



SHORT REPORT

Open Access



A local QSAR model based on the stability of nitrenium ions to support the ICH M7 expert review on the mutagenicity of primary aromatic amines

Ayaka Furukawa^{1*} , Satoshi Ono^{2*} , Katsuya Yamada¹, Nao Torimoto², Mahoko Asayama¹ and Shigeharu Muto¹ 

Abstract

Background: Aromatic amines, often used as intermediates for pharmaceutical synthesis, may be mutagenic and therefore pose a challenge as metabolites or impurities in drug development. However, predicting the mutagenicity of aromatic amines using commercially available, quantitative structure–activity relationship (QSAR) tools is difficult and often requires expert review. In this study, we developed a shareable QSAR tool based on nitrenium ion stability.

Results: The evaluation using in-house aromatic amine intermediates revealed that our model has prediction accuracy of aromatic amine mutagenicity comparable to that of commercial QSAR tools. The effect of changing the number and position of substituents on the mutagenicity of aromatic amines was successfully explained by the change in the nitrenium ion stability. Furthermore, case studies showed that our QSAR tool can support the expert review with quantitative indicators.

Conclusions: This local QSAR tool will be useful as a quantitative support tool to explain the substituent effects on the mutagenicity of primary aromatic amines. By further refinement through method sharing and standardization, our tool can support the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) M7 expert review with quantitative indicators.

Keywords: Primary aromatic amine, Quantitative structure–activity relationship (QSAR), In silico, Nitrenium ion, ICH M7, Expert review

Introduction

The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) M7 guideline describes the process of hazard identification and risk assessment of impurities that may

be present in a drug substance or product [1]. Risk assessment allows the assessment of mutagenicity using quantitative structure–activity relationship (QSAR) tools, which enables the screening of numerous potential impurities. ICH M7 recommends that bacterial mutagenicity should be assessed using two QSAR methodologies, namely, by expert rule-based and statistics-based methodologies. The prediction accuracy of QSAR tools is improved by increasing the number of training sets and modifying algorithms [2, 3]. However, the mutagenicity of several structures, such as aromatic amines, is still difficult to predict when using QSAR tools [4, 5].

* Correspondence: furukawa.ayaka@ma.mt-pharma.co.jp; ono.satoshi@mg.mt-pharma.co.jp

¹Safety Research Laboratories, Innovative Research Division, Mitsubishi Tanabe Pharma Corporation, 26-1, Muraoka-Higashi 2-chome, Fujisawa, Kanagawa 251-8555, Japan

²Discovery Technology Laboratories, Innovative Research Division, Mitsubishi Tanabe Pharma Corporation, 1000, Kamoshida-cho, Aoba-ku, Yokohama, Kanagawa 227-0033, Japan



© The Author(s). 2022 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Because aromatic amines are widely used as intermediates in pharmaceutical synthesis and can remain as impurities, their mutagenicity is a serious safety issue [4, 5]. In addition, aromatic amines can be formed as metabolites, especially for drugs with amide bonds that may easily break down by enzymatic hydrolysis [6, 7].

Predicting the mutagenicity of aromatic amines remains a major challenge for QSAR tools because positive predictions are sometimes false positives. Therefore, an improvement of prediction accuracy is strongly desired from the viewpoint of regulatory science.

In recent years, several efforts have been made to improve the prediction accuracy of QSAR tools. One method is to share nonpublic data held by pharmaceutical companies, as existing QSAR tools have been built based on public knowledge [4, 5]. Another method is to perform expert reviews. The SAR fingerprint, a chemical fingerprint of aromatic amine mutagenicity developed by Ahlberg et al. [4], is a useful approach to performing expert reviews [4, 8–10]. However, the SAR fingerprint approach is difficult to apply to mutagenicity assessment when activating and deactivating substituents are simultaneously present. We considered whether quantitative predictive indices based on mutagenic mechanisms can complement the SAR fingerprinting approach.

The mechanism underlying mutagenic induction of aromatic amines has been well studied (Fig. 1). In the metabolic activation of aromatic amines, first, hydroxylamine is generated by *N*-hydroxylation by cytochrome P450 enzymes (CYPs; mainly CYP1A2). The hydroxylamine is either conjugated by phase II enzymes (*O*-acetyltransferases, *N*-acetyltransferases, or sulfotransferases) or directly hydrolyzed to nitrenium ions, which form covalent bonds with DNA [11]. Several local QSAR tools have been evaluated in terms of ease of hydroxylamine formation [12, 13], ease of nitrenium ion formation [14], and nitrenium ion stability [15–17].

These local QSAR tools are useful, but they are not user friendly for genotoxicologists. Therefore, this study developed a local QSAR model based on nitrenium ion stability that would be easy to use for genotoxicologists

and would support ICH M7 expert reviews with quantitative indicators.

Materials and methods

Data set

Two data sets were used in this study, namely, an in-house primary aromatic amine data set used for evaluating QSAR tool performance and a data set of 23 known primary aromatic amines. The in-house data set contained 85 aromatic amines, of which 51 were mutagenic, and 34 were nonmutagenic, and they were collected using the following procedure. Primary aromatic amines with standard Ames [18], fluctuation Ames [19], or Vito-tox [20] test results were extracted from the in-house compound database. The aromatic amines were classified as follows: mutagenic (clearly positive under any condition in any of the tests) and nonmutagenic (negative in at least three strains (TA98, TA100, TA1537 [or TA2637]) in the standard Ames test (in the absence and presence of S9)). We further narrowed down the aromatic amines using six criteria, as described by Bentzien et al. [16]: (i) no formal charge, (ii) molecular weight of < 500 Da, (iii) maximum one stereocenter, (iv) < 10 rotatable bonds, (v) only one aromatic amine functionality, and (vi) no aromatic nitro groups.

