FTR: Fragment tree reconciliation - Supplements

Nico Domschke[1†], Chen Wang[2†], Richard Golnik[1], Ales Charvat[1], Anatolii Spesyvyi[2,3], Thomas Gatter[1*], Bernd Abel[2,3], Peter F. Stadler[1,4-9*]

[1]Bioinformatics Group, Department of Computer Science & Interdisciplinary Center for Bioinformatics & School for Embedded and Composite Artificial Intelligence (SECAI), Leipzig University, Härtelstraße 16–18, D-04107 Leipzig, Germany.
[2]J. Heyrovský Institute of Physical Chemistry , Czech Academy of Sciences, Dolejškova 2155/3, CZ-18223 Praha, Czech Republic.
[3]Institute of Chemical Technology, Leipzig University, Linnestrasse 3, D-04103 Leipzig, Germany.
[4]Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI), Leipzig University, D-04103 Leipzig, Germany.
[5]Max Planck Institute for Mathematics in the Sciences, Inselstraße 22, D-04103 Leipzig, Germany.
[6]Department of Theoretical Chemistry, University of Vienna, Währingerstraße 17, A-1090 Wien, Austria.
[7]Facultad de Ciencias, Universidad National de Colombia, Bogotá, Colombia.
[8]Center for non-coding RNA in Technology and Health, University of Copenhagen, Ridebanevej 9, DK-1870 Frederiksberg, Denmark.
[9]Santa Fe Institute, 1399 Hyde Park Rd., Santa Fe, NM 87501, USA.

*Corresponding author(s). E-mail(s): thomas@bioinf.uni-leipzig.de; studla@bioinf.uni-leipzig.de;
Contributing authors: dnico@bioinf.uni-leipzig.de; chen.wang@uni-leipzig.de; richard@bioinf.uni-leipzig.de; ales.charvat@uni-leipzig.de; anatolii.spesyvyi@jh-inst.cas.cz; bernd.abel@uni-leipzig.de;
†These authors contributed equally to this work.

1 Energy-dependent fragments

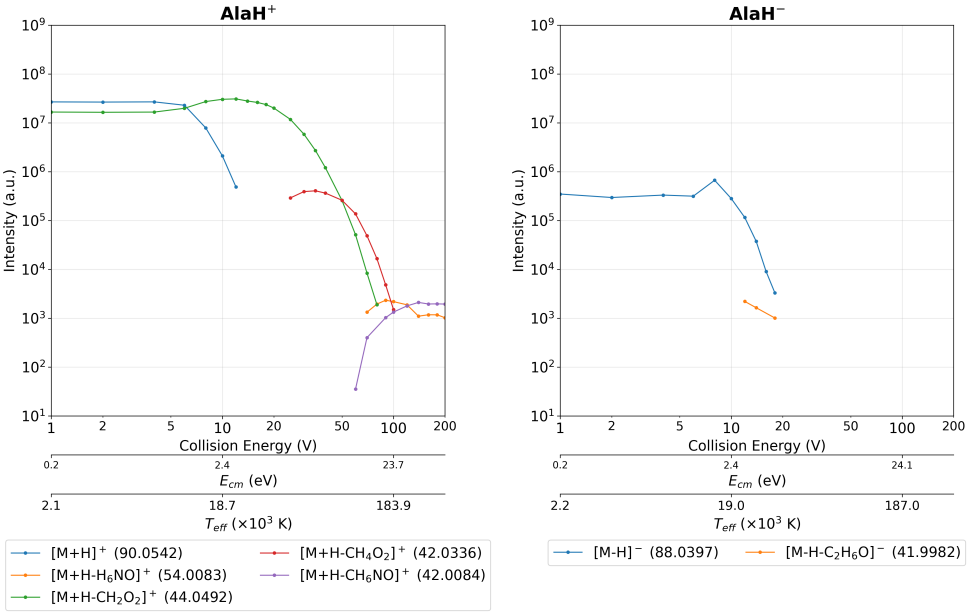

Fig. 1: Fragment emergence sampled over a collision energy range of $1 - 200V$ in both positive (left) and negative (right) modes for alanine.

