

Software User Manual

MOLGEN-MS/MS* Version 1.0.0.3

1. Introduction

The computer program *MOLGEN-MS/MS* generates candidate molecular formulas from high-resolution MS data. It calculates match values (MV) that show how well candidate molecular formulas fit the MS isotope peak distributions (MS MV) and the high-resolution MS/MS fragment peak masses (MS/MS MV). Finally it computes a combined match value from these two scores. This software can be regarded as a further development of the *EICoCo* and *MolForm* modules of *MOLGEN-MS*¹ with a clear specialization towards MS/MS.

2. Installation

To install, just copy the executable *MolgenMsMs.exe* to a destination directory of your choice, e.g. `C:\Programs\MolgenMsMs`

Now your installation is already completed. The next logical step is to enter your license key. This is described in the Section 3. However, if you have not yet purchased a full license, or if you are not familiar with running Windows command line applications, it is useful to skip the next section and continue reading with Section 4.

3. License

MOLGEN-MS/MS has two operation modes. As long as you do not have a full license, *MOLGEN-MS/MS* will run in evaluation mode. This means that only problems up to a maximum weight of about 400 u will be processed. A full license can be obtained from www.molgen.de. In order to compute your license key, your Windows Product ID is required. *MOLGEN-MS/MS* will display this number if you call the program with option `license`. Your screen might look like this:

```
C:\PROGRAMS\MolgenMsMs>MolgenMsMs license
Specify license key for Windows Product ID
50026270432972295102
```

When purchasing a full license you receive a license key that matches your Windows Product ID. Please enter this license key by calling *MOLGEN-MS/MS* with option `license=<license-key>`. Your screen will look like this:

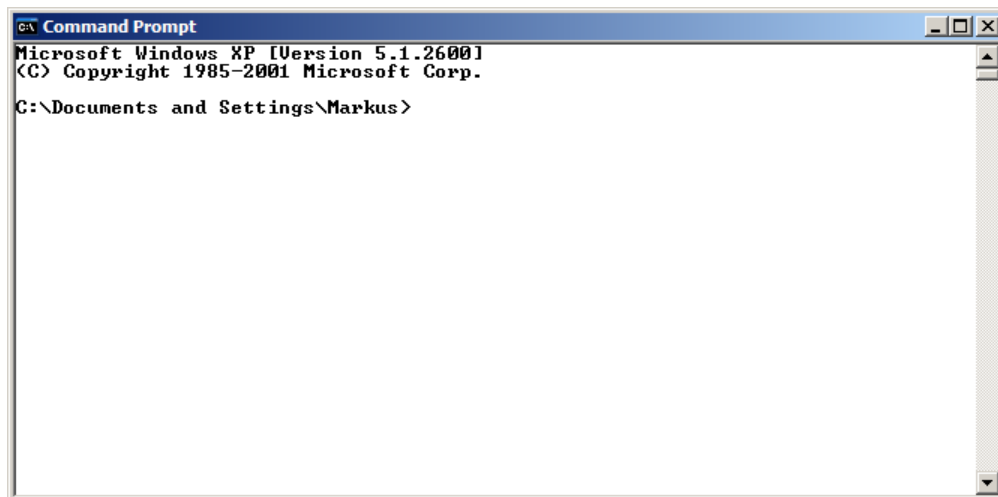
```
C:\PROGRAMS\MolgenMsMs>MolgenMsMs license=1993722487
Checking license key...
Valid license key has been saved
```

Now your *MOLGEN-MS/MS* is fully operational.