The data set of known primary aromatic amines consisted of the following structures: a) aniline as a standard, b) aniline with methyl groups (7 structures), c) aniline with a methoxy group (3 structures), d) aniline with a sulfonate group (3 structures), e) aniline with a sulfonamide group (2 structures), and f) primary aromatic amines for which expert reviews have been reported [4, 8, 9] and their analogs (7 structures).

Commercial QSAR tools

In this study, the following commercial mutagenicity prediction QSAR tools were used: Derek Nexus, version 6.1.0, 2020.1.0 version of the knowledge base (Lhasa Limited, UK); CASE Ultra rule-based, version 1.8.0.5, GT_EXPERT v1.8.0.1.16392.500 (MultiCASE Inc., USA); CASE Ultra statistical based, version 1.8.0.5, GT1_

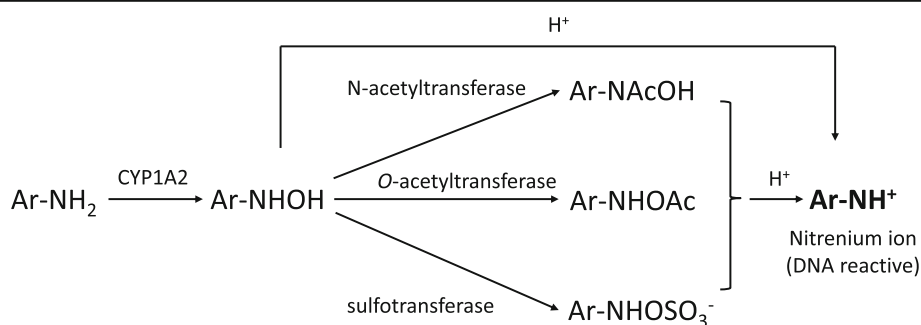


Fig. 1 Mechanism of mutagenicity induced by aromatic amines

BMUT v1.8.0.1.11479.500 (MultiCASE Inc., USA); and Leadscope Model Applier, version 3.0.1–1, Bacterial Mutation v2 (Leadscope Inc., USA).

Calculating the stability of nitrenium ions

Bentzien et al. reported an *in silico* method of predicting the Ames mutagenicity of primary aromatic amines using the nitrenium ion hypothesis of Ford et al. [15, 16]. The mutagenic effect of primary aromatic amines is described as the formation of a reactive nitrenium ion, and the authors concluded that nitrenium ion stability correlates with the mutagenic potential. They calculated nitrenium ion stability using a semi-empirical quantum mechanical method and the relative energy $\Delta\Delta E$:

$$\Delta\Delta E = \Delta E_{ArNH^+} + \Delta E_{PhNH_2} - \Delta E_{ArNH_2} - \Delta E_{PhNH^+}$$

We used the heat of formation energies of Austin model 1 (AM1) [21] optimized structures calculated using MOPAC v7.1 bundled with Molecular Operation Environment (MOE) 2019.01 software (Chemical Computing Group ULC, Canada). To calculate $\Delta\Delta E$, we first prepared a structure-data file of the molecule and constructed a 3D structure using the “rebuild 3D” feature of MOE. Then, we designed and performed conformational sampling to accurately calculate the conformational state using LowModeMD [22] with a force field (e.g., MMFF94x) for each neutral molecule within an energy cutoff of 7 kcal/mol. Subsequently, we performed geometry optimization using semi-empirical quantum mechanical calculations with AM1 Hamiltonian for each neutral conformer. The most stable conformer (i.e., with the lowest heat of formation) was selected to determine the nitrenium ion species, one of the amine hydrogens was replaced by a dummy atom X, and geometry optimization was again performed using the keyword CHARGE = +1. The lowest $\Delta\Delta E$ value was adopted for this molecule. The $\Delta\Delta E$ of aniline was set to 0 kcal/mol. If the geometry optimization was not converged, a not-a-number (NaN) was assigned for that formula. A compound with a negative $\Delta\Delta E$ was predicted to be mutagenic, and a compound with a positive $\Delta\Delta E$ was predicted to be nonmutagenic. The main output file was in csv format; therefore, we could easily use spreadsheet programs. The csv file also had a remark column for errors, for example, a NaN, an aromatic ring opening during the geometry optimization, or no aromatic amine.

This procedure was written in Scientific Vector Language (SVL), which is integrated in the MOE modeling package and is freely available to MOE users from MOL-SIS Inc. upon request.

Analysis of QSAR tool performance

To evaluate the predictive performance of the QSAR model, we used compounds that are not included in known databases or commercial QSAR training sets. Therefore, the performance of the QSAR tools was evaluated using in-house aromatic amine data only. The performance metrics employed were accuracy, sensitivity, specificity, positive prediction value, negative prediction value, Matthews correlation coefficient (MCC), and coverage [23].

Results

Analysis of QSAR tool performance using in-house aromatic amines

Figure 2 shows the $\Delta\Delta E$ of each in-house aromatic amine and Ames test results. The $\Delta\Delta E$ values were small for mutagens and large for nonmutagens. Table 1 shows the predictive performance of commercial QSAR tools and $\Delta\Delta E$ against the in-house aromatic amines.