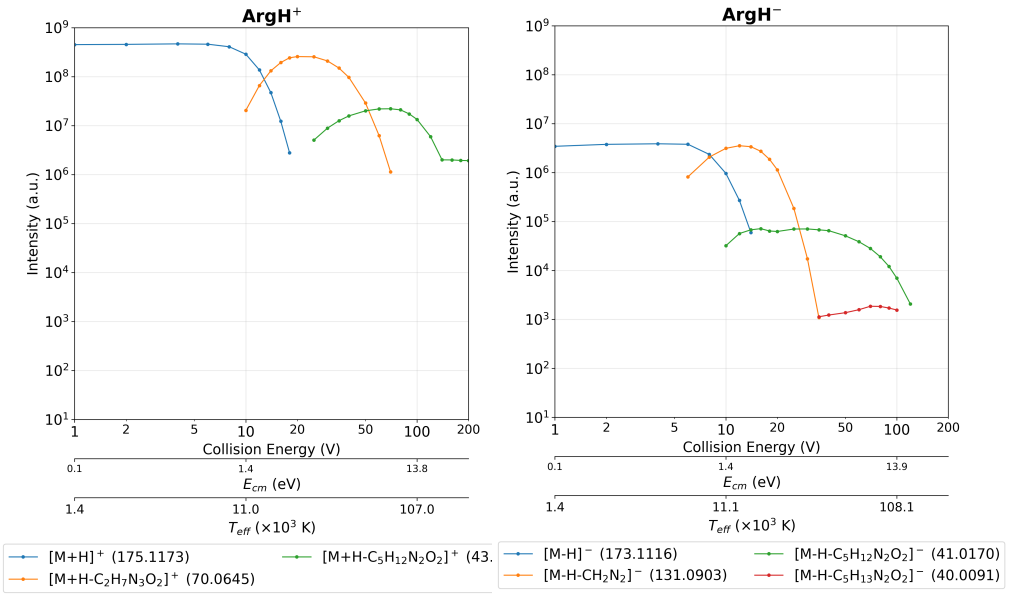

Fig. 2: Fragment emergence sampled over a collision energy range of $1 - 200V$ in both positive (left) and negative (right) modes for arginine.

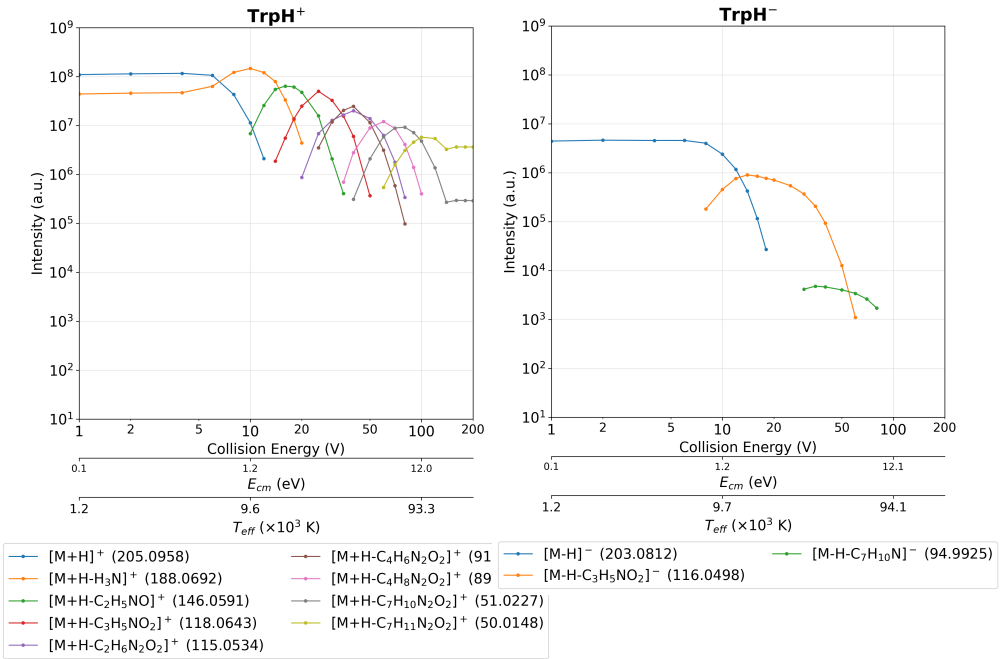

Fig. 3: Fragment emergence sampled over a collision energy range of $1 - 200V$ in both positive (left) and negative (right) modes for tryptophan.