* Author: Markus Meringer, Munich, March 2009, revised: Christoph Rucker, Freiburg, May 2009, last updated: Markus Meringer, Hof, July 2009

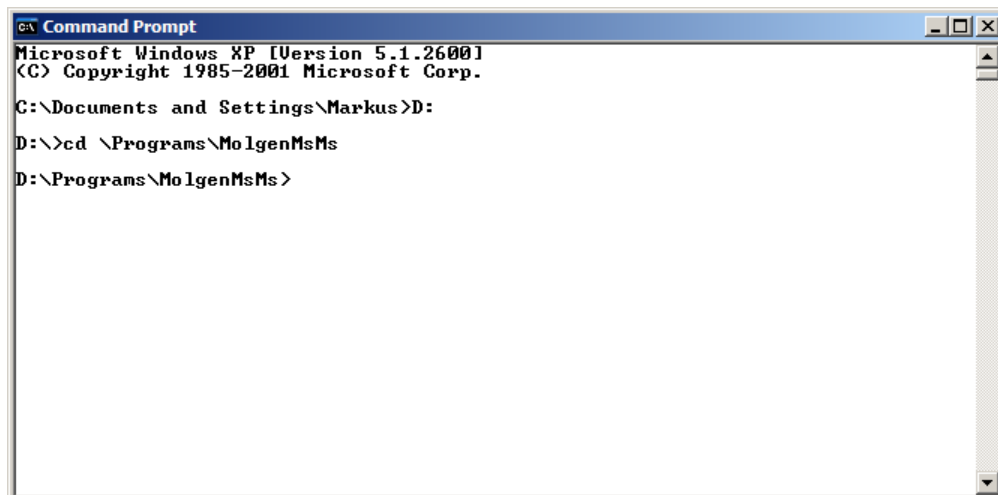
4. Getting Started

Command line applications are executed using the Windows *Command Prompt*. You can access *Command Prompt* via *Start* | *Run...* and typing `cmd.exe`. A window similar to the following will appear.



```
GA Command Prompt
Microsoft Windows XP [Version 5.1.2600]
(C) Copyright 1985-2001 Microsoft Corp.
C:\Documents and Settings\Markus>
```

To run the program, change to the drive and directory where the program is installed, using the `cd` command. In the situation shown below the program was installed on drive `D:` and directory `\Programs\MolgenMsMs`. Note that you can change the drive just by typing the drive letter followed by a colon.



```
GA Command Prompt
Microsoft Windows XP [Version 5.1.2600]
(C) Copyright 1985-2001 Microsoft Corp.
C:\Documents and Settings\Markus>D:
D:\>cd \Programs\MolgenMsMs
D:\Programs\MolgenMsMs>
```

