

## Journal of Medicinal Chemistry

### Creating a Molecular Formula Strings Spreadsheet

- 1. Create a new, blank spreadsheet with your preferred spreadsheet program.** (Note that there are free, high-quality alternatives to commercial spreadsheets; see, for example, Gnumeric (<http://www.gnumeric.org/>.)
- 2. List your compounds.** Type the header “Compound” at the top of the first column of the spreadsheet, and then type the name (e.g., **CORP1357**) or ID (e.g., **1a, 1b, 2**) of each new compound in your article into the first column.
- 3. Add the SMILES string for each compound in your spreadsheet.**
  - Type the header “SMILES” at the top of the second column of the spreadsheet
  - For each compound added in step 2:
    - Draw the entire molecule with your chemical drawing program of choice or on a freely available webpage. (Examples of suitable programs include ChemDraw and ChemSketch, or you may use the chemical draw program on a free webpage. See below for specific instructions.) Include the same level of detail used in the main text of the article, such as stereochemistry where known, tautomers and charge states. Omit counterions and coions. It is not necessary for the molecule to be beautifully drawn, because the graphics will not go into the spreadsheet.
    - Use the software or webpage to generate the SMILES string for your compound (see below for specific instructions). Copy and paste it into the second column of your spreadsheet, next to the correct compound name or ID.
- 4. Fill out the additional columns to provide more information about each compound from your article.** For example, a third column might list the activity of each compound against a protein target. Be sure to type in a column header which clearly references the matching data in the article and states units.
- 5. Save the spreadsheet disk as a comma-separated variable (CSV) file.** Specific instructions follow for several spreadsheet programs.
  - Microsoft Excel 2013. use File→Save As and choose “Save as type: CSV (Comma delimited) (\*.csv)”. Choose location and name of your file and click Save.
  - Apple Numbers. File→Export→CSV. (Leave Text Encoding set to Unicode (UTF-8).) Click Next and specify your desired location and filename.
  - Gnumeric. Data→Export Data→Export as CSV File.... Leave File type as Comma separated values (CSV). Choose location and name of your file and click Save.
  - LibreOffice Calc. File→Save As... and choose “Save as type Text CSV (.csv)((\*.csv).” Choose name and location of your file and click Save.

**Creating SMILES strings for your compounds.** Instructions are provided here for three different chemical drawing programs or websites.

- 1. ChemDraw.** Draw the compound in full, as described above. Select the whole molecule, and then use Edit→Copy As→SMILES to copy the required SMILES string to the clipboard. Return to your spreadsheet and Paste the SMILES string into column 2 of the row containing the compound name or identifier. Clear the draw screen and repeat for each compound.
- 2. ACD ChemSketch.** Draw the compound in full, as described above. Select the entire molecule, and use Tools→Generate→ SMILES Notation to create the SMILES string, which will appear in the draw area. Copy the SMILES string and paste it into the appropriate cell of your spreadsheet. (This may not work in LibreOffice, due to a bug. If so, try pasting into another document and then copy/paste into LibreOffice from there. Alternatively, use Gnumeric as an alternative free spreadsheet program.) Clear the draw screen and repeat for each compound.
- 3. PubChem Chemical Sketcher.** Navigate to this page: <https://pubchem.ncbi.nlm.nih.gov/upload/sketcher/index.html?smiles=&cnt=0> Draw the compound in full, as described above. The corresponding SMILES will appear in the top text window. Select and copy it (e.g., Ctrl-C in Windows), and paste it into the appropriate cell of your spreadsheet. Clear the draw screen and repeat for each compound.

**Upload the CSV document in ACS Paragon Plus at time of manuscript submission.** The document should be uploaded as “Supporting Information for Publication.”

1. These instructions were prepared by Michael K. Gilson.