# ChemLoader

A plugin for Cinema 4D

## Introduction

ChemLoader is a plugin intended to parse a chemical molecule definition file then create all necessary objects to display and render the molecule. Results such as this are possible, this is the molecule of the antibiotic bacitracin, rendered using the basic settings (no materials):

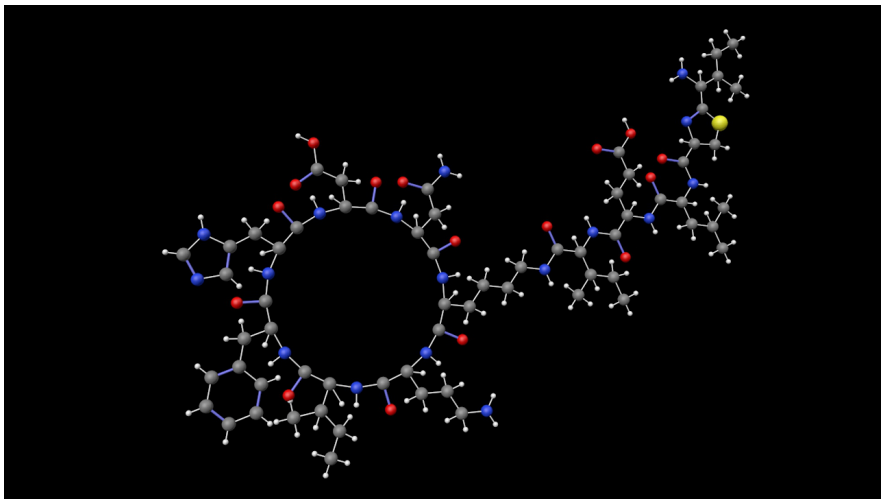

*Figure 1. Bacitracin molecule*

The created object is fully parametric and can be changed at any time. Settings are available for atom size, chemical bonds, colours, adding materials to specific elements, etc.

Important: this plugin is NOT intended to produce a physically or chemically accurate output! It is intended for the easy inclusion of molecule objects in a scene but the settings are there mostly for aesthetic reasons and shouldn't be taken as scientifically accurate representations of any molecule.

## Installation

ChemLoader requires Cinema 4D R26 or 2023. It is not available for earlier versions of Cinema.

Unzip the downloaded archive into your user folder plugins folder as with any other plugin. Note that the subfolder named 'res' contains a file named 'ptable.csv'. Do not delete or rename this file. If you do, the plugin will not work at all since this file contains values for colour, atomic radius and mass for each element.

## Using the plugin

The first thing you need is a definition file. Several are supplied with this plugin in the folder 'files'. Each file has a name indicating what chemical it is and has the extension 'sdf'. This is a Structured Data File format. Don't change the extension because ChemLoader relies on it to load the correct file format. Unfortunately there is nothing within the file to show that it really is an SDF file, so the extension must be relied upon for this.

To load the file, click the folder icon to the right of the 'File' field and choose an SDF file. Once loaded, nothing immediately happens; to display the molecule, click the 'Build Molecule' button. From now on, you only need to click that again if you load a different SDF file. The name of the new object will be the same as the name in the first line of the SDF file, unless that is empty, in which case the 'ChemLoader' name is retained. Note that the 'name' can actually be a chemical formula, a catalogue number, or even (occasionally!) the actual name of the chemical compound.

Having done this you will see the molecule in the viewport. This uses some basic settings as described below and you can change a lot of these. See the 'Settings' section for more details and the 'Reference' section for details of all the parameters in the plugin.

## General Concepts

There are a number of things to change which have a profound effect on the result. Before looking at the settings in detail, here are some important concepts to remember.

## Molecule Size

If you look at the definition file for a simple molecule such as acetone, you can see that each atom has a 3D coordinate (in some files the Z-coordinate may be zeroed to give a flat, 2D representation). However, the actual coordinates are very small, meaning that the atoms will be very close together and the overall molecule is tiny. You can adjust this with the molecule scale setting. Internally, the coordinates from the file are multiplied by 100, then again by the molecule scale (which is 1.0 by default). The more you increase the molecule scale the farther apart atoms will be from one another and the molecule will be a larger object. Note that the molecule scale does not affect the size of the atom or radius of the bond; there are separate settings for these.

## Atom Size

What size should the atoms be? Atoms are tiny (their radius is measured in picometres) so clearly they are scaled up hugely for display purposes. One thing we might want is to have the spheres representing atoms to be relative in size to another as they physically are. For example, if in reality some atom is 10 times as big as a hydrogen atom we might want the sphere representing that atom be 10 times as big as the sphere representing a hydrogen atom.

It turns out however that this doesn't necessarily produce the most pleasing results in an image. For example, we are accustomed to thinking of a hydrogen atom as being much smaller than a carbon atom. However, this isn't the case. A carbon atom is only just over twice as big (that is, has just over twice the radius) of a hydrogen atom. We might also expect that elements with higher atomic numbers are larger than those with lower numbers, but that isn't so either. For example, a carbon atom (atomic number 6) has a larger radius than nitrogen, oxygen or fluorine atoms (atomic numbers 7, 8, and 9 respectively).

Another problem is that there are different ways of measuring the radius of atoms, which give different results, and for a surprisingly large number of elements the radius is an estimate made by extrapolating between atoms with lower and higher atomic numbers. ChemLoader uses the covalent radius of the atoms from this source as there seems to be more complete data using this method than others.

If atomic radius doesn't give the result we want, what else could we use? There is more and better data for atomic mass compared to radius and has the advantage of giving results closer to what we might expect. Mass usually increases in line with atomic number (though there are a small number of exceptions to this) and the differences in mass are more marked. For example, a carbon atom has about 12 times the mass of a hydrogen atom. A possible downside of this method is that we lose physical accuracy in favour of appearance.

Yet another method is to base the size entirely on atomic number, taking a base size then scaling it up using the atomic number for each atom. This gives results that are similar to the mass but with more consistency; but again, physical accuracy is lost. Finally, we could simply abandon all pretence at accuracy and give all atoms the same size; this actually produces better (more visually pleasing) results than might at first be expected.

ChemLoader lets you choose any of these methods; the default is to use a constant radius. The plugin's default settings will produce the following results for each method of setting the size of atoms (the molecule is

acetone, C3H6O):

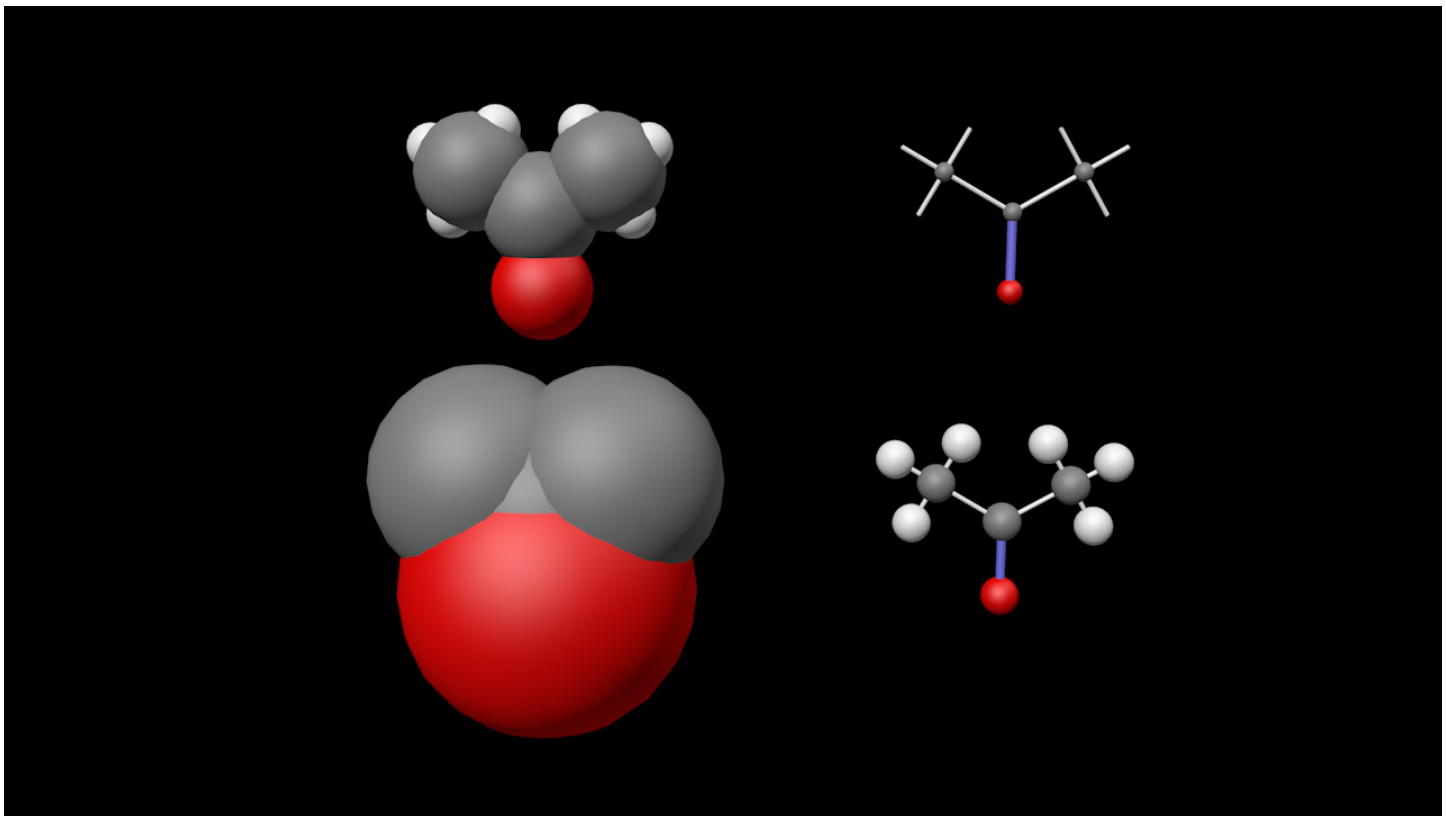

*Figure 2. The four atom size modes.*

Clearly, these could do with some adjustment to make the results look better, except for the constant radius mode. ChemLoader uses constant radius as its default mode so as to produce an acceptable result with little or no adjustment. Changing the various settings to give a better appearance results in this:

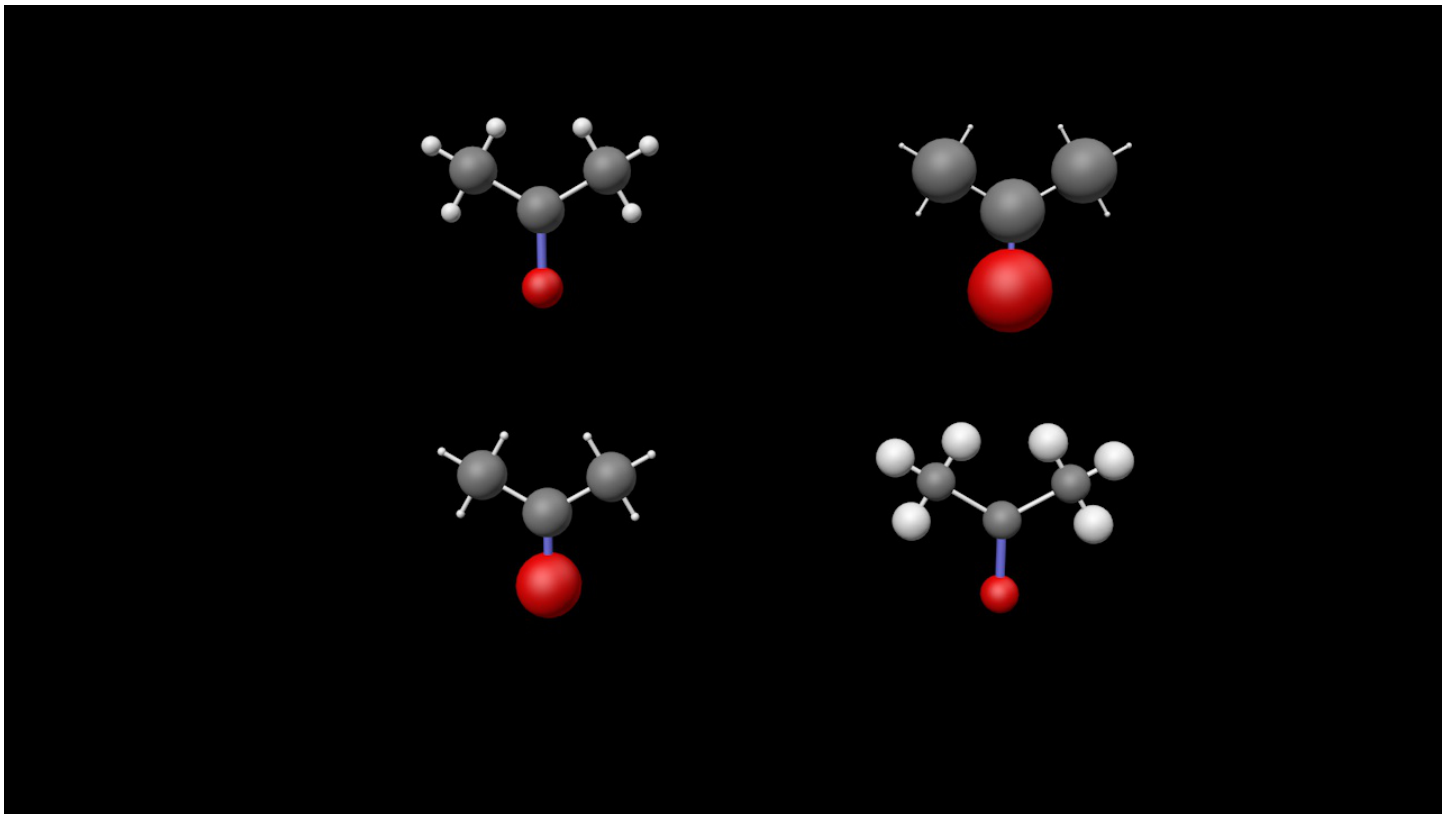

*Figure 3. Adjusted atom size modes*

Now the results look much better. They all look slightly different though; choose whichever mode seems to you to be the best for your scene. In general, constant radius and using the atomic radius probably give the

best results.

## Bonds

ChemLoader can display bonds of different kinds (single, double, triple or aromatic) with different size and colour. Again, we're left with the question of how long a bond should be. This can actually be calculated quite easily, because the bond length can be estimated as the sum of the atomic radius for the two connected atoms. Although this is feasible it's not aesthetically very pleasing and it doesn't allow the generation of the ball-and-stick molecule figures that are so commonly seen. ChemLoader will simply show bonds as drawn between the centres of the two atoms. The more you increase the molecule scale the farther apart atoms will be from one another, and the longer the bonds will be. This image shows the same molecule with a scale of 100 (the default) on the left and 200 on the right:

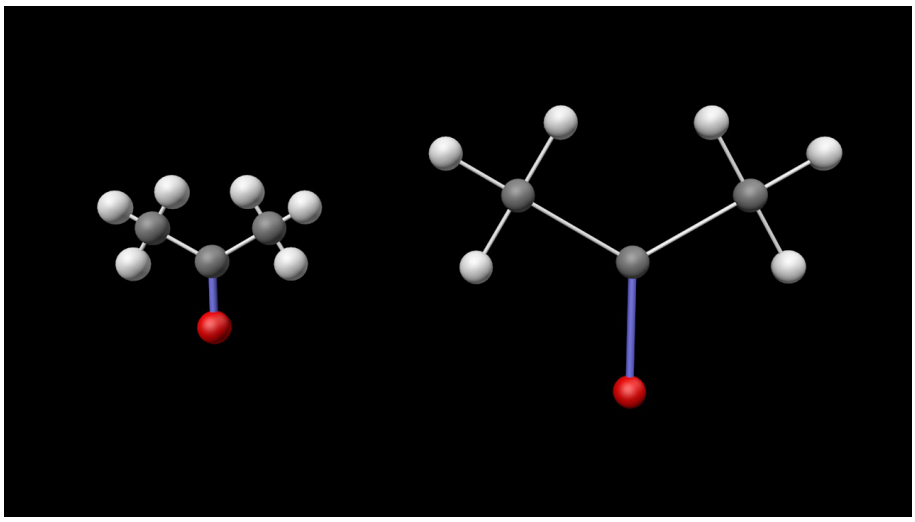

*Figure 4. Effect of changing the molecule scale factor*

## Colours

These are purely for display and have no physical equivalent. First, you can set the atom colours. ChemLoader contains details (colour, radius, mass) for the first 94 elements in the periodic table and each one has a different colour. The colour scheme used here is the CPK colour scheme (https://en.wikipedia.org/wiki/CPK_coloring) which is a semi-standardised set of colours widely used in molecular visualisation software. There are several different versions of this set, but the one used here is that used by Jmol (https://jmol.sourceforge.net/jscolors/) a popular standalone program used to visualise molecules.

Unfortunately, since there are 94 different colours it's impractical to edit this scheme within the plugin. You can edit it yourself using a text editor or spreadsheet; details are given in Appendix 1. In fact, the vast majority of molecules you will visualise will contain only four elements, hydrogen, carbon, oxygen and nitrogen. Therefore you can select a scheme which lets you edit these four colours within the plugin so you can set whatever colours you like. There are several preset four-colour schemes selectable.

Do remember however, that these colours are used to set the object colour of the various objects making up the atoms and bonds. The colours will render, but they only give a basic diffuse colour. If you want anything else, such as transparency, reflection, etc. you will need to add one or more materials. This is discussed below.

As well as setting the atom colours you can also set separate colours for the bonds, and for the nucleus if you add that. These are not part of any colour scheme so you can change them to whatever you like.

## File Formats

ChemLoader is designed to use molecule definition files in the SDF format. This is a development of an earlier format, 'MOL'. The only difference which concerns us is that a MOL file contains one molecule definition per

file whereas SDF files can contain multiple molecules each in the MOL format. However, if you have an SDF file containing more than one molecule definition, please note that in the current version of ChemLoader only the first definition in the file will be used; all others are ignored. It seems that MOL files are less common than SDF but ChemLoader can use MOL files as well as SDF.

A more recent format is 'MOL2' which is a more complex format containing a good deal more information. Perhaps because they are more complex to write, MOL2 files are less common than SDF/MOL and ChemLoader does not at present load these files. If you find a good source of MOL2 files, let me know and I may add MOL2 support.

There are a number of other formats but ChemLoader doesn't handle them. Perhaps the commonest is the SMILES format; this produces a very compact definition but it is so limited in content that it isn't worth supporting them in this plugin.

# Reference

This section describes what each parameter in the plugin interface does and how/when to use it.