Now start the program by typing its name *MolgenMsMs*. If your installation was successful and you have a valid license for *MOLGEN-MS/MS*, your screen should look like this:

```

C:\ Command Prompt
D:\Programs\MolgenMsMs>MolgenMsMs
Error: missing argument ms
Usage: MolgenMsMs ms=<filename> [msms=<filename>] [out=<filename>] [exist]
      [m=<number>] [ion=+H|-H|+Na|-e] [cha=<number>] [ppm=<number>]
      [msmv=ndp|nsse|nsae] [acc=<number>] [rej=<number>] [thms=<number>]
      [thmsms=<number>] [thcomb=<number>] [sort[=ppm|msmv|msmsmv|combm]]
      [el=<elements>] [ff=<fuzzy formula>] [vsp[=<even|odd>]]
      [vsm2mv[=<value>]] [vsm2ap2[=<value>]] [hcf] [wm[=lin|sqrt|log]]
      [wi[=lin|sqrt|log]] [exp=<number>] [oei] [dbeexc=<number>]
      [oms[=<filename>]] [omsms[=<filename>]] [analyze] [dbe] [cm] [pc] [sc]

Explanation:
ms      : filename of MS data (*.txt)
msms   : filename of MS/MS data (*.txt)
out     : output generated formulas
exist  : allow only molecular formulas for that at least one
         structural formula exists; supersets vsp, vsm2mv, vsm2ap2
m       : experimental molecular mass (default: mass of MS basepeak)
ion     : type of ion measured (default: +H)
cha     : charge of ion measured (default: +1)
ppm     : accuracy of measurement in parts per million (default: 5)
msmv    : MS match value based on normalized dot product, normalized
         sum of squared or absolute errors (default: nsae)
acc     : allowed deviation for full acceptance of MS/MS peak in ppm
         (default: 2)
rej     : allowed deviation for total rejection of MS/MS peak in ppm
         (default: 4)
thms    : threshold for the MS match value
thmsms  : threshold for the MS/MS match value
thcomb  : threshold for the combined match value
sort    : sort generated formulas according to mass deviation in ppm,
         MS match value, MS/MS match value or combined match value
el      : used chemical elements (default: CHBrClFINOPSSi)
ff      : supersets el and uses fuzzy formula for limits of
         element multiplicities
vsp     : valency sum parity (even for graphical formulas)
vsm2mv  : lower bound for valency sum - 2 * maximum valency
         (<=>0 for graphical formulas)
vsm2ap2 : lower bound for valency sum - 2 * number of atoms + 2
         (<=>0 for graphical connected formulas)
hcf     : apply Heuerding-Clerc filter
wm      : m/z weighting for MS/MS match value
wi      : intensity weighting for MS/MS match value
exp     : exponent used, when wi is set to log
oei     : allow odd electron ions for explaining MS/MS peaks
dbeexc  : excess of double bond equivalent for ions
oms     : write scaled MS peaks to output
omsms   : write weighted MS/MS peaks to output
analyze : write explanations for MS/MS peaks to output
dbe     : write double bond equivalents to output
cm      : write calculated ion masses to output
pc      : output match values in percent
sc      : strip calculated isotope distributions

Description:
Generate molecular formulas, calculate match values by MS isotope
peak intensities and MS/MS fragment peak masses

D:\Programs\MolgenMsMs>

```

Do not worry about the error message "missing argument ms". The program needs a mass spectral file as an obligatory argument.

5. Description

MOLGEN-MS/MS is provided as command line application and is called using the following syntax:

```

Usage: molgenmsms ms=<filename> [msms=<filename>] [out=<filename>]
      [exist[=mv]] [m=<number>] [ion=+H|-H|+Na|-e] [cha=<number>]
      [ppm=<number>] [msmv=ndp|nsse|nsae] [acc=<number>] [rej=<number>]
      [thms=<number>] [thmsms=<number>] [thcomb=<number>]
      [sort[=ppm|msmv|msmsmv|combm]] [el=<elements>] [ff=<fuzzy formula>]
      [vsp[=<even|odd>]] [vsm2mv[=<value>]] [vsm2ap2[=<value>]] [hcf]
      [wm[=lin|sqrt|log]] [wi[=lin|sqrt|log]] [exp=<number>] [oei]
      [dbeexc=<number>] [oms[=<filename>]] [omsms[=<filename>]]
      [analyze] [dbe] [cm] [pc] [sc]

```

Explanation:
ms : filename of MS data (*.txt)

msms : filename of MS/MS data (*.txt)
out : output generated formulas
exist : allow only molecular formulas for that at least one structural formula exists; overrides vsp, vsm2mv, vsm2ap2; argument mv enables multiple valencies for P and S
m : experimental molecular mass (default: mass of MS basepeak)
ion : type of ion measured (default: +H)
cha : charge of ion measured (default: +1)
ppm : accuracy of measurement in parts per million (default: 5)
msmv : MS match value based on normalized dot product, normalized sum of squared or absolute errors (default: nsae)
acc : allowed deviation for full acceptance of MS/MS peak in ppm (default: 2)
rej : allowed deviation for total rejection of MS/MS peak in ppm (default: 4)
thms : threshold for the MS match value
thmsms : threshold for the MS/MS match value
thcomb : threshold for the combined match value
sort : sort generated formulas according to mass deviation in ppm, MS match value, MS/MS match value or combined match value
el : used chemical elements (default: CHBrClFINOPSSi)
ff : overrides el and uses fuzzy formula for limits of element multiplicities
het : formulas must have at least one hetero atom
vsp : valency sum parity (even for graphical formulas)
vsm2mv : lower bound for valency sum - 2 * maximum valency (>=0 for graphical formulas)
vsm2ap2 : lower bound for valency sum - 2 * number of atoms + 2 (>=0 for graphical connected formulas)
hcf : apply Heuerding-Clerc filter
wm : m/z weighting for MS/MS match value
wi : intensity weighting for MS/MS match value
exp : exponent used, when wi is set to log
oei : allow odd electron ions for explaining MS/MS peaks
dbeexc : excess of double bond equivalent for ions
oms : write scaled MS peaks to output
omsms : write weighted MS/MS peaks to output
analyze : write explanations for MS/MS peaks to output
dbe : write double bond equivalents to output
cm : write calculated ion masses to output
pc : output match values in percent
sc : strip calculated isotope distributions

