Regulatory Toxicology and Pharmacology 134 (2022) 105237

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

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

Sub-structure-based category formation for the prioritisation of genotoxicity hazard assessment for pesticide residues (part 2): Triazoles

S.J. Enoch [a,*], Z. Hasarova [a], M.T.D. Cronin [a], M. Frericks [b]

[a] School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, Liverpool, England, United kingdom
[b] BASF, Limburgerhof, Germany

ARTICLE INFO

Handling Editor: Dr. Lesa Aylward

Keywords:
Read-across
Mutagenicity
Triazole
Metabolism
Category formation
Structural alerts
Genotoxicity
Hazard assessment
Pesticide residues

ABSTRACT

In dietary risk assessment, residues of pesticidal ingredients or their metabolites need to be evaluated for their genotoxic potential. The European Food Safety Authority recommend a tiered approach focussing assessment and testing on classes of similar chemicals. To characterise similarity and to identify structural alerts associated with genotoxic concern, a set of chemical sub-structures was derived for an example dataset of 66 triazole agrochemicals for which either Ames, chromosomal aberration or micronucleus test results are publicly available. This analysis resulted in a set of ten structural alerts that define the chemical space, in terms of the common parent and metabolic scaffolds, associated with the triazole chemical class. An analysis of the available profiling schemes for DNA and protein reactivity shows the importance of investigating the predictivity of such schemes within a well-defined area of structural space. Structural space alerts, covalent chemistry profiling and physicochemistry properties were combined to develop chemical categories suitable for chemical prioritisation. The method is a robust and reproducible approach to such read-across predictions, with the potential to reduce unnecessary testing. The key challenge in the approach was identified as being the need for pesticide-class specific metabolism data as the basis for structural space alert development.

1. Introduction

The European Food Safety Authority (EFSA) guidance on the establishment of residue definition for the dietary risk assessment for genotoxicity specifically outlines the usage of category formation and read-across (EFSA, 2016). However, at the time of writing this guidance has not been agreed between stakeholders and EFSA. Thus, there is currently no established legal framework for the use of read-across (although documentation has been published outlining guidance for the application of read-across generically (ECHA, 2017)). Therefore, a robust scientific weight of evidence needs to be established for such methods to become commonplace for the prediction of genotoxicity of pesticide residues. For these purposes, the term "residue" is defined as any compound associated with the active ingredient that may result in risk to human and/or livestock following the application of a pesticide. EFSA have published a workflow to enable the use of read-across to predict either the presence or absence of genotoxicity within a category of similar chemicals where data may be missing or incomplete (EFSA, 2016). In cases where genotoxicity is predicted, further testing is required to confirm the read-across prediction. The test strategy needs to ensure that a representative number of the chemicals in the category are tested for gene mutation as well as structural and numerical chromosomal aberration. A battery of *in vitro* and *in vivo* tests is recommended by EFSA to cover the three key genotoxicity endpoints with minimal animal usage (EFSA, 2011). The initial battery is typically the Ames test (gene mutation) and an *in vitro* micronucleus test (structural and numerical chromosomal aberration). This combination of testing is considered state of the art within most regulatory guidelines, with *in vivo* testing only being conducted as a higher tier to evaluate positive *in vitro* micronucleus assay results (EFSA, 2011). In contrast to this process for positive predictions, the absence of genotoxicity within a category (with data gaps being filled via read-across) requires no further genotoxicity testing.

Importantly, generating information metabolism is mandatory for pesticides. Therefore, data exist within the publicly available Draft Assessment Report (DAR) documents that detail the metabolism for compound classes used as pesticides. From a genotoxicity point of view, the toxicity of major metabolites is taken to be the same as the parent active ingredient. In the current version of the EFSA draft guidance a major metabolite is defined as one that is present at a value of 10% of the

* Corresponding author.
E-mail address: s.j.enoch@ljmu.ac.uk (S.J. Enoch).

https://doi.org/10.1016/j.yrtph.2022.105237
Received 9 March 2022; Received in revised form 7 July 2022; Accepted 19 July 2022
Available online 31 July 2022

administered dose in the urine in a repeated low dose study (EFSA, 2016). However, since genotoxicity studies are performed up to the maximum tolerable dose or limit dose, the metabolism that they cover is better reflected by the single high dose data available in the DAR documents. It should also be noted that the captured metabolism information is only taken at a single timepoint and that there is a flow between the different metabolites within a metabolic pathway. Therefore, rather than focussing on the amount of a single metabolite to define 'major' it may be preferable to include information from a complete metabolic pathway in which the sum of the metabolites in the pathway exceeds 10%.

Recent research has shown how so-called structural space alerts (where the term 'alert' is used to define the presence of a sub-structure within a chemical and not as an indication of toxicity. This is the same definition as is implemented in the profiling schemes within the OECD QSAR Toolbox) can be defined from an analysis of the genotoxicity and metabolism data available in the DAR/Renewal Assessment Report (RAR) documents (available from the EFSA website) (Enoch et al., 2022). This analysis showed how metabolic information could be used to drive the development of the structural space alerts – enabling chemical groupings to be defined in which common metabolic pathways were present in the analogues – something that has been identified as being a key measure of similarity (Gadaleta et al., 2020; Yordanova et al., 2021; Boyce et al., 2022). The analysis also showed how these structural space alerts could be used in conjunction with other profiling schemes (for example, those available in the OECD QSAR Toolbox) to build a weight of evidence for the prediction of Ames, chromosomal aberration, and the micronucleus assays via read-across. Importantly, this study showed how the weight of evidence approach could predict both *in vitro* and *in vivo* genotoxicity for sulphonyl urea residues. The aim of the current study was to extend the structural space alert concept to the triazole pesticides and to show how the metabolic pathway information as described above could be used as a source of secondary information to support read-across predictions.

2. Method

2.1. Dataset

A dataset of 66 triazole agrochemical active ingredients and metabolites with either Ames, *in vitro* chromosomal aberration or *in vivo* micronucleus test results were extracted from the 19 publicly available DAR/RAR documents and the recently published EFSA dataset of pesticide residues (Draft, 2022; Metruccio et al., 2017). Genotoxicity data were extracted for both chemicals that had been directly tested or for those chemicals that satisfied the definition of a major metabolite as outlined in the EFSA guidance (EFSA, 2016). The dataset, termed the 'triazole genotoxicity dataset' contained the following test results (*in vitro* assays with S9 fraction, Ames tests in the standard battery):

- Ames - 66 chemicals (all negative)
- *in vitro* chromosomal aberration - 42 chemicals (33 negative, 7 positive, 2 equivocal)
- *in vivo* micronucleus - 48 chemicals (all negative)

All chemical structures and associated toxicological data are available in the Supplementary Information.