The prediction accuracy of our in-house aromatic amines using $\Delta\Delta E$ was about 70%, which is comparable to the reported prediction accuracy of commercial QSAR tools for aromatic primary amines. The prediction accuracies of commercial QSAR tools for primary aromatic amines are 59 to 64% for 599 test compounds and can be improved to about 63 to 79% by adding new data sets and revising the rules [4]. For another 268 test compounds, the reported accuracy ranged from 57 to 73% [5].

The prediction accuracies of commercially available QSAR tools for in-house developed aromatic amines were about 45–58%, which are slightly lower than the reported value. The performance of commercial QSAR tools depends on the quantity, quality, and diversity of the training set. Therefore, we considered the possibility of improving the accuracy by training with our proprietary compounds. On the other hand, since the $\Delta\Delta E$ calculation lacks the concept of learning by structure sets, predictions for new structures are not expected to affect the accuracy.

“Quantitative” SAR fingerprint approach

Figures 3 show the results of quantitative chemical fingerprinting (termed “SAR fingerprint”) of aromatic amines for the representative substituents. According to Ahlberg et al. [4], methyl groups in ortho-, meta-, and para-positions are activating substituents. One methyl group each in the ortho-, meta-, and para-positions decreased $\Delta\Delta E$. The $\Delta\Delta E$ was lower for aromatic amines with two or three methyl groups, which tested positive (mutagenic) in the Ames test. For the methoxy group, which is an activating substituent, the $\Delta\Delta E$ decreased, except when the methoxy group was in the meta-position, and the aromatic amines with these groups tested positive in the Ames test. Sulfonate and

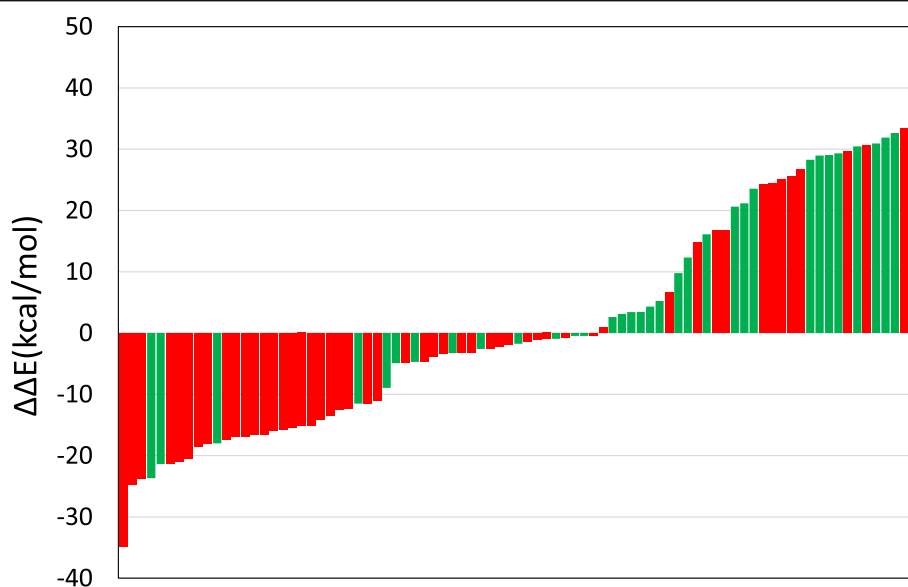


Fig. 2 Mutagenicity and $\Delta\Delta E$ distribution of in-house primary aromatic amines. Each bar indicates a compound. Mutagens are shown in red and nonmutagens in green. The aromatic amines are arranged in order of their $\Delta\Delta E$ values (kcal/mol)

sulfonamide groups, which are considered deactivating substituents, increased the $\Delta\Delta E$ in the ortho-, meta-, and para-positions, and aromatic amines with these groups tested negative in the Ames test.

ICH M7 expert review case studies

The applicability of the $\Delta\Delta E$ method was examined for several compounds for which expert reviews based on ICH M7 have been reported (Table 2).

Case study 1

4-Amino-3-methylbenzenesulfonic acid, which is No. 17 in Table 2, is compound D in the report by Ahlberg et al. [4]. Given that this compound has strong activating (methyl in the ortho-position) and deactivating (sulfonate in the para-position) substituents simultaneously, we opted

to use it for quantitative evaluation. In addition, their paper did not specify the results predicted by commercial QSAR tools for this compound, and our study showed conflicting results with those from commercial QSAR tools (Table 2).

The methyl group in the ortho-position slightly decreased the $\Delta\Delta E$, whereas the sulfonate group in the para-position caused a significant increase. The $\Delta\Delta E$ of this compound was approximately the sum of the $\Delta\Delta E$ values of aromatic amines with a methyl group in the ortho-position and a sulfonic acid group in the para-position, and it was predicted to be nonmutagenic and tested negative in the Ames test (Fig. 3, Table 2) [33]. The results suggest that the quantitative index complements the SAR fingerprint approach of aromatic amine mutagenicity developed by Ahlberg et al. [4].

