Molecule Set Comparator (MSC)

A short introduction

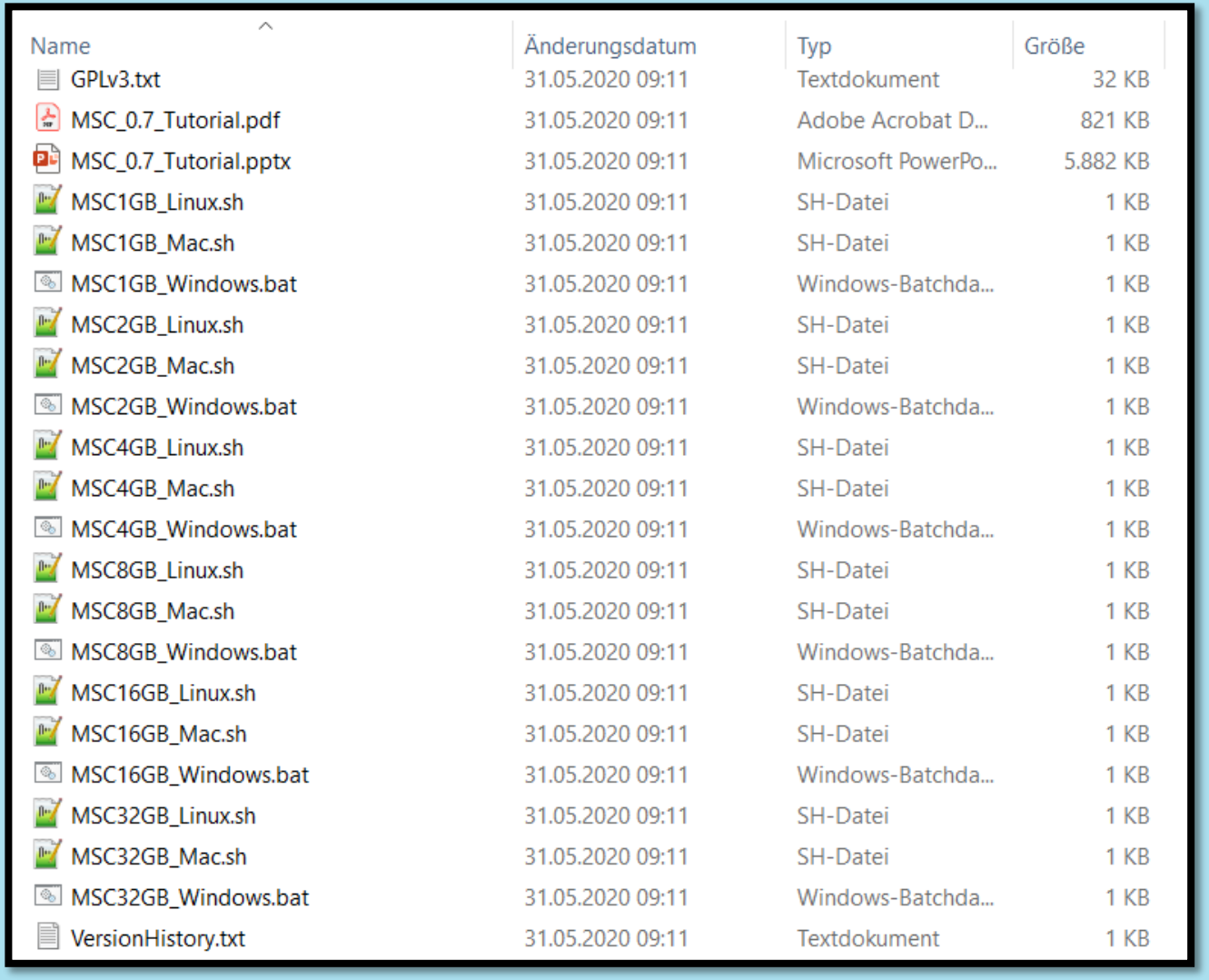
Jan-Mathis Hein

Westfälische Hochschule

FRIEDRICH-SCHILLER-UNIVERSITÄT JENA

Starting the application

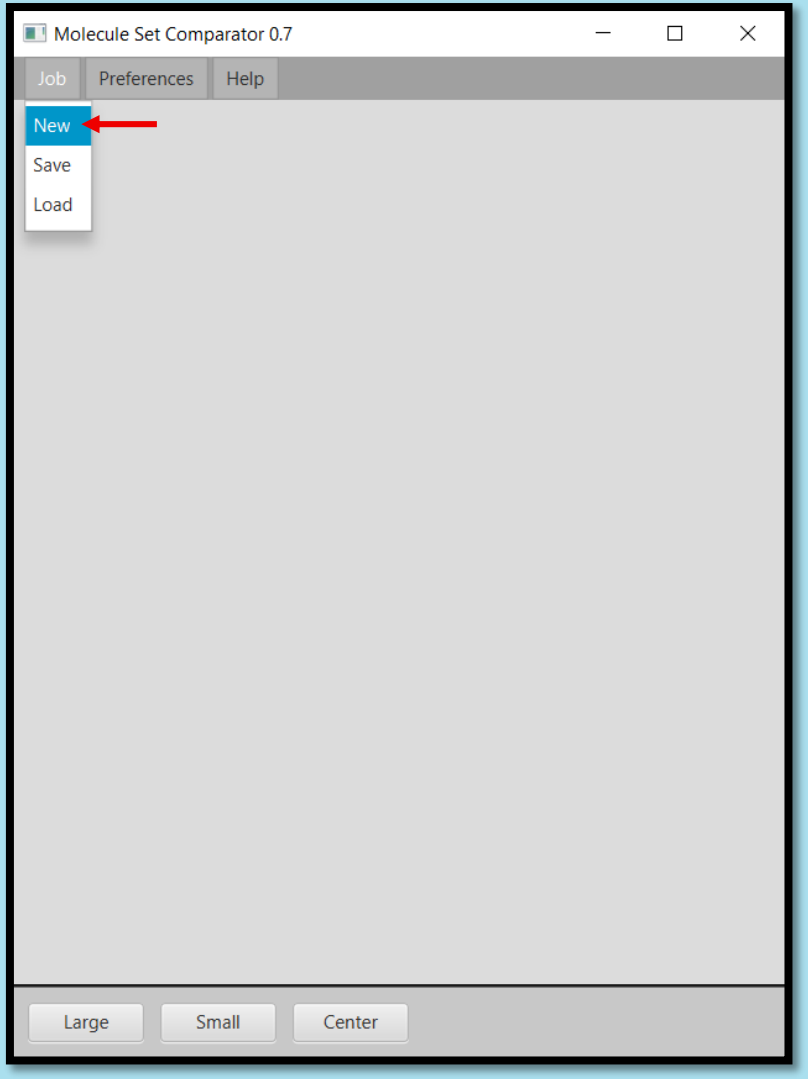
- Open the *“MSC_1.0”* folder
- Select and execute one of *“MSC_?GB_<OS>”* files where <OS> should be your operating system and ? specifies the RAM that is allocated for the JVM in gigabytes.