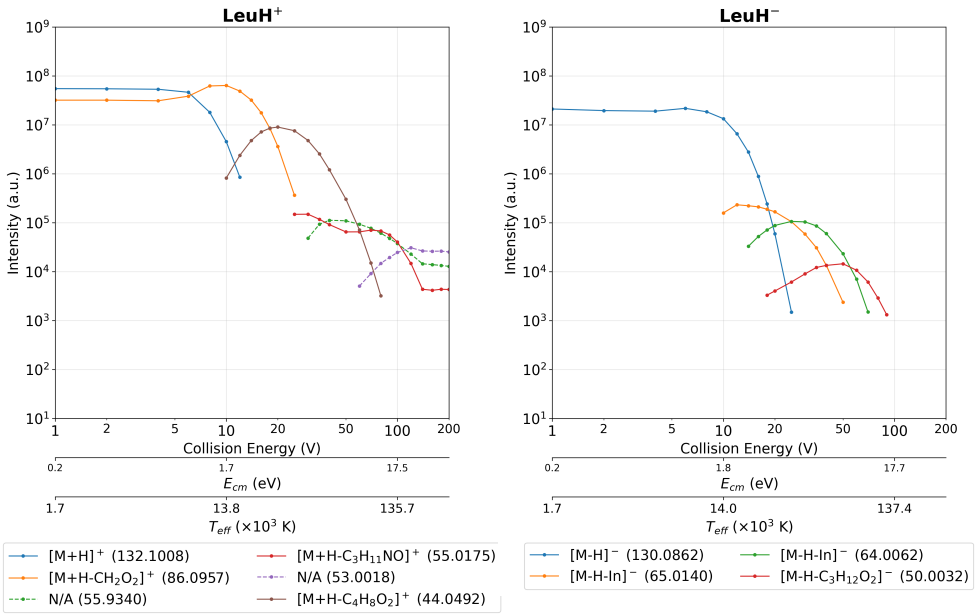

Fig. 4: Fragment emergence sampled over a collision energy range of $1 - 200V$ in both positive (left) and negative (right) modes for leucine.

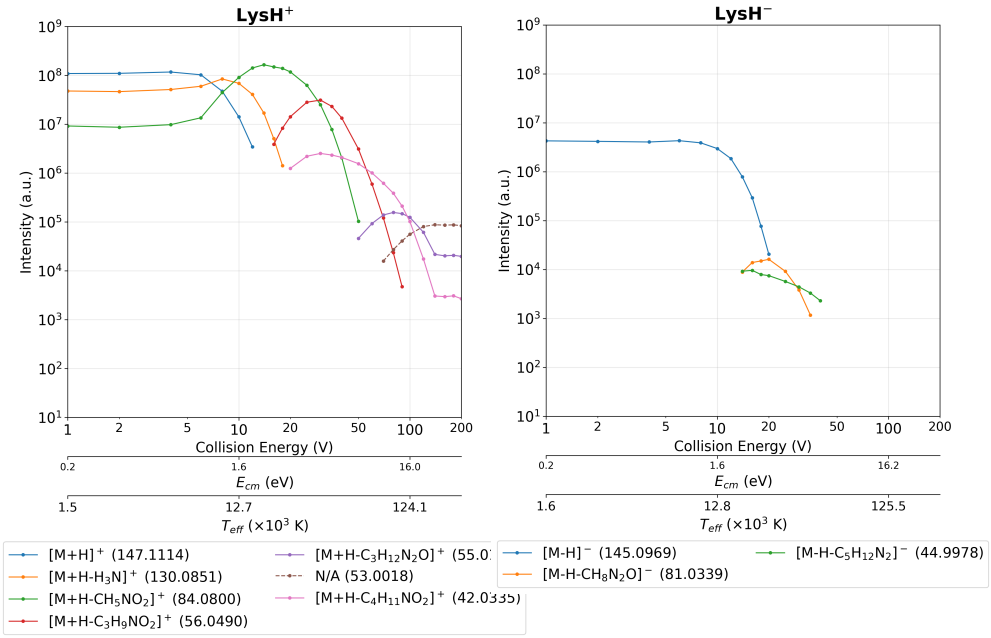

Fig. 5: Fragment emergence sampled over a collision energy range of $1 - 200V$ in both positive (left) and negative (right) modes for lysine.

2 Additional benchmarking figures for single first fragmentation events

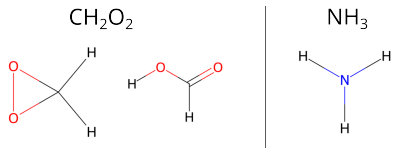

Fig. 6: Structural isomers of the neutral loss of leucine and alanine (CH_2O_2, left) and of arginine (NH_3, right). Due to the small difference in mass and thus a limited elemental composition, the losses only show very low variety. In order to save computation overhead for such small neutral losses, a dictionary containing commonly found neutral losses was was used.