Description:

Generate molecular formulas, calculate match values by MS isotope peak intensities and MS/MS fragment peak masses

5.1 Input

The only obligatory argument is a mass spectrum file, which is specified using parameter `ms`. The MS file should be an ASCII file with one peak per line, specified by the m/z value and the intensity. An example MS file could look as follows:

```
225.07578    5357930210.9161
226.07578    363441756.2500
227.06588    34586811.0915
```

The first column contains the measured m/z values, followed by the intensities. It does not matter whether the measured intensities are given in absolute counts, or as some relative value. The three MS peaks shown above are stored in the example file *SinapinicAcidMs.txt*, which is delivered with the software. We can start the program with this MS file as argument using the command:

```
MolgenMsMs ms=SinapinicAcidMs.txt
```

The program gives the following response:

```
finished reading MS file (3 peaks)
232/232/232/232/232 (final/MSMS/MS-filter/valid/total) formula(s) in 0.3s
```

At first *MOLGEN-MS/MS* confirms that it has read the MS file and displays the number of peaks found in the file. In the second line it shows the number of formulas computed and the time required. The exact meaning of this output and how to display generated formulas will be described in Section 5.2. A second, optional but nevertheless important input file may contain MS/MS data. The example MS/MS file delivered with this software is named *SinapinicAcidMsMs.txt* and contains the following data:

147.0442	111087.1387
155.0704	141098.0469
175.0390	3453872.2031
178.0581	61648.4158
181.0860	58958.1887
183.0653	431582.9688
207.0652	47440342.5000
210.0264	83937.9844
224.0635	189674.1650
225.0758	1469526.1797

It is structured as the MS data: The first column contains the measured m/z values, followed by the intensities. The MS/MS data file can be specified using the parameter `msms`, e.g. `msms=SinapinicAcidMsMs.txt`.

5.2 Output

After input files are read, the program shows the numbers of peaks in the MS and the MS/MS file. Using option `out` the generated formulas are written to the standard output. For instance the first five lines look as follows:

C4H137F2	-1.5	0.97412	0.00000	0.00000
C5H21F2C13	3.6	0.44050	0.00000	0.00000
C12H9NF3	-0.9	0.94161	0.00000	0.00000
C4H10NF8	-0.2	0.97766	0.00000	0.00000
C2H21NF2I	-0.8	0.95635	0.00000	0.00000

You see that most of the formulas generated this way are not valid in terms of their double bond equivalents. Invalid formulas can be eliminated by using option `exist`. Then only such molecular formulas are generated for which at least one valid structural formula exists. Using this option the output is decreased to 24 valid formulas. The first five lines look like this:

C7H8N4F4	-0.0	0.97851	0.14347	0.14039
C5H5N10F	1.1	0.97848	0.20000	0.19570

```
C2H6N10OF2      -4.0  0.99045  0.00286  0.00283
C6H13N4O3Cl     3.9  0.75703  0.00000  0.00000
C11H12O5         0.1  0.94402  0.70000  0.66081
```

and in the summary line some numbers have changed to

```
24/24/24/24/232 (final/MSMS/MS-filter/valid/total) formula(s) in 0.3s
```

Each generated formula is followed by four numbers. The first value denotes the relative deviation of the calculated mass of the molecular ion from the mass of the basepeak (basemass) in the MS in parts per million (ppm). A negative signum indicates that the calculated mass is less than the measured mass.

