TOXICOLOGICAL ASSESSMENT OF OECD TG 443 STUDIES

General VCI Comment on the Proposal for Harmonised Classification and Labelling of "Propyl 4-hydroxybenzoate"

Belgium submitted a harmonised classification and labelling (CLH) dossier for Propyl 4-hydroxybenzoate for inclusion in Annex VI of the CLP Regulation.[1] The justification of the dossier submitter for a classification of Propyl 4-hydroxybenzoate as Repr. 2 (H361fd) is based on the assessment of an Extended One-Generation Reproductive Toxicity (EOGRT, OECD TG 443) Study, as well as individual parameters from non-GLP and non-OECD studies.
Here, VCI would like to make a general comment firstly on fundamental toxicological aspects during CLH evaluation and secondly on the evaluation of OECD TG 443 studies, which are still under discussion because of an ongoing ECHA project. Substance-specific comments are outside the scope of this consultation submission. Rather, VCI would like to provide the following key statements:

- Adherence of scientific principles for toxicological evaluation considering all relevant data (statistically significance and biological relevance of effects, use of historical data, consideration of dose dependency).
- Functional parameters related to observations for isolated values within the OECD TG 443 should be considered in order to evaluate the adversity of effects, evidence of adversity, evidence of causality.
- Isolated evaluation of single biological parameters without consideration of the whole database and other toxicological OECD-conform studies is not scientifically justified and should be performed in a Weight of Evidence approach.
- Relevance of toxicological observation for classification and labelling.
- Results of the ongoing ECHA project assessing OECD TG 443 studies should be taken into account.

In the **Annex** we refer to the relevant legal texts from the CLP Regulation.

Scientific principles for toxicological evaluation as basis for the hazard classification

Toxicological assessments (hazard identification and characterisation) are carried out by using among others structure-activity relationships (SAR), cell cultures (in vitro) or animal experiments

[1] ECHA, Harmonisierte Einstufung und Kennzeichnung, 03.06.2022 (08:30)
https://echa.europa.eu/de/harmonised-classification-and-labelling-consultation/-/substance-rev/69402/term

(in vivo).[2] For a scientifically justified toxicological evaluation of higher tier studies (in vivo studies) especially the OECD TG 443 following basic considerations in the assessment of toxicological relevant effects should be considered:

- Statistically significance vs. Biological relevance,
- Comparison of parameters with historical control data,
- Evidence of adversity,
- Evidence of causality,
- Weight-of-evidence evaluation.

Statistical Significance vs. Biological relevance

Firstly, statistical evaluations and methods are part of the tools used to identify potential hazards. Statistically significant changes in biological parameters could be considered as a first approach to identify toxicologically relevant effects. However, a statistical significance is not equivalent to biological relevance. A statistically "significant" effect can be biologically irrelevant and vice versa.[3] Working with biological systems makes data evaluation complex. Use of historical data allows to follow the biological response of the test system and is therefore an important tool in effect interpretation. Due to biological variance, data have a natural range that still allows for robust assessment. The evaluation of isolated effects that are even statistically significant relative to the concurrent control, is nevertheless not justified for classification and labelling decisions without consideration of the overall biological variability by taking the range of historical control data into account.

Historical control data

Secondly, historical control data are used to assess biological variability with increased power, as compared to the relatively small concurrent control group. For ethical reasons a concurrent control group size should be kept reasonable low in vertebrate studies. To overcome this problem, historical control data are an important tool to increase the strength of the investigation and thus allowing to distinguish true responses from chance findings. Such considerations are especially relevant if there are regulatory hazard categories like “present” or “not present” as it is the case for reproductive toxicity chemical hazard classification according to the globally harmonised system (GHS) or CLP.

