. For solvation calculations, the Conductor Screening Polarizable Continuum Model (C-PCM) solvation[74] was used.

The Schrodinger LigPrep[75] program was used to generate initial 3D structures optimized using the MMFFs force field, remove salts, neutralize, and generate up to eight tautomers using Tautomerizer, for the aryl–amines taken from the database. If there was an unspecified chiral center, one possible stereoisomer was generated. A conformational search of each of these initial structures, including each of the tautomers, was performed using MacroModel[76] and the MCMM[77,78] protocol with a default setup using the MMFFs[79] force field. The 10 lowest energy conformations of each aryl–amine and of its tautomers were then optimized using density functional theory.

Two planar nitrenium ion configurations were then created from the lowest energy amine for each aryl–amine group, equivalent to the syn/anti configurations of an imine. These were then optimized in the singlet and triplet states. Because the quantum mechanics calculations allow bonds to form or break, it was necessary to check that the connectivity was the same in the amine and the nitrenium ion. As a further check, for sets A–D, the lowest energy amine, singlet nitrenium, and triplet nitrenium structures were visually inspected to ensure the correct reaction was represented. The difference between the lowest amine energy and the lowest nitrenium energy plus the energy of the hydride yielded the nitrenium formation energy.

All statistics and the beanplot graphics were generated using R or Excel. TOPKAT was accessed through the TOPKAT model component provided in Pipeline Pilot.[80]

Acknowledgments

This work would not have been possible without the dedicated resources of the high performance computing team within our internal NIBR IT group, specifically the enthusiastic participation of Steven Litster and Michael Derby. P.M. is a NIBR postdoctoral fellow supported financially by the NIBR Education Office. J.K. was a summer student from Newton South High School, MA.

Supplementary data

Supplementary data (results and geometries for all reaction pathways studied for set A, effect of thermodynamic corrections and solvation in set A, calculations of path A for the larger set B, singlet–triplet gaps for set A nitrenium ions, details of the workflow, set overlap Venn diagram, and a summary of other reported nitrenium formation energy correlations to Ames results. Calculated nitrenium formation energies, Toxcheck/Derek, Toxtree, and TOPKAT predictions for sets A, C, and D with Ames test results are provided with SMILES and InChI 2D descriptors associated with this article can be found, in the online version, at doi:10.1016/j.bmc.2011.03.066. Additionally nitrenium formation energy values for the 14,000 commercially available aryl-amines are available on request.