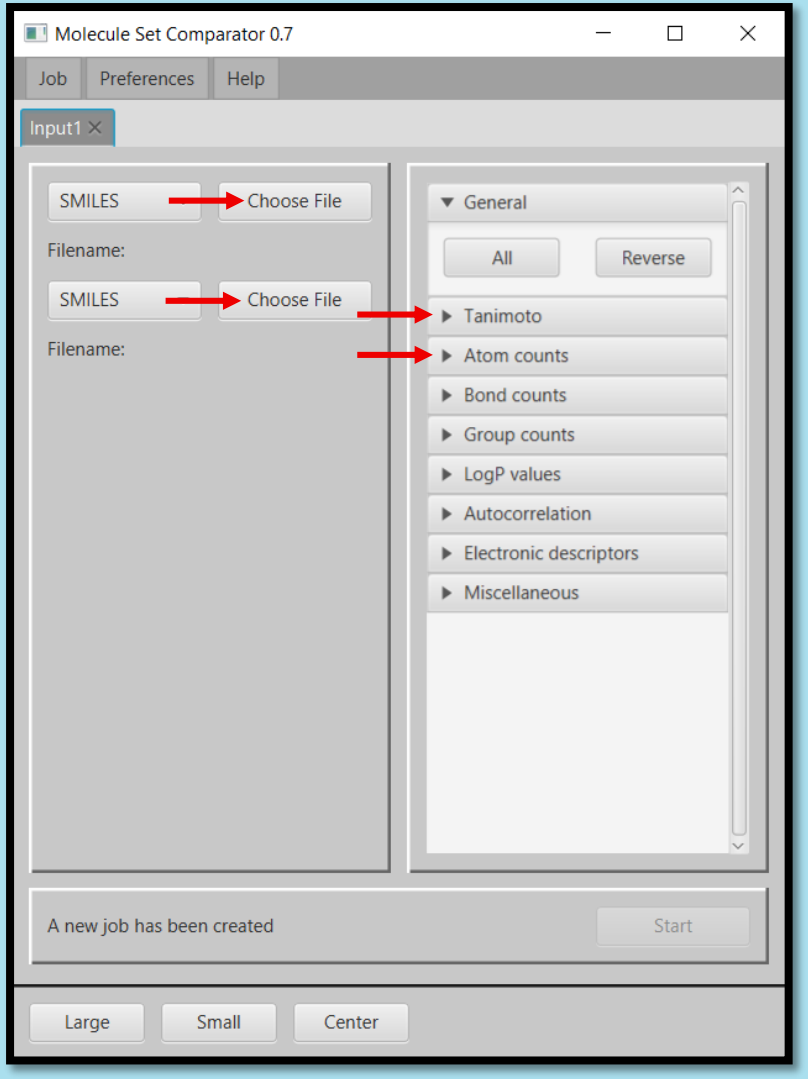


Name	Änderungsdatum	Typ	Größe
GPLv3.txt	31.05.2020 09:11	Textdokument	32 KB
MSC_0.7_Tutorial.pdf	31.05.2020 09:11	Adobe Acrobat D...	821 KB
MSC_0.7_Tutorial.pptx	31.05.2020 09:11	Microsoft PowerPo...	5.882 KB
MSC1GB_Linux.sh	31.05.2020 09:11	SH-Datei	1 KB
MSC1GB_Mac.sh	31.05.2020 09:11	SH-Datei	1 KB
MSC1GB_Windows.bat	31.05.2020 09:11	Windows-Batchda...	1 KB
MSC2GB_Linux.sh	31.05.2020 09:11	SH-Datei	1 KB
MSC2GB_Mac.sh	31.05.2020 09:11	SH-Datei	1 KB
MSC2GB_Windows.bat	31.05.2020 09:11	Windows-Batchda...	1 KB
MSC4GB_Linux.sh	31.05.2020 09:11	SH-Datei	1 KB
MSC4GB_Mac.sh	31.05.2020 09:11	SH-Datei	1 KB
MSC4GB_Windows.bat	31.05.2020 09:11	Windows-Batchda...	1 KB
MSC8GB_Linux.sh	31.05.2020 09:11	SH-Datei	1 KB
MSC8GB_Mac.sh	31.05.2020 09:11	SH-Datei	1 KB
MSC8GB_Windows.bat	31.05.2020 09:11	Windows-Batchda...	1 KB
