Computational Toxicology 7 (2018) 27–35

Contents lists available at ScienceDirect

Computational Toxicology

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

A pharma-wide approach to address the genotoxicity prediction of primary aromatic amines

Mukesh Patel[a,*], Michael Kranz[b], Jordi Munoz-Muriedas[b], James S. Harvey[b], Amanda Giddings[b], Steve Swallow[c], Mick Fellows[d], Russell Naven[e], Anne-Laure Werner[a], David J. Yeo[a], Frank Bringezu[f], Joerg Wichard[g], Andreas Sutter[g], Susanne Glowienke[h], Lewis Whitehead[i], Matt Selby[j], James Reuberson[j], Franck Atienzar[k], Helga Gerets[k], Michelle O. Kenyon[l], Krista L. Dobo[l], Magnus W. Walter[m], Robert A. Jolly[n], Alexander Amberg[o], Hans-Peter Spirkl[o], Wolfgang Muster[p], Jacques Van Gompel[q]

[a] Lhasa Limited, Granary Wharf House, Leeds West Yorkshire LS11 5PS, United Kingdom
[b] GlaxoSmithKline R&D, Park Road, Ware, Hertfordshire SG12 0DP, United Kingdom
[c] AstraZeneca, Charter Way, Macclesfield SK10 2NA, United Kingdom
[d] New Drug Modalities, Drug Safety and Metabolism, IMED, AstraZeneca, Cambridge, United Kingdom
[e] Takeda Pharmaceuticals International Co, Cambridge, MA 02139, United States
[f] Merck KGaA, Biopharma, Non-Clinical Safety, HPC: U009/101, 64271 Darmstadt, Germany
[g] Bayer AG, Pharmaceuticals Division, Genetic Toxicology, Müllerstr. 178, D-13353 Berlin, Germany
[h] Novartis Institutes for Biomedical Research, Novartis AG, Werk Klybeck 135-118, 4057 Basel, Switzerland
[i] Novartis Institutes for Biomedical Research, 181 Massachusetts Avenue, Cambridge, MA 02139, United States
[j] UCB, 216 Bath Road, Slough SL1 3WE, United Kingdom
[k] UCB BioPharma SPRL, Investigative Toxicology, Chemin du Foriest, B-1420 Braine-l'Alleud, Belgium
[l] Genetic Toxicology, Drug Safety Research and Development, Pfizer Worldwide Research and Development, Groton, CT 06340, United States
[m] Eli Lilly and Company, Erl Wood Manor Research Centre, Windlesham, Surrey GU20 6PH, United Kingdom
[n] Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285, United States
[o] Sanofi R&D Preclinical Safety, Industriepark Hoechst, 65926 Frankfurt am Main, Germany
[p] F. Hoffmann-La Roche Ltd, Roche Pharma Research & Early Development, Pharmaceutical Sciences, Roche Innovation Center, Basel, Grenzacherstrasse 124, Switzerland
[q] EU Head Genetic Toxicology, Mechanistic & Investigative Toxicology, Discovery Sciences, Janssen Pharmaceutica, Turnhoutseweg 30, B-2340 Beerse, Belgium

ARTICLE INFO

Keywords:
Primary aromatic amines
Mutagenicity
Ames test
Expert review

ABSTRACT

Primary aromatic amines (pAAs) are attractive building blocks in medicinal chemistry programmes yet their potential for mutagenic activity causes real concern owing to the risk of genotoxicity-related drug attrition. In addition, despite the existence of a substantial body of experimental data, the prediction of aromatic amine mutagenicity still poses a significant challenge for *in silico* tools. Major contributors to this dilemma are the stability and physicochemical properties of a subset of aromatic amines that affords them capricious mutagenic properties in the Ames test. Such inconsistent mutagenic potential is also compounded by the inherent variability with the assay itself and underscores the need for a rigorous approach in executing the experimental protocol. In order to understand the utility of the *in silico* approach towards the prediction of pAAs mutagenicity and to widen the availability of mutagenicity data, a group of pharmaceutical companies has formed a consortium with the aim of exchanging their in-house data and making them publicly available for the first time. Summary data compiled during the first phase of this effort is disclosed here and its utility in conjunction with *in silico* prediction is discussed. Conclusions from this analysis highlight the critical role of expert judgement in