Evidence of adversity

Thirdly, after checking for significance and variance of observations, the effects must be related into correlation with functional parameters and other related effects. Single Alterations in

[2] BfR, Toxikologische Beurteilung von chemischen Stoffen, 03.06.2022 (08:30), https://www.bfr.bund.de/de/toxikologische_beurteilung_von_chemischen_stoffen-70287.html

[3] EFSA, Statistical Significance and Biological Relevance, EFSA Journal, 2011, https://efsa.onlinelibrary.wiley.com/doi/pdf/10.2903/j.efsa.2011.2372

clinical chemistry or organ weights alone (especially in moderate extent) does not bear major adversity. Are there several parameters altered which are plausibly linked (functional parameters, clinical chemistry with organ weights, histopathology, etc.) then adversity should be discussed.

Evidence of causality

Fourthly, a further aspect of assessment of effects within reproductive studies is the question, whether effects are probably caused by a primary mechanism or a secondary mechanism due to maternal toxicity or due to overload phenomenon where physiological processes are out of biological ranges. Such effects may be discriminated by dose response relationship. Non-significant alterations at the highest dose level in isolation should not be taken for classification and labelling without taken all available data into account.

Weight-of-evidence evaluation

Finally, a weight-of-evidence should be conducted taking all available data into account and not only some potential positive observations in single studies. Therefore, all data should be taken into account to evaluate individual endpoints relevant for classification. This is of particular importance for statistically non-significant observations considered as relevant for classification. In this case the biological relevance should consider historical data provided by the laboratory, coherence of cohorts, absence or presence of dose response and related functional endpoints.

Examples within the CLH-Dossier Propyl 4-hydroxybenzoate to highlight the above mentioned points:

- **Statistical Significance vs. Biological relevance:** The concerns raised in the submitted CLH-dossier regarding developmental toxicity based on slightly increased post-implantation loss in an OECD TG 443 are not statistically significant. Moreover, they could not be revealed in both generations of the OECD TG 443 indicating that these effects are due to biological variability rather than test item related.
- **Historical control data:** The submitted CLH-dossier raised a concern regarding developmental toxicity based on slightly increased post-implantation loss and anogenital distance (AGD) in an OECD TG 443 and a concern regarding reproductive toxicity based on slight effects on sperm morphology. However, all values were within the range of historical control data which was not considered by the evaluating member state.
- **Evidence of adversity:** The submitted CLH-dossier raised a concern regarding developmental toxicity based on slightly increased post-implantation loss and AGD in an OECD TG 443 and a concern regarding reproductive toxicity based on slight effects on sperm morphology. As mentioned, none of these effects were statistically significant and all values were within the range of historical control data. More importantly functional parameters related to these observation (e.g. mating index, fertility index, Pup parameters, gender specific reproduce and developmental endpoints etc.) were not affected. However, such

considerations are essential to avoid misrepresentative analyses and to reach objective and reasonable conclusions from the overall toxicological database.

- **Evidence of causality**: The submitted CLH-dossier raised a concern regarding developmental toxicity based on slightly increased post-implantation loss and AGD in an EOGRTS and a concern regarding reproductive toxicity based on slight effects on sperm morphology. However, none of these observations follow a dose response relationship.
- **Weight-of-evidence:** These examples demonstrate that a robust and transparent weight of evidence evaluation approach as required by the CLP Regulation has not been carried out.

Ongoing ECHA project to be taken into account in decision-making

The justification of the dossier submitter for a classification of Propyl 4-hydroxybenzoate as Repr. 2 (H361fd) is based, among others, on the assessment of an OECD TG 443. This Test Guideline is designed to provide an evaluation of reproductive and developmental effects that may occur as a result of pre- and postnatal chemical exposure as well as an evaluation of systemic toxicity in pregnant and lactating females and young and adult offspring.[4]
In this context, we would like to highlight the ongoing ECHA project that examines OECD TG 443 studies.[5] Aim of the project is to obtain a picture that will further improve the understanding of this complex study and its interlocking parameters. However, to achieve this aim, an intensive exchange with all relevant stakeholders must take place within the framework of the project. Such an exchange takes time, of course, but should enable scientifically valid results which allow an improved evaluation of OECD TG 443 studies related for CLH dossiers.

