Quantum Chemistry EI Mass Spectrometry Database for Exposome Research

W. Rojas[a,*], A. Křenek[b], H. Hecht[a,*], E. J. Price[a,*]

[a] RECETOX Centre, Faculty of Science, Masaryk University, Kotlarska 2, Brno 60200, Czech Republic

[b] Institute of Computer Science, Masaryk University, Botanická 554/68a, Brno 60200, Czech Republic

[*] *wudmir.rojas@recetox.muni.cz, helge.hecht@recetox.muni.cz, elliott.price@recetox.muni.cz*

Aim

Evaluate the accuracy of in-silico generation of spectra for chemicals with experimental high-resolution electron ionization mass spectra (HR-[EI+]-MS) available in the RECETOX Exposome library[1] to assess potential of spectral predictions for chemicals with no reference spectra.

Method

Used the Quantum Chemical Electron Ionization Mass Spectrometry package (QCxMS v5.1.3)[2] (**Figure 1**) for spectra prediction at the semiempirical quantum mechanical and density functional theory (DFT) levels & evaluated accuracy against experimental spectra of coumarin & estragole.

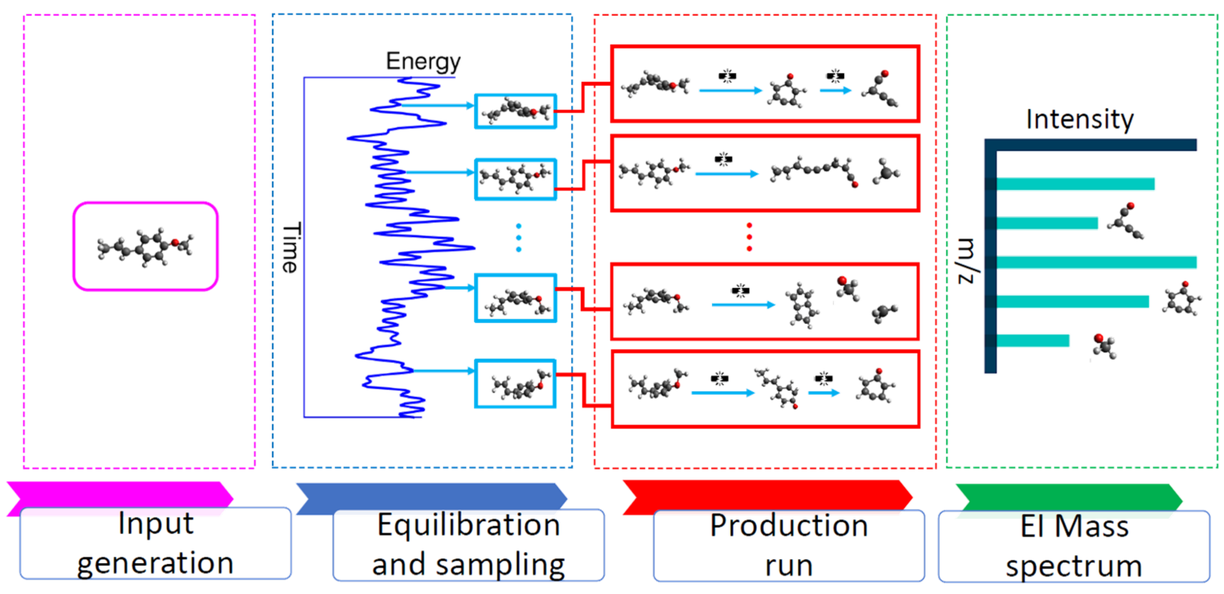
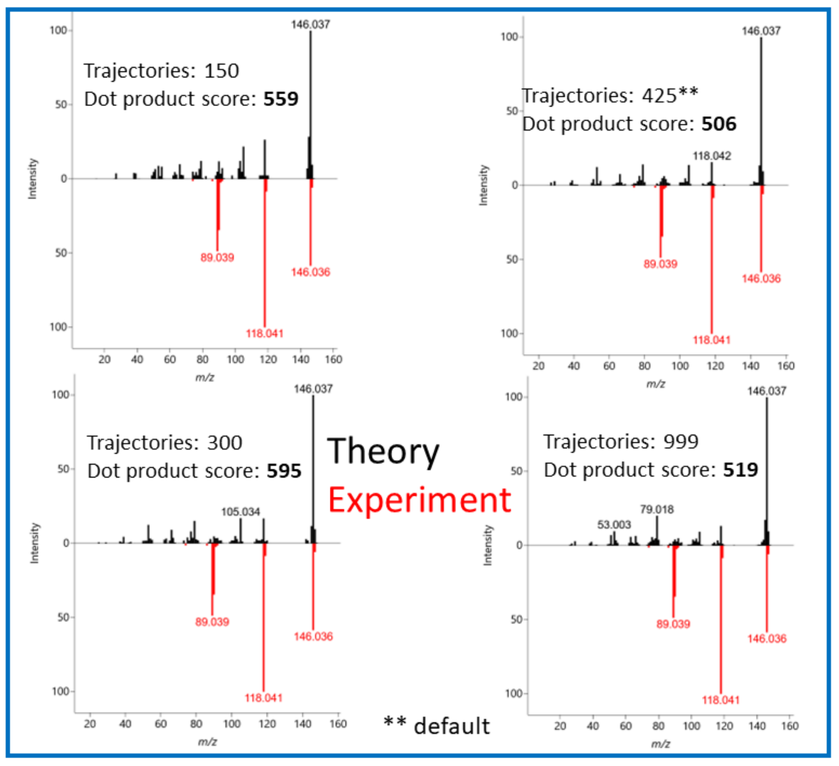