2.2. Structural space alert development

The development of the structural space alerts utilised the following protocol (the overall process is summarised in Fig. 1):

1. Definition of the metabolic map for the triazole pesticides: This analysis involved inspection of the available metabolism data in the 19 DAR/RAR documents to identify metabolic transformations common to the triazole active ingredients (Draft, 2022). In terms of the triazole pesticides, these metabolic transformations were the hydrolysis reaction on either the triazole or aromatic rings followed by conjugation reactions. In addition, the majority of the triazole compounds undergo cleavage of the triazole moiety and functionalisation of the functional groups at the common bridge position. The metabolic map was developed, with the significance of these three key transformations for each triazole pesticide summarised.
2. Definition of significance of metabolic pathways: The metabolic pathways identified in step 1 of Fig. 1 were assigned to one of three levels of significance – minor, significant, or major. This involved summing the % dose values for metabolites within these pathways using the following data:
   a. Metabolic pathways were assigned as minor if the cumulative % dose information for the metabolites in the pathway was less than 5%, significant between 5 and 10%, and major when more than 10%
   b. Metabolic pathways were analysed in both the urine and faeces
   c. Where available, % dose data from male and female rats were averaged
   d. Data were taken from low repeated dose experiments
3. Metabolic scaffold identification: Common metabolic scaffolds were then identified from the metabolic map developed in step 1 of Fig. 1. This involved applying the metabolic transformations defined in the metabolic map to the triazole active ingredients. The resulting structures from this analysis were defined as metabolic scaffolds. These metabolic scaffolds were grouped together based on maximum common sub-substructures for development into structural space alerts.
4. Structural space alert development: The metabolic scaffolds identified in step 2 of Fig. 1 were used to profile the triazole genotoxicity dataset – metabolic scaffolds that had genotoxicity data associated with them (either Ames, chromosomal aberration, or micronucleus test data) were denoted as structural space alerts and encoded as SMARTS patterns. No additional physico-chemical boundaries were

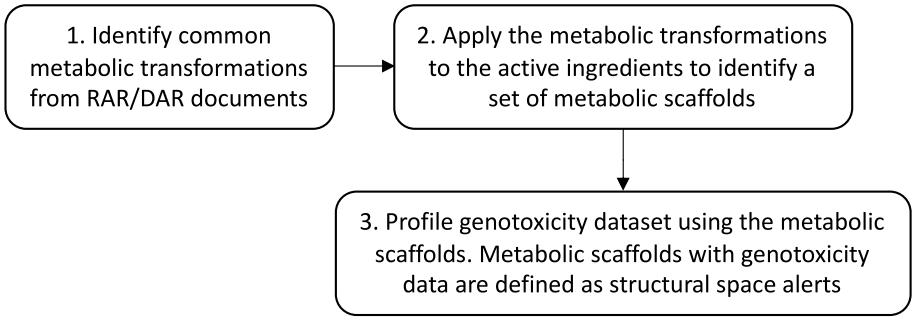


Fig. 1. Flow chart outlining the protocol for the development of structural space alerts.

imposed on the structural space alerts (enabling then to identify any chemical containing the alert sub-structure).

2.3. Chemical profiling

Chemicals in both datasets were profiled using the profiling schemes within the OECD QSAR Toolbox (V4.1.1). A subset of the available profilers was utilised based on the results of a previous study into their suitability for read-across predictions within the pesticide chemical space (Enoch et al., 2022). These profilers being (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
- Bioavailability (Lipinski) by OASIS

2.4. Read-across case study

Problem Formulation: Genotoxicity data for metabolites of epoxiconazole are lacking. The aim of this read-across was to fill data gaps this endpoint for 19 metabolites of epoxiconazole (Benigni et al., 2019). The hypothesis to be applied was that the metabolites are similar in terms of structural features and the absence of groups that may cause genotoxicity due to metabolism. Whilst this is an illustrative case study, the intention is to develop a read-across approach that would be suitable for regulatory consideration. Due to the importance of this endpoint and the regulatory consideration, the acceptable level of uncertainty would be low.

Endpoint: The endpoints investigated in the case study were the Ames test, *in vitro* chromosomal aberration, and *in vivo* micronucleus test. These three endpoints are important from a regulatory point of view as between them they cover gene mutation, as well as structural and numerical chromosomal damage.

Method: Each metabolite (the 'target' chemical) was assigned to a structural space category based on the presence of one or more alerts. Potential analogues, containing the same structural space alert (or alerts) were then identified from the triazole genotoxicity dataset. This set of chemicals formed the initial category. The structural domain of the category was assessed in terms of the presence (or absence) of structural alerts for DNA and protein reactivity relevant to genotoxicity. Analogues were removed from the category if they had a different profile from the target metabolite. Bioavailability was then assessed, again analogues with a different profile to the target metabolite were removed from the category. Finally, the metabolic similarity was assessed using the available experimental data in the DAR documents. Analogues with significantly different experimental pathways to the target metabolite were identified as metabolically dissimilar and were removed from the category. This analysis was primarily focused on phase one metabolism only due to the available data. Only analogues within both the structural and metabolic domain were used to make read-across predictions.

Uncertainty: The main source of uncertainty in the protocol outlined above is within the use of metabolism data to define metabolic similarity within a category. Specifically, what was considered as a different metabolic pathway leading to the exclusion of potential analogues. In the case study, experimental metabolism data were available for many of the parent chemicals.

3. Results and discussion

The aim of this study was to develop a series of structural alerts to define the structural space associated with a set of triazole pharmaceutical and agrochemical active ingredients, along with their metabolites, for which genotoxicity data exist. The development of these structural space alerts was driven by the need to prioritise potential metabolites (coming from plants and other animals) for further testing using a read-across approach and builds upon the previous work in this area (Enoch et al., 2022). The application of these structural space alerts to predict set of previously published metabolites from the pesticide, epoxiconazole, is also outlined (Benigni et al., 2019).

3.1. Structural space alert development

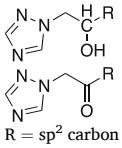
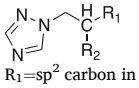
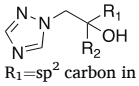
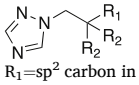
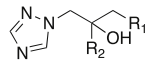
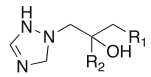
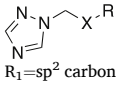
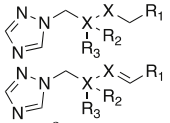
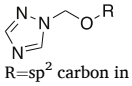
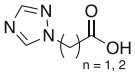
The structural space associated with the triazole genotoxicity dataset was defined through a set of ten structural space alerts as shown in Table 1. This set of structural space alerts identified 55 of the 66 chemicals in the triazole training dataset from which they were developed. Inspection of the remaining 11 chemicals showed them to have no common sub-structures suitable for further alert development (highlighted in the Supplementary Information). The sub-structures of these 10 structural space alerts are depicted in Table 1. All the *in vitro* Ames test, and *in vivo* micronucleus test results were negative. However, the *in vitro* chromosomal aberration test results showed at least one positive result within each structural space categories 2, 3, 5, 7, 8 and 10. Inspection of these positive results showed that in all cases a follow-up *in vivo* micronucleus test had been performed – the result of these tests being negative.

An important aspect in the development of structural space alerts is that they are linked to the key metabolic transformations present in the chemical class. In terms of the triazole dataset in the current study, many of these are related to the functional groups present at the branch point. For example, the presence or absence of the hydroxyl group in alerts 1–6 is related to the potential for this group to undergo glucuronidation. In addition, structural space alert one also covers the metabolic transformation that converts a secondary alcohol at this position into a ketone. In contrast to these well-defined structural features, the branch point for structural space alerts 2–6 was left undefined. These alerts capture a broad range of chemicals that can be subsequently separated into those that undergo chain cleavage reactions and those that do not. However, this metabolism is more complex structurally and is not easily encoded in a SMARTS pattern. Thus, an expert analysis of metabolic similarity of the potential category members is an important step to subcategorise this initial category into chemicals capable of such metabolism and those that are not.