Table 1 Comparison of the predictive performance of QSAR tools for in-house aromatic amines

	Derek Nexus	GT_EXPERT	GT1_BMUT	Leadscope Model Applier	$\Delta\Delta E$
Accuracy	57.6%	44.7%	57.6%	50.0%	69.4%
Sensitivity	35.3%	39.2%	62.5%	43.2%	74.5%
Specificity	91.2%	52.9%	50.0%	61.5%	61.8%
Positive prediction value	85.7%	55.6%	65.8%	65.5%	74.5%
Negative prediction value	48.4%	36.7%	46.4%	39.0%	61.8%
Matthews correlation coefficient	0.3	-0.1	0.1	0.0	0.4
Coverage	100%	100%	77.6%	82.3%	100%
Inconclusive	N/A	N/A	22.4%	10.6%	N/A
Out of domain	N/A	N/A	N/A	7.1%	N/A

QSAR Quantitative structure-activity relationship, N/A Not applicable

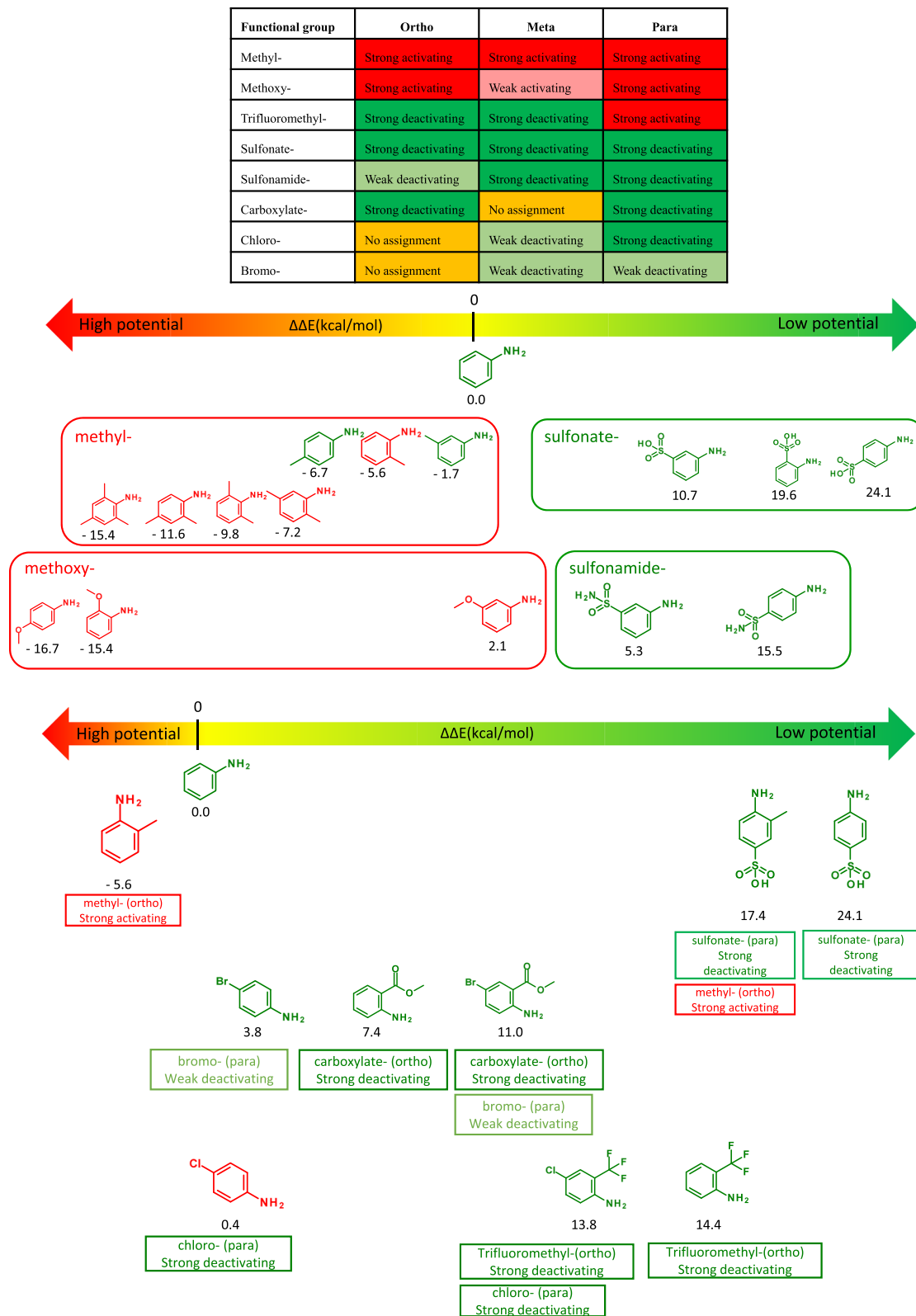


Fig. 3 Effect of activating and deactivating substituents on $\Delta\Delta E$ and mutagenicity of aromatic amines. The structures of mutagens are shown in red and those of nonmutagens in green. The numbers below the structures indicate $\Delta\Delta E$ values. The text in the squares shows the classification by Ahlberg et al

