Read-across for the prediction of genotoxicity of dimethenamid-P groundwater metabolites: A case study using synthetic chemistry and computational toxicology

S. J. Enoch[1], M. Frericks[2], M. Kemeny[2], B. Noller[2], M. Bartel[2]

[1]School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, Liverpool, England

[2]BASF SE, Limburgerhof, Germany

Corresponding author: s.j.enoch@ljmu.ac.uk

Abstract

This work presents a read-across case study to predict the genotoxicity of two previously untested groundwater metabolites of dimethenamid-P (α-chloroacetamide herbicide). Due to synthesis challenges, direct testing of the metabolites was not feasible. Structural space alerts and matched molecular pairs analysis were employed to identify analogues and assess structural features. These analyses indicated that these features were not associated with genotoxicity. However, neither method provided evidence regarding the genotoxic potential of the lactone/hemiacetal moiety unique to the metabolites. To address this data gap, a surrogate compound containing the relevant moiety was synthesized and tested negative for genotoxicity. Computational profiling and metabolic simulation further demonstrated structural and metabolic similarity between the surrogate and the target metabolites.

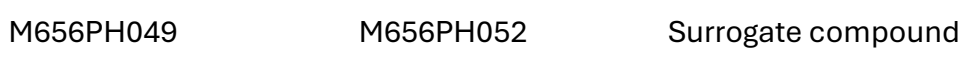
This work shows that structural data gaps can be resolved through the synthesis of a surrogate compound combined with experimental toxicological data. The integrated weight of evidence from synthetic and computational chemistry supports the conclusion that the two dimethenamid-P groundwater metabolites are unlikely to be genotoxic.

Overall, this study demonstrates that the use of synthetic chemistry and computational toxicology can provide a robust strategy to address structural moiety data gaps in the genotoxicity evaluation of plant protection products, their active ingredients, and their metabolites.

Keywords: Matched Molecular Pairs, Structural Space Alerts, Genotoxicity, Read-Across, Surrogate Compound, Computational Toxicology

Introduction

Dimethenamid-P (DMTA-P) is a thiophene containing compound in the α-chloroacetamide class of herbicide compounds. Research has shown that the α-chloroacetamides undergo two major metabolic transformations in the rat, these being: hydrolysis and glutathione conjugation of the α-chloroacetamide moiety. In addition, the methyl substituents of the thiophene aromatic ring were also shown to undergo hydroxylation reactions, although these were reported as minor pathways. Importantly, the thiophene ring present in DMTA-P was not documented to undergo significant metabolism in the rat (1). However, two metabolites, M656PH049 and M656PH052, involving the transformation of the thiophene ring into a lactone/hemiacetal moiety have been only found in groundwater – with these metabolites needing to be assessed for genotoxicity (2). Structures shown in Figure 1.



M656PH049 M656PH052 Surrogate compound

Figure 1: Structures of M656PH049, M656PH052 and the surrogate compound. Acidic groups highlighted.

Unfortunately, extensive efforts over several years (involving several laboratories) to synthesise these groundwater metabolites have not been successful due to a lack of stability/decomposition and purity, and thus, it has not been possible to test them directly for genotoxicity.

The difficulties experienced with the synthesis and purification of M656PH049 and M656PH052 can, in part, be attributed to the hemiacetal moiety (lactone) found in both metabolite structures. Hemiacetals are chemical compounds with the chemical moiety $HO\text{-}CH_2\text{-}O\text{-}CH_2\text{-}R$, which typically exist as transient intermediates in chemical reactions (3). For this reason, they are not typically isolated in their pure form. Instead, both above mentioned metabolites exist in an equilibrium of the ring-closed (lactone/hemiacetal) and ring-open form, with the position of equilibrium being dependent on the pH value. Thus, synthesis of M656PH049 and M656PH052 is extremely challenging, since these molecules contain acidic groups (Figure 1).

Figure 2: Equilibrium for the ring-closed (left hand structure) and ring-open (right hand structure) forms of M656PH049 (R = carboxylic acid) and M656PH052 (R = sulfonic acid).

Despite the difficulty in synthesising the two metabolites of DMTA-P, a successful synthesis of a 'surrogate' compound has been carried out (structure as shown in Figure 1). This compound does not contain either of the acidic groups, making the synthesis and purification possible. Importantly, the surrogate contains all functional groups of the metabolites of interest with exception of the acidic group.