MSC16GB_Linux.sh	31.05.2020 09:11	SH-Datei	1 KB
MSC16GB_Mac.sh	31.05.2020 09:11	SH-Datei	1 KB
MSC16GB_Windows.bat	31.05.2020 09:11	Windows-Batchda...	1 KB
MSC32GB_Linux.sh	31.05.2020 09:11	SH-Datei	1 KB
MSC32GB_Mac.sh	31.05.2020 09:11	SH-Datei	1 KB
MSC32GB_Windows.bat	31.05.2020 09:11	Windows-Batchda...	1 KB
VersionHistory.txt	31.05.2020 09:11	Textdokument	1 KB

Initializing a new job

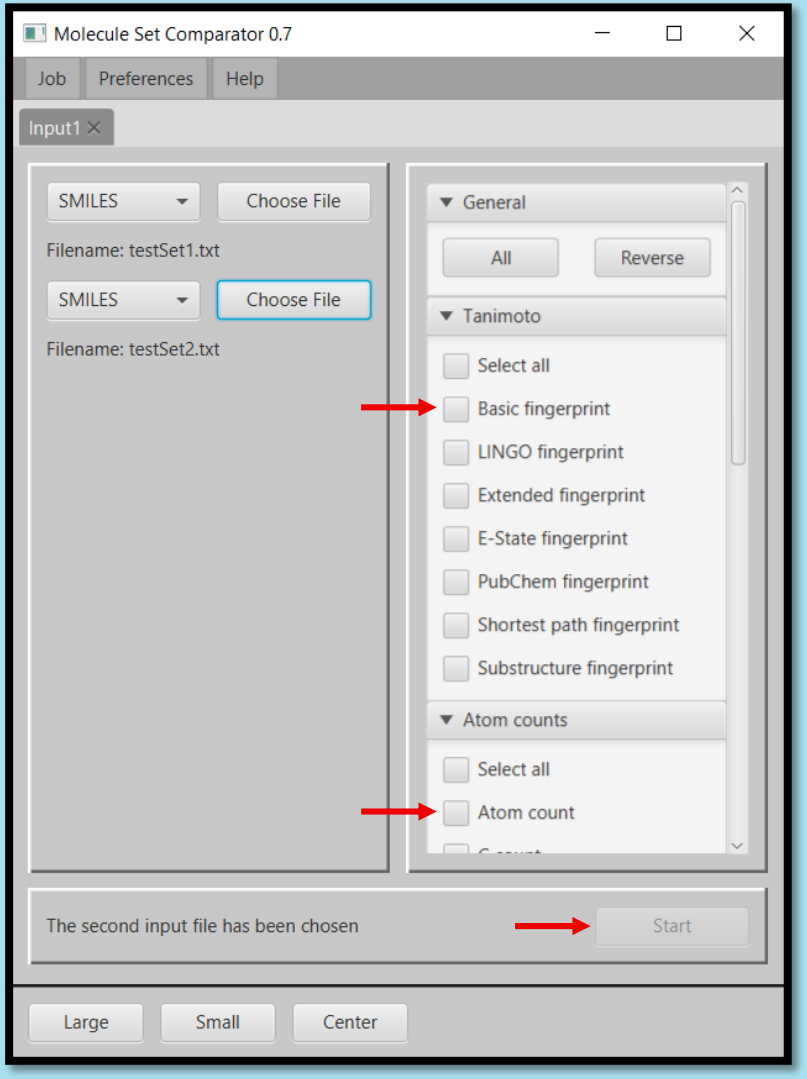
- Create a new job by choosing **Job | New** in the menu bar
- Press the first **Choose File** button
- Use the file chooser to navigate to the *"TestSets"* folder contained in the MSC folder and select the *"testSet1.txt"* file
- Repeat the same procedure with the second **Choose File** button and the *"testSet2.txt"* file
- Then press the **Tanimoto** and **Atom counts** titled panes (marked with red arrows) to expand them





