Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2017-574-AC2, 2017

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Evaluating the mutagenic potential of aerosol organic compounds using informatics based screening" *by* Stefano Decesari et al.

Stefano Decesari et al.

s.decesari@isac.cnr.it

Received and published: 3 October 2017

REPLY TO REFEREE'S 2 COMMENTS

We thank the Referee for his/her useful comments which give us the opportunity to clarify some aspects of the methodology.

COMMENT: Have the compounds which were tested previously been measured in atmospheric SOA, and if they have at what concentrations and which types of SOA (e.g. anthropogenic, biogenic, etc)? If they haven't been specifically measured can the authors give a rough estimation of the concentrations that could be expected in

Printer-friendly version

Discussion paper

different types of SOA? Is there a trend of more of the toxic compounds being likely to be present in a specific type of SOA?

REPLY: We have screened the existing literature and compiled a new table (Table S10, in attach to the present document, and to be included in the Supplementary Information of the revised paper) with concentration data observed for a subset of SOA compounds from the most relevant clusters of mutagenic species in our list. When direct observations are not available, we provided a rough estimate of ambient concentrations on the basis of observed/predicted yields in laboratory setups and on source apportionment results for ambient organic aerosols. We found that the new SOA compounds predicted to be mutagenic can occur in ambient air in concentrations of 10E-2 to 10E1 ng m-3. In respect to the Referee's question about possible trends in the abundance of toxic compounds between different classes of SOA, we think that Figure 2 of the present version of the manuscript already addresses this. The figure shows that a greater fraction of species predicted to be mutagenic is found in anthropogenic SOA systems than in the natural ones. As observed by Referee 1, this finding - though limited to the pool of compounds considered in our study - is qualitatively in line with the current evidence of clearer toxicological and epidemiological effects of anthropogenic combustion-related aerosol with respect to biogenic particles. We will add a note about this in the Discussion section of the revised version of the manuscript.

COMMENT: Are the concentrations which would be expected in the respiratory tract upon inhalation of atmospheric SOA similar to the concentrations used during the Ames test and assumed in the models? When both positive and negative Ames tests results are reported due to testing conditions (Page 7, Line 2) are there any trends that have been observed for the range of results (e.g. a dependence on concentration, or a different experimental condition) and could some of the experimental results be discarded due to those conditions being unlikely to occur in the respiratory tract/ human body after inhalation?

REPLY: We are happy to re-iterate and further clarify the role of presented data in our

study using three key points related to the questions posed.

1) First of all, it is worth re-iterating that mutagenicity is often related to long-term, chronic exposures in environmental toxicology. Even trace air concentrations of mutagens cumulate to mg-level doses during a lifetime exposure. That is why the WHO air quality guidelines recommend sub-ng/m3 threshold values for the concentrations of benzo[a]pyrene (WHO, 2000). The concentrations of the new SOA compounds predicted to be mutagenic in the Ames test on the basis of our QSAR approach are not always known. The ranges of concentrations of the compounds predicted to be mutagenic in our study (new Table S10, in attach to the present document) compare well with those observed for known atmospheric organic mutagens such as PAHs and nitro-PAHs (Alves et al., 2017). In conclusion, the newly identified mutagenic compounds can occur in ambient air in appreciable concentrations and be inhaled in mg amounts during a lifetime exposure. We therefore recommend them for in vitro toxicological screening for confirmation of their mutagenic effect and for determination of dose-response functions.

2) As far as the comparison with the doses used in the Ames test is concerned, we would like to emphasize that the Ames test (or bacterial reverse mutation test) is only a screening in vitro test performed to support a preliminary hazard assessment of chemicals, which are screened for their mutagenic potential. In more detail, the Ames test uses amino-acid requiring strains of Salmonella typhimurium and Escherichia coli to detect point mutations (OECD, 1997). Such mutations lead to revertant bacteria in which the functional capability to synthesize the essential amino acid is restored. Revertant bacteria are then detected by their ability to grow in absence of the amino acid necessary for the growth of the parent test strain. A wide range of concentrations are tested, with an upper limit which mainly depends on the solubility and cytotoxicity of the test compound in the final treatment mixture. The tested concentrations are expressed in either μg/plate or μL/plate, where “plates” are samples of bacterial strains, which makes a direct comparison with exposure concentrations for humans not properly fea-

sible. It is important to emphasize that screening tests, such as the Ames test, are supposed to be used in the preliminary phases of hazard assessment, and especially for hazard identification, and not to mimic exposure conditions. We will add this to the text of the new version of the manuscript. Finally, as for the appropriateness of conditions used in models to replicate conditions in the respiratory tract, this is not within scope of the paper. Indeed, as we iterate, the driver for this work is to demonstrate the efficacy of a methodology to highlight potential hazardous compounds that might be measured in future epidemiological studies.