Read-across is a non-testing approach in chemical risk assessment where data from a structurally and mechanistically similar substance (source substance) are used to predict the toxicological properties of another substance (target substance) for a specific endpoint, e.g. genotoxicity (4). To this end, recent research has developed the Structural Space Alert (SPA) concept to address the similarity between plant protection product metabolites that lack an alert for either covalent protein or DNA reactivity (1, 5-7). This concept utilises the rat metabolism data available in the Draft Assessment Report/Renewal Assessment Report (DAR/RAR) documents of plant protection products (available from the EFSA website) to identify a set of common metabolic pathways for a given compound class. These common metabolic pathways are then used to define structural space alerts to enable compounds to be grouped, ensuring metabolic similarity between the category members. This approach to grouping has been included in recent EFSA read-across guidance (4, 8).

In addition to the research on SPAs, Matched Molecular Pairs Analysis (MMPA) has also been utilised to identify analogues suitable for making read-across predictions of the genotoxicity of plant protection product metabolites (9). This approach relies on the identification of a target-analogue pair in which an analogue differs from the target at (least) a single substituent position. Importantly, the remainder of the structure is common between the target and analogue. This analysis leads to a set of matched molecular pairs in which the structural domain of the target compound is covered by the combined set of the identified analogues. However, to be successful this approach is data intensive as sufficient analogues must be available to cover the structural features present in the target. Whilst a MMPA does not directly identify metabolically similar compounds, it does offer a robust method for covering the structural domain of the

target. This work extended the use of MMPA from the pharmaceutical (10) and cosmetic (11) sectors, applying it to plant protection products for the first time.

Thus, the aim of this study was to apply weight of evidence approach for the prediction of genotoxicity for the two groundwater metabolites of DMTA-P. As lines of evidence the following computational methods for the prediction of genotoxicity were used: the application of the α-chloroacetamide SPAs, an MMPA to identify suitable analogues for read-across, and the use of quantitative structure-activity models. In addition, evidence of the experimental toxicity data for a surrogate compound was also utilised and the approach was supported by synthetic chemistry to have a well-designed surrogate to support a robust and practical solution for genotoxicity assessment.

Dataset

The publicly available DAR documents were used to compile six α-chloroacetamide herbicide active ingredients and their metabolites as identified in the rat. In addition, data for a further five active ingredients and their metabolites were compiled from historical data from the industrial co-authors of this study. This led to a dataset of 69 α-chloroacetamide herbicides for which genotoxicity data were extracted for all compounds that had been directly tested in either the Ames test, *in vitro* chromosomal aberration test or the *in vivo* micronucleus test. All compounds that were identified as being positive in the *in vitro* chromosomal aberration test had additional data from the *in vivo* micronucleus test showing them to be negative. Importantly, the final dataset expanded the chemical space for the α-chloroacetamide herbicides compared to the available data in the publicly available EFSA genotoxicity dataset (available from https://doi.org/10.5281/zenodo.14946069). The dataset, termed the 'α-

chloroacetamide genotoxicity dataset' contained the following test results (*in vitro* assays with S9 fraction):

- Ames: 69 compounds (all negative)
- *In vitro* chromosomal aberration: 56 compounds (39 negative, 16 positive, 1 equivocal)
- *In vivo* micronucleus: 39 compounds (all negative)

All chemical structures and associated genotoxicity data are available in the Table S1 in the Supplementary Information.

Structural space alerts (SPA)

The previously published set of structural space alerts for the α-chloroacetamides were utilised to investigate the metabolic similarity between the target and potential analogues (1). Seventeen structural space alerts defined three key scaffolds (benzene, thiophene and styrene) covering six metabolic pathways defined from an analysis of rat metabolism data. These being two major pathways: glutathione conjugation and hydrolysis/oxidation of the chloroacetamide moiety. In addition, four further minor pathways are also covered: aromatic ring hydroxylation, aliphatic hydroxylation, dealkylation, and dehalogenation.

Matched molecular pairs analysis (MMPA)

Matched molecular pairs analysis was carried out using the previously published KNIME workflow (9). Briefly, this workflow identifies target-analogue pairs based on the presence of common sub-structures. This is achieved by sequentially 'cutting' each acyclic single bond in the target and potential analogues. This process generates target-

analogue pairs with a common scaffold that differ at a single substituent point. Given sufficient data, the method generates a range of target-analogue pairs in which the in-common scaffold varies sufficiently to ensure that all the structural features of the target are covered by the identified analogues. Assuming all the analogues show a lack of toxicity, this structural coverage enables confidence that none of the structural features in the target compound are linked to toxicity. In the published KNIME workflow only one acyclic bond can be cut at a time when generating; however, in the current work this setting was modified to allow two acyclic bonds to be cut at the same time. This leads to common scaffold (between target and analogue) that varies at two substituent positions.