3.2. Epoxiconazole read-across case study

The ten structural space alerts defined in Table 1 were used to profile the 21 epoxiconazole metabolites identified in a previous study (Benigni et al., 2019). Importantly, five of these metabolites represent the possible structures associated with hydroxylation events on either the para-fluorinated or the ortho-chlorinated aromatic rings (two and four structures, respectively). These multiple structures enumerate the various positions on these rings where metabolism could add the hydroxyl group. This being because the exact hydroxylation position on these rings was not (or could not) be determined in the original experimental metabolism studies. Thus, the dataset consisted of 17 unique metabolites (the case study dataset consists of two structures both denoted as M03 that represent potential hydroxylation on the *para*-fluorinated ring and four structures all denoted M01 that represent hydroxylation on the *ortho*-chlorinated ring). Of these unique metabolites, 15 were assigned to at least one structural space category – the exception being metabolites M08 and M52. In addition, six metabolites were assigned to two structural space categories – these being metabolites M04, M05, M12, M18, M19 and M56. In keeping with a previous study (Enoch et al., 2022), these chemicals were also profiled for potential DNA and protein reactivity relevant to genotoxicity and for their bioavailability (based on the Lipinski rule of five) using the OECD QSAR Toolbox V4.1.1. This profiling resulted in three of the 17 unique metabolites triggering an epoxide alert for protein binding relevant to chromosomal aberration (the two groups of metabolites representing the aromatic hydroxylation structures outlined above, and M62). Interestingly, no chemicals triggered any alert for DNA binding relevant

Table 1
Structural space alerts developed from the triazole genotoxicity dataset (explicitly defined hydrogen atoms as shown). CA = chromosomal aberration, MNT = micronucleus test. *In vitro* tests carried out in the presence of S9. Positive = number of chemicals containing the structural space alert that tested positive in each assay. Negative = number of chemicals containing the structural space alert that tested negative in each assay.

			In vitro		In vivo
Alert	Example sub-structure	Test	Ames	CA	MNT
1	 R = sp^2 carbon in a six-membered ring	Positive	0	0	0
		Negative	2	0	0
2	 R_1=sp^2 carbon in a six-membered ring R_2=any atom except hydrogen or hydroxyl	Positive	0	1	0
		Negative	10	6	9
3	 R_1=sp^2 carbon in a six-membered ring R_2=any atom except hydrogen or hydroxyl	Positive	0	1	0
		Negative	8	4	6
4	 R_1=sp^2 carbon in a six-membered ring R_2 = any atom except hydrogen or hydroxyl	Positive	0	0	0
		Negative	6	6	6
5	 R_1=sp^2 atom in a five/six-membered ring R_2=any atom except hydrogen or hydroxyl	Positive	0	2	0
		Negative	4	1	3
6	 R_1=sp^2 atom in a five/six-membered ring R_2=any atom except hydrogen or hydroxyl	Positive	0	0	0
		Negative	3	1	1
7	 R_1=sp^2 carbon in a six-membered ring X=sp^3 carbon atom in a ring	Positive	0	2	0
		Negative	9	4	6
8	 R_1=sp^2 carbon in a six-membered ring R_2/R_3=any atom except hydrogen X=sp^3/sp^2 carbon (ring or non-ring)	Positive	0	2	0
		Negative	11	7	9
9	 R=sp^2 carbon in a six-membered ring	Positive	0	0	0
		Negative	4	1	4
10		Positive	0	1	0
		Negative	3	1	2

to Ames, chromosomal aberration, or the micronucleus test. Finally, all chemicals were profiled as being bioavailable. Full profiling results are available in the Supplementary Information.

3.3. Genotoxicity read-across: structural space alert one

Structural similarity: Metabolites M06 and M07 were assigned to the structural space category defined by alert one. The structural space category associated with this alert consisted of two analogues (labelled (1) and (2) in Table 2), both of which had only been tested in the Ames assay (negative, no data for available for either the *in vitro* chromosomal aberration or *in vivo* micronucleus tests). All chemicals (targets and analogues) within the category contained no alerts for either DNA or protein binding relevant to genotoxicity and were predicted to be bioavailable. Thus, the analogues were within the same structural domain as the two target metabolites (all structures shown in Table 2).

Metabolic similarity: Analysis of available experimental metabolism data for the category members (targets and analogues) showed them to be metabolically related to one another due to the oxidation of the alcohol moiety to a ketone (Fig. 2). Thus, the analogues were considered within the same metabolic domain as the target metabolites.

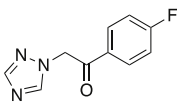
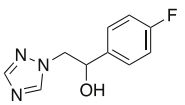
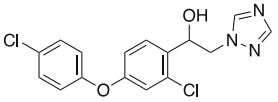
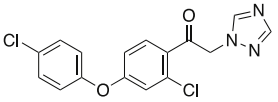
Read-across prediction: The result of the structural and metabolic similarity analysis enabled a many-to-many read-across prediction to be made for metabolites M06 and M07 based on the available experimental data. This prediction was for both chemicals to be negative in the Ames test. However, no data were available for the prediction of either *in vitro* chromosomal aberration or *in vivo* micronucleus test. These data gaps would need to be filled via targeted experimental testing as the Ames test alone does not assess structural and numerical chromosomal aberration.

3.4. Genotoxicity read-across: structural space alert two

Structural similarity: A single metabolite, M09 was assigned to structural space category two. This category consisted of an initial set of eight potential analogues with associated Ames, *in vitro* chromosomal aberration, and *in vivo* micronucleus test results. The profiling results for the category members showed none of them to contain a structural alert for DNA or protein reactivity relevant to the biological endpoints of interest. Chemical structures and associated biological data are as shown in Table 3.

Metabolic similarity: The available experimental metabolism data for the chemicals in this category showed them all to undergo metabolism of the alkyl chains resulting in sequential shortening of this chain. This metabolism involves sequential hydroxylation of the terminal position followed by oxidation to the carboxylic acid, which is then cleaved shortening the chain by one carbon atom (summarised in Fig. 3 for a simple alkyl chain). However, the target chemical features a thiol moiety at this position that cannot undergo this type of metabolism. Thus, the

Table 2
Chemical structures of the category defined by structural space alert 1. Available experimental data for the identified analogues (chemicals 1 and 2) as shown. No experimental genotoxicity data were available for M06 or M07. Data gaps filled by read-across as indicated (CA = *in vitro* chromosomal aberration, MNT = *in vivo* micronucleus test, R/A = read-across).

