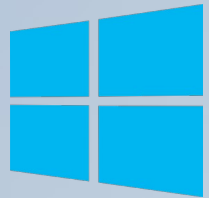


Getting Started with SIR



Download SIR software from:

<https://www.ba.ic.cnr.it/softwareic/sir/sir-download/>



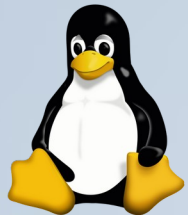
Windows: 64-bit: Windows 10, 11

File: sir-xx.yy.zz-win64.exe



Mac OS X: macOS 12 (Monterey) and later, Apple Silicon (arm64)

File: sir-xx.yy.zz-arm64.dmg



GNU/Linux: Ubuntu 24.04 LTS, Ubuntu 22.04 LTS, other Debian-based distributions, Red Hat-based distributions (e.g., RHEL, Fedora)

File: sir-xx.yy.zz-codename_amd64.deb, expo2-xx.yy.zz-Source.tar.gz

Getting Started with Expo



Installation notes:

<https://www.ba.ic.cnr.it/softwareic/sir/installing-the-program-2/>



The documentation is available at the link:

<https://www.ba.ic.cnr.it/softwareic/sir/>

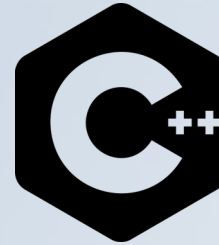


Tutorials:

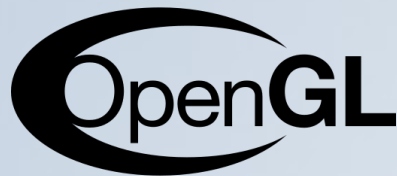
<https://www.ba.ic.cnr.it/softwareic/sir/tutorial-and-lectures/>

Technical Details

- Program written in Fortran and C++



- Program is built using OpenGL and Qt 6 libraries



- Licensed under terms of GNU General Public License



SIR: A Program for Structure Solution

SIR is designed for the automatic *ab initio* and non *ab initio* **crystal structure solution** of small, medium and macromolecules according to the following scheme:

Ab initio phasing approaches:

- Modern Direct Methods (MDM)
- Standard Direct Methods (SDM) ✓
- Vive la Difference (VLD)
- Patterson techniques

Non ab initio phasing approaches:

- Simulated Annealing ✓
- Molecular Replacement
- SAD/MAD

Ab initio methods



No starting structural model (no coordinates)

Non-ab initio methods



A model is supplied (known fragment, homologous structure, known molecular geometry)

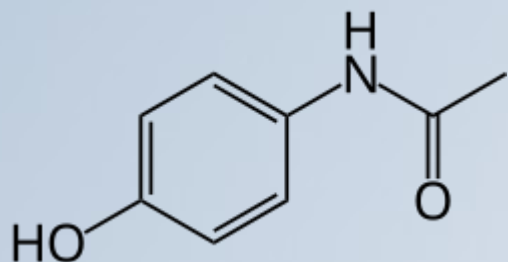
But in both cases, you still need:

- Experimental diffraction data
- Crystallographic symmetry
- Basic chemical information

Structure of the SIR Input File

The input consists of a sequence of **comments**, **commands** and **directives**. The commands start with a '%' character and directives must follow the related command. Sir2024 input file can be created using a text editor or by means of the GUI provided by the program.

Crystal structure solution of paracetamol



Space group: $P 2_1/c$

Cell: $a=7.08$ $b=9.37$ $c=11.68$ $\beta=97.47$

Chemical formula: $C_8H_9NO_2$

Cell Formula Units: 4

Reflection file name: paracetamol.hkl

Format file: (3i4, 2f8.2)

Reflection type: F^2

Electrons

paracetamol.sir

```
!TEST on Paracetamol
%structure paracetamol
%job paracetamol
%data
  cell 7.08 9.37 11.68 90 97.47 90
  space P 21/n
  content (C8H9NO2)4
  reflections paracetamol.hkl
  electrons
%continue
```

Main SIR Commands

%STRUCTURE *string*

This command is used to specify the project name and to define the names of the output files generated by the program (e.g. *loganin*, *paracetamol_form2*, *sample01*).

%JOB

A caption is printed in the output (e.g. *my beautiful structure, data provided by Mauro Gemmi, sample no. 3*).

%DATA

Data input routine.

%PHASE

Phasing routine and Direct Space Refinement routines.

%FRAGMENT *file*

Used to import a structural model (for Simulated Annealing).