## Object Tab

The interface looks like this:

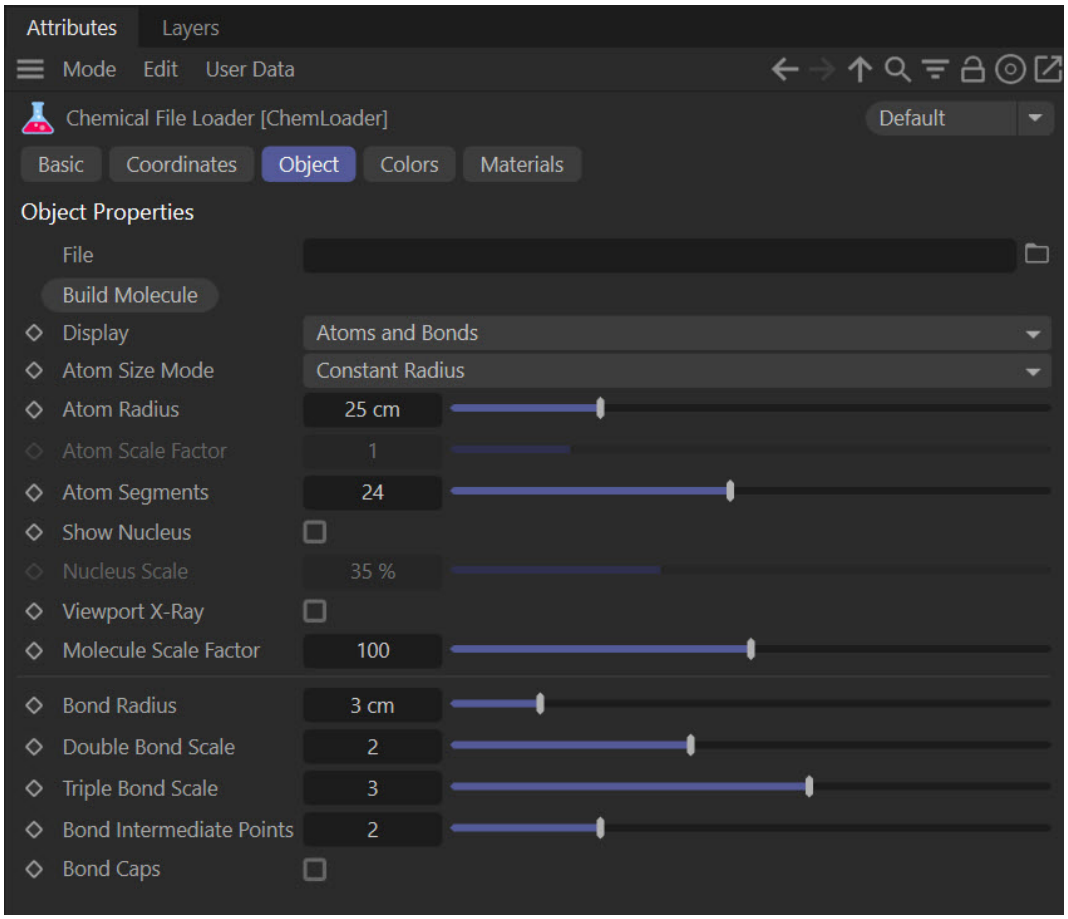


*Figure 5. Object tab interface*

### *File*

This is the molecule definition file to load, in SDF or MOL format. Once loaded, the molecule isn't immediately displayed – click the 'Build Molecule' button to actually generate the molecule. At any point, should you need to do so, you can click this to rebuild the molecule but that shouldn't be necessary until you load a different definition file.

### *Build Molecule*

Click this button to build the molecule. If there is no file in the 'File' field, or there is but it isn't a recognised format, nothing will happen apart from an error message in the case of an incorrect format.

### *Display*

This menu lets you control which components you want to display when the molecule is built. The options are:

- Atoms and Bonds: both atoms and bonds are shown; this is the default setting
- Atoms Only: only atoms are shown, no bonds
- Bonds Only: only bonds are shown, no atoms

### *Atom Size Mode*

This is explained in the 'Atom Size' section in 'General Concepts' above. This menu lets you choose which mode to use, but you can alter this at any time. The options are:

- Use Atomic Radius: the atom spheres have a radius derived from the covalent atomic radius of the atom. The atoms are quite large with this mode and will often overlap (itself quite a nice effect) but you can reduce the size by reducing the 'Atom Scale Factor' setting, or you can separate them by increasing the molecule scale factor.
- Use Atomic Mass: the same as using the atomic radius, but the sphere radius is derived from the physical atomic mass. The difference in size between atoms is much more marked than when using the atomic radius and you may need to increase the 'Atom Scale Factor' value as hydrogen atoms especially will be very small.
- Scale by Atomic Number: this takes the 'Atom Radius' setting and scales it up by the atomic number of the element, with hydrogen = 1, carbon = 6, etc. This produces similar results to using atomic mass but the size differences between elements are even more marked. You will almost certainly need to reduce the 'Atom Radius' setting to a much smaller value than the default setting as that will lead to very large atoms especially with high atomic numbers.
- Constant Radius: all atoms have the same size, which is found in the 'Atom Radius' setting.

### *Atom Radius*

Only used when 'Atom Size Mode' is set to 'Scale by Atomic Number' or 'Constant Radius'. It is the radius of the atom sphere used either as it is or scaled up by the atomic number.

### *Atom Scale Factor*

Only used when the 'Atom Size Mode' is set to 'Use Atomic Radius' or 'Use Atomic Mass'. The calculated radius of the atom sphere is then scaled up or down by this factor.

### *Atom Segments*

The number of segments in the atom sphere (and the nucleus sphere if using that). By default it is set to 24, which is high but good for close-ups of single molecules. If you have a lot of molecules in the scene they will probably be quite small and you can reduce the polygon count considerably by lowering this value, which could be important with hundred of atoms in a cloner or generated from a particle emitter.

### *Show Nucleus*

If this is turned on, a smaller sphere representing the atomic nucleus will be shown at the centre of the atom. There are a few things to remember about this. There is no point at all in turning this on unless you are

intending to add a transparent material to the atoms. If you don't, the nucleus will never be visible and is just wasted polygons. You can see the nucleus in the viewport if you turn on 'Viewport X-Ray' but again, this won't render so you won't see the nucleus in the final render anyway. Also note that the x-ray setting is never turned on for the nucleus spheres since if it was, you wouldn't see them (an object inside another object with x-ray enabled will be invisible if it also has x-ray enabled).

### *Nucleus Scale*

By default the nucleus is set to have 35% the radius of the atom (however that is determined). You can change that with this setting.

### *Viewport X-Ray*

If this switch is turned on, the 'X-Ray' setting in the basic object data for the sphere and bonds will be turned on. This is for convenience only since it doesn't render (unless you select the 'Viewport Renderer' in the render settings and render to the picture viewer).

### *Molecule Scale Factor*

You can use this setting to increase or decrease the overall size of the molecule as discussed in 'Molecule Size' in the 'General Concepts' section. Note that this is not dependent on the atom size mode and will not change the size of the atoms, only how far apart they are from each other.

### *Bond Radius*

Chemical bonds are created using a Sweep object and a Circle spline as the profile. The radius of the circle is set by this  parameter.

### *Double Bond Scale*

Double bonds are shown differently from single bonds in two ways: a different colour and a different radius. The colour is defined in the 'Colors' tab but the size is set by scaling the 'Bond Radius' by this value. The default setting of 2 will double the radius for double bonds.

Note that aromatic bonds will have the same size as double bonds but use a different colour.

### *Triple Bond Scale*

The same as for the double bond scale but for triple bonds. The default setting of 3 will give a bond radius of three times the 'Bond Radius' setting. They also have their own colour.

Note:  if you prefer that the bonds of all types have the same size, simply set these two scale values to 1. Then all types of bond will have the same radius.

### *Bond Intermediate Points*

This is the same as the intermediate points setting found in Cinema 4D splines. It is used for the circle spline in the Sweep object used to generate a bond. The default setting is 2, which is low but sufficient in most cases since the bond radius is usually small. If you have a very large bond radius or a bond is shown in close-up you might need to increase this value to avoid faceting.

### *Bond Caps*

By default, caps on the Sweep object used for bonds are turned off, to keep the polygon count down. If you want to see the caps – for example, if you only show the bonds and not the atoms – you can turn this switch on.