	
M06 (R/A. Ames: ve)	M07 (R/A. Ames: ve)
	
(1) Ames: ve	(2) Ames: ve

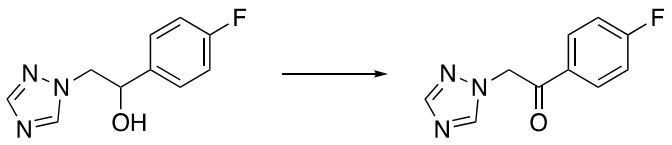
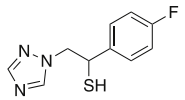
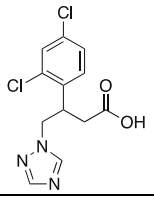
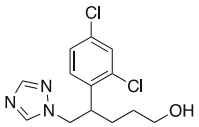
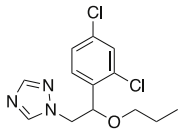
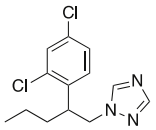
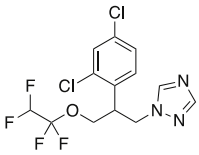
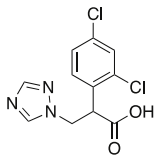
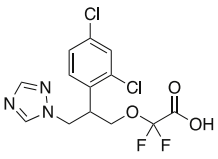
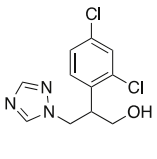


Fig. 2. Metabolic transformation of an alcohol to a ketone that defines the metabolic similarity for the category defined by structural space alert one.

Table 3
Chemical structures of the category defined by structural space alert two. Available experimental data for the identified analogues (chemicals 1–8) as shown. No experimental genotoxicity data were available for M09. (CA = *in vitro* chromosomal aberration, MNT = *in vivo* micronucleus test, R/A = read-across).

		
M09 (No prediction)	(1) Ames, CA, MNT: ve	(2) Ames, CA, MNT: ve
		
(3) Ames, CA, MNT: ve	(4) Ames, CA, MNT: ve	(5) Ames, CA, MNT: ve
		
(6) Ames, MNT: ve; CA: +ve	(7) Ames: ve	(8) Ames, MNT: ve

eight analogues are metabolically similar to one another, but not to the target chemical of interest. Given the difference in the key metabolic transformations between target and analogues, the target was considered to be out of the metabolic domain of the category.

Read-across prediction: No read-across prediction was possible for metabolite M08 due to this chemical being out of the metabolic domain of the category defined by structural space alert two. Further *in vitro* testing (in the first instance) would be required to assess the genotoxicity of this chemical.

3.5. Genotoxicity read-across: structural space alerts three and five

Structural similarity: Seven metabolites were assigned to structural space category defined by alerts three and five. These being M04, M05, M12, M18, M19 and M56. This category contained two potential analogues, all chemicals within the category were profiled as having no structural alerts related to DNA or protein reactivity and were predicted to be bioavailable. Thus, the analogues and target chemicals were identified as being within the same structural domain. Chemical structures and toxicological data are as shown in Table 4.

Metabolic similarity: In keeping with the other chemicals discussed, the experimental metabolism data for the chemicals in this category showed the aromatic rings to undergo hydroxylation. Metabolites M12, M18, M19 and M56 were also shown to undergo functionalisation of the thiol moieties, suggesting the presence of an additional metabolic pathway not present in the remaining category members (Fig. 4). Thus, the available experimental data suggested the four thiol containing metabolites to be out of the domain of the remaining category members (due to the presence of an additional key metabolic pathway).

Read-across prediction: The structural and metabolic similarity analysis outlined above enabled the genotoxicity of metabolites M04 and M05 to be predicted via a many-to-many read-across. These predictions being negative in the Ames and *in vivo* micronucleus assays. In contrast, no prediction was possible for the *in vitro* chromosomal aberration test due to the equivocal nature of the available data. No read-across predictions were possible metabolites M12, M18, M19 and M56 due to them being out of the category domain. These chemicals would require further analysis and/or experimental testing to establish their

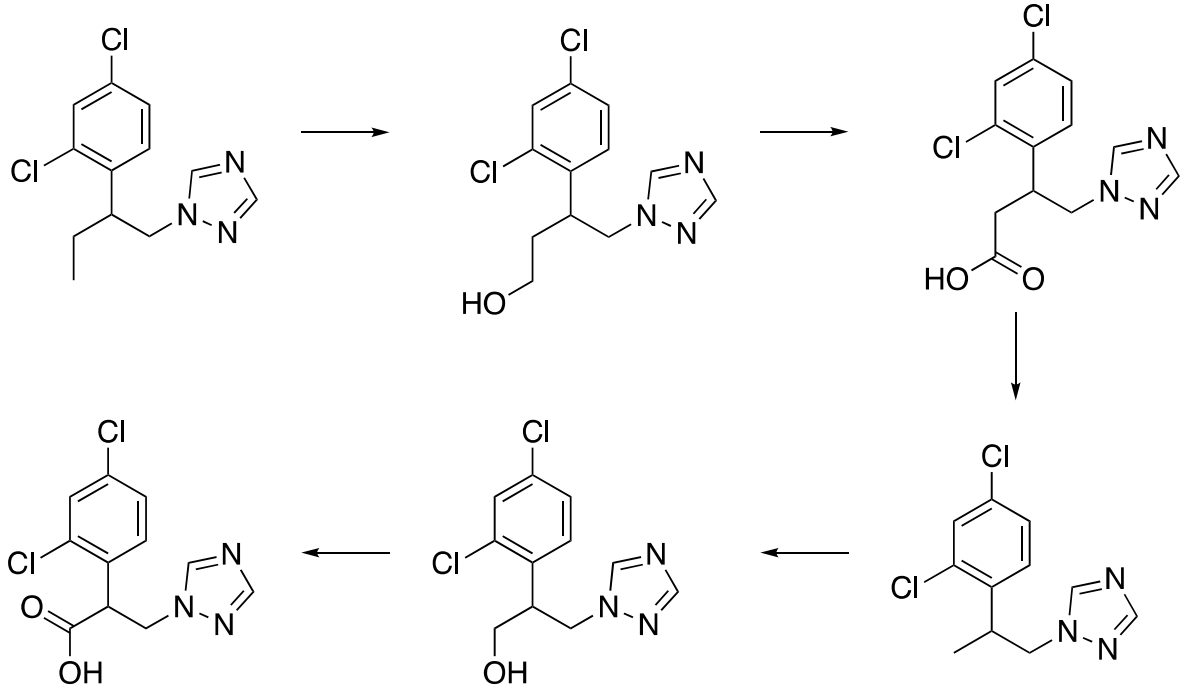
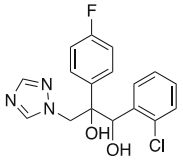
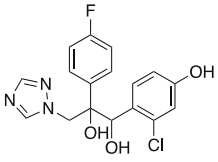
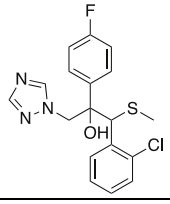
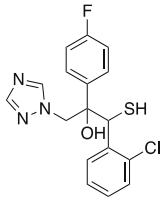
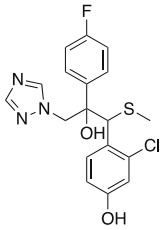
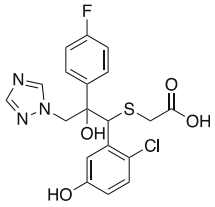
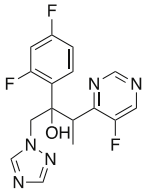
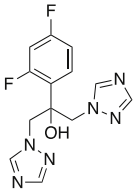


Fig. 3. Alkyl chain metabolism resulting in sequential shortening of the types of alkyl chain present in the category defined by structural space alert two.