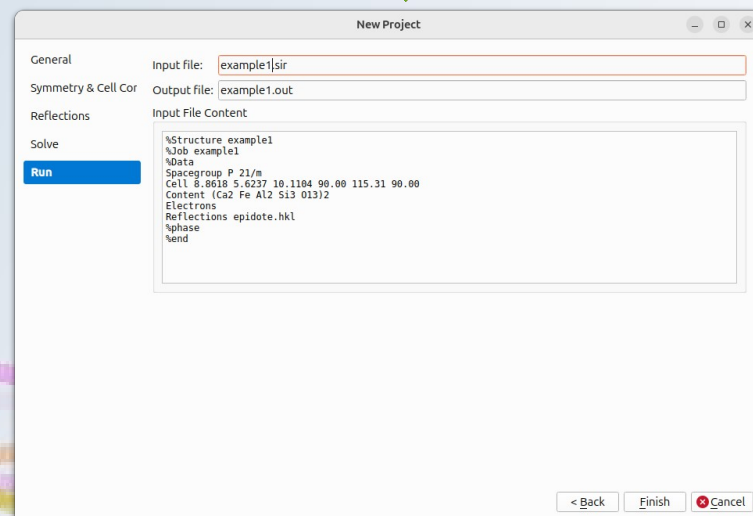
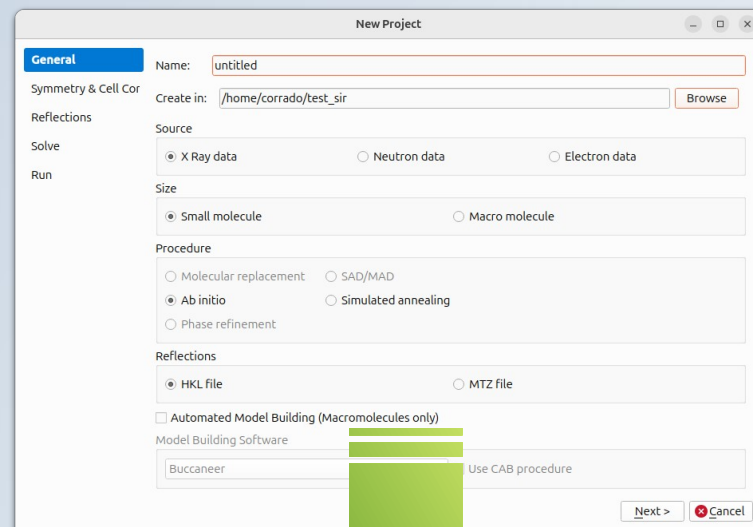
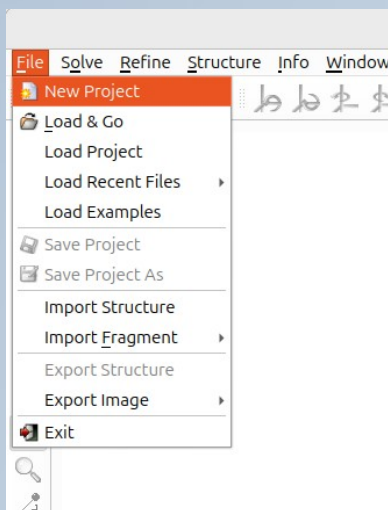
%SANNEAL

Simulated Annealing routine.

%END

End of the input file

How to Make a SIR Input File with the GUI



Example 1: Direct-Methods Structure Solution of $\text{Ca}_2\text{FeAl}_2\text{Si}_3\text{O}_{13}$

```
%Structure epidote
```

```
%Job Test on Epidote
```

```
%Data
```

```
Cell      8.8618  5.6237 10.1104  90  115.313  90
```

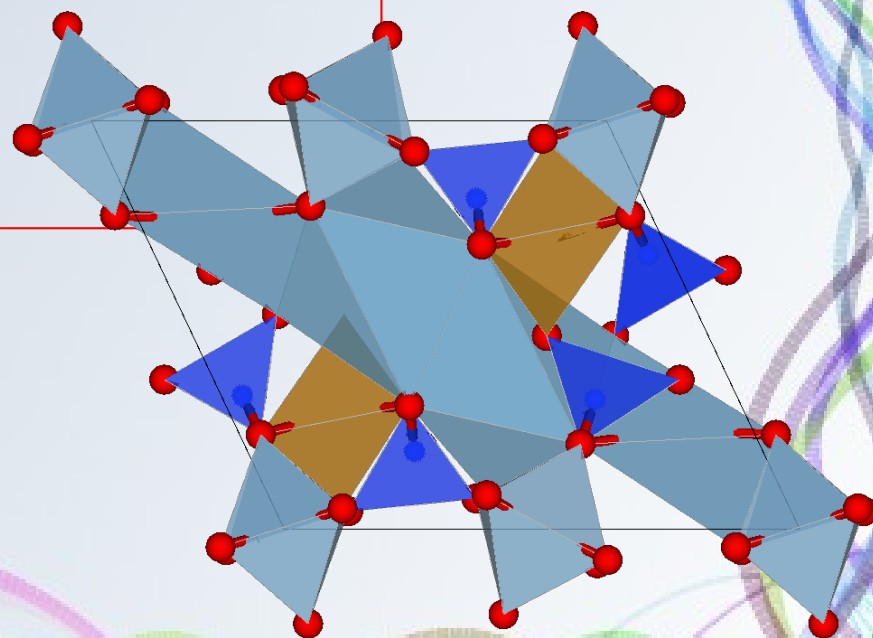
```
Content (Ca2 Fe Al2 Si3 O13)2
```

```
Space P 21/m
```

```
Reflections epidote.hkl
```

```
Electrons
```

```
%continue
```