## Colors Tab

All colours used by the plugin are set in this tab. Remember that these only set the object's basic object colour; for more features you will have to use a material. The interface looks like this:

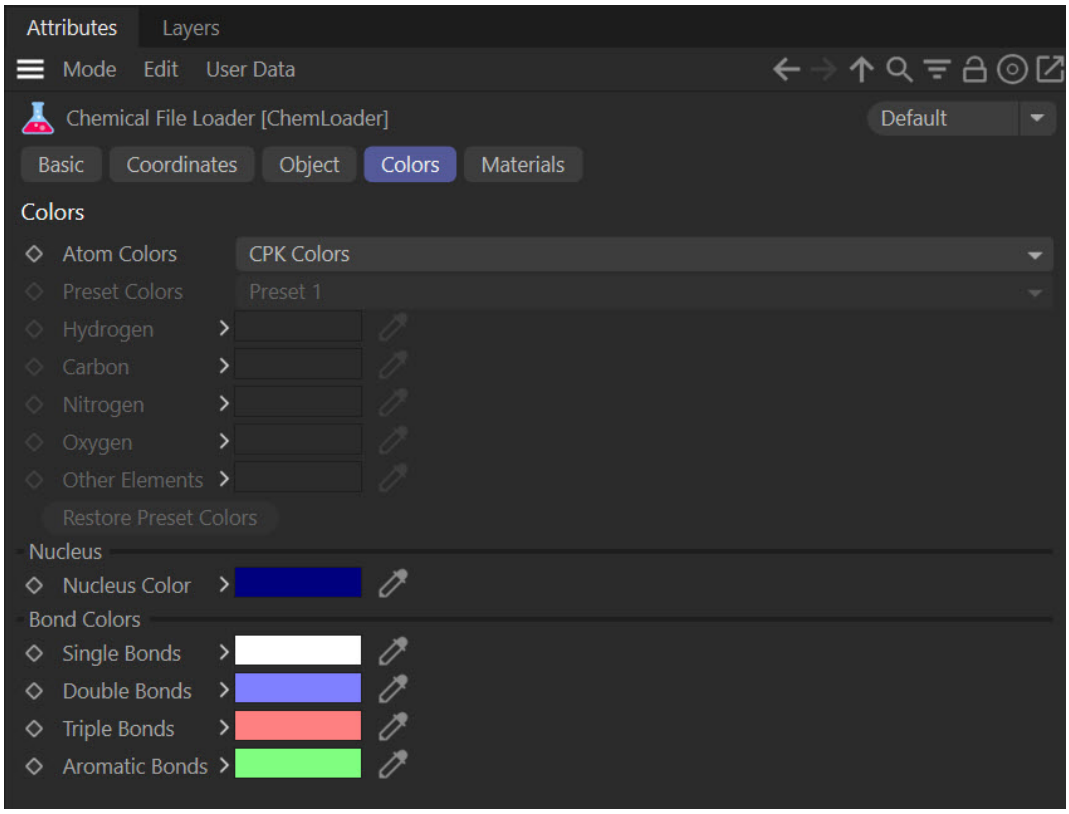


*Figure 6. Colors tab interface*

### *Atom Colors*

This menu has two entries which let you choose the colour mode. These modes are discussed in 'Colours' in the 'General Concepts' section above. The menu options are:

- CPK Colors: use the semi-standardised colour scheme used in other software. These colours cannot be edited (but see Appendix 1 for more information)
- H-C-N-O: this option lets you set the colours for the four commonest elements likely to be seen in the molecules you generate, namely hydrogen, carbon, nitrogen and oxygen. These colours are editable, and there is a fifth colour which is assigned to the atoms of all other elements.

### *Preset Colors*

This is a list of several colour presets, available if you select the 'H-C-O-N' colour mode. You can select any of these presets and edit them as required. Note that the presets are hard-coded into the plugin and if you edit the colours those changes will only apply to the scene file you are working on.

### *Colours (Hydrogen, Carbon, Nitrogen, Oxygen, and Other Elements)*

There are five colour controls which set the atom colours. These are only available in the 'H-C-N-O' colour mode. You can set the colours for the four named elements but atoms of all other elements receive the 'Other Elements' colour.

### *Restore Preset*

If you have edited the colours from a preset, you can restore the original colours with this button. It only

affects the current preset colours. You will be asked to confirm this before the colours are restored.

### *Nucleus Color*

The colour used by the atom's nucleus. Only used if the 'Show Nucleus' switch is on.

### *Bond Colors*

These four colours are used to colour the different types of bond in the molecule. You can set them to be all the same colour if you don't want different-coloured bonds.

## Materials Tab

Although ChemLoader has extensive support for changing colours of atoms and bonds, often you need more than a simple colour, such as transparency or reflection. For that, you need to add a material to the atoms, which you do in this tab.

If you want to add a material, and you want the same material to be used for everything – atoms, bonds, and nucleus – you don't need this tab. Simply drop the material onto the ChemLoader object and all its constituent objects will get the same material. Usually though, you will want to add different materials for different elements and others for the bonds.

The tab looks like this before materials are added:

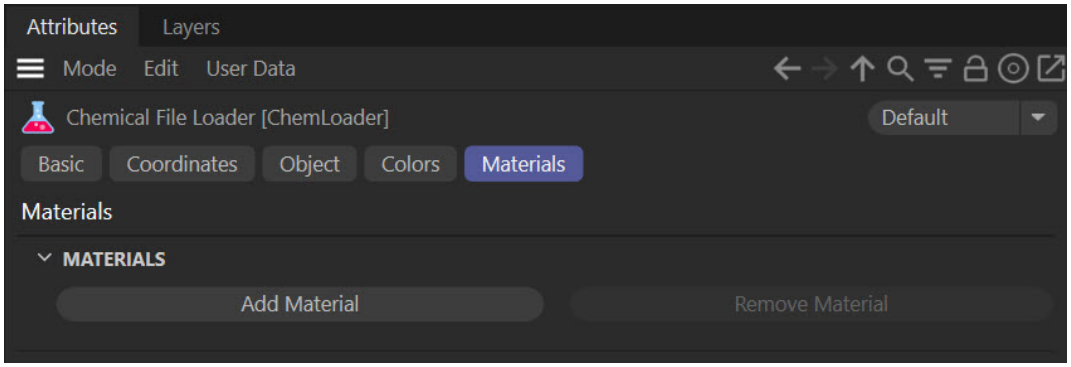


*Figure 7. Materials tab interface*

### *Material & Atomic Symbol*

To add a material, you need to create one first. Then click the 'Add Material' button. The interface will change to as shown in Figure 8.

To assign the material, drag and drop the material into the 'Material' link field. Now you need to tell ChemLoader which element should receive this material. To do that, enter the element's atomic symbol into the 'Atomic Symbol' field, so you would enter C for carbon, N for nitrogen, Mg for magnesium and so on. You can find these symbols in any periodic table of elements, such as https://pubchem.ncbi.nlm.nih.gov/periodic-table/. The material will be assigned to all atoms of that element.

Bonds, of course, do not have atomic symbols. Therefore you can enter one of four special symbols to denote a material to be assigned to a bond. These are:

- BD1: single bonds
- BD2: double bonds
- BD3: triple bonds
- BD4: aromatic bonds

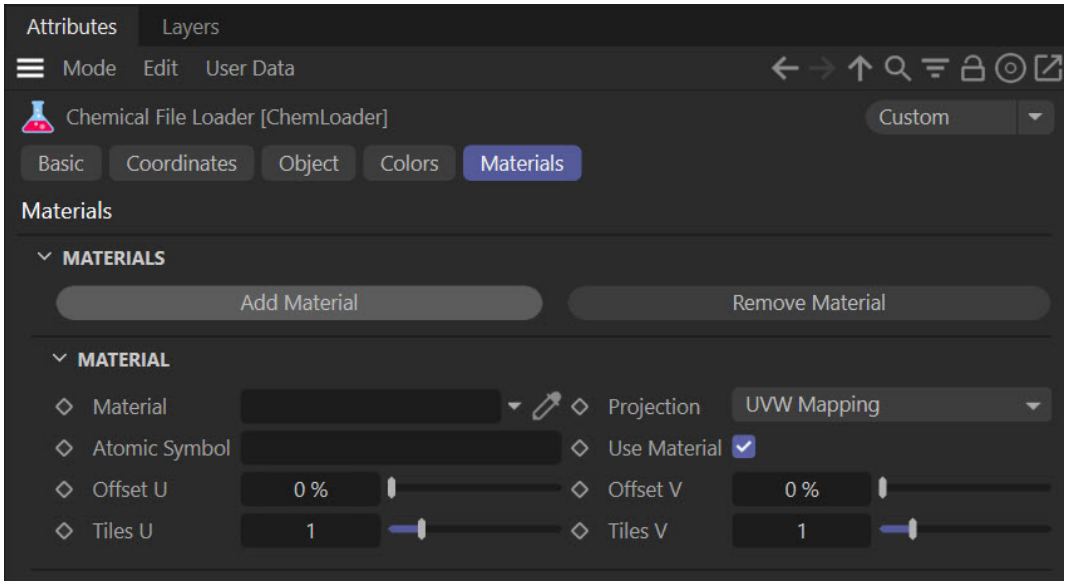


*Figure 8. Material entry added*

The nucleus, if shown, also has no atomic symbol. You can only assign one material to all nuclei, no matter what element it is. To do that, enter 'NUC' in the atomic symbol field (without the apostrophes).

### *Same material assigned to more than one element*

This is straightforward. Simply add a material for as many atoms as you need, drag the same material into all the link fields, and enter one atomic symbol into each symbol field. You cannot add multiple symbols to a field, and if you do that the material will be ignored.