Abbreviations: CIGAA, consortium for the investigation of genotoxicity of aromatic amines; GLP, good laboratory practice; KNIME, the Konstanz Information Miner; pAAs, primary aromatic amines; (Q)SAR, quantitative structure-activity relationship; SMARTS, SMiles ARbitrary Target Specification; TP, true positive; TN, true negative; FP, false positive; FN, false negative; Sens, sensitivity; Spec, specificity; PP, positive predictivity; NP, negative predictivity, Acc, accuracy; Bal acc, balanced accuracy; FPR, false positive rate; FNR, false negative rate; LL, liability load (= TP + FP + no classification calls); ATR, assays to run (= Liability Load/Total no of compounds); OECD, organisation for economic co-operation and development; 2AA, 2′-aminoacetophenone; 2A3F, 2-amino-3-fluorophenol; 2PTA, 2-propyl-2*H*-1,2,3-triazol-4-amine; M3AB, methyl 3-aminobenzoate

* Corresponding author.
E-mail address: mukesh.patel@lhasalimited.org (M. Patel).

https://doi.org/10.1016/j.comtox.2018.06.002
Received 25 April 2018; Received in revised form 13 June 2018; Accepted 18 June 2018
Available online 19 June 2018

rationalizing the experimental activity seen in the Ames test with predictions from *in silico* models. This collaboration demonstrates the value of sharing such data pre-competitively to aid in both the selection of Ames negative building blocks for drug development while simultaneously helping to develop better *in silico* tools.

1. Introduction

Aromatic amines remain popular building blocks in medicinal chemistry owing to their synthetic versatility coupled with their ability to confer diverse properties on a chemical series. While this compound class is ubiquitous in drug molecules, their potential for genotoxicity (i.e. mutagenicity) is perceived as a serious safety issue. In many cases, however, the potentially-mutagenic primary aromatic amines (pAAs) are coupled to a chemical group such that the bioactivation pathway associated with their genotoxic mechanism is no longer available. Nevertheless compounds containing aromatic amines must be screened in order to assess the risk from the drug product itself as well as any potential synthetic impurities and/or metabolites. In 2009 [1] it was estimated that 14% of marketed drugs contain aromatic amines (primary and secondary), and an analysis of one of the consortium member's compound collection suggests that around 25% of compounds synthesized for medicinal chemistry programmes contain pAAs or their potentially-metabolically labile derivatives. Within this proprietary compound set, the number of heteroaromatic amines exceeded anilines by a ratio of approximately 2–1. Examples of marketed drugs that contain primary and functionalised aromatic amines are included in Fig. 1.

In 2010 there was an increasing awareness amongst drug discovery chemists that an extensive array of Ames test data (derived from synthetic intermediates, starting materials, impurities, etc.) was being generated within the pharmaceutical industry that could both be shared and used to improve the prediction of pAA-mediated mutagenicity. To address this issue, in February 2011 the Royal Society of Chemistry, prompted by Pfizer scientists, invited a number of pharmaceutical companies to discuss a joint approach to further understand the liability of this common building block in drug discovery. A precompetitive collaborative group was established to share data and Lhasa Limited was tasked with coordinating this Consortium for the Investigation of Genotoxicity of Aromatic Amines (CIGAA). The aims of this group were to collect and publish proprietary summary Ames test data for publically disclosed aromatic amines and to use this data as a warehouse to prevent unnecessary repeat testing and to improve *in silico* models.