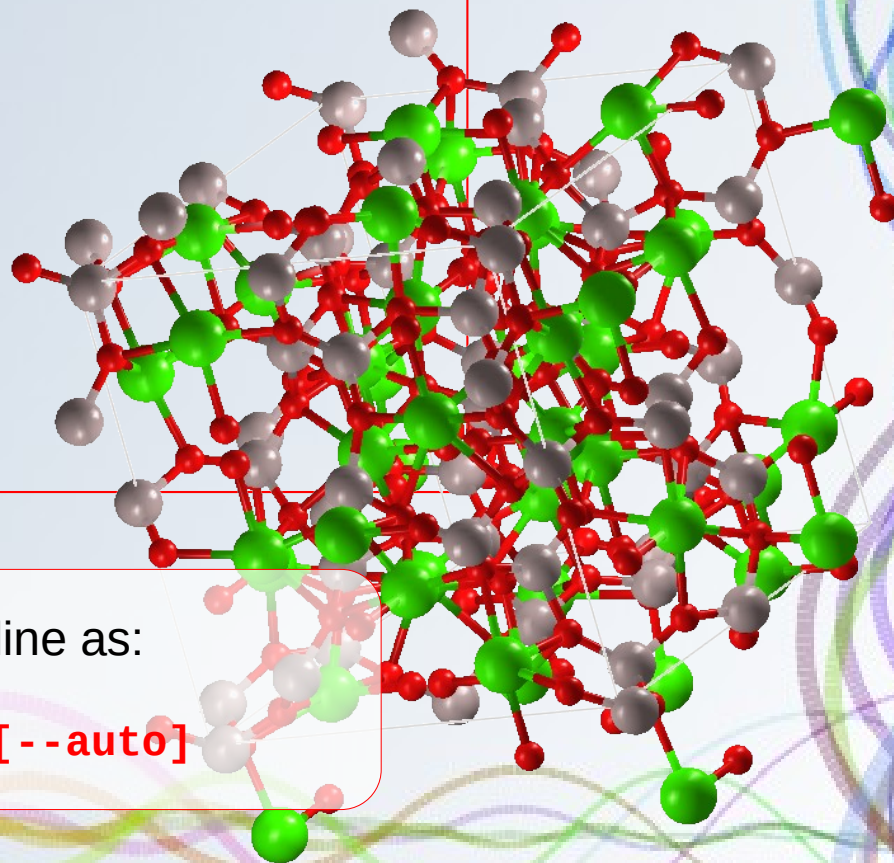
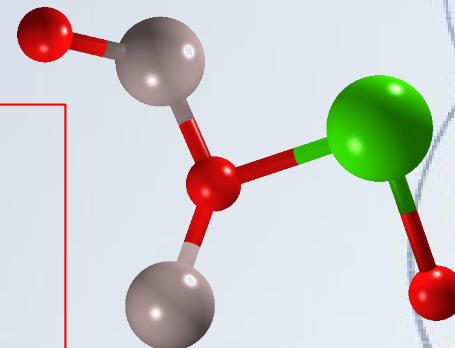


Example 2: Structure Solution of Mayenite, $(\text{CaO})_{12}(\text{Al}_2\text{O}_3)_7$

```
%structure mayenite
%job Mayenite - Calcium Oxide - Aluminium Oxide
%data
  cell 11.989 11.989 11.989 90.000 90.000 90.000
  space I -4 3 d
  cont [(CaO)12(Al2 O3)7]2
  refl mayenite.hkl
  fobs
  electrons
%continue
```

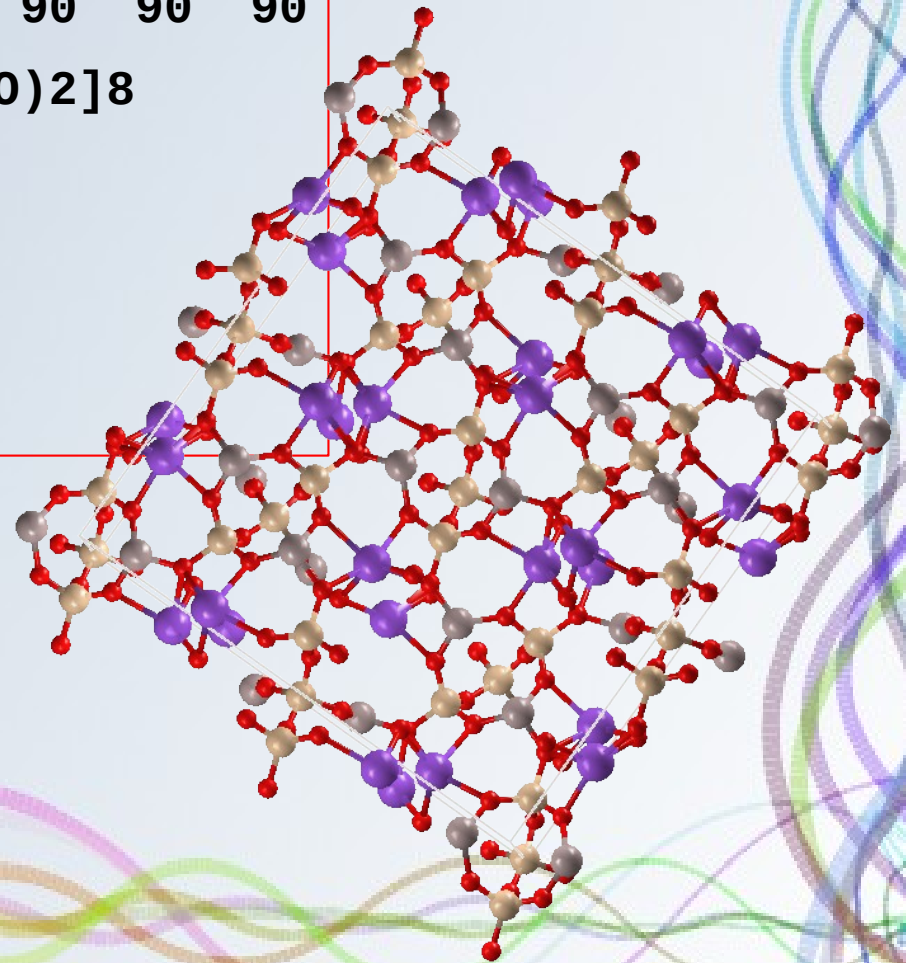
SIR can also be launched from the command line as:

```
sir input_file [output_file] [--nogui][--auto]
```



Example 3: Structure Solution of Natrolite, $\text{Na}_2\text{Al}_2\text{Si}_3\text{O}_{10} \cdot 2\text{H}_2\text{O}$

```
%Structure natro
%Job Test on Natrolite
%Data
  Cell      18.8278 18.667 6.6414 90  90  90
  Content   [Na2 Al2 Si3 O10 (H2O)2]8
  Space F d d 2
  Reflections natrolite.hkl
  Electrons
%continue
```