Table 4
Chemical structures of category defined by structural space alert three. Available experimental data for the identified analogues (chemicals 1 and 2) as shown. No experimental genotoxicity data were available for chemicals M04, M05, M12, M18, M19 and M56. Data-gaps filled by read-across as indicated. (CA = *in vitro* chromosomal aberration, MNT = *in vivo* micronucleus test, R/A = read-across).

		
M04 (R/A. Ames, MNT: ve)	M05 (R/A. Ames, MNT: ve)	M12 (No prediction)
		
M18 (No prediction)	M19 (No prediction)	M56 (No prediction)
		
(1) Ames, MNT: ve; CA: +ve	(2) Ames, CA, MNT: ve	

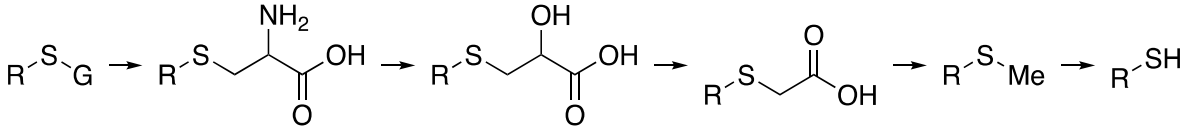


Fig. 4. Experimentally observed metabolic pathway relating metabolites M12, M18, M19 and M56 to one another (G = glutathione, R = structure of M12, M18, M19 and M56).

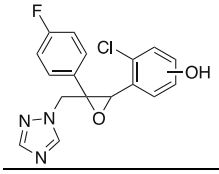
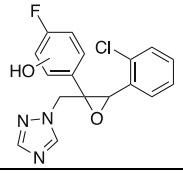
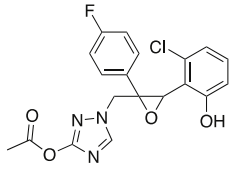
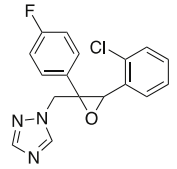
genotoxicity, specifically in relation to the potential effect of the thiol metabolic pathway.

3.6. Genotoxicity read-across: structural space alert seven

Structural similarity: Three metabolites were assigned to structural space category seven, these being metabolites M01, M03, and M62. In contrast to all the other metabolites of epoxiconazole, these chemicals triggered an alert for protein binding relevant to chromosomal aberration. Initial inspection of this structural space category showed it to contain nine analogues; however, only one of these analogues triggered the same protein binding alert. Sub-categorising the structural space category for the presence of this protein binding alert resulted in a single analogue within the category (no alert for DNA binding was triggered in any of the chemicals). The three target metabolites and single analogue were profiled as being bioavailable, suggesting this set of category members to be within the same structural space. Chemical structures and toxicological data are as shown in Table 5.

Metabolic similarity: The available metabolic data for the four category members showed them all to undergo the same common ring opening hydrolysis reaction (Fig. 5). This hydrolysis reaction presumably occurs rapidly preventing the epoxide ring from being able to react with nucleophilic centres in either DNA or proteins – thus rendering it unreactive in this set of chemicals. As expected, the other key transformations were hydroxylation reactions on the aromatic ring systems.

Table 5
Chemical structures of category defined by structural space alert three. Available experimental data for the identified analogues (chemical 1) as shown. No experimental genotoxicity data were available for chemicals M01, M03 and M62. Data-gaps filled by read-across as indicated (CA = *in vitro* chromosomal aberration, MNT = *in vivo* micronucleus test, R/A = read-across).

	
M01 (R/A. Ames, CA, MNT: ve)	M03 (R/A. Ames, CA, MNT: ve)
	
M62 (R/A. Ames, CA, MNT: ve)	(1) Ames, CA, MNT: ve

The presence of this common set of transformations suggested the target and analogues within the category to be in the same metabolic domain.

Read-across prediction: The result of the structural and metabolic

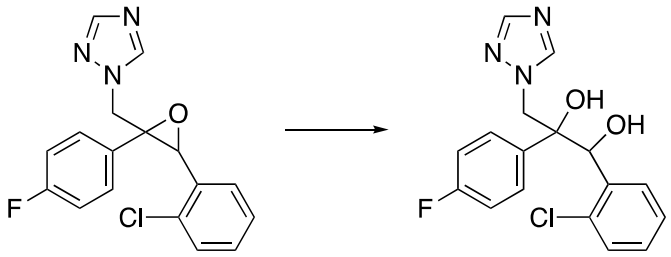


Fig. 5. Ring opening hydrolysis reaction leading to the detoxification of the epoxide moiety present in the metabolites M01, M03, M62 and epoxiconazole (the structure on the left is epoxiconazole).

similarity analysis enabled a one-to-many read-across prediction to be made for metabolites M01, M03 and M62 based on the available experimental data. The three target metabolites being predicted as negative in the Ames, *in vitro* chromosomal aberration and *in vivo* micronucleus test assay via a one-to-many read-across.

3.7. Genotoxicity read-across: structural space alert eight

Structural similarity: The final category, defined based on the structural space alert eight, contained three metabolites. All chemicals within the category were profiled as having no structural alerts related to DNA or protein reactivity and were predicted to be bioavailable. Thus, the analogues and target chemicals were identified as being within the same structural domain. Chemical structures and toxicological data are as shown in Table 6.

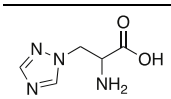
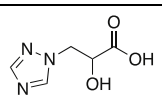
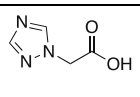
Metabolic similarity: The available experimental metabolism data showed the three category members to be related to be metabolically related to one another with M49 being converted into M51 via metabolite M50 (Fig. 6). Thus, the category members were considered to within a single metabolic domain.

Read-across prediction: The well-defined structural and metabolic domains for the category enabled the genotoxicity data-gaps for these chemicals to be filled via many-to-one read-across. These predictions being negative for the Ames, *in vitro* chromosomal aberration, and *in vivo* micronucleus tests for metabolite M50. Negative predictions were also made for metabolite M49 (*in vitro* chromosomal aberration) and M51 (*in vivo* micronucleus test).

3.8. Secondary lines of evidence and available ADME data

The key challenge in the structural space alert approach is the availability of metabolites with genotoxicity data. In the triazole dataset defined in the current study the published EFSA guidelines were followed in the collection of these data (EFSA, 2016). However, strict adherence to these guidelines prevents the inclusion of additional metabolic information that can be potentially used as secondary lines of evidence to support a read-across prediction. The key available data being:

Table 6
Chemical structures of category defined by structural space alert three. Available experimental data for the identified for chemicals M49, M50 and M51 as shown. Data-gaps filled by read-across as indicated (CA = *in vitro* chromosomal aberration, MNT = *in vivo* micronucleus test, R/A = read-across).