### *Different materials assigned to the same element*

In this screenshot:

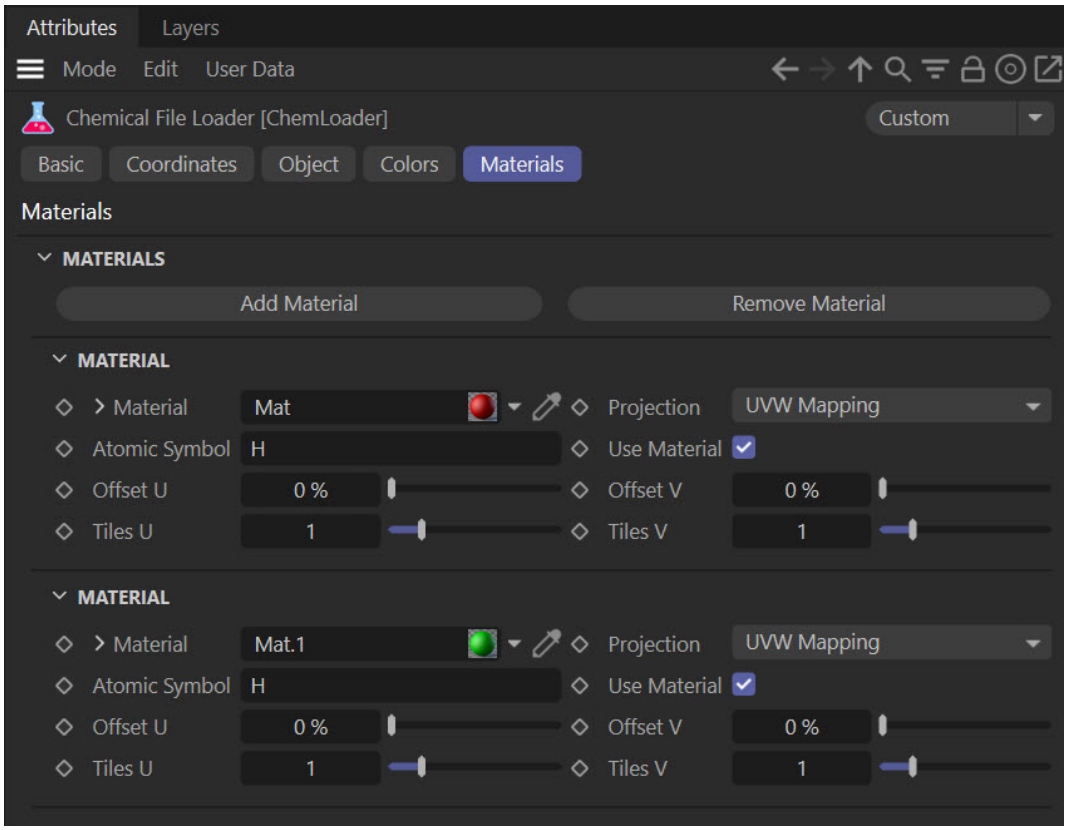


*Figure 9. Different materials assigned to same element*

Two different materials, red and green, are assigned to hydrogen atoms. Which one is used? Internally, ChemLoader builds a list of all materials and the elements they are assigned to. When building the molecule, it looks for the first material assigned to the element and stops if it find one. So in this case, since the red material is the first one in the list the hydrogen atoms all receive the red material.

If you wanted them to have the green material, you could drag the green material into the link field currently occupied by the red material. That would work fine, but there is a better way. Simply turn off (uncheck) the 'Use Material' switch for the red material. Doing that means that the red material won't be included in the internal materials list so will no longer be available, and all the hydrogen atoms will be green. You can add multiple materials and swap between them using this technique.

### *Use Materials*

Each material assignment has this switch, which is on by default. Turn it off to stop using that material without having to remove it.

### *Remove Material*

This button does the opposite of the 'Add Material' button – it removes the last entry in the list. For example, if you have a list of five materials, clicking this button will remove the last one (the lowest one in the list). Clearly, if what you wanted was to remove the first material in the list, you would need to remove all those below it first, which is hardly user-friendly. For this reason, if you find that you don't need a material after all, and it's not the last one in the list, rather than removing it you can turn off the 'Use Material' switch instead.

### *Other Settings*

The remaining settings – OffsetU, OffsetV, TilesU, TilesV and Projection – are the same as you see in the Cinema 4D texture tag. Use these settings in the same way. Note that only a subset of projection types is currently supported.

### *Displaying atomic symbols*

Sometimes, you might want to display the atomic symbol of the element on each atom. There are two ways to do this. It would be possible to generate text objects, such as MoText, and attach one to each atom. This causes many additional complications, such as keeping the size correct, changing the position and rotation, and adding numerous additional complex objects to the molecule. For these reasons, ChemLoader doesn't do this and the way to add atomic symbols is to add a material to each atom with a bitmap to display the symbol. The results might look as shown in Figure 10.

The easiest way to do this is create bitmap overlays and use them in the desired channel. ChemLoader comes with six overlays with black text on white backgrounds and the same again but with reversed colours. The six were chosen as the six most common elements likely to be found in molecule definitions (hydrogen, carbon, nitrogen, oxygen, sulphur and phosphorus). If you need others, you can easily make them in an image editor. ChemLoader uses 200x200 pixel bitmaps but you can choose whatever size, font and colours work for you.

This was achieved using the Color channel in the material mixing a bitmap with the colour (Figure 12).

Note that when using the black background overlays, which you might need to do with very dark colours in the material, you would probably set ‘Mix Mode’ to ‘Add’.

Figure 12 shows the material settings used in ChemLoader.

