

## Practice Session: Structure Generation and Tricky Cases

### Suggested Exercises - ANSWERS

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.

- a) How many structural isomers are there of C<sub>6</sub>H<sub>6</sub> using MOLGEN Online? **217 structures.**
- b) How many are in ChemSpider? **22 hits.**
- c) How many halogenated ethanes? **630.**
- d) How many chlorinated benzenes? **13.**
- e) How many do you generate with example 4? Why? **1. Only one isomer possible.**
- f) How many do you generate with example 5? Are you surprised? **45.**

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

This is GC-MS data. What are the right fragmentation settings? **I suggest e.g. mzabs 0.4, mzppm 5, M, pos. You can vary these and see what happens!**

What do the results look like? What do the structures look like? **A lot of strange structures at the top, right?**

Which is your "favourite"? **Personal choice!**

How many do you think exist? What do you think is the right answer? How many others do you think are "possible"? **These answers are very subjective...**

Figure out the formula and search ChemSpider – do the results change? **7 structures in ChemSpider. But does presence in ChemSpider guarantee that it exists? If it's not in ChemSpider, does it exist?**

What do you think is the right answer? How many others do you think are "possible"? **The correct answer is <http://www.chemspider.com/Chemical-Structure.6552.html>**

Are your answers different now? **How many would have wanted to ignore all structures with a cyclobutyl in them? With a cyclopropyl? A bridge over two aromatic atoms?**

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?

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? **The first person to get this to work with the substructures in my slides wins a prize!**

What happens if you try to include diclofenac as a possible answer? **There are some fragments in common, many not.**

Try uploading the generated structures in DCMD\_16m6752\_run3\_export\_modified.sdf to MetFrag. How many are in the file? **6.** How many are displayed in MetFrag? **5.** Why? **Two identical structures are clustered.** Which is the correct answer? **Structure 3, unfortunately not the best score.** Look at the slides.... Which substructure wasn't added correctly into the MOLGEN restrictions in this file? **The NH group wasn't defined properly, adding this in would reduce the set to the two structures in my slides.**