Instead of the basemass the user can define the reference mass via parameter `m`. By default the calculated mass is computed for protonated, simply positively charged molecular ions $[M+H]^+$. The type and charge of the molecular ion can be modified using parameters `ion` and `cha`, respectively. For instance, you can switch to electron impact ionization using `ion=-e`. By default a relative difference of at most 5 ppm between calculated and measured mass is allowed. This maximum mass deviation can be changed using argument `ppm`.

The next number following the mass deviation is the MS/MS MV. It is calculated from measured and theoretical isotope peak intensities. The MV ranges from 0 to 1. The better theoretical and measured isotope distributions fit, the higher is the MV. There are three different algorithms implemented for calculating the MS MV, which can be accessed via parameter `msmv`: the normalized dot product (`msmv=ndp`), the normalized sum of squared errors (`msmv=nsse`) or the normalized sum of absolute errors (`msmv=nsae`). The latter is also the default.

The third numerical value in each line of output is the MS/MS MV. It shows how well the masses of the MS/MS fit possible fragments of the candidate molecular formula. Again, the MS/MS MV is a value between 0 and 1, whereby higher values reflect a better fit. The main two parameters for calculating this MV indicate deviations of measured and calculated fragment mass in ppm as follows: Parameter `acc` specifies the (maximum) allowed deviation for full acceptance of a peak in the MS/MS. I.e. whenever the candidate formula has a subformula with mass deviation below this value then the MS/MS peak is considered as *fully explained*. Parameter `rej` on the other hand defines the (minimum) necessary deviation for total rejection of an MS/MS peak. I.e. if the candidate formula has no subformula with mass deviation below this value then the MS/MS peak is considered as *fully unexplained*. Default values are `acc=2` and `rej=4`. In case parameter `rej` is greater than `acc` it is possible that *partly explained* peaks occur. How fully explained, partly explained and unexplained peaks contribute to the MV is explained in detail in Reference 2. Further options that have influence on the calculation of the MS/MS MV are described in Section 6.

Finally, the last value in every line is a combined MV which is simply obtained by multiplying the MS MV and the MS/MS MV. Thus the combined MV has values between 0 and 1, too. In most analyses of unknown samples of medium molecular weight, it should be possible to distinguish the true formula from false candidates by its higher combined MV.

In order to reduce the amount of formulas in the output, it is possible to prescribe thresholds for the MS MV, the MS/MS MV and the combined MV by using parameters `thms`, `thmsms` and `thcomb`, respectively. These filters are applied successively, and the numbers of remaining formulas after each step are written to the results line. For instance the command

```
MolgenMsMs ms=SinapinicAcidMs.txt msms=SinapinicAcidMsMs.txt exist out
thms=0.9 thmsms=0.5 thcomb=0.5
```

produces the output

```

finished reading MS file (3 peaks)
finished reading MS/MS file (10 peaks)
C11H12O5          0.1  0.94402 0.70000 0.66081
C3H12N6O4Si      -1.9  0.93338 0.55949 0.52221
2/3/13/24/232 (final/MSMS/MS-filter/valid/total) formula(s) in 0.2s

```

This means that 24 of the 232 candidate molecular formulas passed the criteria that at least one structural formula exists, 13 of the 24 received a MS MV of at least 0.9, three of the remaining 13 formulas have a MS/MS MV greater or equal to 0.5, and two of them have a combined MV greater or equal 0.5. For this example the true formula $C_{11}H_{12}O_5$ receives the highest MS, MS/MS and combined match value.