Initializing a new job

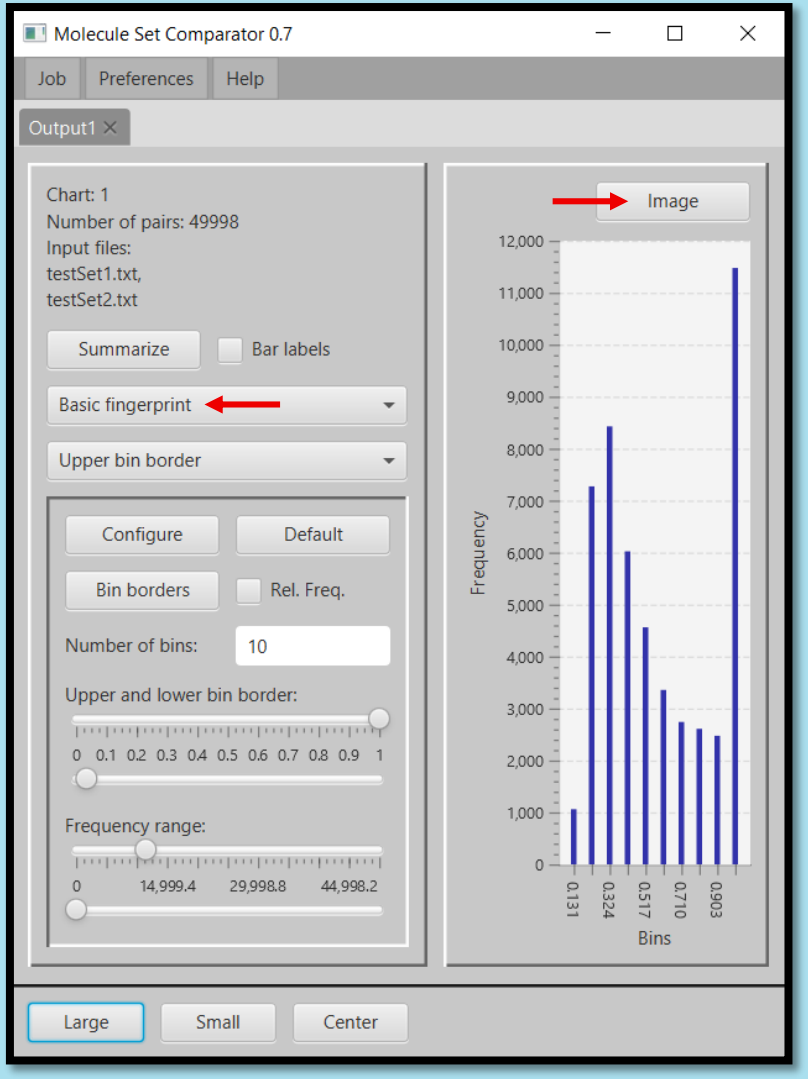
- Select the **Basic fingerprint** and **Atom count** checkboxes
- Then start the execution of the job by pressing the **Start** button and wait until the job finished its execution