The most reliable assay correlating mutagenicity to animal carcinogenicity is the widely used Bacterial Reverse Mutation Test or "Ames" mutagenicity test [2]. Typically, medicinal chemists will aim to de-risk their projects by assessing the mutagenicity of aromatic amine building blocks and intermediate compounds through an *in silico* assessment or by using a pre-screen assay before investing heavily in a compound series. Unfortunately, *in vitro* testing at the early stages of drug discovery can be demanding as material is often scarce, particularly for proprietary building blocks and lead compounds, and relatively large quantities of high purity test material are required for robust OECD compliant Ames testing. A large amount of Ames test results are available in the public domain and form the basis of training sets for the available predictive Quantitative Structure-Activity Relationship ((Q)SAR) tools, [3,4,5] however each drug company holds Ames test data for many more drug-related chemical structures in their proprietary repositories. These proprietary datasets may often cover quite unique areas of chemical space and can pose a challenge for predictive models built only on public domain data. Not only would the availability of more Ames test data provide the opportunity to improve (Q)SAR models, it would be highly advantageous to medicinal chemists as they would be able to minimize mutagenicity risk through the selection of Ames negative building blocks without recourse to extensive synthesis, purification and testing of scarce materials.

The metabolism of pAAs occurs through their activation to aryl hydroxylamines by the cytochrome P450 family of oxygenases, [6,7] followed by additional Phase II metabolism to acetate, sulphonate or glucoronides (Scheme 1). Whilst the intention of this metabolic modification is detoxification, these are relatively unstable species that have the potential to bind covalently to nucleophilic bases in DNA and thus lead to a potential mutagenic event [8]. The leading theory for the mutagenicity of pAAs is the generation of a reactive nitrenium ion by the heterolytic scission of the aryl nitrogen oxygen bond [9,10,11]. In the context of the *in silico* prediction of mutagenic potency, the nitrenium ion hypothesis is an attractive mechanism as it is possible to base predictions on the ease of formation of this electrophilic species [3,12,13]. Indeed, in 2012, Birch and co-workers [14] used the energy of the dissociation reaction between the aryl amino ester and the nitrenium ion to prioritise the synthesis of safer biphenyl amines, a class of compounds associated with high mutagenic potency. While this is a plausible mechanism for the origin of mutagenic activity, it must not be forgotten that other interconnecting pathways may also contribute to the observed activity, including intercalation in DNA and radical generation. Whilst not discussed in detail here, relevant publications where the reader can find more information include Beland et al. [7] which contains several examples of different routes of activation for different aromatic amines and nitro compounds. Skipper et al. [8] describe thiol/GSH adducts and Chao et al. discuss the role of reactive oxygen species [15].

Notwithstanding additional mechanisms of mutagenic activity, variations in the reproducibility of the Ames test have been reported; both with respect to inter-laboratory studies [16] and within chemical class [17]. Contributing factors towards these inconsistencies include differences in test conditions, particularly with respect to the application of metabolic activation, as well as the source and purity of the test material.

For example, anthraquinone has been shown to be non-mutagenic when synthesized via the Diels-Alder reaction between 1,4-naphthoquinone and 1,3-butadiene or via the Friedel-Crafts reaction between benzene and phthalic anhydride [18]. In contrast, anthraquinone sourced through the oxidation of anthracene has been shown to possess direct mutagenic activity in *Salmonella* strains TA98, TA100 and TA1537. Despite the high purity that was reported for this batch at the time (99%), the observed mutagenic activity was associated with low-

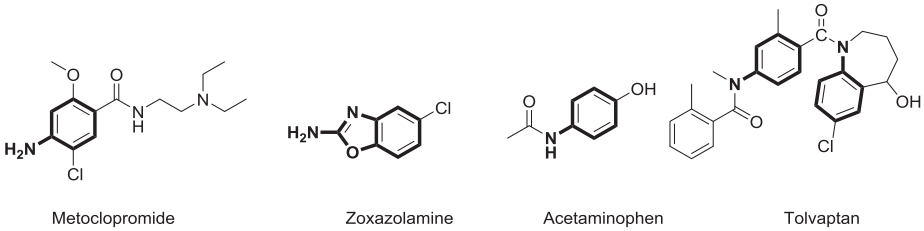

Fig. 1. Example of drugs containing primary and functionalised aromatic amines (highlighted in bold).