Table 2 Aromatic amines used in the “Quantitative” SAR fingerprint approach

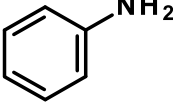
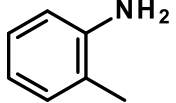
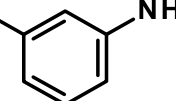
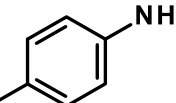
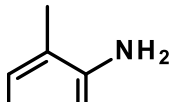
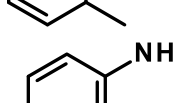
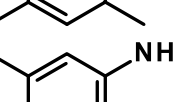
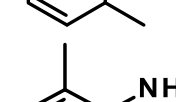
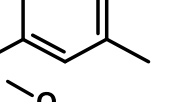
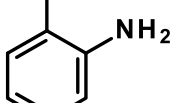
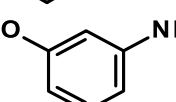
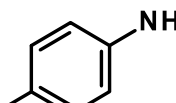
No.	Structure	CAS number	Ames	$\Delta\Delta E$ (kcal/mol)	GT1_BMUT	Leadscope	Derek	GT_Expert	Ref.
1		142-04-1	Non mutagen	0	Inconclusive	Indeterminate	Inactive	Known Negative	[24]
2		95-53-4	Mutagen	-5.6	Positive	Indeterminate	Inactive	Known Positive	[25]
3		108-44-1	Non mutagen	-1.7	Inconclusive	Indeterminate	Inactive	Known Negative	[26]
4		106-49-0	Non mutagen	-6.7	Positive	Active	Inactive	Known Negative	[26]
5		87-62-7	Mutagen	-9.8	Negative	Indeterminate	Plausible	Known Positive	[27]
6		95-68-1	Mutagen	-11.6	Positive	Active	Plausible	Known Positive	[27]
7		95-78-3	Mutagen	-7.2	Positive	Indeterminate	Plausible	Known Positive	[27]
8		88-05-1	Mutagen	-15.4	Inconclusive	Active	Plausible	Known Positive	[28]
9		134-29-2	Mutagen	-15.4	Positive	Active	Plausible	Known Positive	[29]
10		536-90-3	Mutagen	2.1	Positive	Inactive	Plausible	Known Positive	[30]
11		104-94-9	Mutagen	-16.7	Positive	Active	Plausible	Known Positive	[30]
12		88-21-1	Non mutagen	19.6	Positive	Inactive	Inactive	Known Negative	[31]

Table 2 Aromatic amines used in the “Quantitative” SAR fingerprint approach (Continued)

No.	Structure	CAS number	Ames	$\Delta\Delta E$ (kcal/mol)	GT1_BMUT	Leadscope	Derek	GT_Expert	Ref.
13		121-47-1	Non mutagen	10.7	Negative	Inactive	Inactive	Known Negative	[27]
14		121-57-3	Non mutagen	24.1	Negative	Inactive	Inactive	Known Negative	[27]
15		98-18-0	Non mutagen	5.3	Negative	Inactive	Inactive	Negative	[5]
16		63-74-1	Non mutagen	15.5	Negative	Inactive	Inactive	Known Negative	[32]
17		98-33-9	Non mutagen	17.4	Positive	Inactive	Inactive	Known Negative	[4, 33]
18		52727-57-8	Non mutagen	11	Negative	Inactive	Inactive	Known Negative	[8, 34]
19		445-03-4	Non mutagen	13.8	Negative	Inactive	Inactive	Known Negative	[9]
20		106-40-1	Non mutagen	3.8	Positive	Inactive	Inactive	Known Negative	[27]
21		134-20-3	Non mutagen	7.4	Negative	Inactive	Inactive	Known Negative	[35]
22		88-17-5	Non mutagen	14.4	Negative	Inactive	Inactive	Known Negative	[36]
23		106-47-8	Mutagen	0.4	Positive	Inactive	Inactive	Known Positive	[29]

Case study 2

Methyl 2-amino-5-bromobenzoate, which is No. 18 in Table 2, is example 10 in the expert review by Amberg et al. [8]. We used this compound as a case study because their paper reported that commercial QSAR tools give conflicting results. This compound is reported to be nonmutagenic by the expert rule-based model and inconclusive by the statistics-based model.

The bromo group in the para-position and carboxylate in the ortho-position increased the $\Delta\Delta E$, and the compound was predicted to be less mutagenic (Fig. 3, Table 2) and tested negative in the Ames test [34]. The bromo group in the para-position and carboxylate in the ortho-position are reported to be deactivating substituents [4]; thus, the results support the expert review with quantitative indicators.

Case study 3

2-Amino-5-chlorobenzotrifluoride, which is No. 19 in Table 2, is case 5 in the report by Mishima et al. [9]. This compound was also used as a case study because their paper reported conflicting results with a commercial QSAR tool. This compound is reported to be nonmutagenic by the expert rule-based model and mutagenic by the statistics-based model.

The chloro group in the para-position slightly increased the $\Delta\Delta E$, and the compound was reported to be mutagenic. The compound with a trifluoromethyl group in the ortho-position had a large $\Delta\Delta E$ and was nonmutagenic. The $\Delta\Delta E$ of this compound with both substituents was approximately the sum of the $\Delta\Delta E$ values of aromatic amines with a chloro group in the para-position and a trifluoromethyl group in the ortho-position and was predicted to be nonmutagenic and tested negative in the Ames test (Fig. 3, Table 2) [37]. The chloro group in the para-position and carboxylate in the ortho-position are reported to be deactivating substituents [4]; thus, the results also support the expert review with quantitative indicators.

In the known primary aromatic amine data set, several compounds gave conflicting results in the commercial QSAR tools (Table 2). The mutagens, namely compounds No. 5, 7, and 8, were accurately predicted in expert rule-based models but were predicted negative, inconclusive, or indeterminate in statistics-based models. These compounds were predicted to be mutagenic with $\Delta\Delta E$ values of -9.8 , -7.2 , and -15.4 kcal/mol, respectively, supporting the prediction by expert rule-based models. Compounds No. 12 and 17, which are nonmutagenic, were correctly predicted by the expert rule-based models but were predicted as positive in the statistics-based model. These compounds were predicted to be nonmutagenic with $\Delta\Delta E$ values of 19.6 and 17.4 kcal/mol, respectively, supporting the prediction by the

expert rule-based model. Accurate prediction of the mutagenicity of compounds with $\Delta\Delta E$ values in the range of about ± 5 kcal/mol was difficult, suggesting that an appropriate cutoff value should be set [16].