With option `sort` the formulas are sent to the output once again after generation, but now in a sorted manner. The formulas can be sorted in ascending order of the absolute value of the relative mass deviation in ppm (`sort=ppm`), or in descending order of MS MV (`sort=msmv`), MS/MS MV (`sort=msmsmv`), or of the combined match value (`sort=combmv`). If option `sort` is used without any value, sorting by the combined MV is applied.

6. Advanced Features

As seen in the beginning of Section 5 there are many more arguments available that help to change the behavior of the program. These advanced features are described below.

6.1 Chemical Elements and their Multiplicities

By default the program uses eleven elements typically occurring in organic chemistry: C, H, N, O, Br, Cl, F, I, P, S and Si. Atomic masses and isotopic abundances are taken from Reference 3. The user should be aware of the fact that the number of available elements is a crucial factor considering the number of candidate formulas. In general, the more elements can be excluded the fewer candidates will be generated. The set of elements can be passed to the program by argument `e1`. It must be followed by a sequence of element symbols. E.g. if you define `e1=CHNO` only formulas consisting of no other atoms than C, H, N and O will be generated.

It can also be useful if you know already the number of atoms for some elements, or if you can at least give lower and/or upper bounds. This might reduce the number of candidates (and also the computation time) dramatically. This kind of information can be passed to the program as a so-called *fuzzy formula*. In a fuzzy formula each element symbol is either followed by an integer (denoting the exact number of atoms for this element) or by an interval of integers. An interval is specified by two integers, the lower and the upper bound, separated by a dash. For instance `ff=C5H6-10` would comprise C_5H_6 , C_5H_8 and C_5H_{10} . Elements not specified in the fuzzy formula may not occur. If argument `ff` is used, specifications via argument `e1` are ignored.

Compounds containing only C and H cannot be detected by electrospray ionization MS. For this reason the user can force the program to generate formulas that contain at least one hetero atom using option `het`.

6.2 Mathematical Rules for Molecular Formulas

In Section 5 the option `exist` was already introduced. This option forces the program to check some mathematical rules that have to be fulfilled for valid molecular formulas. These rules are based on the

elements' valences. The program uses valence 1 for H, Br, Cl, F and I, valence 2 for O and S, valence 3 for N and P, and valence 4 for C and Si. If option `exists=mv` is used, multiple valences are considered for P (3, 5) and S (2, 4, 6). Altogether there are three rules that have to be fulfilled:

1. The parity of the sum of valences has to be even,
2. the valence sum minus two times the maximum valence must not be negative and
3. the valence sum minus two times the number of atoms plus two must not be negative.

For a more mathematical formulation of these rules and explanations on their meanings see Reference 2 and literature cited therein. As already mentioned, option `exist` forces the program to check all three rules. But it is also possible to use only subsets of these rules or to modify them slightly. Using argument `vsp` you can specify the valence sum parity. For instance it is possible to set `vsp=odd` in order to generate only formulas with odd parity valence sums. With argument `vsm2mv` you can specify the lower bound for the valence sum minus two times the maximum valence, which must be `vsm2mv=0` in order to apply Rule 2 properly. Argument `vsm2ap2` allows defining the lower bound for the valence sum minus two times the number of atoms plus two, i.e. Rule 3 can be independently enabled or modified. If option `exist` is used the above three arguments `vsp`, `vsm2mv` and `vsm2ap2` are ignored.