Example 4: Structure Solution of Orthacetamol

%Structure orthacetamol

%Job Test Structure: Orthacetamol

%Data

Cell 10.3 10.3 13.9 90 90 90

Content (C8 H9 N O2)8

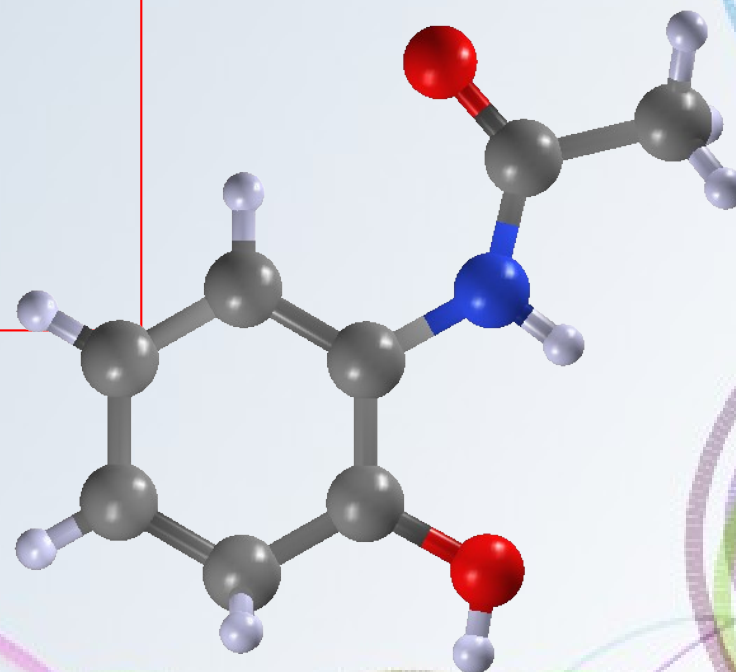
Space C 2/c

Reflections orthacetamol.hkl

Electrons

resm 0.9

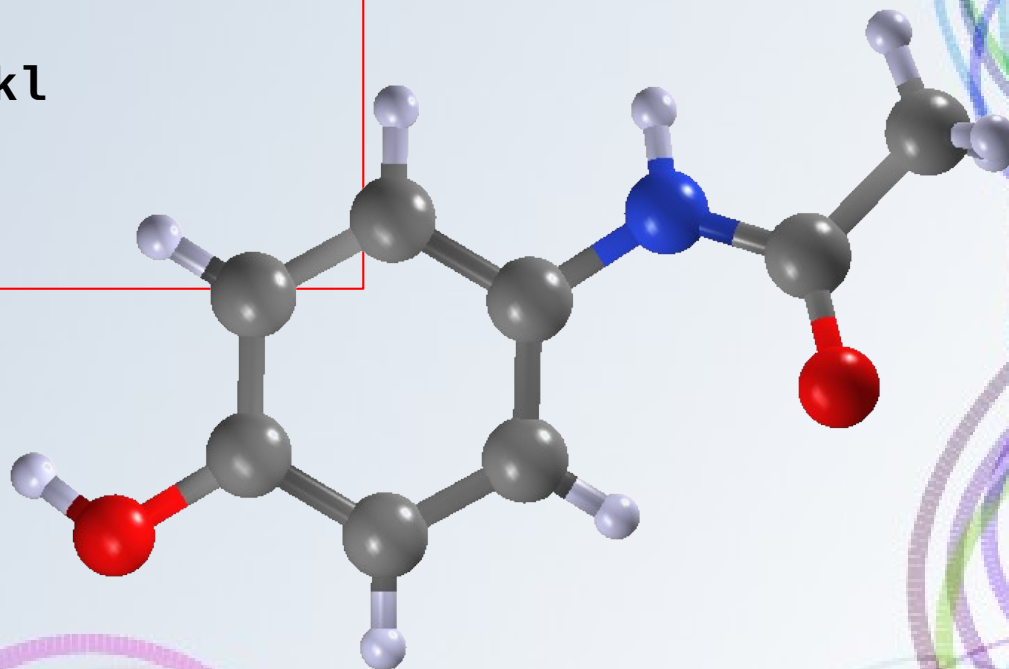
%continue



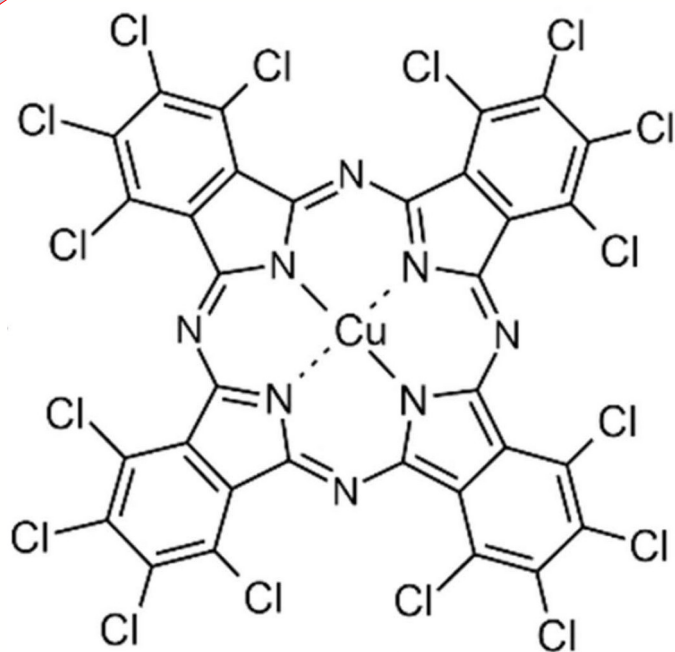
*Andrusenko et al., Angew. Chem. Int. Ed. 2019, 58, 10919 –10922