Discussion

Mechanism-based local QSAR tools are useful; however, they are not user friendly for genotoxicologists and have not been standardized. To make such tools widely applicable to safety assessment for industry and regulatory purposes, we developed a shareable procedure in SVL for predicting aromatic amine mutagenicity based on the concept described by Bentzien et al. [16]. The difference is that in our method, $\Delta\Delta E$ is calculated after selecting the most stable conformer by a conformational search using a force field. This difference makes it easier for genotoxicologists to use our method, as they can now perform everything from compound structure preprocessing to $\Delta\Delta E$ calculations using the MOE software.

The evaluation using in-house aromatic amine intermediates showed that our QSAR tool has a mutagenicity prediction accuracy comparable to that of commercial QSAR tools. Given the complex structure of in-house compounds, characterization of all the compounds that were not predicted correctly was difficult. However, the compounds that resulted in false positive predictions had large substituents. The presence of bulky substituents can inhibit the metabolic activation of aromatic amines by CYP1A2 due to steric hindrance. Alternatively, electron-withdrawing substituents can have a resonance effect on aromatic rings, reducing electron density and disrupting the electron distribution necessary for metabolic activation [13]. These factors are not reflected in the calculation of $\Delta\Delta E$ values, which may lead to false positive predictions.

We also investigated the possibility of supporting the SAR fingerprint of aromatic amines [4], which involves a qualitative evaluation. The effect of changing the number and position of substituents on the mutagenicity of aromatic amines is successfully explained by the change in $\Delta\Delta E$. Furthermore, case studies using several aromatic amines evaluated by the ICH M7 expert review showed that our QSAR tool can support the expert review with quantitative indicators. In particular, when activating and deactivating substituents coexist, as in case study 1, it is difficult to predict mutagenicity by qualitative evaluation; however, we believe that our method can provide important information.

Using in-house compound data set, we examined the cutoff value of $\Delta\Delta E$ in increments of 2.5 kcal/mol between -10 and $+10$ kcal/mol and found that accuracy and MMC were the highest when the cutoff value was based on the value of $+2.5$ kcal/mol (Table 3). The cutoff value may change depending on the validation

Table 3 Examination of the cutoff value of $\Delta\Delta E$ using in-house compound data set

Cutoff value (kcal/mol)	-10.0	-7.5	-5.0	-2.5	0	+2.5	+5.0	+7.5	+10.0
Accuracy	63.5%	62.4%	62.4%	65.9%	69.4%	70.6%	64.7%	64.7%	63.5%
Sensitivity	85.7%	82.8%	82.8%	77.5%	74.5%	75.0%	68.4%	67.8%	66.7%
Specificity	52.6%	51.8%	51.8%	55.6%	61.8%	63.6%	57.1%	57.7%	56.0%
Positive prediction value	47.1%	47.1%	47.1%	60.8%	74.5%	76.5%	76.5%	78.4%	78.4%
Negative prediction value	88.2%	85.3%	85.3%	73.5%	61.8%	61.8%	47.1%	44.1%	41.2%
Matthews correlation coefficient	-0.2	-0.2	-0.2	-0.2	0.4	0.4	0.2	0.2	0.2
Coverage	100%	100%	100%	100%	100%	100%	100%	100%	100%
Inconclusive	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Out of domain	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

N/A Not applicable

compound used, and further investigation using more compounds is necessary.

In this paper, we have shown the usefulness of the model focusing on aromatic amines. However, in the mutagenic mechanism of aromatic nitros, the nitro group is reduced by nitroreductase to produce nitrenium ions [38]. Although the potential of our method for the application to aromatic nitros has not been investigated, our method can be possibly applied.

This local QSAR model is different from the two complementary QSARs (expert rule-based and statistics-based) recommended in ICH M7. ICH M7 recommends that analyses using these QSAR models should be performed first, and expert review should be considered when necessary. We hope that this local QSAR model is useful in making decisions during expert reviews.

Further validation using more known and undisclosed aromatic amines may be necessary to clarify the applicability of our method. To facilitate its standardization, the SVL script for the method is provided free of charge to MOE users by MOLSIS Inc. We hope that further refinement of this method will contribute to the standardization of expert reviews on the prediction of mutagenicity of aromatic amines under the ICH M7 guideline.

Conclusions

A shareable MOE SVL script was developed for predicting the mutagenicity of primary aromatic amines based on nitrenium ion stability. This local QSAR tool will be useful as a quantitative support tool to explain the substituent effects on the mutagenicity of primary aromatic amines. By further refinement, this tool can support the ICH M7 expert review with quantitative indicators.

Abbreviations

ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; MOE: Molecular Operation Environment; AM1: Austin model 1; QSAR: Quantitative structure–activity relationship; SAR: structure–activity relationship; SVL: Scientific Vector Language;

MCC: Matthews correlation coefficient; CAS: Chemical Abstracts Service; NaN: Not-a-number

Acknowledgements

We thank Dr. Hiroki Kano for his useful advice on this research. We would also like to thank Mr. Yoshirou Kimura of MOLSIS Inc. for his help in creating the MOE SVL script.

MOE script

The SVL script "mut_nitre.svl" can be obtained from MOLSIS Inc. Please contact ccg@molsis.co.jp to request it.