References and notes

1. Mirza, A.; Desai, R.; Reynisson, J. *Eur. J. Med. Chem.* **2009**, *44*, 5006.
2. Robinson, D. I. *Org. Process Res. Dev.* **2010**, *14*, 946.
3. Skipper, P. L.; Kim, M. Y.; Sun, H. L. P.; Wogan, G. N.; Tannenbaum, S. R. *Carcinogenesis* **2010**, *31*, 50.
4. Weisburger, J. H. *Mutat. Res., Fundam. Mol. Mech. Mutagen.* **2002**, *506-507*, 9.
5. Miller, J. A. *Cancer Res.* **1970**, *30*, 559.
6. Kalgutkar, A. S.; Gardner, I.; Obach, R. S.; Shaffer, C. L.; Callegari, E.; Henne, K. R.; Mutlib, A. E.; Dalvie, D. K.; Lee, J. S.; Nakai, Y.; O'Donnell, J. P.; Boer, J.; Harriman, S. P. *Curr. Drug Metab.* **2005**, *6*, 161.
7. Srivastava, S.; Ruane, P.; Toscano, J.; Sullivan, M.; Cramer, C. J.; Chiapperino, D.; Reed, E.; Falvey, D. *J. Am. Chem. Soc.* **2000**, *122*, 8271.
8. Chakraborty, M.; Jin, K. J.; Brewer, S. C.; Peng, H.-L.; Platz, M. S.; Novak, M. *Org. Lett.* **2009**, *11*, 4862.
9. Winter, A.; Falvey, D.; Cramer, C. J. *J. Am. Chem. Soc.* **2004**, *126*, 9661.
10. McCann, J.; Choi, E.; Yamasaki, E.; Ames, B. N. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 5135.
11. Ames, B. N.; Durston, W. E.; Yamasaki, E.; Lee, F. D. *Proc. Natl. Acad. Sci. U.S.A.* **1973**, *70*, 2281.
12. Fetterman, B. A.; Kim, B. S.; Margolin, B. H.; Schildcrout, J. S.; Smith, M. G.; Wagner, S. M.; Zeiger, E. *Environ. Mol. Mutagen.* **1997**, *29*, 312.
13. International Council on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. Guidance on genotoxicity testing and data interpretation for pharmaceuticals intended for human use. S2 (R1) 2008, p 28.
14. Mortelmans, K. *Mutat. Res., Rev. Mutat. Res.* **2006**, *612*, 151.
15. Westerink, W. M. A.; Stevenson, J. C. R.; Lauwers, A.; Griffioen, G.; Horbach, G. J.; Schoonen, W. G. E. J. *Mutat. Res., Genet. Toxicol. Environ. Mutagen.* **2009**, *676*, 113.
16. Knight, A. W.; Little, S.; Houck, K.; Dix, D.; Judson, R.; Richard, A.; McCarroll, N.; Akerman, G.; Yang, C.; Birrell, L.; Walmsley, R. M. *Regul. Toxicol. Pharmacol.* **2009**, *55*, 188.
17. Bentzien, J.; Hickey, E. R.; Kemper, R. A.; Brewer, M. L. *J. Chem. Inf. Model.* **2010**, *50*, 274.
18. Leach, A. G.; Cann, R.; Tomasi, S. *Chem. Commun.* **2009**, 1094.
19. Borosky, G. *J. Mol. Graphics Model.* **2008**, *27*, 459.
20. Borosky, G. L. *Chem. Res. Toxicol.* **2007**, *20*, 171.
21. Bebernitz, G. R.; Beaulieu, V.; Dale, B. A.; Deacon, R.; Duttaroy, A.; Gao, J.; Grondine, M. S.; Gupta, R. C.; Kakmak, M.; Kavana, M.; Kirman, L. C.; Liang, J.; Maniara, W. M.; Munshi, S.; Nadkarni, S. S.; Schuster, H. F.; Stams, T.; St Denny, I.; Taslimi, P. M.; Vash, B.; Caplan, S. L. *J. Med. Chem.* **2009**, *52*, 6142.