Example 5: Structure Solution of Paracetamol

```
%structure paracetamol  
%job paracetamol  
%data  
  cell 7.08 9.37 11.68 90 97.47 90  
  space P21/n  
  content (C8H9NO2)4  
  reflections paracetamol.hkl  
  electrons  
%continue
```



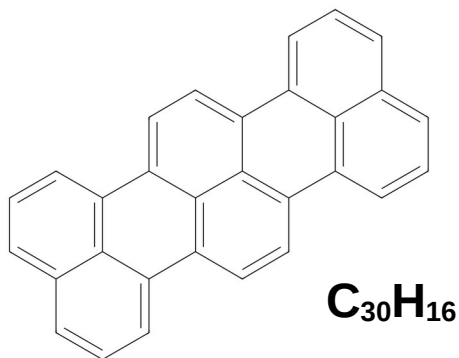
Example 6: Structure Solution of Copper Perchlorophthalocyanine



CuPcCl₁₆
CuC₃₂N₈Cl₁₆
Pigment Green 7

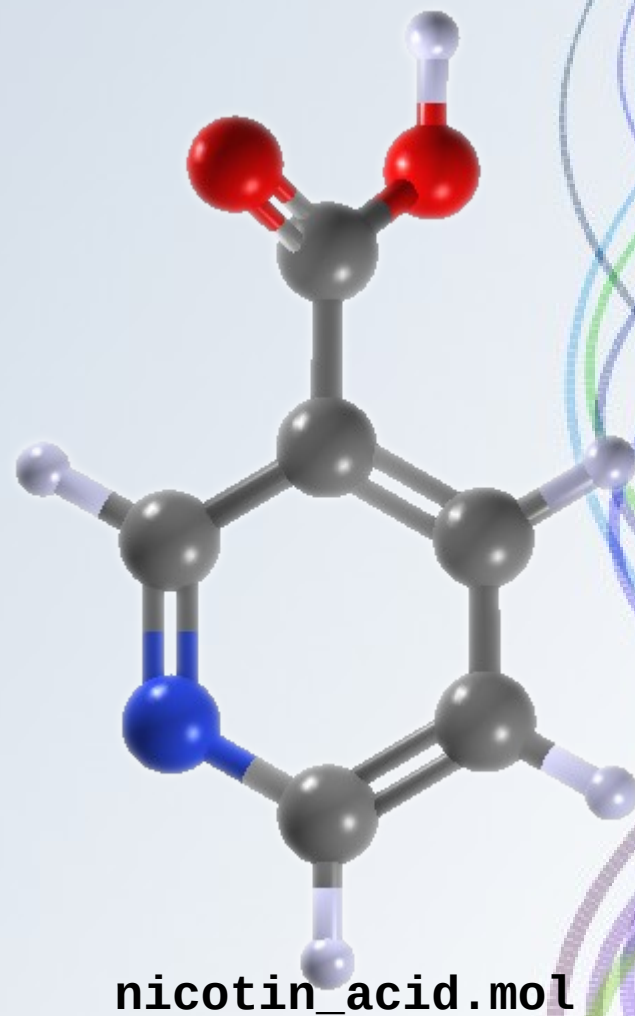
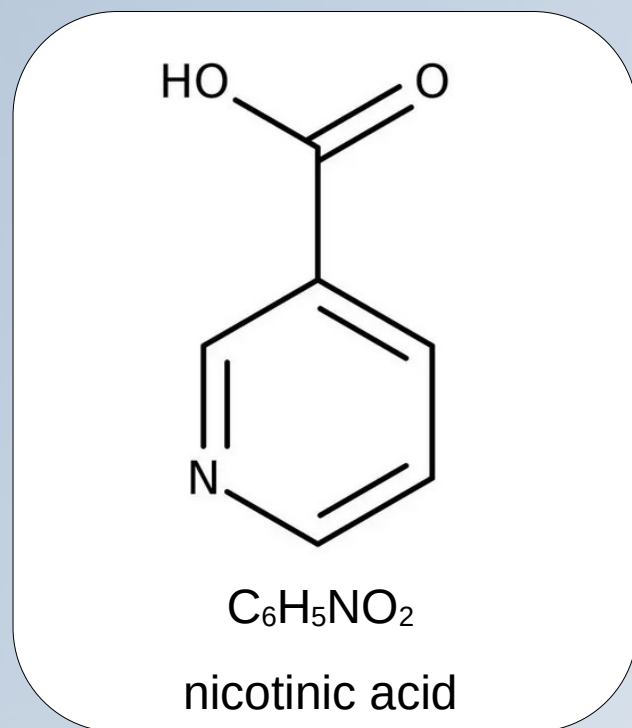
```
%Structure pigmentg7
%job copper perchlorophthalocyanine
%Data
  Cell 17.685 25.918 3.8330 90 95.05 90
  Content (C32 N8 Cl16 Cu)2
  Spacegroup C 2/m
  Reflections pigmentg7.hkl
  electrons
%Continue
```


Example 7: Structure Solution of Terrylene



```
%Structure terrylene
%Job Test on Terrylene
%Data
    Cell 11.4 10.4 14.4 90 95.6 90
    Content C 120 H 64
    Space P 21/a
    Reflections terrylene.hkl
    Electrons
%Continue
```