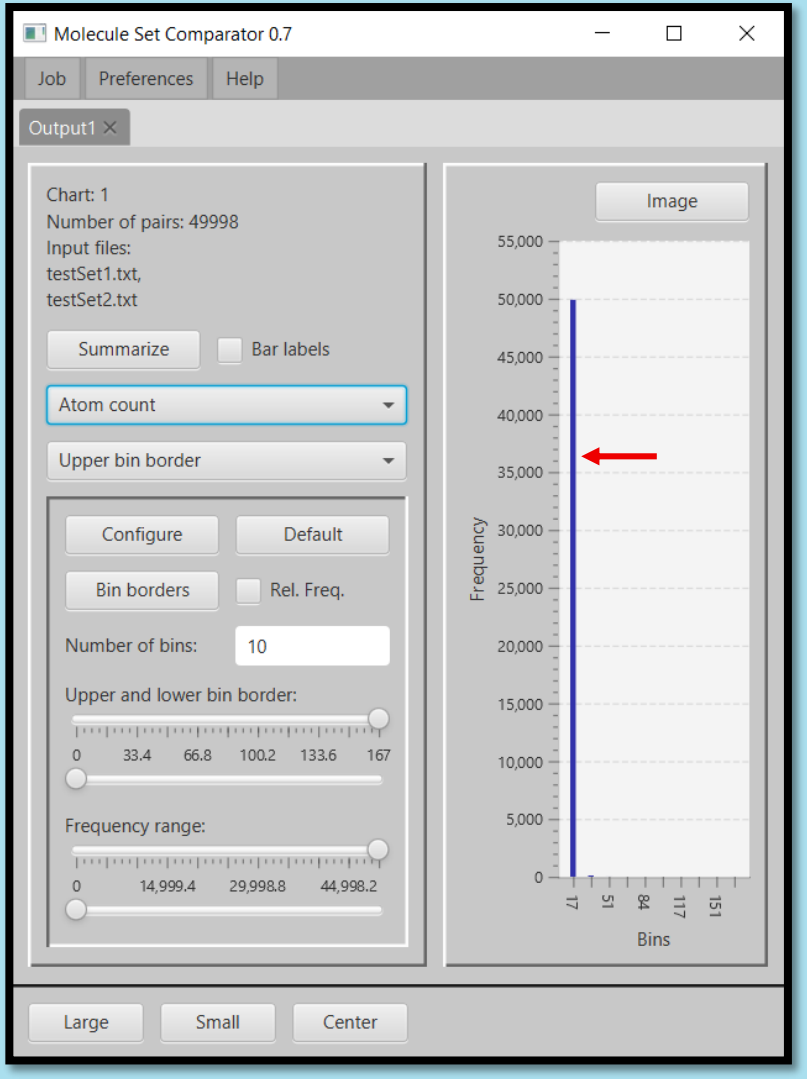


Jan-Mathis Hein
MSC – A short introduction
4

The output tab

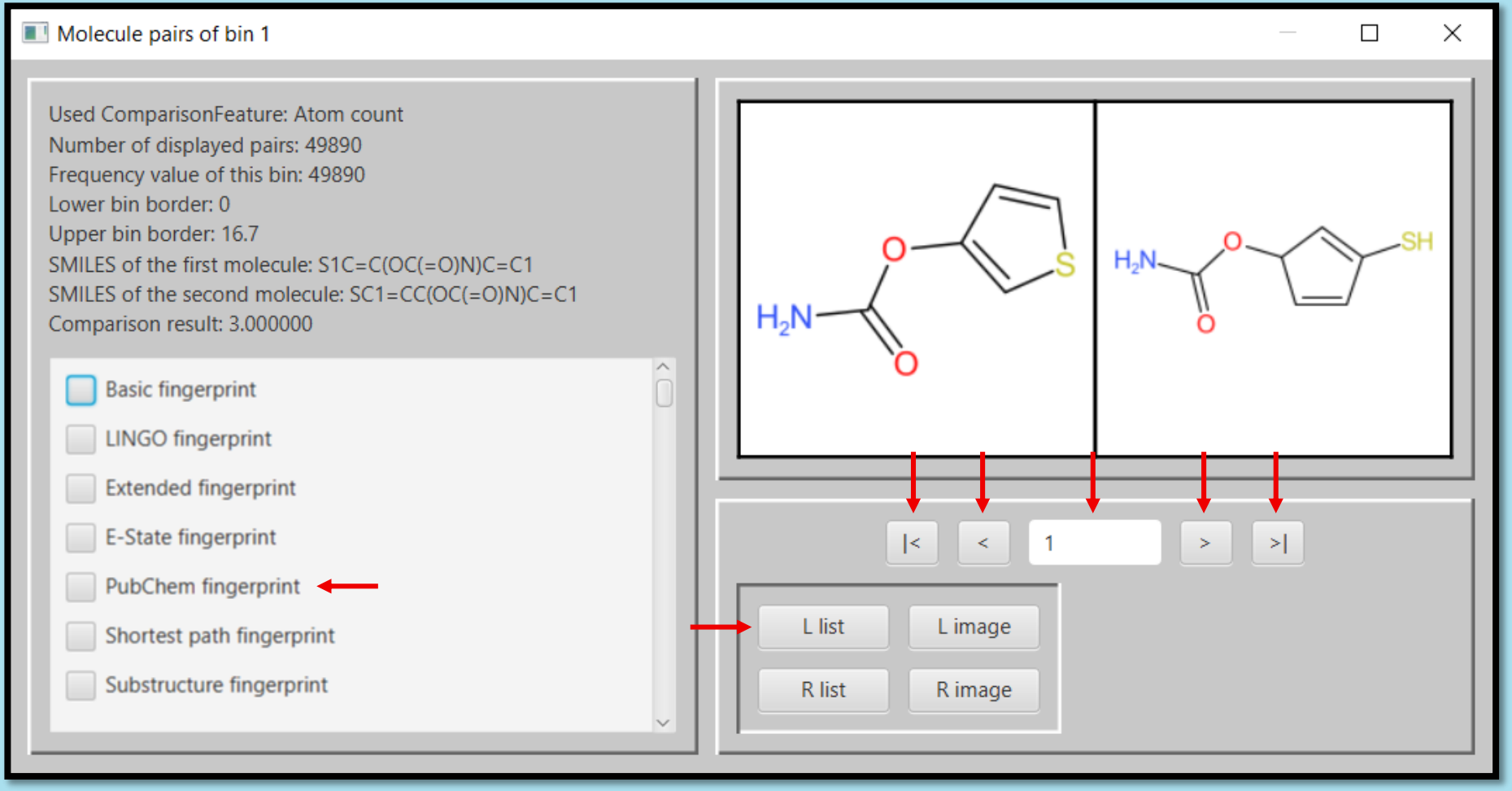
- Save the chart as an PDF file by pressing the **Image** button and using the file chooser
- Switch to the chart of the *“Atom count”* descriptor by pressing the upper choice box and choosing *“Atom count”*
- Open the detail window for the marked bin by clicking on it





