UVCB substances: Methodology for substance identification, generation of constituents and selection of representative samples

Stela Kutsarova[a], Darina Yordanova[a], Yordan Karakolev[a] and Ovanes Mekenyan[a]

[a]Laboratory of Mathematical Chemistry, Prof. Asen Zlatarov University, Bourgas, Bulgaria

INTRODUCTION

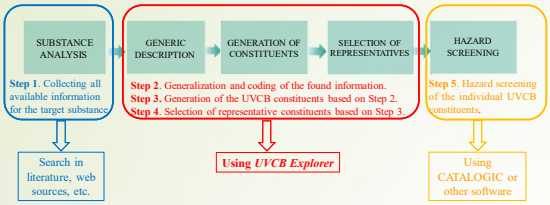
The complete characterization of UVCBs (substances of Unknown or Variable composition, Complex reaction products, or Biological materials) remains challenging due to their complex and variable composition, often involving a large number of constituents. This complexity complicates their (eco)toxicological assessment and regulatory compliance. To address this, a methodology for UVCB characterization based on common substance identifiers, allows coding, generation, storage and hazard assessment was developed by LMC in cooperation with the ECHA [1]. The methodology is implemented in the *UVCB Explorer* software. It was further updated by allowing endpoint (non-)specific representatives to be selected for testing purposes [2-4]. Real case study is used to exemplified the UVCB characterization process using the software.

MATERIALS AND METHODS

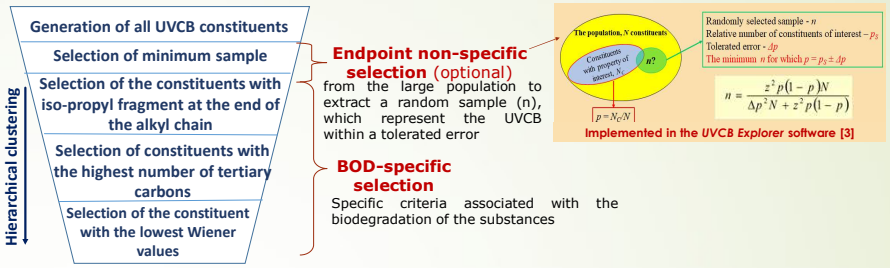
Case study: Selection of UVCB representatives for BOD modelling

❖ **Case study**: Selection of the least degradable constituents of *Diisotridecyl adipate (DITA), CAS 26401-35-4*

Methodology for UVCB characterisation

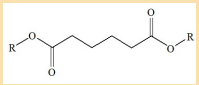


General scheme for selection of representatives



General information on the target substance

Generic description of DITA

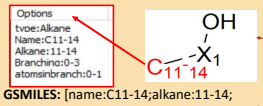


Where:
R = C11-C14 alkyl chain with up to three methyl branches. One or two ethyl branches are allowed only for C13

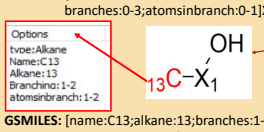
Produced as a result of reaction between adipic acid and C11-14 branched alcohols

Building G GRAPH (Generic Graph): *UVCB explorer* software implementation

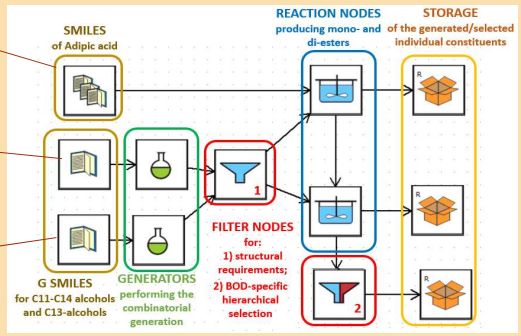
SMILES: OC(=O)CCCCC(O)=O



GSMILES: [name:C11-14;alkane:11-14; branches:0-3;atomsinbranch:0-1]X1O

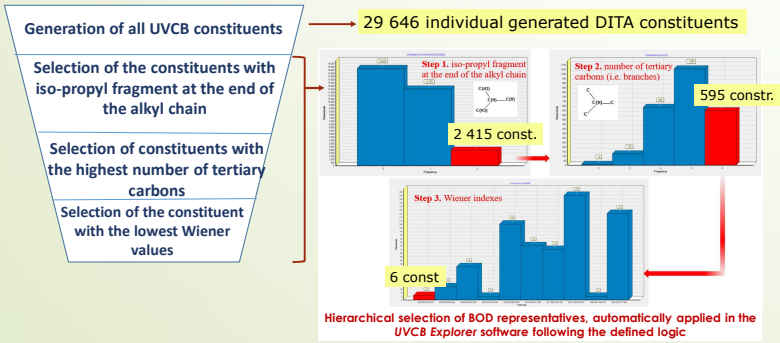


GSMILES: [name:C13;alkane:13;branches:1-2; atomsinbranch:1-2]X1O



RESULTS AND DISCUSSION

❖ Selected representatives



Six out of almost 30 000 structures are selected as representatives for worst-case scenario biodegradation.

❖ Validation of the selection

- To validate the adequacy of the selection, the ranges of variation of parameters associated with the branching and BOD ranges for the large number of initially generated DITA constituents and those of the final selection are compared

Property	Ranges of property variation for the whole UVCB /29646 constituents/	Ranges of property variation for the selected sample /6 constituents/
Number of iso-propyl fragments	0 to 2	2
Number of branches (i.e. number of tertiary carbons)	0 to 6	6
Wiener indexes	4310 to 8509	4364 to 4705
BOD, %*	21 – 92	25 to 29

* BOD values are estimated by the CATALOGIC 301C model. Due to the unmanageable number of constituents, BOD ranges representing the whole UVCB are defined based on statistically selected sample of 382 constituents.

- Based on the comparative analysis, it could be claimed that the selected six chemicals could be taken as representatives for DITA with respect to the most conservative biodegradation scenario, namely not degradable

CONCLUSIONS

- A methodology for coding and generation of UVCB constituents and selection of representative samples has been developed and implemented in the *UVCB Explorer* software.
- The methodology is demonstrated through a case study on a UVCB, named diisopropyl adipates, focusing on the selection of the most non-degradable constituents.
- All possible individual constituents were generated and a few representatives were selected by automatically applied criteria associated with biodegradability.
- The adequacy of the selection of the final sets with representative constituents was validated by comparing the predicted properties of the sample set to the entire target UVCB set.
- The results confirm that the selected representatives are appropriate and can support prioritization for further analysis or experimental validation.
- Similar approaches could be developed and applied to human health endpoints.

References

[1] ECHA. 2013. ECHA/2011/172 Characterization, chemical representation and modelling of UVCB substances
[2] UVCB I : Dimitrov et al, UVCB substances: Methodology for structural description and application to fate and hazard assessment, Environmental Toxicology and Chemistry, 2015, November, 3 (11): 2450-62
[3] UVCB II: Kutsarova et al., UVCB substances II: Development of an endpoint-nonspecific procedure for selection of computationally generated representative constituents. Environmental Toxicology and Chemistry, 38(3),2019, 682–694.
[4] UVCB III: Yordanova et al., Selection of Representative Constituents for Unknown, Variable, Complex, or Biological Origin Substance Assessment Based on Hierarchical Clustering, Environmental Toxicology and Chemistry, 40(11), 2021,3205-3218