Authors' contributions

AF and SO designed the project and performed the experiment and data analyses. AF and SO drafted the manuscript. KY, NT, and MA reviewed the manuscript. SM supervised the project. All authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 26 July 2021 Accepted: 14 February 2022

Published online: 21 March 2022

References

1. ICH. ICH harmonized guideline. Assessment and control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals to Limit Potential Carcinogenic Risk. Current Step 4 version dated 31 March 2017. <https://www.ich.org/home.html>. Accessed 31 Jan 2022.
2. Honma M, Kitazawa A, Cayley A, Williams RV, Barber C, Hanser T, et al. Improvement of quantitative structure–activity relationship (QSAR) tools for predicting Ames mutagenicity: outcomes of the Ames/QSAR international challenge project. *Mutagenesis*. 2019;34(1):3–16. <https://doi.org/10.1093/mutage/gey031>.
3. Hasselgren C, Bercu J, Cayley A, Cross K, Glowienke S, Kruhlak N, et al. Management of pharmaceutical ICH M7 (Q)SAR predictions – the impact of

- model updates. *Regul Toxicol Pharmacol.* 2020;118:104807. <https://doi.org/10.1016/j.yrtph.2020.104807>.
4. Ahlberg E, Amberg A, Beilke LD, Bower D, Cross KP, Custer L, et al. Extending (Q)SARs to incorporate proprietary knowledge for regulatory purposes: a case study using aromatic amine mutagenicity. *Regul Toxicol Pharmacol.* 2016;77:1–12. <https://doi.org/10.1016/j.yrtph.2016.02.003>.
 5. Patel M, Kranz M, Munoz-Muriedas J, Harvey JS, Giddings A, Swallow S, et al. A pharma-wide approach to address the genotoxicity prediction of primary aromatic amines. *Comput Toxicol.* 2018;7:27–35. <https://doi.org/10.1016/j.comtox.2018.06.002>.
 6. Galloway SM. International regulatory requirements for genotoxicity testing for pharmaceuticals used in human medicine, and their impurities and metabolites: genotoxicity testing for pharmaceuticals and impurities. *Environ Mol Mutagen.* 2017;58(5):296–324. <https://doi.org/10.1002/em.22077>.
 7. Zeller A, Brigo A, Brink A, Guerard M, Lang D, Muster W, et al. Genotoxicity assessment of drug metabolites in the context of MIST and beyond. *Chem Res Toxicol.* 2020;33(1):10–9. <https://doi.org/10.1021/acs.chemrestox.9b00348>.
 8. Amberg A, Beilke L, Bercu J, Bower D, Brigo A, Cross KP, et al. Principles and procedures for implementation of ICH M7 recommended (Q)SAR analyses. *Regul Toxicol Pharmacol.* 2016;77:13–24. <https://doi.org/10.1016/j.yrtph.2016.02.004>.
 9. Mishima M, Hashizume T, Haranosono Y, Nagato Y, Takeshita K, Fukuchi J, et al. Meeting report, ICH M7 relevant workshop: use of (Q)SAR systems and expert judgment. *Genes Environ.* 2018;40(1):19. <https://doi.org/10.1186/s41021-018-0107-2>.
 10. Norinder U, Myatt G, Ahlberg E. Predicting aromatic amine mutagenicity with confidence: a case study using conformal prediction. *Biomolecules.* 2018;8(3):85. <https://doi.org/10.3390/biom8030085>.
 11. Benigni R, Bossa C. Mechanisms of chemical carcinogenicity and mutagenicity: a review with implications for predictive toxicology. *Chem Rev.* 2011;111(4):2507–36. <https://doi.org/10.1021/cr100222q>.
 12. Loew GH, Poulsen M, Kirkjian E, Ferrell J, Sudhindra BS, Rebagliati M. Computer-assisted mechanistic structure-activity studies: application to diverse classes of chemical carcinogens. *Environ Health Perspect.* 1985;61:69–96. <https://doi.org/10.1289/ehp.856169>.
 13. Shamovsky I, Ripa L, Börjesson L, Mee C, Nordén B, Hansen P, et al. Explanation for Main features of structure–genotoxicity relationships of aromatic amines by theoretical studies of their activation pathways in CYP1A2. *J Am Chem Soc.* 2011;133(40):16168–85. <https://doi.org/10.1021/ja206427u>.
 14. McCarren P, Springer C, Whitehead L. An investigation into pharmaceutically relevant mutagenicity data and the influence on Ames predictive potential. *J Cheminformatics.* 2011;3(1):51. <https://doi.org/10.1186/1758-2946-3-51>.
 15. Ford GP, Griffin GR. Relative stabilities of nitrenium ions derived from heterocyclic amine food carcinogens: relationship to mutagenicity. *Chem Biol Interact.* 1992;81(1–2):19–33. [https://doi.org/10.1016/0009-2797\(92\)90024-f](https://doi.org/10.1016/0009-2797(92)90024-f).
 16. Bentzien J, Hickey ER, Kemper RA, Brewer ML, Dyekjær JD, East SP, et al. An in silico method for predicting Ames activities of primary aromatic amines by calculating the stabilities of Nitrenium ions. *J Chem Inf Model.* 2010;50(2):274–97. <https://doi.org/10.1021/ci900378x>.
 17. Kuhnke L, ter Laak A, Göller AH. Mechanistic reactivity descriptors for the prediction of Ames mutagenicity of primary aromatic amines. *J Chem Inf Model.* 2019;59(2):668–72. <https://doi.org/10.1021/acs.jcim.8b00758>.
 18. Maron DM, Ames BN. Revised methods for the Salmonella mutagenicity test. *Mutat Res.* 1983;113(3–4):173–215. [https://doi.org/10.1016/0165-1161\(83\)90010-9](https://doi.org/10.1016/0165-1161(83)90010-9).