Conformer generation

Conformer generation for M656PH049, M56PH052 and the surrogate were carried out using the on-line Rowan Scientific platform using the (https://www.rowansci.com – accessed November 2025). The following parameters were utilised:

- Conformer generator: ETKDG
- Conformer generation mode: Rapid
- Conformer optimisation method: GFN-FF
- Final method: OMol25's eSEN Con
- Solvent: Gas phase

Chemical profiling

Quantitative structure-activity relationship (QSAR) analysis between M656PH049, M656PH052 and the surrogate compound was carried out using eleven commercially available models available from Lhasa Ltd, MultiCase Inc and Instem Ltd. The models were as follows (MNT is the micronucleus test):

- Lhasa Ltd: Derek Nexus expert system V6.2: Ames and *in vivo* MNT. Sarah Nexus statistical model V3.2: Ames test
- MultiCase Inc: Case Ultra statistical models V1.8.1.6: Ames (GT1_BMUT and PHARM models) and *in vivo* MNT (GT3_MNT_Mouse and GT3_MNT_Mouse PHARMA models).
- Instem Ltd: Leadscope Model Applier expert system V3.1: Ames. Leadscope Model Applier statistical models: Ames and *in vivo* MNT

Chemical categories were profiled using the profiling schemes within the OECD QSAR Toolbox (V4.7). A subset of the available endpoint specific profilers for genotoxicity was utilised based on the results of a previous study into their suitability for read-across predictions within the plant protection chemical space (1, 5-7). These profilers were (CA is chromosomal aberration and MNT is the micronucleus test):

- DNA alerts for AMES, CA and MNT by OASIS
- Protein binding alerts for CA by OASIS

In addition, metabolism prediction was carried out using the *in vitro* rat liver S9 metabolism simulator in the OECD QSAR Toolbox V4.7.

Synthetic chemistry

As extensive synthetic efforts of M656PH049 and M656PH052 were unsuccessful using several chemical synthesis routes or biosynthesis approaches, a surrogate structure was successfully synthesised instead containing the lactone/hemiacetal moiety (Figure 1). The acidic groups of M656PH049 and M656PH052 are “incompatible” with the hemiacetal group within the molecule. Isolating these metabolites in pure form, as needed for toxicological tests, was not possible. The formation of hemiacetals occurs within a dynamic chemical equilibrium, which is postulated to exist in the above mentioned DMTA-P metabolites, where both ring-closed and ring-open forms are present in equilibrium (Figure 2). The specific synthetic route used to obtain the surrogate compound remains confidential and cannot be disclosed. The chemical formula of the surrogate was identified as $C_{12}H_{19}NO_5$, with a purity greater than 99% (GLP-certified material). The surrogate was evaluated for genotoxic potential and found to be negative for mutagenicity (AMES and HPRT) and chromosomal damage (both structural and numerical, as determined by *in vitro* MNT).

Results and Discussion

The aim of this study was to predict the genotoxicity of two groundwater metabolites of DMTA-P via read-across. The research involved in the application of the previously published structural space alerts for the α-chloroacetamides (1), which define metabolic similarity between a target and analogue. In addition, a matched molecular pairs analysis was applied for the identification of structurally related analogues using the recently outlined method (9). The final part of the analysis outlines how the

applicability domain of a read-across prediction can be expanded through the synthesis of chemically accessible analogues.

DMTA-P groundwater metabolites

The key metabolites in this study are two groundwater metabolites arising from DMTA-P, compounds M656PH049 and M656PH052 (Figure 1). Metabolism due to oxidation or glutathione conjugation results in either a carboxylic acid or sulphonic acid functional group replacing the chlorine atom (both pathways being identified as being major routes of metabolism in the rat for this class of compound (1)). In addition, the thiophene ring is oxidised to a lactone, which is in equilibrium with a ring-opened moiety (structures as shown in Figure 3, equilibrium shown in Figure 2). Interestingly, this oxidation reaction was not reported in the rat metabolism data (1).

DMTA-P

Ring-closed Oxidation

Ring-open Oxidation

M656PH049

Ring-closed Glutathione

Ring-open Glutathione

M656PH052

Figure 3: Structures of DMTA-P and its two groundwater metabolites (shown in both ring-closed and ring-open forms, oxidation and glutathione labels refer to the reaction having taken place at the chlorine present in DMTA-P).