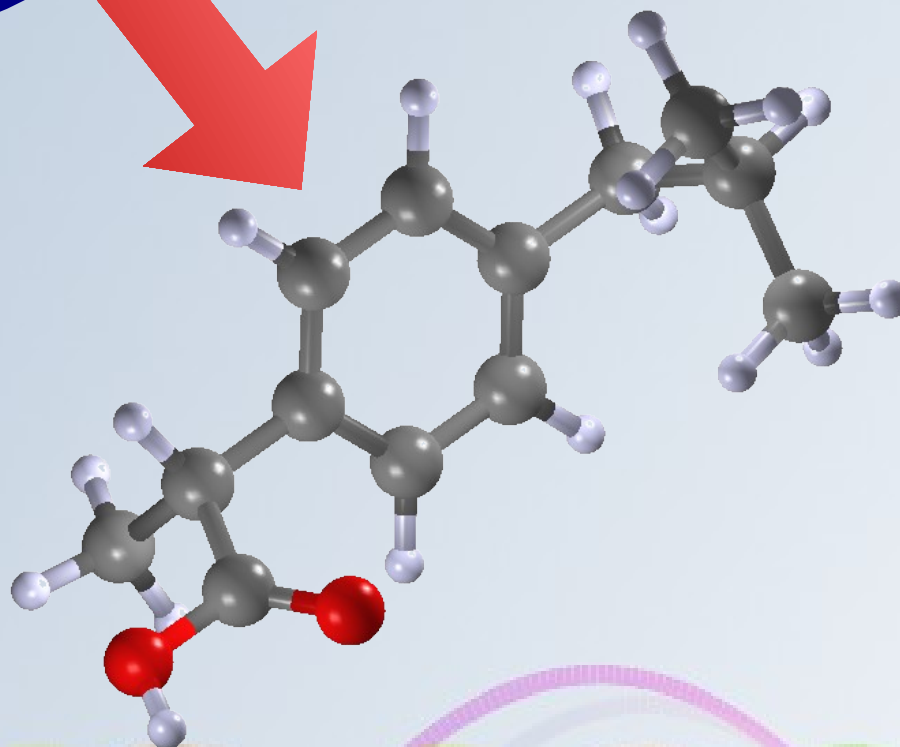

Example 8: crystal structure determination of nicotinic acid *



(*) *Acta Cryst.* (2016). A72, 236–242

Building starting model

Check for
similar
molecules in
databases



Crystal Structure Databases*

Non-commercial databases are in *red*

- **CSD** (Cambridge Structural Database) (organics & organometallics):
<http://www.ccdc.cam.ac.uk/>
- **ICSD** (Inorganic Crystal Structure Database)
(inorganics, elements, minerals & intermetallics): <http://icsd.ill.fr/>
- **COD** (Crystallography Open Database) (general database):
<http://www.crystallography.net/>

Other databases: ICDD PDF-4+, **American Mineralogist Crystal Structure Database**, **MINCRYST**, **Zeolite Structures Database**, ...

File format: CIF (Crystallographic Information File)

*Joint special issue: *Acta Cryst. B58*, 317-422 (2002)

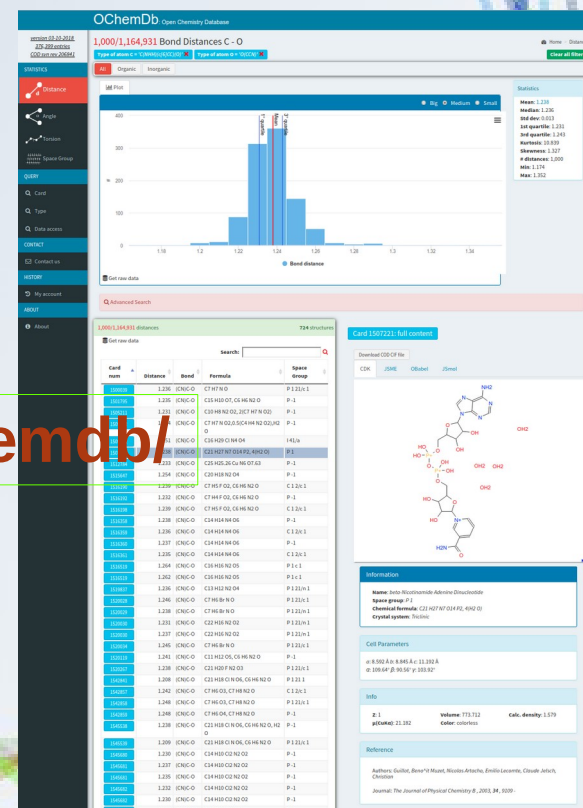
Free Chemistry Databases

- PubChem: <https://pubchem.ncbi.nlm.nih.gov/>
- NIST Chemistry WebBook: <http://webbook.nist.gov/chemistry/>
- Drugbank: <http://www.drugbank.ca/>

Other databases: ZINC, eMolecules, ChEBI, NMRShiftDB, ...

Chemical file formats: *sdf*, *mol*, *mol2*, *cml*, *SMILES*, ...