6.3 Heuristic Rules for Molecular Formulas

In addition to these mathematical rules, there are several heuristic rules, which were obtained by statistical examinations of large compound databases. Such heuristic rules were compiled by Heuerding and Clerc and were published in Reference 4. They are implemented in *MOLGEN-MS/MS* and can be activated as filters by option `hcf`. If these filters are applied, there will be a new number added to the results line, which states how many formulas passed this *Heuerding-Clerc filter* successfully. For instance the command

```
MolgenMsMs.exe ms=SinapinicAcidMs.txt exist hcf
```

produces the result line

```
12/12/12/12/24/232 (final/MSMS/MS-/HC-filter/valid/total) formula(s) in 0.3s
```

It states that twelve formulas passed the Heuerding-Clerc filter. Although this might appear like a handy tool to narrow down the number of candidate formulas, usage of this option is discouraged. During the test phase it turned out that for many examples the true formula failed to pass this filter with the fatal consequence that the true formula is missing in the output (cf. Note 5).

6.4 Refinements of the MS/MS Match Value

By default, peak intensities in the MS/MS have no influence on the calculation of the MS/MS MV (cf. definition of MV_1 in Reference 2). However, it might appear useful that peaks with high intensity should be weighted stronger than small peaks. Such a weighting by intensity can be activated with option `wi`. Then the MS/MS MV is calculated according to the definition of MV_2 in Reference 2. The weighting by peak intensity can be done linearly (`wi=lin`, default), by square root (`wi=sqrt`) or logarithmically (`wi=log`) in terms of the intensity (cf. definition of MV in Reference 2). If logarithmic weighting by intensity is used, then additionally an exponent a can be specified using parameter `exp`. The weighting is then done by the function $f(I(m_i)) = \log_{10}(1 + I(m_i) \cdot 10^a)$, where $I(m_i)$ are the intensities of the peaks in the MS/MS. The default value for a is 5.

Another idea is motivated by the fact that peaks of higher mass might be more important for the identification of the molecular formula than peaks of small masses. Weighting by peak masses can be enabled using option `wm` (cf. definition of MV_3 in Reference 2). Also the weighting by peak masses can be done linearly (default), by square root, or logarithmically.

Since radical ions rarely appear in MS/MS spectra, by default the calculation of the MS/MS MV only accounts for subformulas that belong to even electron ions. However, using option `oei` it is possible to include subformulas that belong to odd electron ions short OEI, i.e. radical ions) into the MS/MS MV calculation as well. In some examples this option turned out to be quite helpful (see Reference 2).

Another criterion for ions in a mass spectrometer is that they should not significantly exceed the number of double bond equivalents (DBE) of the precursor ion. In MOLGEN-MS/MS this excess in DBE is set to three by default. This number can be modified using parameter `dbeexc`.

6.5 Extending and Changing the Output

The rescaled MS and the MS/MS as modified for intensity weighting by arguments `wi` and `exp` can be written to the standard output by options `oms` and `omsms`, respectively. These options can be followed by filenames in order to write the preprocessed MS or MS/MS to file. So if you call

```
MolgenMsMs ms=SinapinicAcidMs.txt msms=SinapinicAcidMsMs.txt omsms
```

you should not be surprised about

```
147.044200    1.0000
155.070400    1.0000
175.039000    1.0000
178.058100    1.0000
181.086000    1.0000
183.065300    1.0000
207.065200    1.0000
210.026400    1.0000
224.063500    1.0000
225.075800    1.0000
```

in the output. By default there is no weighting by intensity, and therefore all peaks are internally treated with equal intensity 1.0.

Using option `analyze` allows to get a deeper insight into the calculation of the MS/MS MV. After each formula candidate a list of MS/MS peaks and subformulas explaining these peaks is added to the output. For instance

```
MolgenMsMs ms=SinapinicAcidMs.txt msms=SinapinicAcidMsMs.txt exist el=CHNO analyze
out
```

produces the output

```
finished reading MS file (3 peaks)
finished reading MS/MS file (10 peaks)
C11H12O5          0.1  0.94402  0.70000  0.66081
 147.04420      C9H7O2          -1.0
 155.07040      C8H11O3          -0.8
 175.03900      C10H7O3          -0.2
 181.08600      C10H13O3         -0.4
 183.06530      C9H11O4          -0.6
```

```

207.06520      C11H11O4      -0.1
225.07580      C11H13O5      -0.2

```

```
1/1/1/1/2 (final/MSMS/MS-filter/valid/total) formula(s) in 0.0s
```

Below the line with the formula $C_{11}H_{12}O_5$ we see in each line a peak mass of the MS/MS followed by the subformula of $C_{11}H_{12}O_5$ explaining this peak and the relative mass deviation of calculated ion masses and measured peak masses in ppm. A negative signum means that the calculated mass is less than the measured mass.