Structural space alerts – metabolic similarity

An initial analysis was carried out using the previously published set of structural space alerts for the α-chloroacetamides (1). This set of structural space alerts defined the five common metabolic pathways for the α-chloroacetamides class of herbicides, consisting of three key scaffolds as shown in Figure 4 (leading to a total of 18 structural space alerts, six per scaffold). The results of this analysis showed that neither of the target compounds (M656PH049 and M656PH052) triggered any of the structural space

alerts. This being even though the structural space alerts encode both the oxidation and glutathione metabolic pathways. However, the rat metabolism data from which the structural space alerts were developed did not feature the oxidation of the thiophene ring to a lactone/hemiacetal conformation. Thus, the two target compounds were out of the metabolic domain of the previously published α-chloroacetamide structural space alerts due to this metabolic change (which, given how the structural space alerts were designed, would be an additional scaffold when comparing to the scaffolds outlined in Figure 4).

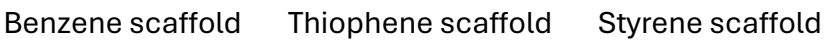


Figure 4: The key scaffolds present in the structural space alerts defined for the α-chloroacetamides (R_1 = alkyl ether chain, R_2 = alkyl chloride, products of oxidation, or products of glutathione conjugation)

Matched molecular pairs

Given that the target compounds were out of the metabolic domain of the α-chloroacetamides structural space alerts a follow up matched molecular pairs (MMP) analysis was undertaken. This utilised the previously published approach and KNIME workflow, in which only a single acyclic bond can be broken to create pairs of compounds (9). This analysis identified four analogues with *in vivo* MNT data for the target compound, M656PH049 (analogues 1-4 in Table 1). The common fragment linking M656PH049 and the four analogues being the presence of branched alkyl ether and carboxylic acid

(defined by MMP 1 in Table 1). A similar analysis identified three analogues with *in* vivo MNT data for the target compound, M656PH052 (analogues 5-7 in Table 1). The common fragment linking M656PH052 with these three analogues being the presence of the same branched alkyl ether as previously and a sulfonic acid moiety (defined by MMP 2 in Table 1). These data show that compounds containing a combination of the branched alkyl ether chain with either a carboxylic acid or sulfonic acid moiety are not associated with positive results in either the Ames test or *in vivo* MNT. However, this MMPA did not identify any analogues containing the lactone/hemiacetal moiety present in either target compound. As a result, it was not possible to reach a definitive conclusion regarding the genotoxicity of this moiety.

Table 1 here

A further MMP analysis in which the previously published method was altered to enable two acyclic bonds to be broken to create pairs of molecules. This has the effect of increasing the number of pairs that are generated as only one functional group/moiety needs to remain constant for two compounds to be considered a pair. This analysis identified further four analogues featuring the branched alkyl ether moiety (defined by MMP 3, analogues 8-11 in Table 2), two analogues featuring the carboxylic acid moiety (defined by MMP 4, analogues 12 and 13 in Table 2), and four analogues featuring the sulfonic acid moiety (defined by MMP 5, analogues 14-17 in Table 2). These compounds tested negative in both the Ames and *in vivo* MNT test – adding further weight of evidence that neither the branched alkyl ether, carboxylic acid or sulfonic acid moieties are associated with genotoxicity. As with the previous MMP analysis, this analysis also did

not identify any analogues containing either the ring-open or closed moieties present in either target compound.

Table 2 here

Surrogate compound

As already outlined, a surrogate compound was synthesised and tested for genotoxicity. This compound features both the branched alkyl ether and ring-open/closed moieties present in the target compounds (M656PH049 and M656PH052, Figure 1). However, the carboxylic and sulfonic acid moieties have been removed. Genotoxicity testing of this compound showed it to be negative in the Ames test and the *in vitro* MNT. These results show that the combination of the branched alkyl ether and the lactone/hemiacetal form do not lead to *in vitro* genotoxicity.

The presence of the carbonyl in the surrogate is an important structural feature due to its ability to form a conjugated π-system with the nitrogen atom (creating a planar amide). This conformation is in competition with the extended π-system that can form upon conjugation between the nitrogen lone pair of electrons and the conjugated alkene present in both the ring-open and closed forms (Figure 3). It is not possible for both functional groups (carbonyl and conjugated alkene) to be fully conjugated (*i.e.*, planar) with the nitrogen, due to steric hindrance. This competition between planar amide and the conjugated alkene is also present in the two target compounds making them electronically and geometrically extremely similar. Conformer calculations for M656PH049, M656PH052, and surrogate in both ring-open and closed forms showed all

low energy conformers to have a planar amide moiety, with the conjugated alkene being approximately 90° to the plane of the amide. Given the influence the electronic structure has on genotoxicity (12-17) and metabolism (18-22), this analysis adds further weight to the hypothesis that the ring-open/closed system will behave in the same way in all three compounds.