Detail window

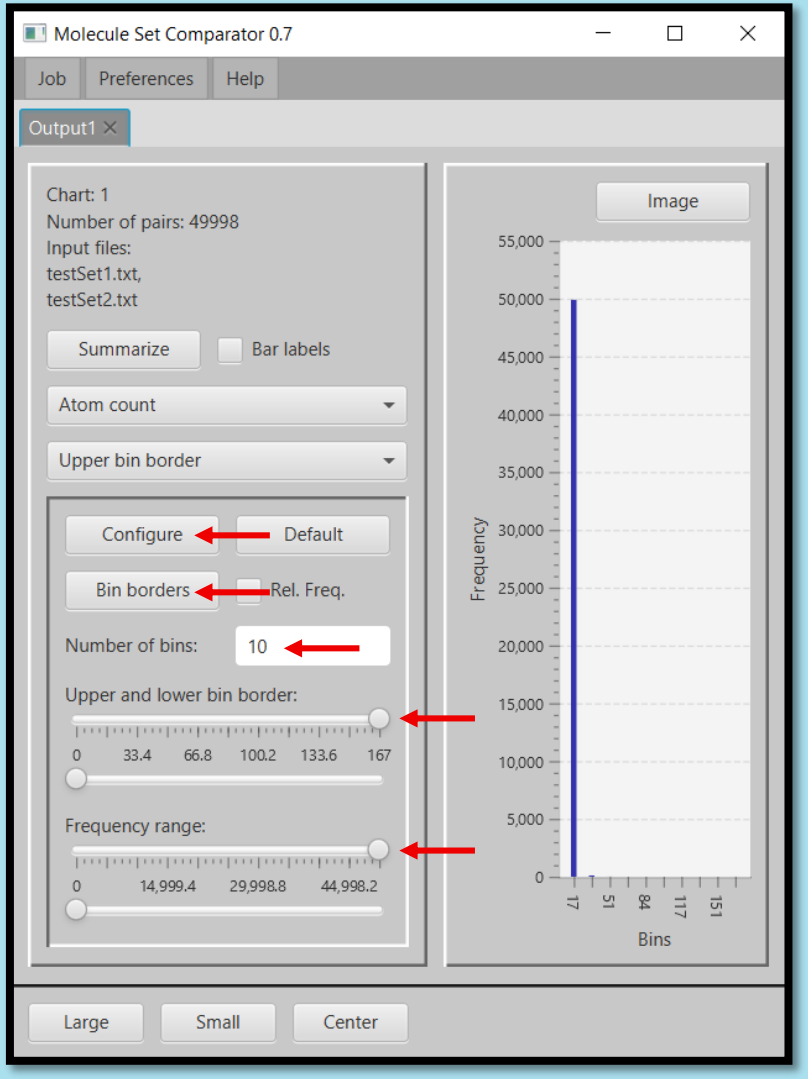
- Use the arrow buttons to browse through the molecule pair list or use the text field to go to a specific position in the list
- Save a list that contains the left molecule of every pair by pressing the **L list** button
- Select the **PubChem fingerprint** checkbox to calculate an additional Tanimoto coefficient
- Close the *"Molecule pairs of bin 1"* window