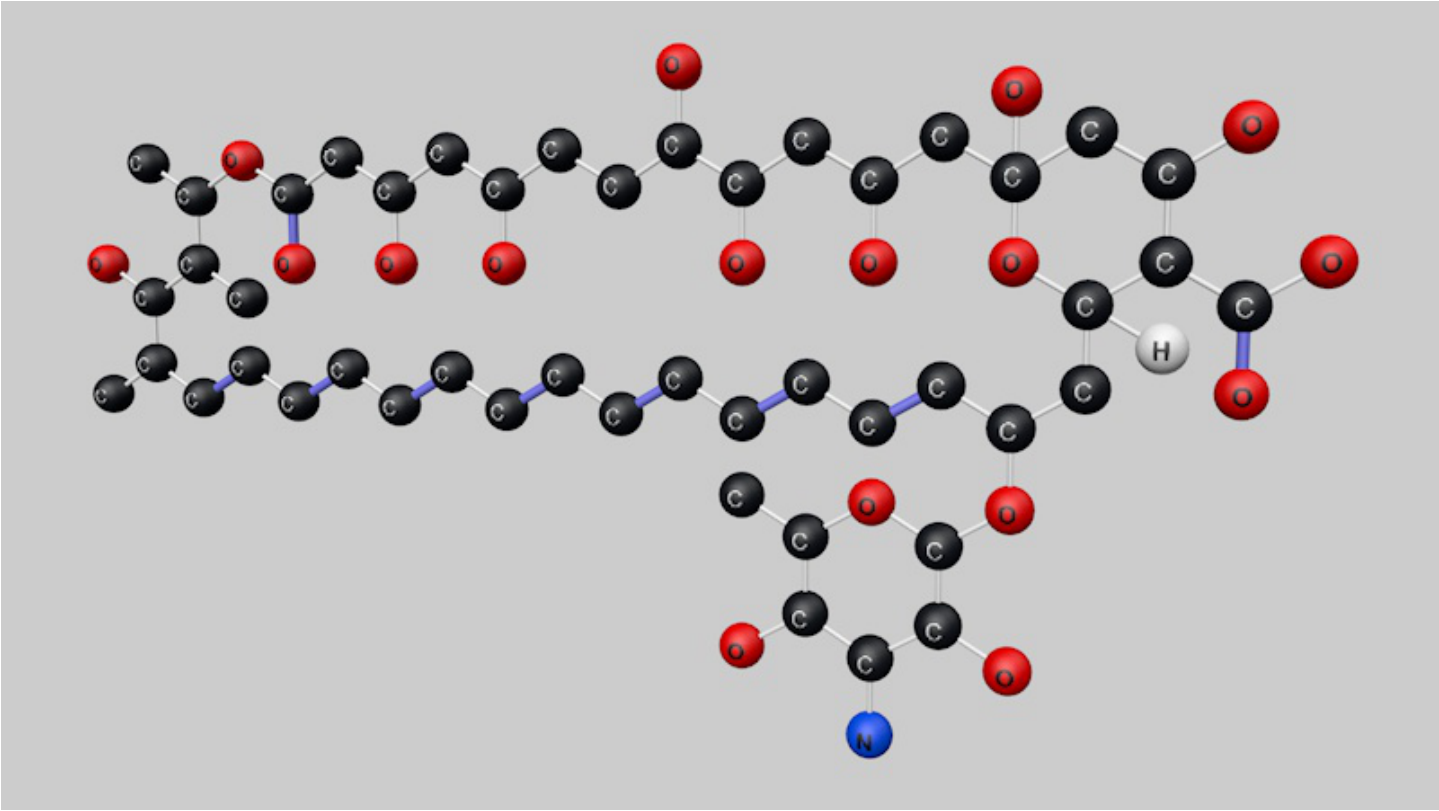

*Figure 10. Molecule with atomic symbols*

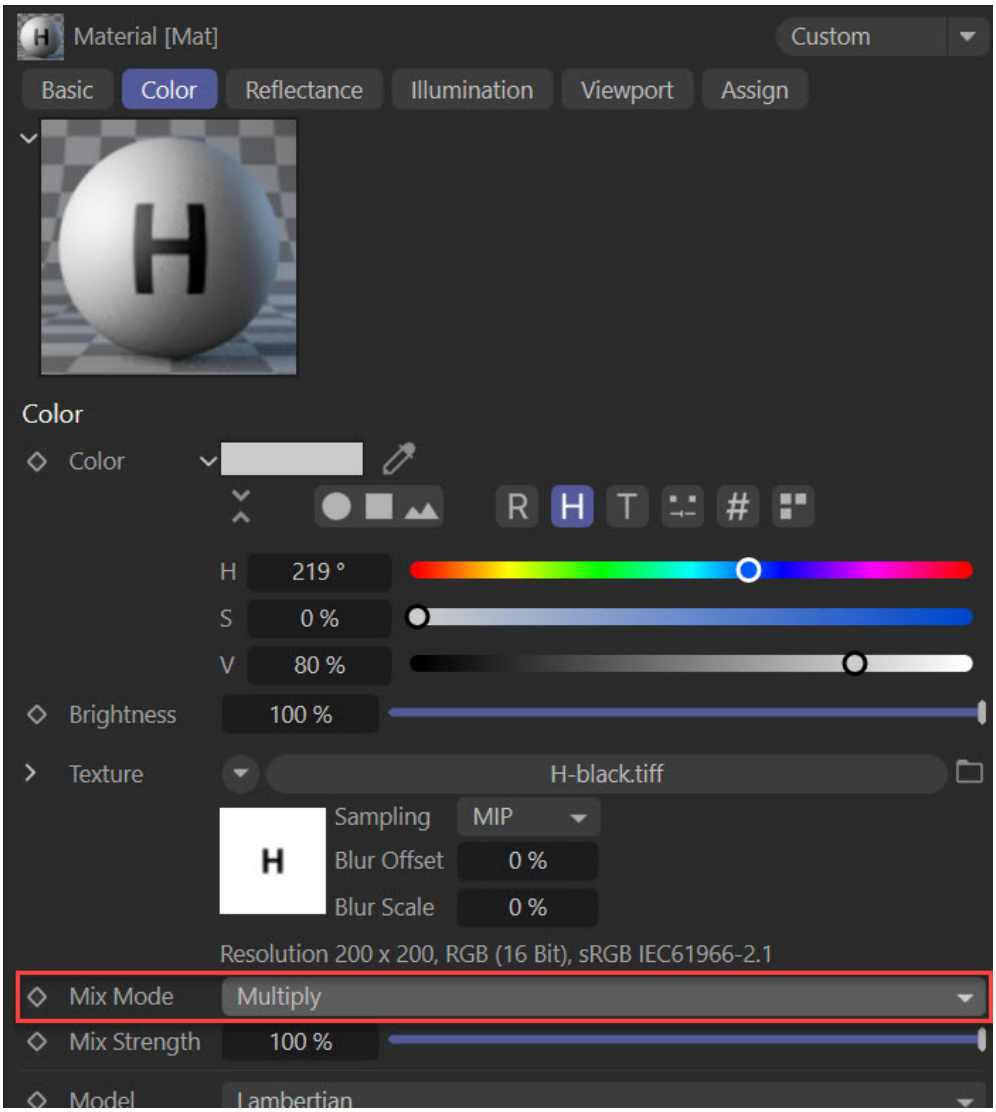


*Figure 11. Using a bitmap overlay to show atomic symbols*

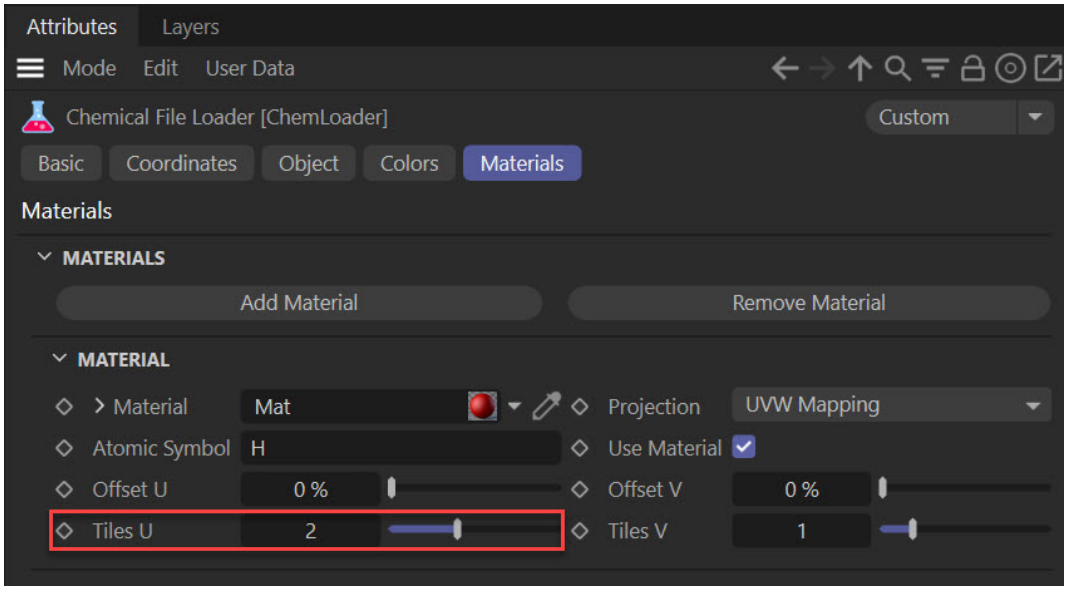


*Figure 12. ChemLoader material settings for atomic symbol display*

# Limitations

This plugin does have some limitations, which are listed here.

### File formats

Only SDF and MOL files are currently supported. If an SDF file contains more than one molecule definition, only the first one is loaded and built. If you need a molecule other than the first one, you will need to create your own SDF file from the original file. See Appendix 2 for details.

### 2D vs. 3D

The atoms of a molecule have coordinates to specify the location of the atom. In many SDF/MOL files these are 3D coordinates but some are 2D, and the Z axis component of the coordinate is zeroed. The acetone.sdf file supplied with the plugin is a 2D file; most of the others are 3D files. If you want to know if a file is 2D or 3D, probably the easiest way to find out is to load and build it, though in some cases the difference between a 2D and 3D representation may be small.

Alternatively, you can look at the file in a text editor. Look at the second line in the file acetone.sdf, which will look something like this:

```
APtclcactv10252206142D 0   0.00000     0.00000
```

The 21st and 22nd characters will be '2D' (as shown above) or '3D'. However, this line is optional and may be absent, so then you will have to load it and see what it looks like.

### No hydrogen atoms

It's a chemistry convention that hydrogen atoms are sometime not included in chemical definition files, to keep the file size down. You can still load such a file but it won't show any hydrogen atoms, which may or may not be a good thing, depending on what you want to show. It’s true that lots of hydrogen atoms can really get in the way with large molecules.

### Different files...different results

If you download an .sdf file for the same molecule from different sources, they may give very different results even though the chemical formula and bonds are identical. For example, Figure 10 and Figure 11 show the

antifungal agent amphotericin B using files from two different sources:

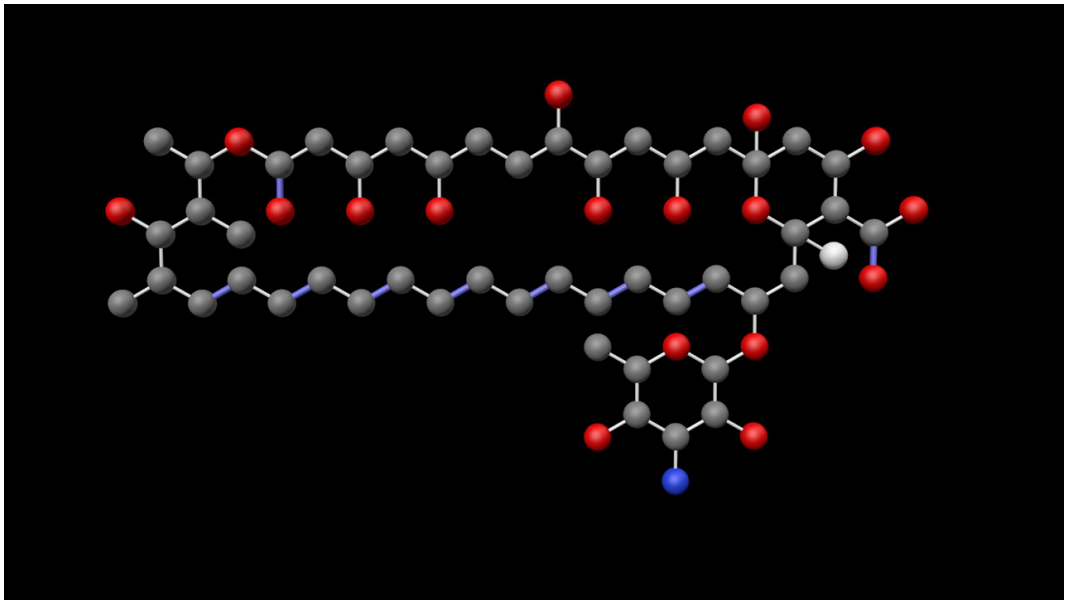

*Figure 13. Amphotericin B (source: CheEBI)*

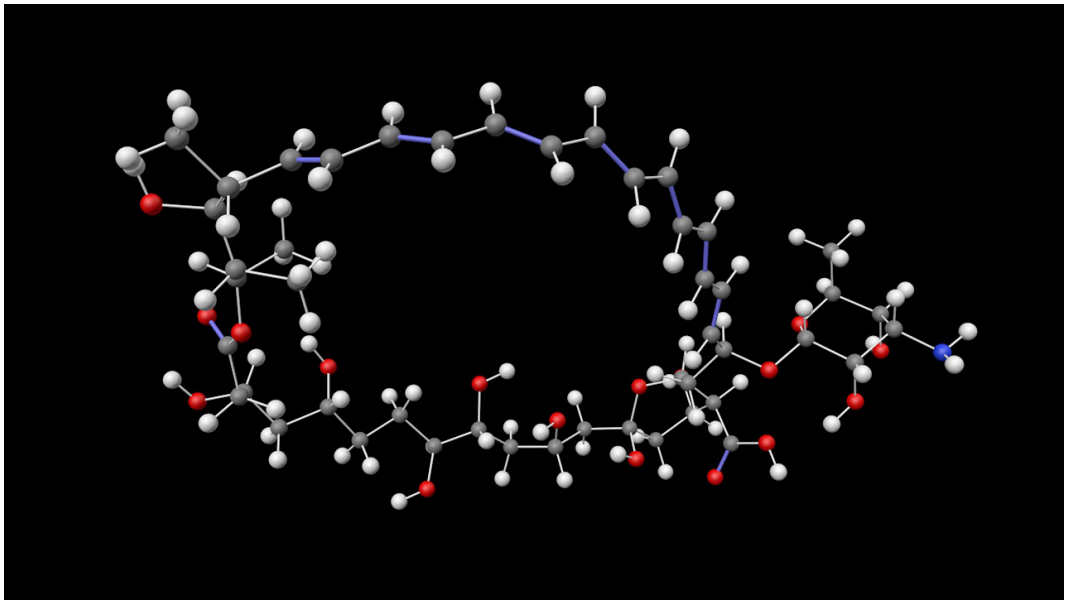

*Figure 14. Amphotericin B (source: Mol Instincts)*

As you see, they appear to be wildly different but closer inspection shows that they are the same molecule. One is a very conventional 2D structure without hydrogen atoms, the other shows the hydrogen atoms and is in 3D. The moral is that, if you don't like the appearance generated by a particular definition file, see if you can find another one!