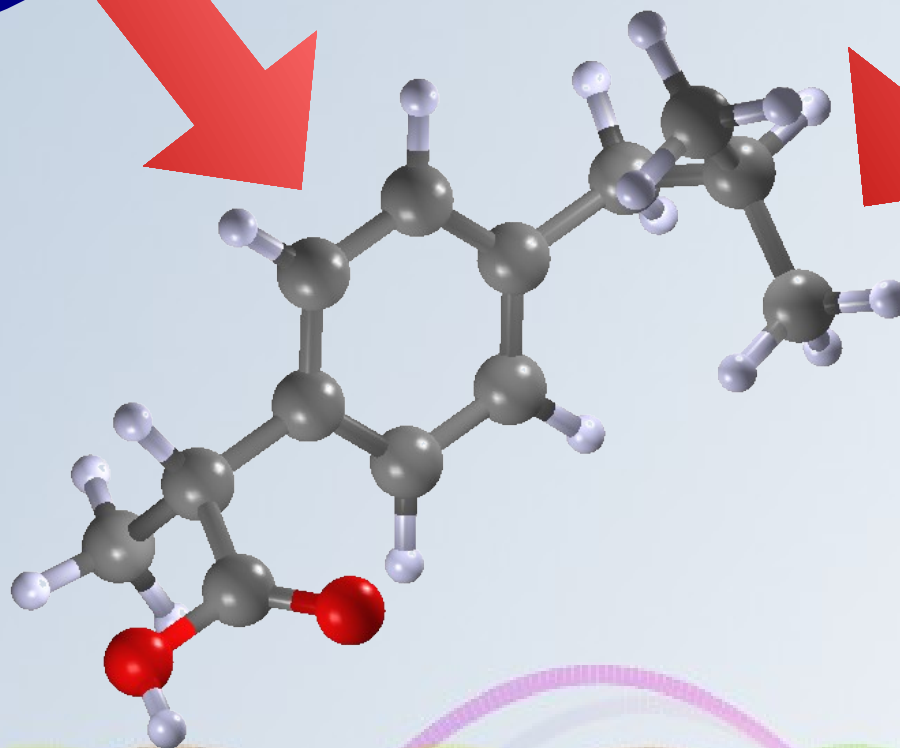
<http://www.ba.ic.cnr.it/ochemdb/>



Building starting model

Check for
similar
molecules in
databases

Optimize
Molecular
geometry by
computational
chemistry



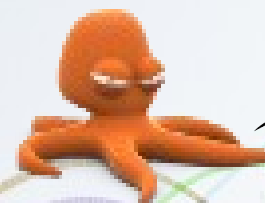
Geometry optimization

Three levels of theory

- Molecular-mechanics force fields (**MM**)
- Semi-empirical methods (**SE**)
- *Ab initio* methods: Hartree–Fock methods, density functional theory (**DFT**)



Strategy: MM → SE → DFT



Molecule editor

A molecule editor allows

- Sketch molecules in 2D or 3D format
- Optimize the geometry by force field method
- Create input file for the quantum-chemistry calculations
- Read output files of the most common computational packages

Some free available software

- ACD/ChemSketch

<https://www.acdlabs.com/resources/free-chemistry-software-apps/chemsketch-freeware/>

- Avogadro

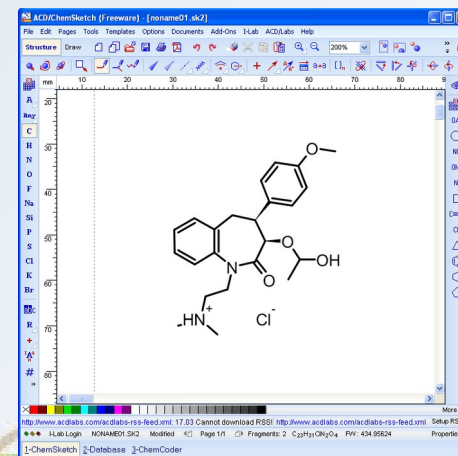
http://avogadro.openmolecules.net/wiki/Main_Page