The last subformula, $C_{11}H_{13}O_5$ has one hydrogen atom more than the proposed formula $C_{11}H_{12}O_5$. This is caused by the fact that protonation was used as ionization mode in the MS, i.e. the ion $C_{11}H_{13}O_5^+$ was subject for the analysis of MS/MS fragments.

Often, it is interesting to see the number of double bond equivalents for the formula candidates and subformulas. The DBE can be added to the output by option `dbe`. Also the calculated ion masses can be included to the output. This can be achieved by option `cm`. If you prefer MV in per cent, you should add option `pc`. The command

```
MolgenMsMs ms=SinapinicAcidMs.txt msms=SinapinicAcidMsMs.txt exist el=CHNO analyze
out dbe cm pc
```

produces the output

```

finished reading MS file (3 peaks)
finished reading MS/MS file (10 peaks)
C11H12O5      6.0      225.07575      0.1      94.402      70.000      66.081
147.04420     C9H7O2      6.5      147.04406      -1.0
155.07040     C8H11O3     3.5      155.07027      -0.8
175.03900     C10H7O3     7.5      175.03897      -0.2
181.08600     C10H13O3    4.5      181.08592      -0.4
183.06530     C9H11O4     4.5      183.06519      -0.6
207.06520     C11H11O4    6.5      207.06519      -0.1
225.07580     C11H13O5    5.5      225.07575      -0.2

```

```
1/1/1/1/2 (final/MSMS/MS-filter/valid/total) formula(s) in 0.1s
```

that demonstrates all modifications and changes. The first number after each formula is the DBE, followed by the calculated mass of the molecular ion or the fragment ion, respectively. Finally we see that the match values, 94.402, 70.000 and 66.081 are now given in percent.

DBE are calculated as defined in Reference 2. Note that for simply positively charged ions (even electron ions, short EEI) they are non-integers, for radical ions (OEI) the DBE is integer, again.

6.6 Processing Calculated Isotopic Distributions

Using option `sc` removes peaks in calculated isotope distributions that are not present in the experimental MS. The effect of this manipulation is rather small and this option was mainly implemented for testing purposes.

Notes and References

- ¹ A. Kerber, R. Laue, M. Meringer, K. Varmuza. MOLGEN-MS: Evaluation of Low Resolution Electron Impact Mass Spectra with MS Classification and Exhaustive Structure Generation. Volume 15 of *Advances in Mass Spectrometry*, pages 939-940. Wiley, 2001.
- ² M. Meringer, S. Reinker et al: MS/MS Data Improves Automated Determination of Molecular Formulas by Mass Spectrometry. To be published.
- ³ Exact Masses and Isotopic Abundances. <http://www.sisweb.com/referenc/source/exactmaa.htm>
Scientific Instrument Services, Inc.
- ⁴ S. Heuerding and T. Clerc: Simple Tools for the Computer-Aided Interpretation of Mass Spectra. *Chemometr. Intell. Lab. Syst.* 20, 57-69, 1993.
- ⁵ Five out of the eleven compounds used during the development phase did not pass the Heuerding-Clerc filter: peptide MRFA ($C_{23}H_{37}N_7O_5S$), reserpine ($C_{33}H_{40}N_2O_9$), CHAPS ($C_{32}H_{58}N_2O_7S$), maltopentaose ($C_{30}H_{52}O_{26}$) and cyclosporin C ($C_{62}H_{111}N_{11}O_{13}$).