Further analysis into the metabolic similarity between M656PH049, M656PH052, and surrogate was undertaken using the *in vitro* rat liver S9 simulator in the OECD QSAR Toolbox V4.7. The *in vitro* profiler was used as it focuses on phase 1 metabolism (due to an *in vitro* system having limited phase II capability) – this being important as it is phase I reactions that can produce electrophiles capable of reacting covalently with DNA and proteins (12, 14, 22, 23). The results of profiling the open/closed ring forms of M656PH049, M656PH052, and surrogate predicted there to be four common metabolic pathways, these being (also summarised in Figure 5):

- Pathway A: O-glucuronidation – this reaction involves the addition of glucuronic acid to the hydroxyl group of the lactone ring.
- Pathway B: Oxidative O-dealkylation – this reaction removes the methyl group of the alkyl ether chain producing an alcohol.
- Pathway C: Ester hydrolysis – the lactone ring is predicted to undergo a ring opening hydrolysis reaction to produce a geminal diol (where both alcohol groups are present on the same carbon atom).
- Pathway D: Geminal diol dehydration – the geminal diol is predicted to undergo a dehydration reaction to produce a hemiacetal. This is the ring opened structure for M656PH049, M656PH052, and surrogate.

- In addition, M656PH049 was also predicted to undergo a second O-glucuronidation reaction due to the presence of the carboxylic acid moiety (not shown in Figure 5).

Overall, the predicted metabolic pathway information showed that the surrogate undergoes the same set of metabolic transformations as the two target compounds (M656PH049 and M656PH052) – this adds to the weight of evidence that the three compounds behave the same metabolically.

Figure 5: Summary of the predicted metabolic pathways for compounds M656PH049, M656PH052, and surrogate (A = O-glucuronidation, B = oxidative O-dealkylation, C = ester hydrolysis, D = geminal diol dehydration, gluc = glucuronic acid)

Computational profiling of M656PH049 and M656PH052

M656PH049 and M656PH052 were computationally profiled for genotoxicity using eleven commercially available quantitative structure-activity models covering the Ames

test (three expert systems and four statistical models) and *in vivo* MNT (one expert system and three statistical models). Overall, nine of the eleven models predicted neither compound, in either their ring-open or closed forms, to be genotoxic. However, Case Ultra identified the presence of the α,β-unsaturated lactone (in the ring-closed compounds) to be associated with a positive result in the Ames test. Whilst Derek Nexus highlighted the presence of the α,β-unsaturated ketone (in the ring-open form) to be associated with a positive result in the *in vivo* MNT (all data in Table S2 in the Supplementary Information).

Inspection of the data associated with the α,β-unsaturated lactone alert in Case Ultra showed 47 of the 63 compounds that featured this alert to be positive in the Ames test. However, 42 of the 47 Ames positive compounds contained additional structural alerts for genotoxicity related to either the presence of halogens (22 compounds) or poly-aromatic hydrocarbons (20 compounds) (12-14). Removing these 42 positive compounds with additional alerts from the dataset results in only five Ames positive compounds, compared to 14 Ames negative compounds containing an α,β-unsaturated lactone moiety. In addition, the α,β-unsaturated lactone present in M656PH049 and M656PH052 is di-substituted at the β-carbon atom where the Michael addition reaction would take place, leading to a significant reduction in reactivity (23, 24). Thus, it is extremely unlikely that the presence of the α,β-unsaturated lactone structural alert in these compounds would lead to a positive result in the Ames test. This being supported by the negative Ames result for the surrogate compound.

A similar analysis for the α,β-unsaturated ketone alert identified by Derek Nexus in the ring-open compounds is not possible due to the confidential nature of the underlying dataset upon which the alert was developed. However, in M656PH049 and M656PH052 (and the surrogate compound) the β-carbon atom is sterically hindered by the presence of a carboxylic acid and a methyl group. This type of steric hindrance has been shown to prevent such compounds from reacting with biological nucleophiles (24). This suggests that it is extremely unlikely that the presence of the α,β-unsaturated ketone moiety would lead to a positive MNT result. This is supported by the negative *in vitro* MNT result for the surrogate compound.

Read-across predictions for M656PH049 and M656PH052

The MMPA data demonstrate that the presence of the alkyl ether chain, sulfonic acid and carboxylic acid moieties are not associated with a positive result in either the Ames test or the *in vivo* MNT (see Tables 1 and 2). However, the MMPA data could not outline the potential genotoxicity of the ring-open/closed moiety (due to it being absent from the compounds shown in Tables 1 and 2). To address this a computational analysis of the synthesised analogue, the surrogate, which features this moiety, showed the importance of the amide in ensuring structural and electronic similarity between it and M656PH049/M656PH052. A simulated metabolic analysis of ring-open/closed structures for these three compounds showed them to undergo a common set of transformations, including production of the hemiacetal form. This computational analysis demonstrates structural and metabolic similarity between these three compounds, building the weight of evidence that surrogate is a suitable analogue for the prediction of the genotoxicity of the ring-open/closed moiety present in M656PH049 and

M656PH052. Importantly, the surrogate compound was tested negative in the main genotoxicity endpoints, namely mutagenicity (AMES and HPRT) and chromosome damage (structural and numerical via *in vitro* MNT).

