MOLGEN-MS: Evaluation of Low Resolution Electron Impact Mass Spectra with MS Classification and Exhaustive Structure Generation

*[1]A. Kerber, [1]R. Laue, [1]M. Meringer and [2]K. Varmuza**

[1] University of Bayreuth, Mathematical Institute, Bayreuth, Germany
[2] Vienna University of Technology, Laboratory for Chemometrics, Institute of Food Chemistry, Vienna, Austria

Introduction

Most of the computer programs used for automatic spectra interpretation depend on large spectral databases.[1] The experimental spectrum in question is compared with the entries of the database and the structures of the most similar spectra are given as possible solutions. This method has several limitations: (a) The quality of the database spectra restricts the performance; even good experimental spectra may lead to wrong results if the reference spectra are erroneous. (b) If the spectrum of a query substance is not included in the library a reasonable result can only be expected if structures are in the database that are very similar to the unknown and are found be the applied spectral similarity criterion.

Method and Experimental

Aim of the new program system MOLGEN-MS[2] is to support systematic structure elucidation of organic compounds, mainly based on low resolution electron impact mass spectra and automatic isomer generation.[3] MOLGEN-MS includes the following components:

(1) *MSclass* (mass spectra classification) predicts the presence or absence of substructures by multivariate classification methods. In the current version there exist 160 classifiers for 85 substructures or more general structural properties.[4]

(2) *ElCoCo* (elemental composition computation) computes candidates for the molecular formula by the recognition of isotope peak patterns in low resolution mass spectra.[5]

(3) *MOLGEN* (molecular structure generator)[6] is able to compute in an efficient and redundancy free way all the connectivity isomers that correspond to the molecular formulae proposed by *ElCoCo* and which fulfill the substructure restrictions obtained from *MSclass* or other restriction defined by the user.

(4) *ReNeGe* (reaction network generator) simulates MS fragmentations for the structure candidates. A ranking of the candidate structures is based on how good the experimental spectrum can be fit by the theoretical isotope peak patterns of the virtual fragments.

While the generation of chemical structures from a given brutto formula and given structural restrictions is exact and exhaustive, the recognition of substructures from a spectrum is still the crucial part limiting the performance of systematic structure elucidation. The available classifiers typically require a user interaction for a critical check of classification results and adding additional structural information which is often available from other spectroscopic data or the history of a sample. This user interaction is fully supported by the software.

For a demo example the mass spectrum from *n-pentanoic acid, methylester* ($C_6H_{12}O_2$, molecular mass 116) was used. MS classification resulted in five "yes" answers with three of them predicting the substructure methylester, and 63 "no" answers predicting the absence of structural properties. From a mathematical analysis of the isotope peak patterns 14 brutto formulas with molecular masses 101, 102, and 116 were proposed. From these brutto formulas and the structural restriction obtained a total of 176 molecular structures were generated. The ranking of the candidates placed the correct structure in the first position.

Acknowledgements

M.M. thanks for financial support by the BMBF, contract 03-KE7BA 1-4. The valuable contributions from W. Werther and F. Stancl in the development of mass spectral classifiers, and from T. Grüner for the isomer generator are acknowledged.

References

1. K. Varmuza and W. Werther, *in* Advances in mass spectrometry (E. J. Karjalainen, A. E. Hesso, J. E. Jalonen, and U. P. Karjalainen, eds.), Vol. 14, p. 611. Elsevier, Amsterdam, 1998.
2. A. Kerber, R. Laue, and M. Meringer, *MOLGEN and MOLGEN-MS. Information and free demo versions: WWW.MOLGEN.DE* (2000).
3. T. Grüner, A. Kerber, R. Laue, M. Meringer, K. Varmuza, and W. Werther, *Match* **38**, 173 (1998).
4. K. Varmuza and W. Werther, *J. Chem. Inf. Comput. Sci.* **36**, 323 (1996).
5. T. Grüner, A. Kerber, R. Laue, M. Liepelt, M. Meringer, K. Varmuza, and W. Werther, *Match* **37**, 163 (1998).
6. T. Grüner, A. Kerber, R. Laue, and M. Meringer, *Match* **37** (1998).