22. Bebernitz, G. R. (Novartis AG). 3-Cyclyl-2-(4-sulfamoyl-phenyl)-*n*-cyclyl-propionamide derivatives useful in the treatment of impaired glucose tolerance and diabetes. Patent WO2007041365, 2007.
23. Bebernitz, G. R. (Novartis AG). Sulfonamide-thiazolpyridine derivatives as glucokinase activators useful the treatment of type 2 diabetes. Patent WO2005095418, 2005.
24. Bebernitz, G. R.; Gupta, R. C.; Jagtap, V. V.; Mundhare, A. B.; Tuli, D. (Novartis AG). Thiazolopyridine derivates, pharmaceutical conditions containing them and methods of treating glucokinase mediated conditions. Patent WO2005095417, 2005.
25. Bebernitz, G. R.; Kirman, L. C. (Novartis AG). Sulfonamide derivatives as glycokinase activators useful in the treatment of type 2 diabetes. Patent WO2007041366, 2007.
26. Bebernitz, G. R. (Novartis A.G.). Substituted (thiazol-2-yl)-amide or sulfonamide as glycokinase activators useful in the treatment of type 2 diabetes. Patent WO2004050645, 2004.
27. Chung, K.; Kirkovsky, L.; Kirkovsky, A.; Purcell, W. *Mutat. Res., Rev. Mutat.* **1997**, *387*, 1.
28. EPA GENETOX. Database, http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?GENETOX, (Accessed June 2009).
29. Benigni, R.; Bossa, C.; Netzeva, T.; Rodomonte, A.; Tsakovska, I. *Environ. Mol. Mutagen.* **2007**, *48*, 754.
30. Benigni, R.; Passerini, L.; Gallo, G.; Giorgi, F.; Cotta-Ramusino, M. *Environ. Mol. Mutagen.* **1998**, *32*, 75.
31. Kazius, J.; McGuire, R.; Bursi, R. *J. Med. Chem.* **2005**, *48*, 312.
32. Hansen, K.; Mika, S.; Schroeter, T.; Sutter, A.; ter Laak, A.; Steger-Hartmann, T.; Heinrich, N.; Müller, K.-R. *J. Chem. Inf. Model.* **2009**, *49*, 2077.
33. Moonen, H.; Briede, J.; van Maanen, J.; Kleinjans, J.; De Kok, T. *Mol. Carcinog.* **2002**, *35*, 196.
34. Watanabe, M.; Igarashi, T.; Kaminuma, T.; Sofuni, T.; Nohmi, T. *Environ. Health Perspect.* **1994**, *102*, 83.
35. Liu, L.; Von Vett, A.; Zhang, N.; Walters, K.; Wagner, C.; Hanna, P. *Chem. Res. Toxicol.* **2007**, *20*, 1300.
36. Chou, H. C.; Lang, N. P.; Kadlubar, F. F. *Carcinogenesis* **1995**, *16*, 413.
37. Orzechowski, A.; Schrenk, D.; Bock-Hennig, B. S.; Bock, K. W. *Carcinogenesis* **1994**, *15*, 1549.
38. Wilson, I. D.; Nicholson, J. K. *Xenobiotica* **2003**, *33*, 887.
39. Afzelius, L.; Hasselgren Arnby, C.; Broo, A.; Carlsson, L.; Isaksson, C.; Jurva, U.; Kjellander, B.; Kolmodin, K.; Nilsson, K.; Raubacher, F.; Weidolf, L. *Drug Metab. Rev.* **2007**, *39*, 61.
40. Guengerich, F. P. *Chem. Res. Toxicol.* **2008**, *21*, 70.
41. Kampstra, P. *J. Stat. Softw.* **2008**, *28*, http://www.jstatsoft.org/v28/c01/paper.
42. Wilcoxon, F. *Biometrics Bulletin* **1945**, *1*, 80.