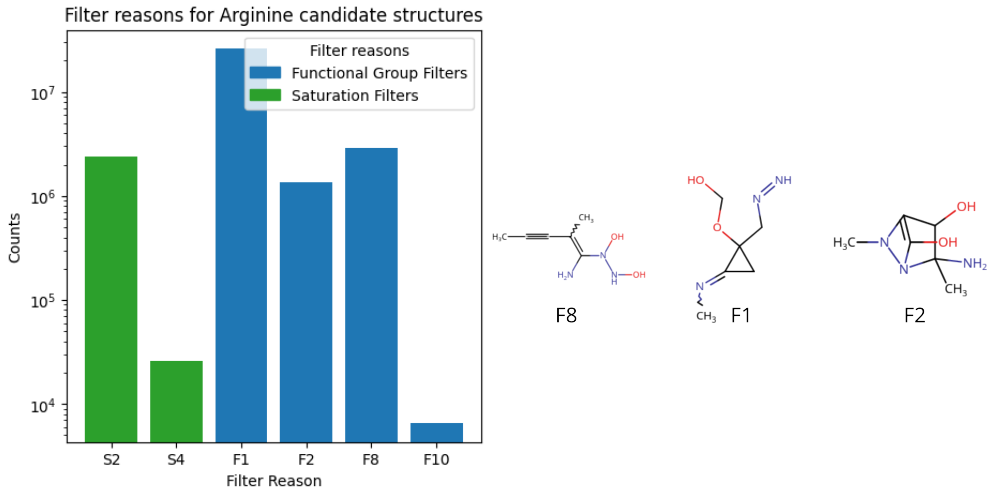

Fig. 7: **Left**: Reasons for discarding potential structure candidates for the primary descendant of arginine, i.e. $C_6H_{11}N_3O_2$, indicated by **s** for *unsaturation*, i.e. to filter for mostly unsaturated hydrocarbons, and **f** for any unstable, reactive, or synthetically infeasible functional group. In particular, **S2**: no unsaturations in 3-membered rings, **S4**: triple bond restrictions, **f1**: only one N or O next to an sp^3 carbon or two oxygens if both oxygens are ring atoms, **f2**: maximal one N or O in small rings, **f8** (enol/enamine): removal of molecule if O or N atoms are adjacent to a non-aromatic C=C, **f10**: carbonic acids ($O-CO_2H$), carbamic acids ($N-CO_2H$), and β-carboxylic acid ($(C=O)-C-CO_2H$) are removed. For a complete overview see table 2 and table 3 full of [1]. **Right**: Example structures filtered by the GDB-17 set of SMART filters.

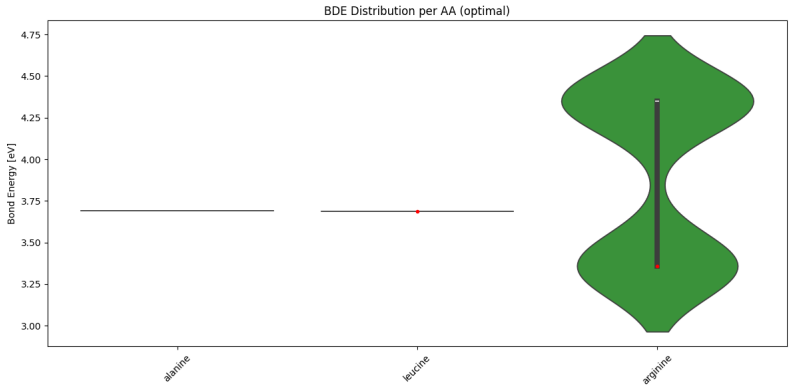

Fig. 8: Distribution of BDE values for all candidates of the first fragmentation reaction with optimal MCES score for alanine, leucine, and arginine. A red dot marks the reference reaction center size calculated for the literature proposed fragmentation. Alanine and leucin show a single energy value, as their structure only allows for one bond dissociation in their respectives parent molecule. In the case of arginine, the lower BDE corresponds to the literature fragmentation.[2]