		
M49 Ames, MNT: ve (R/A. CA: ve)	M50 (R/A. Ames, CA, MNT: ve)	M51 Ames, CA: ve (R/A. MNT: ve)

- Single metabolites observed in excess of 10% of the administered dose that occur in only one sex
- Additional dosing regimens – typically a combination of single low dose, single high dose, and low repeated dose are available in the DAR documents. These data can potentially be used in combination with one another to build a weight of evidence in the identification of key metabolic pathways
- The use of metabolite information drawn from either bile or faecal samples (or a combination of urine and faecal). For example, within the triazoles DAR documents many more metabolites are observed at higher %dose values in the faeces than the urine – these data are especially useful when considering the metabolites that are likely to occur in *vitro* assay using the S9 mix. In addition, inspection of these data can also help in the identification of the key metabolic pathways for a given chemical
- The ability to identify a key metabolic pathway based on the cumulative %doses of the phase I and/or phase II metabolites in the pathway. Such information is not generally captured when metabolite data are collected by inspection of the %dose value associated with a single metabolite (or its conjugates) as is recommended in the EFSA guidance (see Table 8 for a summary of this information for the triazole pesticides)

As an example of the how these secondary data can be useful in adding weight of evidence to a read-across prediction consider metabolites M12, M18, M19 and M56 assigned to the category defined by structural space alert 3 (Table 4). It was not possible to make a read-across prediction for these chemicals based on the data available in the triazole dataset due to the presence of a glutathione metabolic pathway (that was not present in the remaining category members). The key uncertainty being whether this additional metabolic pathway could lead to genotoxicity in these four metabolites. Inspection of the ADME data for epoxiconazole showed there to be no metabolites that meet the currently published EFSA guidance capable of providing suitable evidence (EFSA, 2016; Draft, 2022). However, there is metabolic information from an analysis of the faeces that could be used to build confidence that the glutathione pathway that the four metabolites are part of is unlikely to result in genotoxicity (these metabolites occur as part of the transformation of M41 into M18 in the pathway shown in Fig. 7). These data showed the final metabolite, M43, in this pathway to be present at around 10% of the administered dose in low single dose, and low repeat dose in at least one sex (Table 7). These data suggest that this metabolic pathway is significant, and that it is unlikely to lead to genotoxicity (based on utilising the experimental data available for the parent molecule, epoxiconazole). Utilising this additional information enables metabolites M12, M18, M19 and M56 to be assigned as within the metabolic domain of the category defined by structural space alert 3 (due to evidence of the glutathione pathway not being associated with genotoxicity). In doing so, this enables their genotoxicity to be predicted as negative in the Ames and *in vivo* micronucleus assays.

The above example outlines how the extensive data available in the DAR documents can be utilised to support chemical category formation and read-across. However, there are several difficulties in utilising these data in the development of informatics tools and in the support of regulatory submissions. The key problems being:

- Ill-defined metabolite structures: it is frequently the case that hydroxylation and conjugation reactions upon aromatic ring systems either have not been/or could not be resolved in the ADME studies. Alternatively, such reactions are placed arbitrarily at a single position within the resulting DAR documents. There is a clear need for a standardised approach to the reporting of such structures and/or the development of informatics tools that can encode this type of structural uncertainty (current tools all require a defined chemical structure).

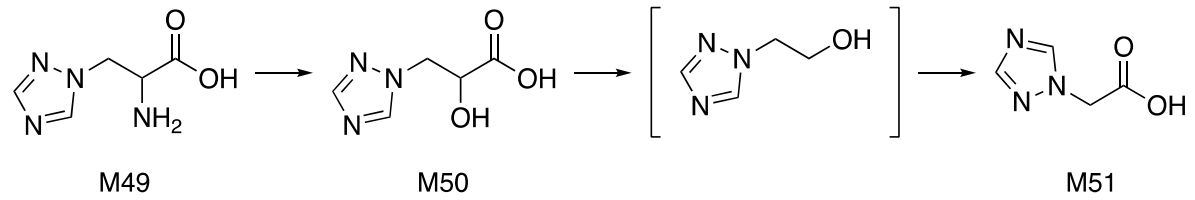


Fig. 6. Metabolic conversion of metabolites M49 into M51 via a chain shortening reaction involving metabolite M50 (structure shown in square brackets shows the predicted intermediate required for the chain shortening reaction linking M50 and M51).

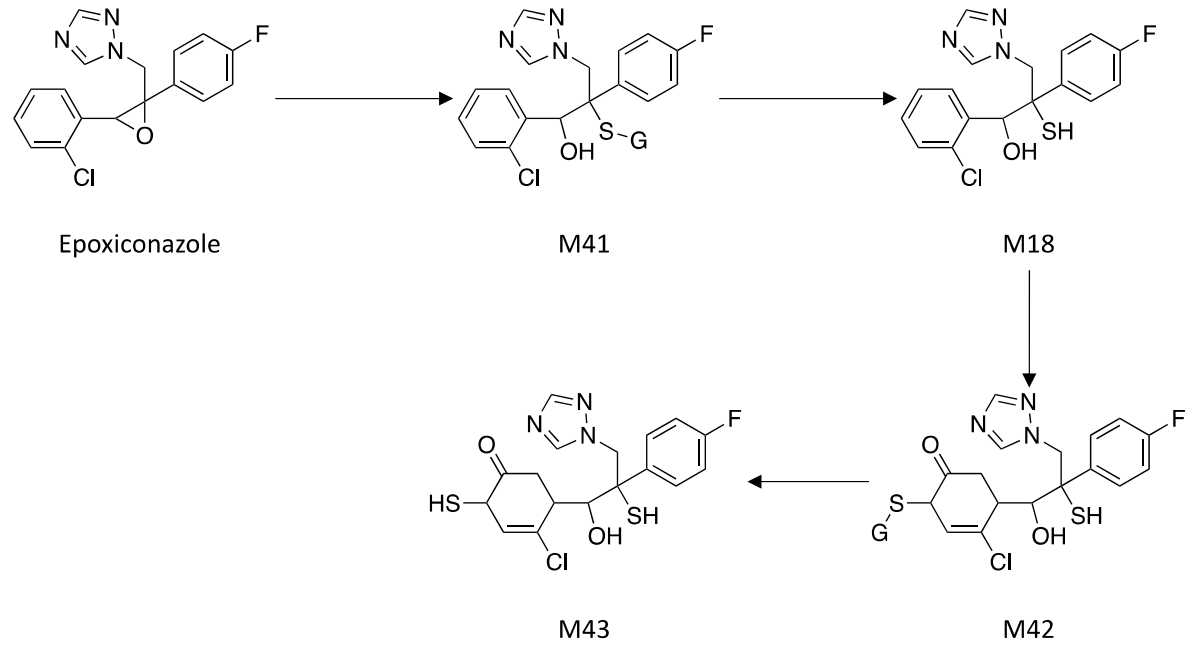


Fig. 7. Glutathione-based metabolic pathway of epoxiconazole involving an initial ring opening reaction to produce metabolites M41 and M18, followed by further reactions on the *ortho*-chlorinated aromatic ring, ultimately leading to the production of metabolite M43.

Table 7
Summary of the identified metabolites shown in Fig. 3 administered at single low dose, single high dose and low repeat dose (values as indicated). Values are expressed as the % dose of orally administered epoxiconazole.