## Downloading more definition files

There are many sources for SDF and/or MOL files on the web. You can search in Google but there are four sites which are especially useful.

### ChEBI (Chemical Entities of Biological Interest)

The URL is https://www.ebi.ac.uk/chebi/init.do. All you do is type the compound you want into the search field and you will be sent to a page which may have numerous compounds which fit the search criteria. For example, if you type 'penicillin' you'll get 78 possible compounds and then you'll need to choose which one

you want. That will take you to the molecule page with download links for the definition files (highlighted in this screenshot):

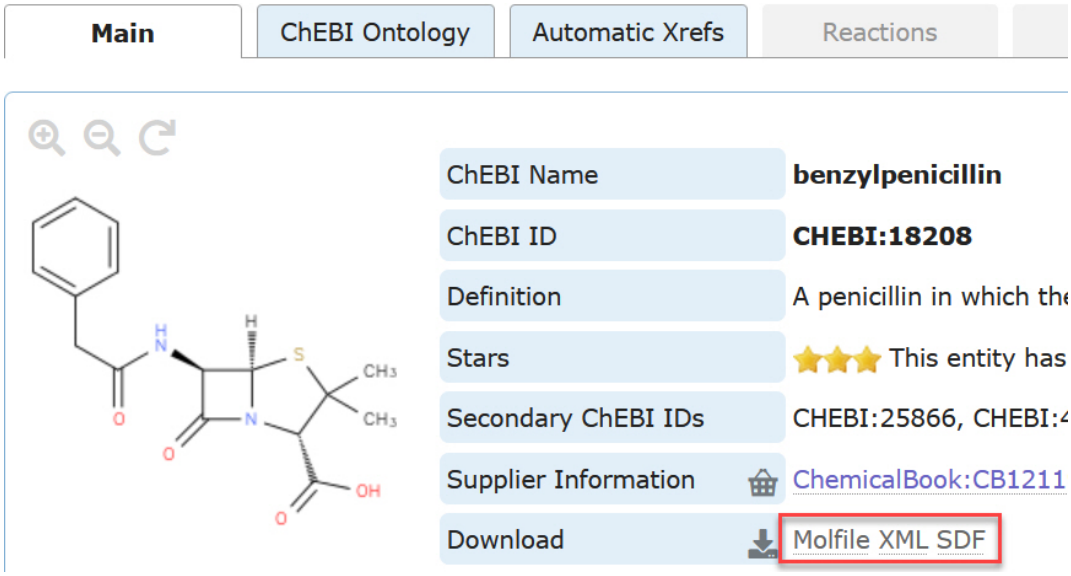


*Figure 15. Downloading from ChEBI*

Choose the Molfile or SDF option. If you download both, you will see that they are identical except that the SDF file has some additional information after the basic molecule data.

## Mol Instincts

Another large database offering SDF files. See the home page here: [https://www.molinstincts.com/sdf-mol-file/](https://www.molinstincts.com/sdf-mol-file/). As you can see, they recommend using Google for searching their site. If you do that, using the method shown on the above page, you see something like this:


*Figure 16. Downloading from Mol Instincts*

Then you can click the link to download the file.

## ZINC database

A massive database of which ZINC-20 is the latest version. If you visit the home page https://zinc20.docking.org/ there is a menu bar and one entry is 'Substances'. Click that, then in the next page enter the substance you want into the 'Search for Substances' field and hit 'Search'. If you enter 'penicillin' you get a page with several entries. Click the one you want and you see this:

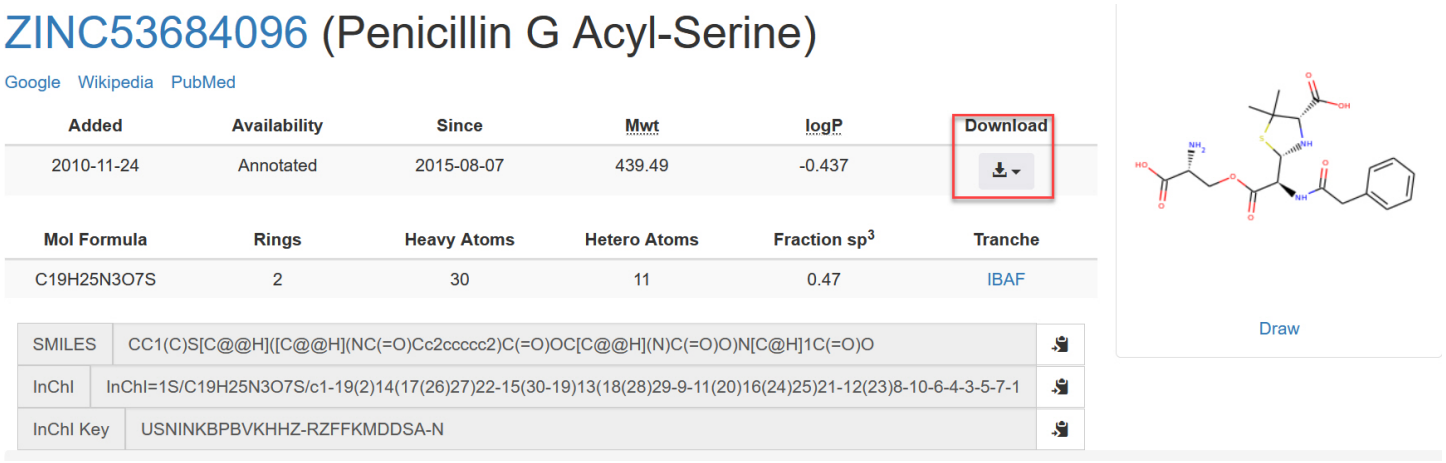


*Figure 17. Downloading from ZINC 20*

The 'Download' button brings up a menu of download formats. Choose 'SDF' and the file will download.

## Chemical Identifier Resolver

A service provided by the NIH in the USA. The home page is https://cactus.nci.nih.gov/chemical/structure. Enter the substance you want into the 'Structure Identifier' field, then change the 'Convert to:' menu to 'SD File' and click 'Submit'. *Don't* click the 'Structure' button, it will show you the structure instead.

When you click 'Submit' a window will open that looks like this:

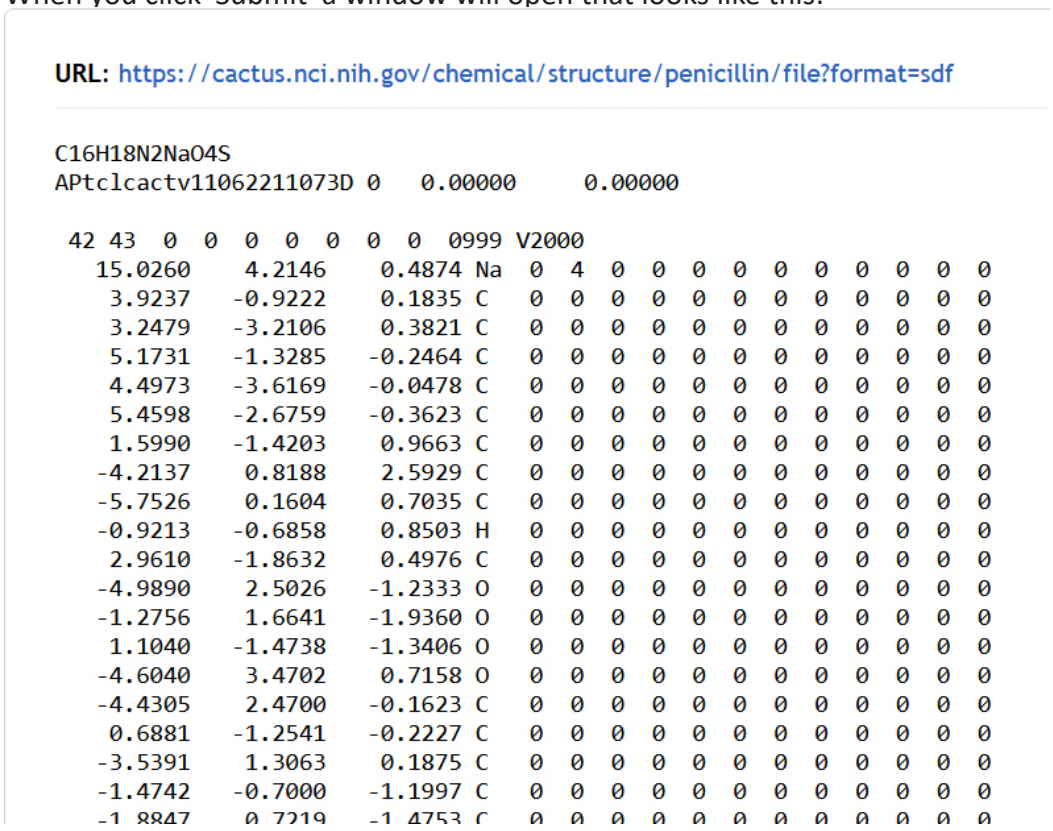


*Figure 18. Downloading from Chemical Identifier Resolver*

Finally, click the URL at the top of that window and a web page opens with the SDF data in it. Save that page from your web browser making sure it has the extension SDF.

