

## Practice Session: Structure Generation and Tricky Cases

### Suggested Exercises

1. Go to <http://molgen.de/?src=documents/molgenonline>. Perform all the examples and click through the structures displayed to see what it does. Look at the NUMBER of structures generated.

- How many structural isomers are there of C<sub>6</sub>H<sub>6</sub> using MOLGEN Online?
- How many C<sub>6</sub>H<sub>6</sub> substances are in ChemSpider?
- How many halogenated ethanes?
- How many chlorinated benzenes?
- How many do you generate with example 4? Why?
- How many do you generate with example 5? Are you surprised?

2. Have a look at some generated structures via MetFrag.

Take the SDF file: BR1A1\_19m208\_run3\_export\_wM\_END.sdf for the structures.

Take the csv file: BR1A1\_19m208.csv for the spectra.

Upload to MetFrag using the "Local SDF" option.

Copy the csv peaks into the spectrum area. You may need to convert to two-column data (mine does this automatically).

This is GC-MS data. What are the right fragmentation settings? Think carefully. If you have no idea – please ask.

What do the results look like? What do the structures look like? Which is your "favourite"?

How many do you think exist? **Think about this before you move on...**

What do you think is the right answer? How many others do you think are "possible"?

Figure out the formula and search ChemSpider – do the results change?

What do you think is the right answer? How many others do you think are "possible"?

Are your answers different now?

3. Have a go at another example: MetFrag with the formula C<sub>13</sub>H<sub>10</sub>ClNO and the spectrum in SW\_DCMD\_MMM\_16m6752m1.csv. How many structures do you get?

You have the substructure information in my slides. Play around a bit with PubChem or ChemSpider. Can you find a way to "pick" a few candidates only with these substructures and fragment them with MetFrag? Remember you can enter specific database identities into MetFrag to "trim" pre-fragmentation.

Try to find all candidates with this formula and the "present" substructures. Does that improve the ranking? Can you get the substructure search to work?

**Keep going over the page....**

What happens if you try to include diclofenac as a possible answer? Can you still explain part of the spectrum? Why? Perform the search with the ChemSpider IDs 26001359 and 2925. What fragments are in common? Which are different? Missing?

Try generating some structures. Go to <http://molgen.de/?src=documents/molgenonline>. Add in Formula C<sub>13</sub>H<sub>10</sub>N[val=3]OCl<sub>1</sub>, try different restrictions that make sense. This job will time-out. There are too many possibilities...Unfortunately, we can't add substructures to the online version.

Try uploading the generated structures in DCMD\_16m6752\_run3\_export\_modified.sdf to MetFrag. How many are in the file? How many are displayed in MetFrag? Why? Which is the correct answer? Look at the slides.... Which substructure wasn't added correctly into the MOLGEN restrictions in this file?