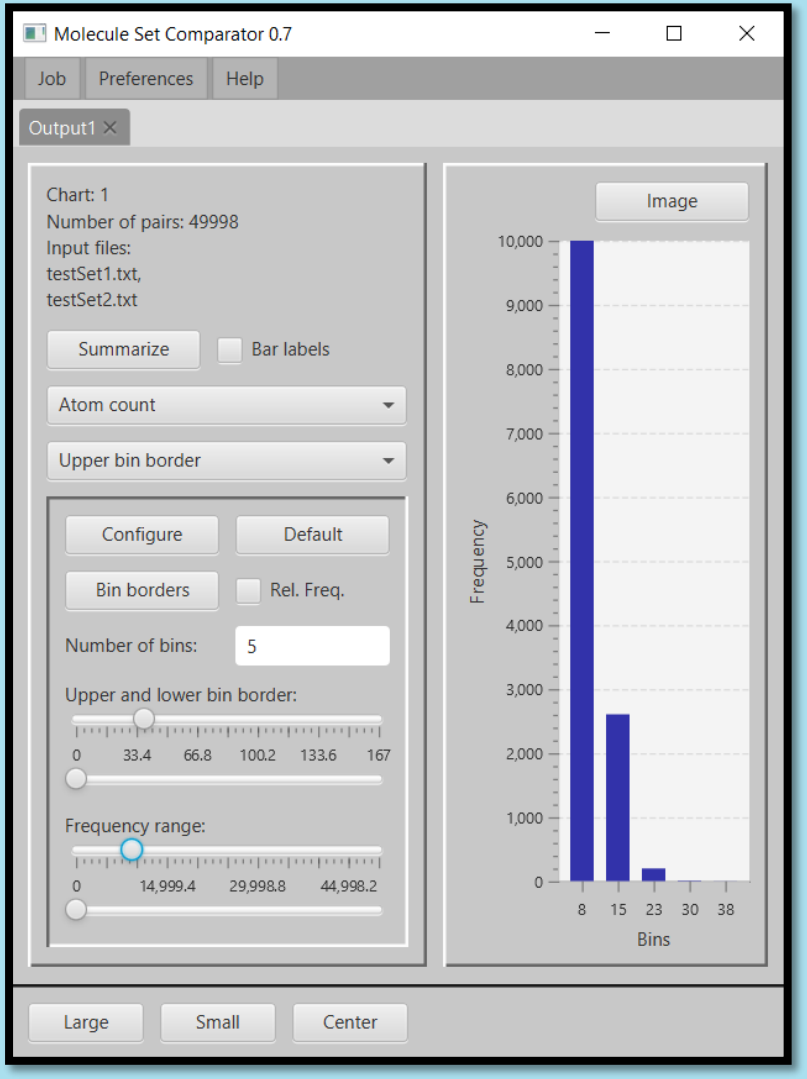


Jan-Mathis Hein
MSC – A short introduction
6

Chart configuration

- Change the upper border of the last bin by using the marked slider
- Change the upper bound of the y-axis range by using the marked slider
- The bin borders and y-axis bounds can be set explicitly by pressing the **Configure** button
- Change the number of bins to 5 by typing in the text field and hitting the enter key
- The bin borders can also be set explicitly by pressing the **Bin borders** button