In addition to the structural and metabolic similarity, all compounds (surrogate, M656PH049, M656PH052 and those shown in Tables 1 and 2) were profiled using the subset of the endpoint specific profiling schemes (as outlined in the methods section) in the OECD QSAR Toolbox V4.7 that have been shown to be predictive of genotoxicity for plant protection products (1, 5-7, 25). This profiling showed that none of the compounds featured any structural features associated with genotoxicity. The overall weight of evidence enables negative predictions for the Ames test and *in vitro* MNT for M656PH049 and M656PH052. Given the EFSA tiered workflow for genotoxicity testing for plant protection products (active ingredients and their metabolites) and according to the EC guidance document SANCO/221/2000, *in vivo* MNT data are not required for metabolites, the negative *in vitro* MNT is considered sufficient to assume no activity in vivo as well (2, 26).

In summary, the computational analysis demonstrated that the surrogate (source substance) is a suitable analogue for the prediction of genotoxicity via read-across for the target substances M656PH049 and M656PH052.

Uncertainty analysis

In keeping with EFSA's recently published guidance on read-across (4), an uncertainty analysis was carried out as follows (headings as in Figure 3 in reference (4)):

- Unambiguous chemical identities of target and source substances are known: All identities are known and published in this study, including all chemical structures of the targets, analogues and the surrogate. All chemical structures and toxicological data are available in Table S1 in the Supplementary Information.
- Analogues of target substances are identified from adequate database(s) by an unbiased system(s): Analogues have been identified from a previously published database of α-chloroacetamide plant protection products (1). Analogues were identified by MMPA following a published protocol (9).
- Selection of close source substances (from the identified analogues) is justified on scientific grounds: An unbiased MMPA was utilised to select analogues covering the structural domains of the two target substances. In addition, a surrogate was synthesised and used to expand the structural domain coverage for the lactone/hemiacetal moiety present in the targets. A detailed computational analysis was undertaken to demonstrate the chemical and metabolic similarity between the surrogate and target substances.
- Toxicological data for use in read-across are derived from valid studies: All toxicological data were taken published DAR/RAR dossiers for the α-chloroacetamide class of plant protection products. Toxicological data for the surrogate was performed under GLP studies following OECD test guidelines (27-29).
- All steps are documented and the use of expert judgement at any stage is justified: The case study presented outlines the steps used to arrive at the read-across conclusion in detail. No expert judgement was used in the final read-across

predictions, with it being based on the outcome from an MMPA, QSAR models and data for the surrogate compound.

Overall, the above analysis enables the read-across predictions for M656PH049 and M656PH052 to be made with low uncertainty.

Conclusions

The aim of this study was to outline a read-across case study for two previously untested groundwater metabolites of DMTA-P, which could not be synthesised for toxicity testing. The approach highlighted a tiered approach to the use of the SPA and MMPA concepts for the identification of analogues. In the case study presented, the two target compounds were out of domain of the available SPA for the α-chloroacetamides, leading to no analogues being identified. In contrast, the MMPA was able to identify a series of analogues featuring two of the three key structural features present in the target compounds. The MMPA was able to build a weight of evidence that these two features were not associated with genotoxicity. However, the MMPA resulted in a lack of structural data for the third feature, meaning that no conclusion could be drawn about its genotoxic potential. Additionally, the case study also showed that a surrogate can be used as a source substance for read-across assessment when metabolites cannot be synthesized for toxicity testing. A detailed computational analysis was carried out to demonstrate the structural, electronic and metabolic similarity between the surrogate and target compounds. This analysis enabled the structural domain of the SPA and MMPA to be expanded to cover the lactone/hemiacetal moiety. The negative results of the experimental *in vitro* genotoxicity studies from the surrogate, in combination with the

computational predictions, led to the conclusion that the two DMTA-P metabolites are unlikely to be genotoxic. In summary, this study demonstrated that integrating synthetic and computational chemistry approaches with toxicological assessment provides a robust strategy to address structural moiety data gaps in the genotoxicity evaluation of plant protection products, their active ingredients, and their metabolites.