Metabolite	Single low dose (3 mg/kg/bw)		Single high dose (100 mg/kg/bw)		Low repeat dose (3 mg/kg/bw)	
	Male	Female	Male	Female	Male	Female
M18	-	-	0.7	0.4	-	-
M43	10.3	4.2	4.7	1.6	9.8	10.2

- Grouping of metabolite structures: the ability to assess a set of metabolites within a metabolic pathway is significantly hindered due to %dose value for every metabolite structure not always being identified. This is due to the difficultly in distinguishing metabolite structures that (for example) co-elute in the HPLC analysis. This difficulty results in groups of metabolites being reported in the DAR documents with a single %dose value. Inspection of these groups of metabolites often shows them to come from differing metabolic pathways making it impossible to assess the cumulative contribution of an individual metabolite within the group to a specific pathway of interest.
- Available matrices and dosing: the availability of consistent data from key matrices at comparable dosing regimens is also problematic. Whilst many of the DAR documents contain data from low repeat dose experiments in the urine (the key data as required by the EFSA guidance), some do not. In addition, data from single high dose experiments that may be more useful as secondary information in support of a read-across prediction are not consistently available. The same is true of the relevant matrices with some chemicals have data for urine, bile, and faeces and others simply detailing a summation of urine and faeces.
- Phase two metabolite identification: there is inconsistency in the DAR documents in the identification and reporting of phase two metabolites. This leads to an inconsistency in the assessment of the contribution of these metabolites to the overall significance of a metabolic pathway. This problem is further complicated by the difficulty in resolving the position of phase one hydroxylation reactions upon aromatic ring systems in the experiments (as detailed above), which leads to ill-defined phase two metabolite structures.
- Data reporting: there is no standardised method for the reporting of ADME data within the DAR documents. This leads to different documents containing similar ADME information that is formatted in different ways – this makes extracting relevant information a challenging and time-consuming process as every document needs to be read in detail.

The solution to these challenges is a move away from the 'paper-based' DAR documents towards a standardised electronic reporting format. The creation of such a tool is already underway in terms of the move towards the use of MetaPath for the reporting of ADME data (Kolanczyk et al., 2012; LMC), 2021). The ability to then search within MetaPath for a metabolite, to examine the metabolic path to which it belongs and to retrieve the corresponding ADME and/or toxicity data will make the development of informatics tools for read-across

Table 8
Summary of pathway significance for the metabolic transformations in the urine, faeces, or both outlined in Fig. 8 (n.d. = not determinable, Alt. = alternative metabolism due to the presence of an aliphatic five-membered ring in the triazole pesticide backbone).

Triazole	Matrix	Pathway A	Pathway B	Pathways C, D, E or F	
Bitertanol	Urine	n.d.	Minor	Minor	E
	Faeces	Significant	Major	Major	
Bromuconazole	Urine + Faeces	n.d.	Major	Major	D
Cyproconazole	Urine	Major	n.d.	Significant	F
	Faeces	Major	n.d.	Major	
Difeconazole	Urine	Major	Minor	Minor	D
	Faeces	Minor	Major	Major	
Epoxiconazole	Urine	n.d.	Minor	Minor	D
	Faeces	n.d.	Major	Major	
Fenbuconazole	Urine	n.d.	Major	Major	F
	Faeces	n.d.	Minor	Major	
Fluquiconazole	Urine	Major	n.d.	Minor	None
	Faeces	n.d.	n.d.	n.d.	
Flutriafol	Urine + Faeces	Minor	n.d.	Major	F
Ipconazole	Urine	Significant	Minor	Minor	Alt.
Mefentrifluconazole	Urine	Major	Major	Major	F
	Faeces	n.d.	Major	n.d.	
Metconazole	Urine	Minor	Significant	Significant	Alt.
	Faeces	n.d.	Significant	Major	
Myclobutanil	Urine + Faeces	n.d.	n.d.	Major	F
Penconazole	Urine	Major	n.d.	Major	E
	Faeces	Minor	n.d.	n.d.	
Propiconazole	Urine	n.d.	n.d.	Major	D
	Faeces	n.d.	Major	Major	
Prothioconazole	Urine	Minor	Minor	Minor	F
	Faeces	n.d.	Major	Major	
Tebuconazole	Urine	Minor	n.d.	Major	F
	Urine + Faeces	Minor	Minor	Major	
Tetraconazole	Urine	Major	n.d.	Major	E
	Faeces	Significant	n.d.	Major	
Triadimenol	Urine + Faeces	n.d.	n.d.	Major	E
Triticonazole	Urine	n.d.	n.d.	Major	Alt.
	Faeces	n.d.	Significant	Major	

significantly easier. In addition, the ability to subsequently identify secondary lines of evidence to support read-across will also become easier and more accepted.

3.9. Metabolic space and structural space alerts

As previously discussed, the key advantage of the structural space alert approach is that structural space alerts are developed from an analysis of parent and metabolite structures for a given chemical class (Enoch et al., 2022). This results in the alerts defining a set of common metabolic scaffolds that define metabolic space for which genotoxicity data exist. This ensures that chemicals grouped together using a specific structural space alert are likely to have a degree of metabolic similarity – an important factor in category formation (Boyce et al., 2022; Schultz and Cronin, 2017; Kuseva et al., 2021). The metabolic space associated with the triazoles chemical class can be summarised as follows (denoted A, B, C, D, E and F in Fig. 8):

A. Triazole cleavage (relevant to structural space alert 10). The majority of the triazole pesticides undergo cleavage of the triazole moiety. This moiety has then been shown to undergo further functionalisation reactions involving the addition of carboxylic acid functional groups.
B. Hydroxylation reactions on the aromatic and triazole rings. The experimental metabolism data showed that six-membered aromatic rings present in the triazole class are readily hydroxylated during phase 1 metabolism. Similar reactions have been observed on the five-membered triazole moiety; however, these are observed in significantly fewer chemicals. In all cases, experimental data showed the hydroxyl groups to be further conjugated during phase 2 metabolism. The commonality in this type of transformation prevented it from being relevant to structural space alert development.
C. Oxidation reactions (relevant to structural space alert 1). This type of metabolic reaction involves the oxidation of the alcohol present at the branch point into a ketone. This reaction cannot occur when the second R group at this position is any other atom expect hydrogen. In addition, the data also showed the hydroxyl group to undergo phase 2 conjugation reactions (not shown in Fig. 8).
D. Ring opening and/or functionalisation reactions (relevant to structural space alert 7). Several of the parent triazoles contain aliphatic heterocyclic ring systems at the branch point (where R_1 and R_2 are part of the same ring – epoxiconazole is shown in Fig. 4). These ring systems are either ring opened (for example, epoxiconazole) or further functionalised and/or undergo rearrangement (bromuconazole, difenoconazole, propiconazole).
E. Alkyl chain cleavage reactions (relevant to structural space alerts 2, 4, 9). The presence of an alkyl chain at the branch position results in a series of cleavage reactions (R_2 = hydrogen in Fig. 8). These reactions typically involve addition of hydroxyl groups, followed by oxidation of this group to a carboxylic acid. Metabolism then removes the carboxylic acid group resulting in the alkyl chain being shortened by a single carbon. This series of reactions can occur repeatedly ultimately converting the chain into a carboxylic acid moiety. Other related reactions are possible for alkyl chains containing ether links. Importantly, the presence of a terminal ring system prevents this metabolic pathway.
F. Alkyl chain cleavage reactions (relevant to structural space alerts 3, 5 and 6). The presence of an alkyl chain at the branch position results in a series of cleavage reactions (R_2 = hydroxyl in Fig. 8). These reactions are analogous to those that occur for pathway E. However, the presence of the hydroxyl group at the branch point also enables conjugation reactions to occur at this position. As for pathway E, the presence of terminal ring systems in position R_1 prevents the chain cleavage reactions.