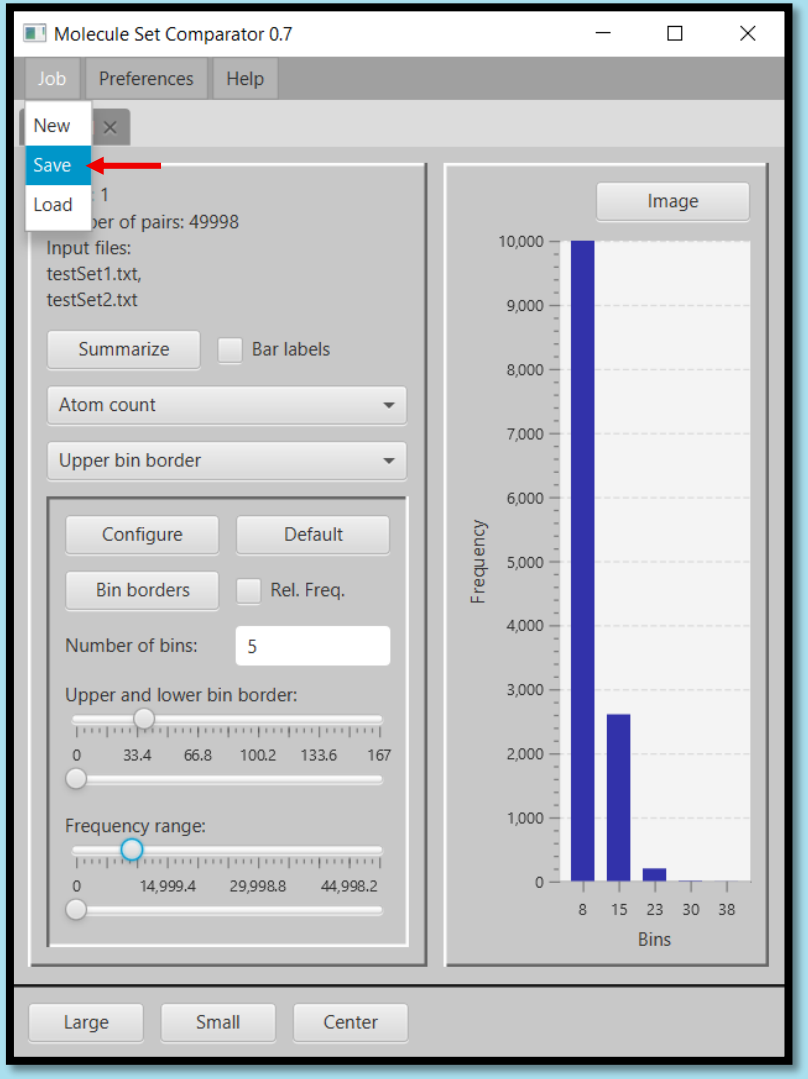


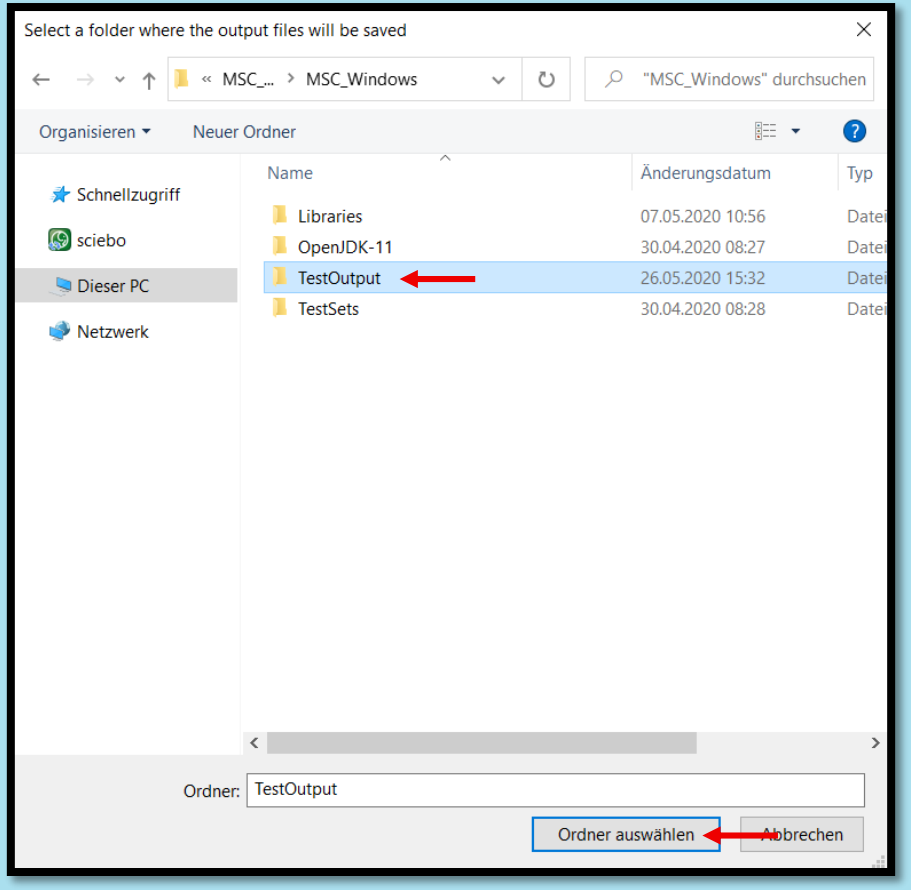


Jan-Mathis Hein
MSC – A short introduction
7

Saving the output

- Save the output by selecting **Job | Save** in the menu bar
- Select or create a folder where the output files will be saved
- The saved output can be loaded later by selecting **Job | Load** in the menu bar and selecting the corresponding folder





Jan-Mathis Hein
MSC – A short introduction
8

Changing the preferences

- Open the *"Other preferences"* dialog by choosing **Preferences | Others** in the menu bar
- Here the preferences for the number of parallel threads to be used for computations, the default number of bins, the number of molecule pairs to be saved when a job is saved and the image quality can be set