3 Annotated negative mode fragmentation trees

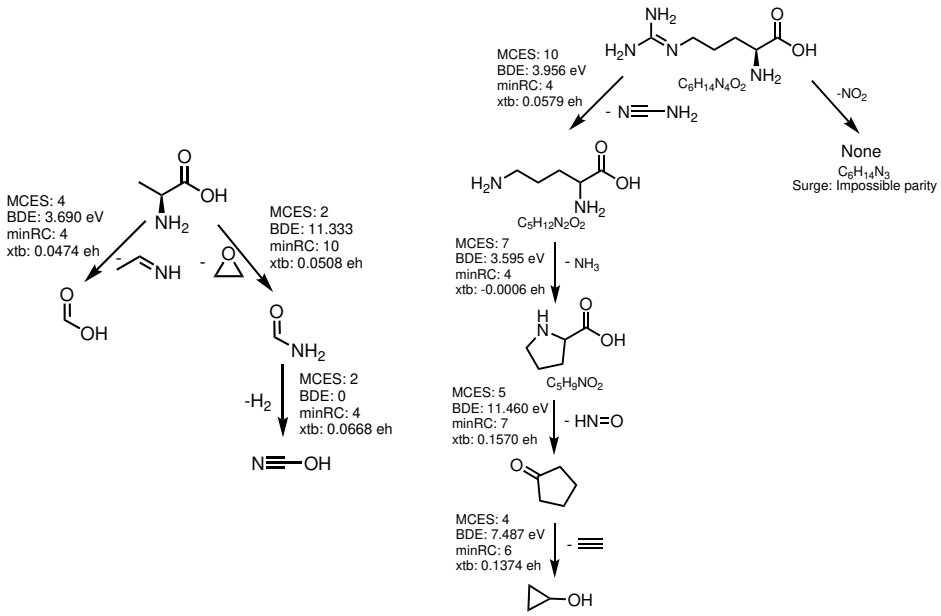

Fig. 9: Annoted fragmentation tree of alanine and arginine measured in negative ion mode.

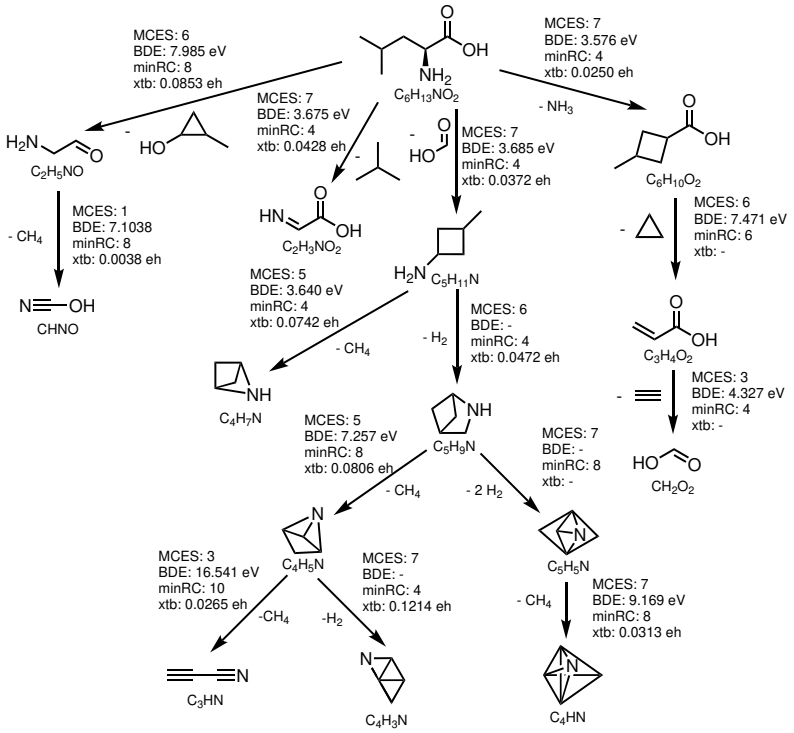

Fig. 10: Annoted fragmentation tree of leucine measured in negative ion mode.

References

[1] Ruddigkeit L, van Deursen R, Blum LC, Reymond JL (2012) Enumeration of 166 Billion Organic Small Molecules in the Chemical Universe Database GDB-17. Journal of Chemical Information and Modeling 52(11):2864–2875, DOI 10.1021/ci300415d

[2] Zhang P, Chan W, Ang IL, Wei R, Lam MMT, Lei KMK, Poon TCW (2019) Revisiting Fragmentation Reactions of Protonated α-Amino Acids by High-Resolution Electrospray Ionization Tandem Mass Spectrometry with Collision-Induced Dissociation. Scientific Reports 9(1):6453, DOI 10.1038/s41598-019-42777-8