The scaffold shown in Fig. 8 can also be extended to include a five-membered aliphatic ring into the backbone of the triazole pesticide (ipconazole, metconazole, and triticonazole). This results in an alternative set of metabolic transformations involving hydroxylation reactions on the five-membered aliphatic ring followed by conjugation reactions (this is covered by structural space alert 8 and is denoted as 'Alt' in Table 8). Finally, the example scaffold can also include an additional carbon atom between branch point (the position at which R_1 and R_2 are defined) and the six-membered aromatic ring (prothioconazole and tebuconazole, covered by structural space alerts 5 and 6).

4. Conclusions

The aim of this study was to develop a set of structural alerts to define the structural space associated with a set of triazole agrochemicals. This analysis resulted in a set of ten structural space alerts developed from a dataset of 66 chemicals for which either Ames, chromosomal aberration of micronucleus test results were publicly available. The structural space alerts were developed based on the common metabolic transformations for the triazole chemical class. This study demonstrated how a combination of these structural space alerts, covalent chemistry profiling and physico-chemistry properties could be used to develop well-defined

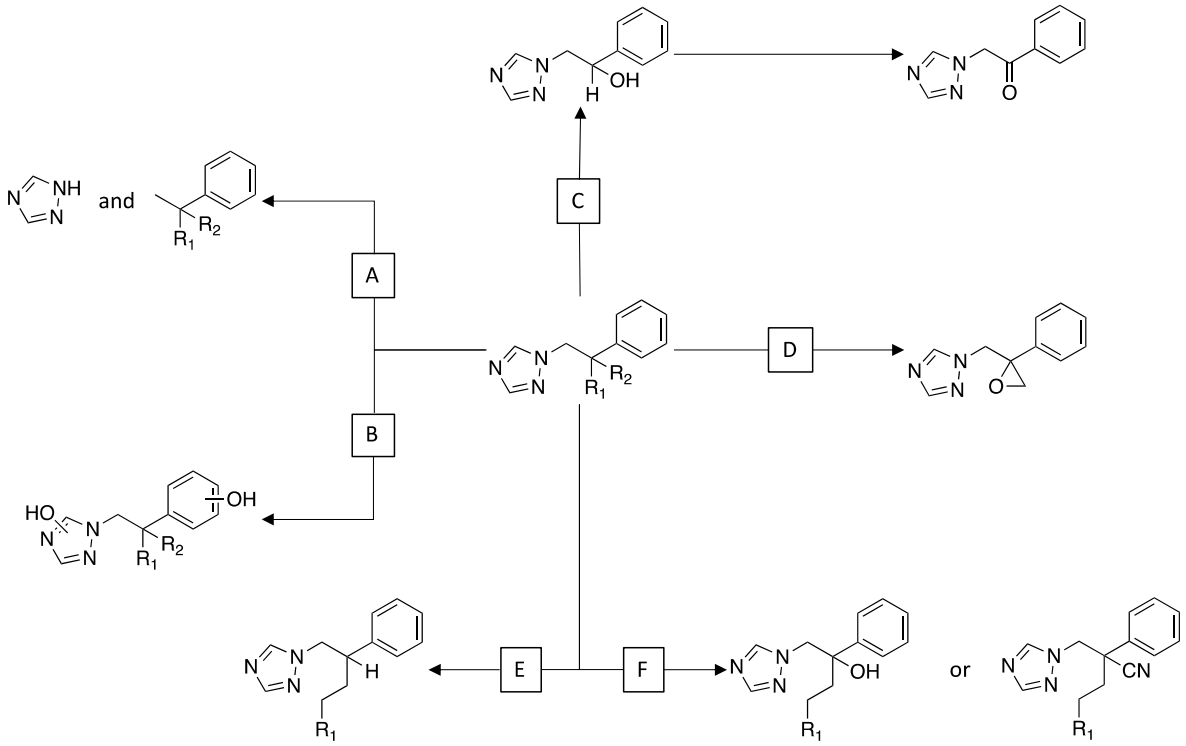


Fig. 8. Overview of the common types of metabolic transformations present in the triazole genotoxicity dataset (R_1 and R_2 as defined in the text).

chemical categories suitable for chemical prioritisation for genotoxicity for triazole pesticides residues. The approach was exemplified using a case study approach that showed these chemical categories could be used to predict genotoxicity or prioritise triazole residues for further targeted testing. Importantly, this approach enabled an assessment gene mutation as well as structural and numerical chromosomal aberration. These case studies also demonstrated how secondary metabolism data could be utilised to support the read-across predictions. Finally, the definition of metabolism-linked structural space alerts have the potential to expand commonly used structural-feature fingerprints such as MACCS keys, expanding their applicability to pesticides.

Funding body

BASF are gratefully acknowledged for funding this research.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research received funding from BASF.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.yrtph.2022.105237.

References

Benigni, R., et al., 2019. Evaluation of the applicability of existing (Q)SAR models for predicting the genotoxicity of pesticides and similarity analysis related with genotoxicity of pesticides for facilitating of grouping and read across. EFSA Supporting Publications 16 (3).

Boyce, M., et al., 2022. Comparing the performance and coverage of selected in silico (liver) metabolism tools relative to reported studies in the literature to inform analogue selection in read-across: a case study. Computational Toxicology 21.

Draft, E.F.S.A., 2022. Assessment Report (DAR)/Renewal Assessment Report (RAR) Documents. Available from. www.efsa.europa.eu/en.

ECHA, 2017. Read-across assessment framework (RAAF). Available from. http://www.echa.europa.eu.

EFSA, 2011. Scientific opinion on genotoxicity testing strategies applicable to food and feed safety assessment. EFSA J. 9 (9).

EFSA, 2016. Guidance on the establishment of the residue definition for dietary risk assessment. EFSA J. 14 (12).

Enoch, S.J., et al., 2022. Sub-structure-based category formation for the prioritisation of genotoxicity hazard assessment for pesticide residues: sulphonyl ureas. Regul. Toxicol. Pharmacol. 129, 105115.

Gadaleta, D., et al., 2020. Automated integration of structural, biological and metabolic similarities to improve read-across. ALTEX 37 (3), 469–481.

Kolanczyk, R.C., et al., 2012. MetaPath: an electronic knowledge base for collating, exchanging and analyzing case studies of xenobiotic metabolism. Regul. Toxicol. Pharmacol. 63 (1), 84–96.

Kuseva, C., et al., 2021. Criteria for quantitative assessment of metabolic similarity between chemicals. II. Application to human health endpoints. Computational Toxicology 19.

Lmc), L.M.C., 2021. MetaPath: Consolidation and Management of Metabolism Data.

Metruccio, F., et al., 2017. Compilation of a database, specific for the pesticide active substance and their metabolites, comprising the main genotoxicity endpoints. EFSA Supporting Publications 14 (5).

Schultz, T.W., Cronin, M.T.D., 2017. Lessons learned from read-across case studies for repeated-dose toxicity. Regul. Toxicol. Pharmacol. 88, 185–191.

Yordanova, D.G., et al., 2021. Assessing metabolic similarity for read-across predictions. Computational Toxicology 18.