Figure 1. *Diagram of the QCxMS workflow & parameters applied.*

- Equilibration: GFN2-xTB[3] semiempirical method (extended tight binding).
- Production run: GFN2-xTB[3] or DFT PBE0-D3/SV(p)[4,5] level.
- DFT calculations with ORCA[6].
- Ground state sampling:
  - Temperature 500 K
  - MD trajectory 20 ps
  - Timestep 0.5 fs
- Production run:
  - Temperature 500 K
  - MD trajectory 20 ps
  - Timestep 0.5 fs
  - Impact excess energy 0.6 eV/atom
  - Electron ionization 70 eV

Results



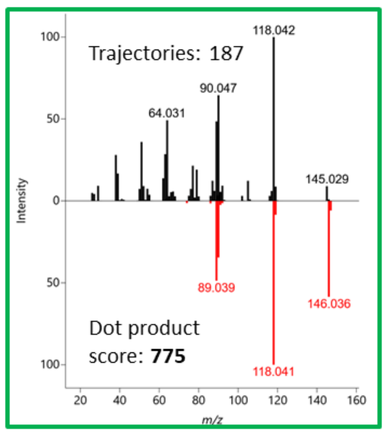
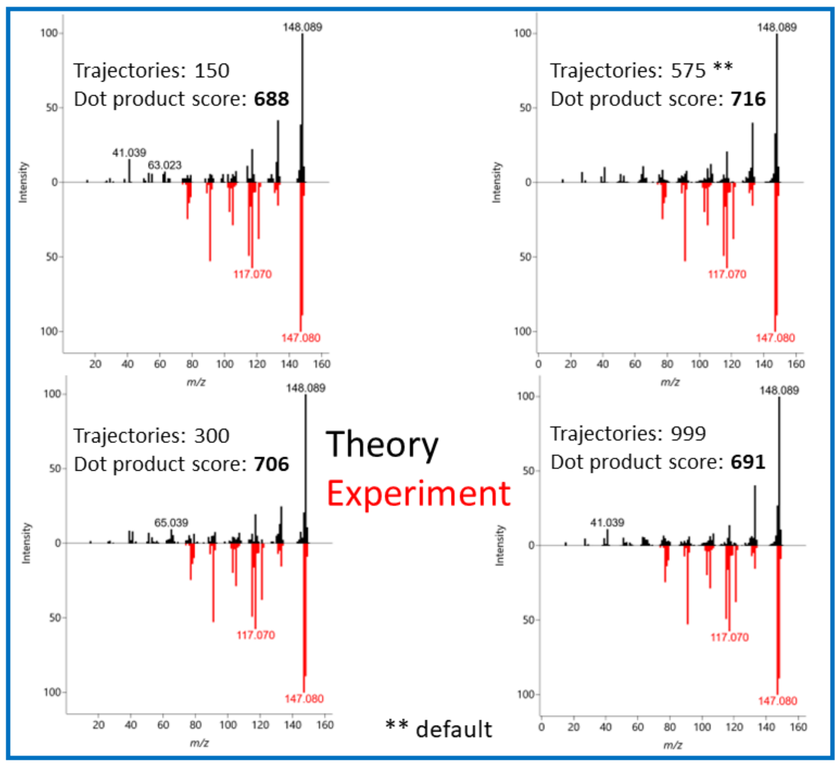