43. Ford, G. P.; Griffin, G. R. *Chem. Biol. Interact.* **1992**, *81*, 19.
44. Hatch, F. T.; Colvin, M. E. *Mutat. Res., Fundam. Mol. Mech. Mutagen.* **1997**, *376*, 87.
45. Debnath, A. K.; Debnath, G.; Shusterman, A. J.; Hansch, C. *Environ. Mol. Mutagen.* **1992**, *19*, 37.
46. Hartman, G. D.; Schlegel, H. B. *Chem. Biol. Interact.* **1981**, *36*, 319.
47. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Delivery Rev.* **2001**, *46*, 3.
48. Sushko, I.; Novotarskyi, S.; Körner, R.; Pandey, A. K.; Cherkasov, A.; Li, J.; Gramatica, P.; Hansen, K.; Schroeter, T.; Müller, K.-R.; Xi, L.; Liu, H.; Yao, X.; Öberg, T.; Hormozdiari, F.; Dao, P.; Sahinalp, C.; Todeschini, R.; Polishchuk, P.; Artemenko, A.; Kuz'min, V.; Martin, T. M.; Young, D. M.; Fourches, D.; Muratov, E.; Tropsha, A.; Baskin, I.; Horvath, D.; Marcou, G.; Muller, C.; Varnek, A.; Prokopenko, V. V.; Tetko, I. V. *J. Chem. Inf. Model.* **2010**, *50*, 2094.
49. Benigni, R.; Bossa, C.; Netzeva, T.; Worth, A. Collection and Evaluation of (Q)SAR Models for Mutagenicity and Carcinogenicity, Report 22772, Office for Official Publications of the European Communities, 2007.
50. Leong, M. K.; Lin, S.-W.; Chen, H.-B.; Tsai, F.-Y. *Toxicol. Sci.* **2010**, *116*, 498.
51. Stankowski, J. L. F.; Sebastian, J. R. S.; Sterner, R. T. *J. Toxicol. Environ. Health, Part A* **1997**, *50*, 451–462.
52. Benigni, R.; Bossa, C.; Tcheremenskaia, O.; Giuliani, A. *Expert Opin. Drug Metab. Toxicol.* **2010**, *6*, 809.
53. Snyder, R.; Pearl, G.; Mandakas, G.; Choy, W.; Goodsaid, F.; Rosenblum, I. *Environ. Mol. Mutagen.* **2004**, *43*, 143.
54. Snyder, R.; Smith, M. *Drug Discovery Today* **2005**, *10*, 1119.
55. Naven, R. T.; Louise-May, S.; Greene, N. *Expert Opin. Drug Metab. Toxicol.* **2010**, *6*, 797.
56. TOPKAT, 6.2; Accelrys: San Diego, CA 92121.
57. Toxtree, 1.60; IdeaConsult, Ltd: Sofia, Bulgaria.
58. 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. *Toxicology* **1996**, *106*, 267.
59. Sanderson, D. M.; Earnshaw, C. G. *Hum. Exp. Toxicol.* **1991**, *10*, 261.
60. Glowienke, S. *Exp. Toxicol. Pathol.* **2009**, *61*, 264.
61. ASDI building block collections, http://web.asdi.net, ASDI , 601 Interchange Blvd., Newark, DE 19711, USA.
62. SciQuest Chemical inventory management and procurement solutions, http://www.sciquest.com. Sciquest, 6501 Weston Parkway, Suite 200, Cary, NC 27513, USA.
63. Diehl, M. S.; Willaby, S. L.; Snyder, R. D. *Environ. Mol. Mutagen.* **2000**, *35*, 72.
64. Brooks, T. M. *Mutagenesis* **1995**, *10*, 447.
65. Name to Structure Batch, v. 11.0.2, 2008, CambridgeSoft, Cambridge, MA 02140.
66. SciFinder 2007, 2006, Chemical Abstracts Service, Columbus, OH 43210.
67. Becke, A. D. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098.
68. Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
69. Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B: Condens. Matter* **1988**, *37*, 785.
70. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cui, J. V. O., Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T. .; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A., Gaussian 03, Rev. E.01; Gaussian: Wallingford, CT, 2004.
71. Hay, P. J.; Wadt, W. R. J. *Chem. Phys.* **1985**, *82*, 270.
72. Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 299.
73. Wadt, W. R.; Hay, P. J. *J. Chem. Phys.* **1985**, *82*, 284.
74. Cossi, M.; Barone, V. *J. Phys. Chem. A* **1998**, *102*, 1995.
75. LigPrep, 2.2; Schrodinger, L.L.C.: New York, NY, 2005.
76. Macromodel, 9.6; Schrodinger, L.L.C.: New York, NY, 2008.
77. Chang, G.; Guida, W.; Still, W. C. *J. Am. Chem. Soc.* **1989**, *111*, 4379.
78. Saunders, M.; Houk, K. N.; Wu, Y.-D.; Still, W. C.; Lipton, M.; Chang, G.; Guida, W. C. *J. Am. Chem. Soc.* **1990**, *112*, 1419.
79. Halgren, T. A. *J. Comput. Chem.* **1999**, *20*, 730.
80. Pipeline Pilot, 7.5; Accelrys Software: San Diego, CA 92121, 2008.