 19. Sui H, Kawakami K, Sakurai N, Hara T, Nohmi T. Improvement and evaluation of high throughput fluctuation Ames test using 384-well plate with Salmonella typhimurium TA100 and TA98. *Genes Environ.* 2009;31(2):47–55. <https://doi.org/10.3123/jemsge.31.47>.
 20. Muto S, Baba H, Uno Y. Evaluation of the Vitotox™ test as a high-throughput genotoxicity assay. *Environ Mutagen Res.* 2003;25(2):69–75. <https://doi.org/10.3123/jems.25.69>.
 21. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP. Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. *J Am Chem Soc.* 1985;107(13):3902–9. <https://doi.org/10.1021/ja00299a024>.
 22. Labute P. LowModeMD—implicit low-mode velocity filtering applied to conformational search of macrocycles and protein loops. *J Chem Inf Model.* 2010;50(5):792–800. <https://doi.org/10.1021/ci900508k>.
 23. Honma M. An assessment of mutagenicity of chemical substances by (quantitative) structure–activity relationship. *Genes Environ.* 2020;42(1):23. <https://doi.org/10.1186/s41021-020-00163-1>.
 24. Dunkel VC, Zeiger E, Brusick D, McCoy E, McGregor D, Mortelmans K, et al. Reproducibility of microbial mutagenicity assays: I. tests with Salmonella typhimurium and Escherichia coli using a standardized protocol. *Environ Mutagen.* 1984;6(S2):1–50. <https://doi.org/10.1002/em.286006702>.
 25. Aiba née Kaneko M, Hirota M, Kouzuki H, Mori M. Prediction of genotoxic potential of cosmetic ingredients by an in silico battery system consisting of a combination of an expert rule-based system and a statistics-based system. *J Toxicol Sci.* 2015;40(1):77–98. <https://doi.org/10.2131/jts.40.77>.
 26. Kamber M, Fluckiger-Isler S, Engelhardt G, Jaech R, Zeiger E. Comparison of the Ames II and traditional Ames test responses with respect to mutagenicity, strain specificities, need for metabolism and correlation with rodent carcinogenicity. *Mutagenesis.* 2009;24(4):359–66. <https://doi.org/10.1093/mutage/geb017>.
 27. Zeiger E, Anderson B, Haworth S, Lawlor T, Mortelmans K. Salmonella mutagenicity tests: IV. Results from the testing of 300 chemicals. *Environ Mol Mutagen.* 1988;11(S12):1–18. <https://doi.org/10.1002/em.2850110602>.
 28. Kugler-Steigmeier ME, Friedrich U, Graf U, Maier P, Schlatter C. Testing of 2,4,5- and 2,4,6-trimethylaniline in the Salmonella assay, in mammalian cell cultures, and in *Drosophila melanogaster* and comparison of the results with carcinogenicity data. *Mutat Res Mutagen Relat Subj.* 1988;203:237–8.
 29. Dunkel VC, Zeiger E, Brusick D, McCoy E, McGregor D, Mortelmans K, et al. Reproducibility of microbial mutagenicity assays: II. Testing of carcinogens and noncarcinogens in Salmonella typhimurium and Escherichia coli. *Environ Mutagen.* 1985;7(S5):1–19. <https://doi.org/10.1002/em.2860070902>.
 30. Zeiger E, Anderson B, Haworth S, Lawlor T, Mortelmans K. Salmonella mutagenicity tests: V. results from the testing of 311 chemicals. *Environ Mol Mutagen.* 1992;19(S21):2–141. <https://doi.org/10.1002/em.2850190603>.
 31. European Chemicals Agency's Dissemination portal with information on chemical substances registered under REACH. Testing status of o-Amino benzenesulfonic acid. Available from: <https://echa.europa.eu/registration-dossier/-/registered-dossier/10511/7/7/2>. Accessed 31 Jan 2022.
 32. National Toxicology Program, Genetic Toxicity Evaluation of Sulfanilamide in Salmonella/E.coli Mutagenicity Test or Ames Test. Study 664069. Available from: <https://cebs.niehs.nih.gov/cebs/study/002-02856-0003-0000-6>. Accessed 31 Jan 2022.
 33. European Chemicals Agency's Dissemination portal with information on chemical substances registered under REACH. Testing status of 3-methylsulphanilic acid. Available from: <https://echa.europa.eu/registration-dossier/-/registered-dossier/17643/7/7/2>. Accessed 31 Jan 2022.
 34. Greene N, Dobo KL, Kenyon MO, Cheung J, Munzner J, Sobol Z, et al. A practical application of two in silico systems for identification of potentially mutagenic impurities. *Regul Toxicol Pharmacol.* 2015;72(2):335–49. <https://doi.org/10.1016/j.yrtph.2015.05.008>.
 35. Mortelmans K, Haworth S, Lawlor T, Speck W, Tainer B, Zeiger E. Salmonella mutagenicity tests: II. Results from the testing of 270 chemicals. *Environ Mutagen.* 1986;8(S7):1–52. <https://doi.org/10.1002/em.2860080802>.
 36. European Chemicals Agency's Dissemination portal with information on chemical substances registered under REACH. Testing status of α,α -trifluoro-o-toluidine. Available from: <https://echa.europa.eu/registration-dossier/-/registered-dossier/11316/7/7/2>. Accessed 31 Jan 2022.
 37. European Chemicals Agency's Dissemination portal with information on chemical substances registered under REACH. Testing status of 4-chloro- α,α -trifluoro-o-toluidine. Available from: <https://echa.europa.eu/registration-dossier/-/registered-dossier/12399/7/7/2>. Accessed 31 Jan 2022.
 38. Purohit V, Basu AK. Mutagenicity of Nitroaromatic compounds. *Chem Res Toxicol.* 2000;13(8):673–92. <https://doi.org/10.1021/tx000002x>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.