Figure 2. *Predicted spectra (black) compared to experimental spectra (red) for coumarin via semiempirical (blue box) and DFT (green box).*



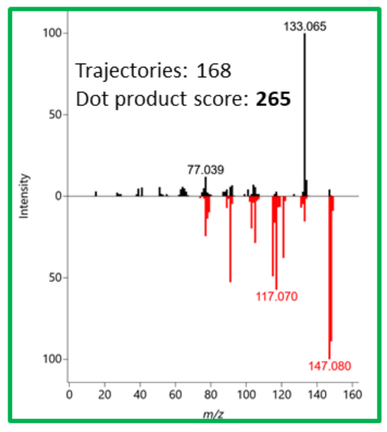


Figure 3. *Predicted spectra (black) compared to experimental spectra (red) for estragole via semiempirical (blue box) and DFT (green box).*

Discussion & Conclusions

- Effect of number of trajectories on spectra evaluated only using the semiempirical method.
- Default number of trajectories (25x#of atoms) gives acceptable score.
- Number of trajectories has low impact on spectra for both molecules.
- Spectra predicted at semiempirical level are sufficient for identification if reverse score match > 500.
- Computation times using semiempirical method are dramatically lower than DFT (using one core -> mins/hours vs weeks, respectively).
- DFT gives highest match score for coumarin (**Figure 2**) but lowest for estragole (**Figure 3**). Low number of trajectories reached for estragole likely reason of low match score.
- Semiempirical method more accurately predicts the highest m/z peak than DFT for these two molecules.
- Generation of in-silico spectral library by semiempirical methods is potentially feasible.

Future work

- Assessment of semiempirical method's performance of molecules classified by common features.

Acknowledgements

This work was supported by RECETOX research infrastructures and financed by Czech Ministry of Education, Youth and Sports (LM2018121) and funding from the CETOCOEN EXCELLENCE Teaming two project supported by Horizon 2020 (857560) and the Ministry of Education, Youth and Sports of the Czech Republic (02.1.01/0.0/0.0/18_046/0015975). E.J.P. acknowledges support from the Czech Operational Programme Research, Development and Education–Project MSCAfellow4@MUNI (No. CZ.02.2.69/0.0/0.0/20_079/0017045). Computational resources were provided by the project e-INFRA LM2018140.

References

1. E. J. Price, et al. High-Resolution EI+Spectral Library of Anthropogenic Compounds. Front Public Health 2021, 9, 622558. doi:10.3389/fpubh.2021.622558; **2.** J. Koopman, S. Grimme J. Am. Soc. Mass Spectrom. 2021, 32, 7, 1735–1751. doi:10.1021/jasms.1c00098; **3.** C. Bannwarth, S. Ehlert, S. Grimme, J. Chem. Theory Comput. 2019, 15, 1652–1671. doi:10.1021/acs.jctc.8b01176; **4.** G. E. Ernzerhof, et al. J. Chem. Phys.1999,110, 5029. doi:10.1063/1.478401; **5.** S. Grimme, et al. J. Comput Chem, (2011), 32, 1456–1465. doi:10.1002/jcc.21759; **6.** F. Neese, The ORCA program system. Wiley Interdiscipl. Rev. Comput. Mol. Sci. 2012, 2, 73–78. doi:10.1002/wcms.81