Table 1: Analogues identified by matched molecular pairs analysis with the 'cut' value set to one (X = ring-open or closed form, CA = chromosomal aberration, MNT = micronucleus test, MMP = matched molecular pair).

	M656PH023 Ames: -ve *In vitro* CA: No data *In vivo* MNT: -ve	Metolachlor OXA CGA51202 Ames: -ve *In vitro* CA: -ve *In vivo* MNT: -ve	SYN542488 Ames: -ve *In vitro* CA: -ve *In vivo* MNT: -ve	M656PH045 Ames: -ve *In vitro* CA: -ve *In vivo* MNT: -ve
MMP 1	Analogue 1	Analogue 2	Analogue 3	Analogue 4
	M656PH027* Ames: -ve *In vitro* CA: No data *In vivo* MNT: -ve	S-Metolachlor ESA CGA376944* Ames: -ve *In vitro* CA: -ve *In vivo* MNT: -ve	M656PH047* Ames: -ve *In vitro* CA: +ve *In vivo* MNT: -ve	No further analogues with *in vivo* MNT data for MMP 2 were in the dataset
MMP 2	Analogue 5	Analogue 6	Analogue 7	

Table 2: Additional analogues identified by matched molecular pairs analysis with the 'cut' value set to two (R = any substituent, X = ring-open or closed form, CA = chromosomal aberration, MNT = micronucleus test, MMP = matched molecular pair).

	M656PH030 Ames: -ve *In vitro* CA: +ve *In vivo* MNT: -ve	M656PH031 Ames: -ve *In vitro* CA: -ve *In vivo* MNT: -ve	M656PH043 Ames: -ve *In vitro* CA: +ve *In vivo* MNT: -ve	SYN542492 Ames: -ve *In vitro* CA: +ve *In vivo* MNT: -ve
MMP 3	Analogue 8	Analogue 9	Analogue 10	Analogue 11
			No further analogues with *in vivo* MNT data for MMP 4 were in the dataset	
	Alachlor carboxylic acid Ames: -ve *In vitro* CA: -ve *In vivo* MNT: -ve	SYN542491 Ames: -ve *In vitro* CA: +ve *In vivo* MNT: -ve		
MMP 4	Analogue 12	Analogue 13		

	M656PH054 Ames: -ve *In vitro* CA: +ve *In vivo* MNT: -ve	M656PH055 Ames: -ve *In vitro* CA: -ve *In vivo* MNT: -ve	Alachlor ESA Ames: -ve *In vitro* CA: -ve *In vivo* MNT: -ve	Sulfonic acid; R290131 Ames: -ve *In vitro* CA: -ve *In vivo* MNT: -ve
MMP 5	Analogue 14	Analogue 15	Analogue 16	Analogue 17

References

1. Enoch SJ, Hasarova Z, Cronin MTD, Bridgwood K, Rao S, Kluxen FM, et al. Metabolism-based category formation for the prioritisation of genotoxicity hazard assessment for plant protection product residues (Part 4): alpha-Chloroacetamides. Regulatory toxicology and pharmacology : RTP. 2024;150:105641.
2. European Commission. Guidance Document on the Assessment of the Relevance of Metabolites in Groundwater of Substances Regulated under Regulation (EC) No 1107/2009. Sanco/221/2000 - rev11. 2021.
3. Hurd CD. Hemiacetals, Aldals and Hemialdals. Journal of Chemical Education. 1966;43(10):527-31.
4. EFSA, Bennekou SH, Allende A, Bearth A, Casacuberta J, Castle L, et al. Guidance on the use of read-across for chemical safety assessment in food and feed. EFSA J. 2025;23(7):e9586.
5. Enoch SJ, Hasarova Z, Cronin MTD, Bridgwood K, Rao S, Kluxen FM, et al. Sub-structure-based category formation for the prioritisation of genotoxicity hazard assessment for pesticide residues: Sulphonyl ureas. Regulatory toxicology and pharmacology : RTP. 2022;129:105115.
6. Enoch SJ, Hasarova Z, Cronin MTD, Bridgwood K, Rao S, Kluxen FM, et al. Metabolism-based category formation for the prioritisation of genotoxicity hazard assessment for plant protection product residues (part 3): Strobilurins. Regulatory toxicology and pharmacology : RTP. 2023;144:105484.
7. Enoch SJ, Hasarova Z, Cronin MTD, Frericks M. Sub-structure-based category formation for the prioritisation of genotoxicity hazard assessment for pesticide residues (part 2): Triazoles. Regulatory toxicology and pharmacology : RTP. 2022;134:105237.
8. ECHA. The use of alternatives to testing on animals for the REACH regulation. ECHA. 2017.
9. Enoch SJ, Hasarova Z, Cronin MTD, Bridgwood K, Rao S, Hueser A, et al. Matched molecular pairs-driven read-across for the prediction of genotoxicity of plant protection product residues. Regulatory toxicology and pharmacology : RTP. 2025;163:105915.
10. Yang Z, Shi S, Fu L, Lu A, Hou T, Cao D. Matched Molecular Pair Analysis in Drug Discovery: Methods and Recent Applications. J Med Chem. 2023;66(7):4361-77.
11. Lester CC, Yan G. A matched molecular pair (MMP) approach for selecting analogs suitable for structure activity relationship (SAR)-based read across. Regulatory toxicology and pharmacology : RTP. 2021;124:104966.
12. Enoch SJ, Cronin MTD. A review of the electrophilic reaction chemistry involved in covalent DNA binding. Critical Reviews in Toxicology. 2010;40:728-48.
13. Benigni R, Bossa C. Structure alerts for carcinogenicity, and the *Salmonella* assay system: A novel insight through the chemical relational databases technology. Mutation Research. 2008;659:248-61.
14. Benigni R, Bossa C. Mechanisms of chemical carcinogenicity and mutagenicity: A review with implications for predictive toxicology. Chemical Reviews. 2011;111:2507-36.
15. Benigni R, Bossa C, Tcheremenskaia O, Worth A. Development of structural alerts for the *in vivo* micronucleus assay in rodents. http://ecbjrceceuropaeu/DOCUMENTS/QSAR/EUR_23844_ENpdf. 2009.
16. Benigni R, Netzeva TI, Benfenati E, Bossa C, Franke R, Helma C, et al. The Expanding Role of Predictive Toxicology: An Update on the (Q)SAR Models for Mutagens and Carcinogens. Journal of Environmental Science and Health, Part C. 2007;25(1):53 - 97.