3) The variation of "testing conditions" mentioned in the manuscript includes factors such as the testing methodology (e.g., plate incorporation method or preincubation method), the tested strains of Salmonella typhimurium and Escherichia coli (at least five different strains are required, as specified in the OECD guideline, in order to detect different mutations), presence or absence of metabolic activation, exposure concentration. No specific trends have been observed among positive and negative results. The Ames test is an in vitro test using prokaryotic cells and, as such, it does not consider important processes in organisms of higher clades (e.g., rodent in vivo toxicological studies). Therefore, there are no experimental conditions that can be considered representative for the mechanisms of uptake of the test substance by humans (e.g., absorption through the respiratory tract), and, as a corollary, no specific testing conditions used in the Ames test could be discarded.

COMMENT: Are there any limitations to the two models used (the ACD/Impurity Profiling model and the Vega/CAESER model), for example, can they reliably predict the toxicity of any compound? What are the differences between them which leads to one giving a positive result whilst the other gives a negative result in some cases (as shown in Table 2)?

REPLY: This is a good point and we are happy to clarify here. All QSAR models are characterized by a defined applicability domain which means that no QSAR model can be guaranteed to provide a reliable prediction for any compound, regardless of

the endpoint. In this specific case, both the ACD/Impurity Profiling QSAR model and Vega/CAESER model are supported with specific parameters providing information on prediction reliability, including applicability domain considerations. As described in the Supplementary Information (Table S4), ACD/Percepta predictions are supplemented with a Reliability Index (RI), which ranges from 0 to 1, and gives an evaluation of whether a submitted compound falls within the Model Applicability Domain. In particular: RI < 0.3 (Not Reliable), RI in range 0.3-0.5 (Borderline Reliability), RI in range 0.5-0.75 (Moderate Reliability), RI >= 0.75 (High Reliability). Estimation of the RI takes into account two main aspects: similarity of the tested compound to the training set and the consistency of experimental values for similar compounds. Similarly, Vega/CAESAR predictions are provided with an Applicability Domain Index (ADI), which ranges from 0 to 1, and gives an evaluation of whether a submitted compound falls within the Model Applicability Domain. In particular: ADI > 0.9 means that the predicted substance is into the AD of the model; ADI < 0.7 means that the predicted substance is out of the AD of the model; ADI in range 0.7-0.9 means that the predicted substance could be out of the AD of the model and further considerations are needed. The ADI is calculated by grouping several other indices, each one taking into account a particular issue, such as: training/test set similar molecules with known experimental value, concordance for similar molecules, accuracy of prediction for similar molecules, atom Centered Fragments similarity check and model descriptors range check.

Predictions obtained by the two QSARs were combined by taking into account the applicability domain of each model (prediction reliability) and consistency between predictions. In case of conflicting reliable predictions, the most conservative outcome, i.e. positive prediction, was assigned. We will add the following text to the revised version of the manuscript:

“Overall, the predictions generated by the two QSAR models were in agreement for the majority of compounds (not consistent predictions obtained for only 4 compounds out of the 53 reliably predicted by both QSAR models) and the combination of the

two tools resulted successful in covering a wider chemical space (59% of compounds reliably predicted by ACD/Percepta alone, 74% of compounds reliably predicted by Vega/CAESAR alone, 82% of compounds reliably predicted combining the two QSAR models). As far the 4 compounds with opposite predictions are concerned, this is mainly due to issues related to the different applicability domain of the two QSAR models."

Minor comments (1): There are a lot of abbreviations throughout the text which should be written out in full for clarity (e.g. PaDEL, nHBDon, nHBAcc, KNIME).

REPLY: These includes names of softwares (PaDEL-Descriptor, and KNIME) for which specific references are already reported in the manuscript. We will instead report explanations for the following acronyms: nHBDon (= number of donor atoms for hydrogen bonds) and nHBAcc (= number of acceptor atoms for hydrogen bonds).

Minor comments (2): There - Section 2 should be renamed from 'Introduction' to 'Methods'.

REPLY: Yes, we thank the Referee for noticing this mistake.

REFERENCES:

1) Alves et al., "Polycyclic aromatic hydrocarbons and their derivatives (nitro-PAHs, oxygenated PAHs, and azaarenes) in PM2.5 from Southern European cities", Science of the Total Environment 595 (2017) 494–504.

2) OECD GUIDELINE FOR TESTING OF CHEMICALS n. 471. Bacterial Reverse Mutation Test. Available at: http://www.oecd-ilibrary.org/environment/test-no-471-bacterial-reverse-mutation-test_9789264071247-en.

3) WHO (2000) Air Quality guidelines for Europe, 2nd edition: WHO Regional Office for Europe (WHO Regional Publications, European Series, No. 91).

Please also note the supplement to this comment:

https://www.atmos-chem-phys-discuss.net/acp-2017-574/acp-2017-574-AC2-supplement.pdf

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-574, 2017.

Printer-friendly version

Discussion paper