Conclusion

The more challenging a toxicological data set is, the more challenging the assessment of these data can be. Therefore, when evaluating (complex) studies such as the OECD TG 443 it is necessary to following fundamental toxicological aspects. As more and more knowledge is gained, it also happens that this knowledge is expanded by new findings.
Therefore, it is necessary that a regulatory assessment of OECD TG 443 studies and resulting CLH dossiers takes the following points into account:

- Adherence of scientific principles for toxicological evaluation considering all relevant data (statistically significance and biological relevance of effects, use of historical data, consideration of dose dependency).

[4] OECD, Test No. 443: Extended One-Generation Reproductive Toxicity Study, 03.06.2022 (08:30), https://www.oecd-ilibrary.org/environment/test-no-443-extended-one-generation-reproductive-toxicity-study_9789264185371-en

[5] ECHA Weekly, 25. November 2020, https://echa.europa.eu/de/view-article/-/journal_content/title/echa-weekly-25-november-20-2

- Functional parameters related to observations for isolated values within the OECD TG 443 should be considered in order to evaluate the adversity of effects, evidence of adversity, evidence of causality.
- Isolated evaluation of single biological parameters without consideration of the whole database and other toxicological OECD-conform studies is not scientifically justified and should be performed in a Weight of Evidence approach.
- Relevance of toxicological observation for classification and labelling.
- Results of the ongoing ECHA project assessing OECD TG 443 studies should be taken into account.

Contact: Dr. [redacted]
Referent, Dept. Science, Technical and Environmental Affairs
Product Stewardship
P +[redacted] | E [redacted]@vci.de

German Chemical Industry Association
Mainzer Landstrasse 55
60329 Frankfurt, Germany

www.vci.de | www.ihre-chemie.de | www.chemiehoch3.de
LinkedIn | Twitter | YouTube | Facebook
Data protection rules | Compliance-Guideline | Transparenz

- Identification no. in the EU Transparency Register: 15423437054-40
- The VCI is registered in the "public list on the registration of associations and their representatives" of German Parliament (Deutscher Bundestag).

The VCI represents the interests of around 1,900 companies from the chemical-pharmaceutical industry and related sectors vis-à-vis politicians, public authorities, other industries, science and media. In 2021, the VCI member companies realized sales of ca. 220 billion euros and employed over 530,000 staff.

Annex

CLP Regulation - Harmonised classification and labelling

CLP Annex I - Classification and labelling requirements for hazardous substances and mixtures

3.7.2. Classification criteria for substances

3.7.2.1. Hazard categories

3.7.2.1.1.

For the purpose of classification for reproductive toxicity, substances are allocated to one of two categories. Within each category, effects on sexual function and fertility, and on development, are considered separately. In addition, effects on lactation are allocated to a separate hazard category.

3.7.2.3. Weight of evidence

3.7.2.3.1.

Classification as a reproductive toxicant is made on the basis of an assessment of the total weight of evidence, see section 1.1.1. This means that all available information that bears on the determination of reproductive toxicity is considered together, such as epidemiological studies and case reports in humans and specific reproduction studies along with sub-chronic, chronic and special study results in animals that provide relevant information regarding toxicity to reproductive and related endocrine organs. Evaluation of substances chemically related to the substance under study may also be included, particularly when information on the substance is scarce. The weight given to the available evidence will be influenced by factors such as the quality of the studies, consistency of results, nature and severity of effects, the presence of maternal toxicity in experimental animal studies, level of statistical significance for inter-group differences, number of endpoints affected, relevance of route of administration to humans and freedom from bias. Both positive and negative results are assembled together into a weight of evidence determination. A single, positive study performed according to good scientific principles and with statistically or biologically significant positive results may justify classification (see also 3.7.2.2.3).

3.7.2.3.3.

If, in some reproductive toxicity studies in experimental animals the only effects recorded are considered to be of low or minimal toxicological significance, classification may not necessarily be the outcome. These effects include small changes in semen parameters or in the incidence of spontaneous defects in the foetus, small changes in the proportions of common foetal variants such as are observed in skeletal examinations, or in foetal weights, or small differences in postnatal developmental assessments.