17. Enoch SJ, Cronin MTD. Development of new structural alerts for chemical category formation for assigning covalent and non-covalent mechanisms relevant to DNA binding. Mutation Research – Genetic Toxicology and Environmental Mutagenesis. 2012;743:10-9.
18. Serafimova R, Todorov M, Pavolo T, Kotov S, Jacob E, Aptula A, et al. Identification of the structural requirements for mutagencitiy, by incorporating molecular flexibility and metabolic activation of chemicals. II. General Ames mutagenicity model. Chemical Research in Toxicology. 2007;20(4):662-76.
19. Mekenyan OG, Dimitrov SD, Pavlov TS, Veith GD. A systematic approach to simulating metabolism in computational toxicology. I. The TIMES heuristic modelling framework. Current Pharmaceutical Design. 2004;10(11):1273-93.
20. Mekenyan O, Todorov M, Serafimova R, Stoeva S, Aptula A, Finking R, et al. Identifying the structural requirements for chromosomal aberration by incorporating molecular flexibility and metabolic activation of chemicals. Chemical Research in Toxicology. 2007;20(12):1927-41.
21. Mekenyan O, Dimitrov S, Dimitrova N, Dimitrova G, Pavlov T, Chankov G, et al. Metabolic activation of chemicals: in-silico simulation. SAR and QSAR in Environmental Research. 2006;17(1):107-20.
22. Kalgutkar AS, Gardner I, Obach RS, Shaffer CL, Callegari E, Henne KR, et al. A comprehensive listing of bioactivation pathways of organic functional groups. Current Drug Metabolism. 2005;6:161-225.
23. Enoch SJ, Ellison CM, Schultz TW, Cronin MTD. A review of the electrophilic reaction chemistry involved in covalent protein binding relevant to toxicity. Critical Reviews in Toxicology. 2011;41:783-802.
24. Schwobel JAH, Koleva YK, Enoch SJ, Bajot F, Hewitt M, Madden JC, et al. Measurement and estimation of electrophilic reactivity for predictive toxicology. Chemical Reviews. 2011;111:2562-96.
25. Kemény M, North CM, Kluxen FM, Frericks M, Vukelic D, Cao S, et al. Validation of OECD QSAR Toolbox profilers for genotoxicity assessment of pesticides using the MultiCase genotoxicity database. Computational Toxicology. 2025;34.
26. Committee) ES. Scientific opinion on genotoxicity testing strategies applicable to food and feed safety assessment. EFSA Journal. 2011;9(9).
27. OECD. Test No. 476: In Vitro Mammalian Cell Gene Mutation Tests using the Hprt and xprt genes2016.
28. OECD. Test No. 471: Bacterial Reverse Mutation Test2020.
29. OECD. Test No. 487: In Vitro Mammalian Cell Micronucleus Test2023.