There are lots of other sources on the web but these four between them should give you all the files you are ever likely to need. Remember that files from different sources may give different results – they may be 2D or3D, they may or may not show the hydrogen atoms, etc., so if you don't like one, try another.

# Summary

ChemLoader makes it easy to create visualisations of chemical molecules – provided you can find the definition file to download! If you have any suggestions for additions or improvements, or find a bug, please let me know. You can contact me through my website at https://www.microbion.co.uk/contact.htm.

Steve Pedler
November 11 2022

# Acknowledgements

Several example SDF files are included with the plugin for you to try. These were found on the following sources and acknowledgement is duly given here:

Chemical Identifier Resolver, https://cactus.nci.nih.gov/chemical/structure

ChEBI (Chemical Entities of Biological Interest), https://www.ebi.ac.uk/chebi/init.do

Mol Instincts, https://www.molinstincts.com/sdf-mol-file/

Test tube icon used in the plugin from Icons8 with thanks.

# Appendix 1

### Editing the file 'ptable.csv'

This file is vital for ChemLoader to work and must be present. It is a text file which has two header lines and then one line for each of 94 elements. Each line contains four data items, separated by commas:

```
ChemLoader v1.0 Periodic table data,,,
Symbol,Colour,Radius,Mass
H,FFFFFF,31,1.0079
He,D9FFFF,28,4.0026
Li,CC80FF,128,6.941
Be,C2FF00,96,9.0122
B,FFB5BF,84,10.811
C,909090,76,12.107
N,3050F8,71,14.0067
O,FF0D0D,66,15.9994
F,90E050,57,18.9984
Ne,B3E3F5,58,20.1797
Na,AB5CF2,166,22.9897
Mg,8AFF00,141,24.305
```

You can edit this file in any text editor but it's easier to do it in a spreadsheet such as OpenCalc which is free at https://www.openoffice.org/product/calc.html. If you use a spreadsheet, make sure the file is saved as a CSV (comma-separated values) file. It is strongly recommended that if you do edit the file, you edit a copy and not the original in case you make a mistake.

The first item is the atomic symbol; don't change this, because then ChemLoader won't work as it relies on these symbols being constant.

*Figure 19. Part of the file ptable.csv*

The second item is the only one you should change. This is the element colour from the CPK colour scheme. It is in a 6-digit hexadecimal format, and if you want to change the colours the new colour must be in that format. There are many colour conversion sites on the web which can generate hex values for colours, for example https://htmlcolorcodes.com/

The last two data items are atomic radius and mass respectively. You can edit these if you like although I'm not sure why you would.

# Appendix 2

## Writing your own SDF files

If you happen to find an SDF file with multiple molecules in it, and you want to use one that isn't the first one, you will need to copy the relevant lines from that file and paste them into a new one.

In the 'files' folder of the plugin, you will see a file named 'multimol.sdf'. This is one I made to illustrate this appendix. If you open it in a text editor, you should be able to see that there are three molecules; we can tell this as there are three blocks of atom coordinates.

The first molecule is helpfully named 'acetone', as you see in the first line, and the second line shows it is a 2D file. There are two things to note. After the first three lines there is a line which among other things shows how many atoms and bonds there are in the molecule (10 atoms and 9 bonds). Then there is a block of data with one line per atom and at the end of the atoms block there is the bonds data block. Below that a line which reads 'M END'. This is NOT the end of the molecule data. The end of a complete molecule is denoted by a line with four dollar signs, i.e. '$$$$'. In this molecule these lines come one after the other but in some molecules there can be a lot of extra data between the two. The complete molecule looks like this; the various parts are highlighted.:

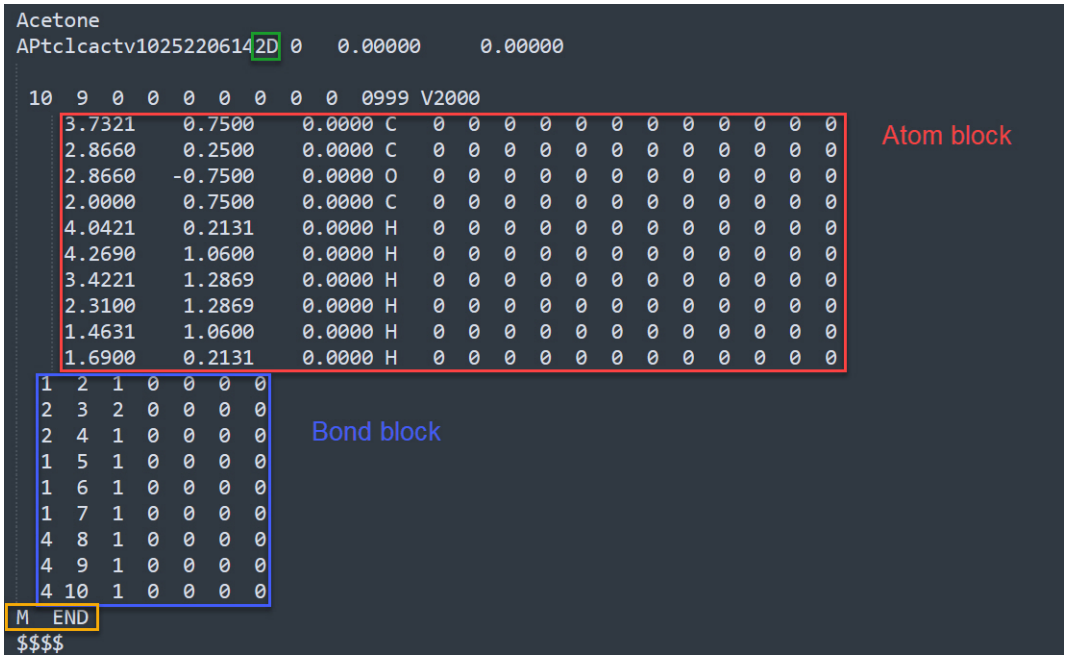


*Figure 20. A complete molecule in MOL/SDF format*

Immediately after the dollar signs, the next molecule starts. This one doesn't have a name (the first line is blank, but it's actually benzene) but from the second line we can see it is a 3D file.

After the atom and bond data there is the 'M END' again, but now there is more information which ChemLoader doesn't need, and then the four dollar signs again to terminate the molecule. Then the third molecule starts. Its name is a catalogue number, but it is in fact paracetamol (acetominophen). The next two lines are blank so we don't know if this is 2D or 3D, although we can see that the Z-coordinates are non-zero, so it's 3D. It's terminated again by 'M END' and then '$$$$'.

Suppose you want just the second (benzene) molecule. You need to copy every line starting with the line immediately after the '$$$$' line which denotes the end of the acetone molecule, up to and including the '$$$$' line. Note that anything after 'M END' is currently ignored by the plugin. Create a new text file and paste

these lines into the new file, then save it with any name - but it must have the extension .'sdf'. You can even edit the new file and add the name 'benzene' yourself if you want, but the name MUST be in the first line for ChemLoader to use it.

That's all there is to it. The critical points, without which ChemLoader will not be able to use the file correctly, are that:

- the file must start with three (no more, no fewer) lines, any or all of which can be completely blank
- it must contain the line preceding the atom block, as that tells ChemLoader how many atoms and bonds there are in the molecule
- it must contain the atom and bond data blocks
- after the bond data it must have the line 'M END' (without the apostrophes of course)
- ideally it would terminate with the '$$$$' line; this isn't strictly necessary for this version of ChemLoader, but it might be for future versions

The SDF file format is explained in detail at [https://www.nonlinear.com/progenesis/sdf-studio/v0.9/faq/sdf-file-format-guidance.aspx](https://www.nonlinear.com/progenesis/sdf-studio/v0.9/faq/sdf-file-format-guidance.aspx).

The new benzene file will look like this:

```
     RDKit          3D

  6  6  0  0  0  0  0  0  0  0999 V2000
   -1.0999    0.8257    0.0500 C   0  0  0  0  0  0  0  0  0  0  0  0
    0.1655    1.3667   -0.1438 C   0  0  0  0  0  0  0  0  0  0  0  0
    1.2367    0.5432    0.1765 C   0  0  0  0  0  0  0  0  0  0  0  0
    1.0910   -0.8063   -0.1151 C   0  0  0  0  0  0  0  0  0  0  0  0
   -0.1418   -1.3735    0.0107 C   0  0  0  0  0  0  0  0  0  0  0  0
   -1.2515   -0.5557    0.0217 C   0  0  0  0  0  0  0  0  0  0  0  0
  1  2  2  0
  2  3  1  0
  3  4  2  0
  4  5  1  0
  5  6  2  0
  6  1  1  0
M  END
>  <zinc_id>  (1)
ZINC000000967532

>  <smiles>  (1)
c1ccccc1

$$$$
```

*Figure 21. Benzene molecule extracted from a multi-molecule SDF file*