- MarvinSketch

<http://www.chemaxon.com/products/marvin/>

- Gabedit

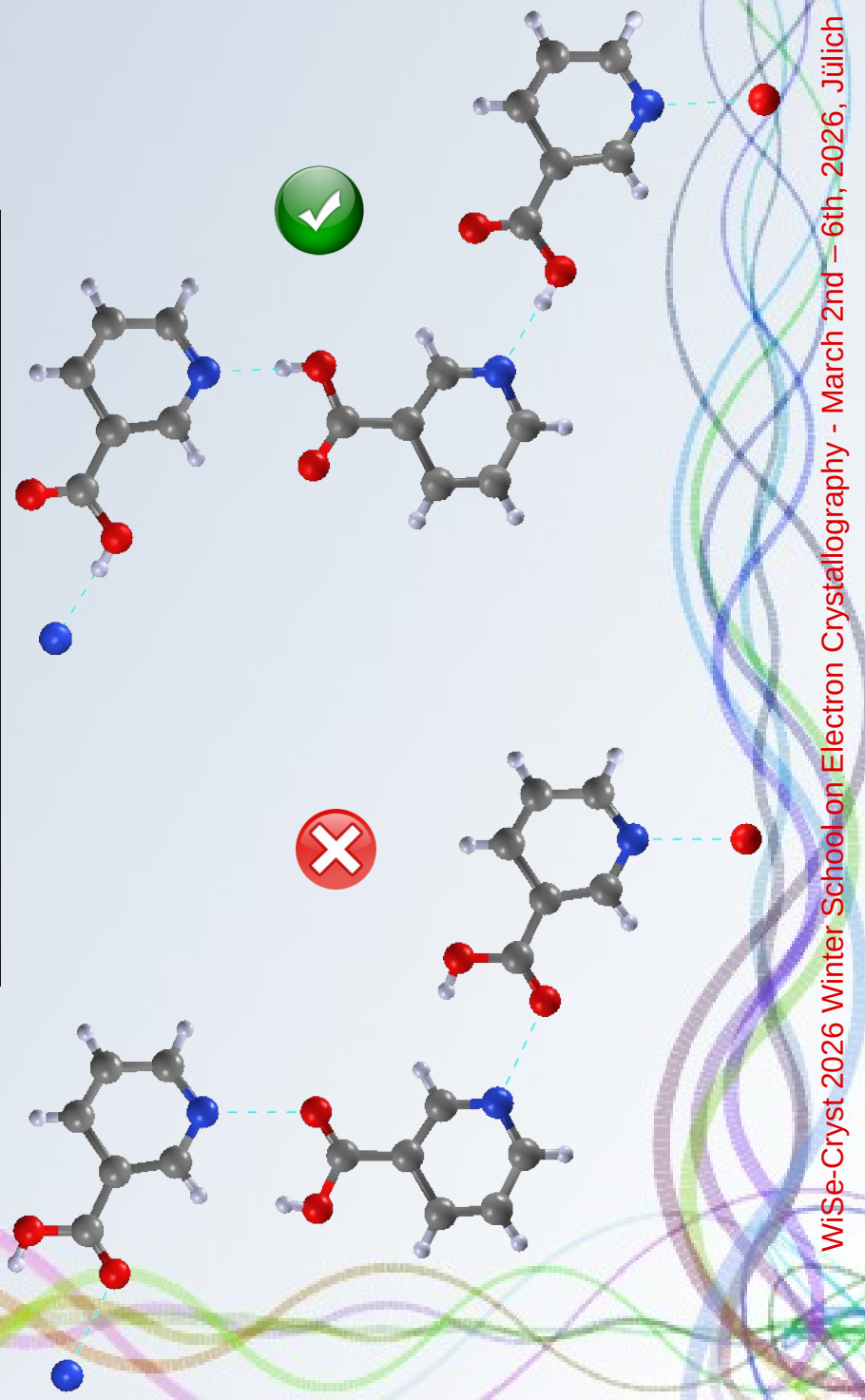
<http://gabedit.sourceforge.net/>



Example 8: crystal structure determination of nicotinic acid

Example 8/ nicotin_acid.sir

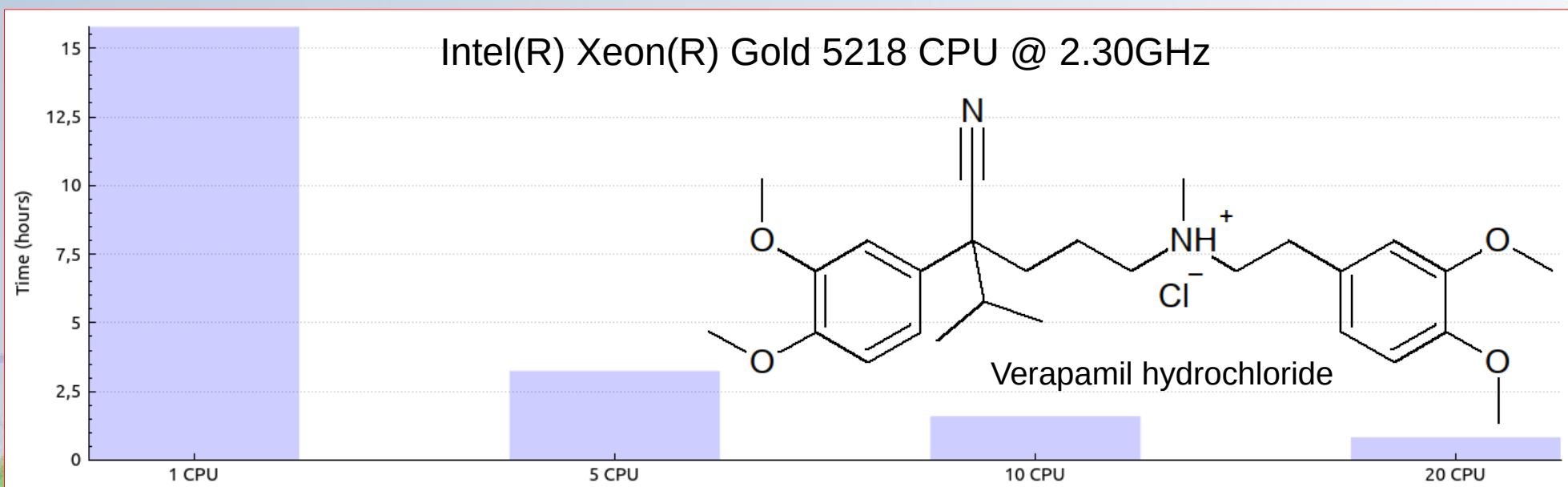
```
%window
%structure nicotin_acid
%job Acta Cryst. (2016). A72, 236–242
%data
cell 7.303 11.693 7.33 90 113.68 90
space P21/c
formula (C6 H5 N O2)4
refl nicotin_acid.hkl
format (3I4,2F8.2)
electrons
!wave 0.02508
%fragment nicotin_acid.mol
%anneal
%end
```



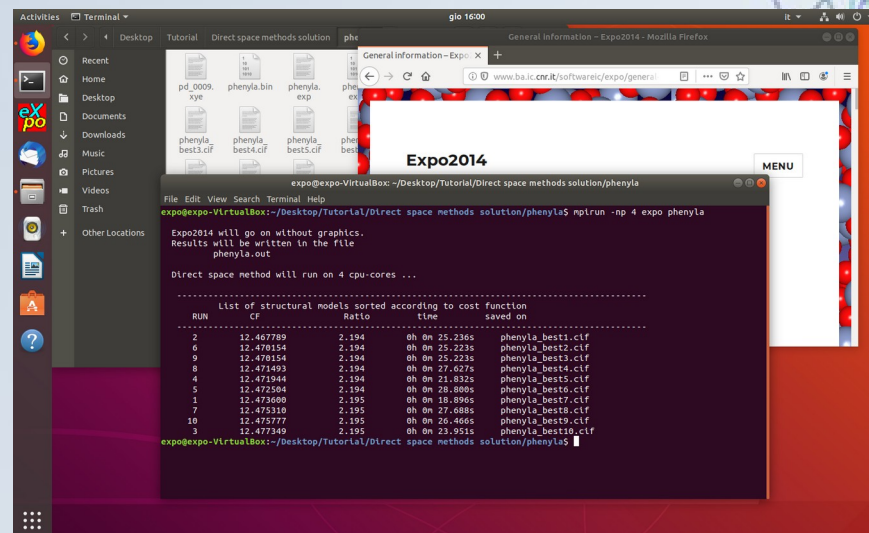
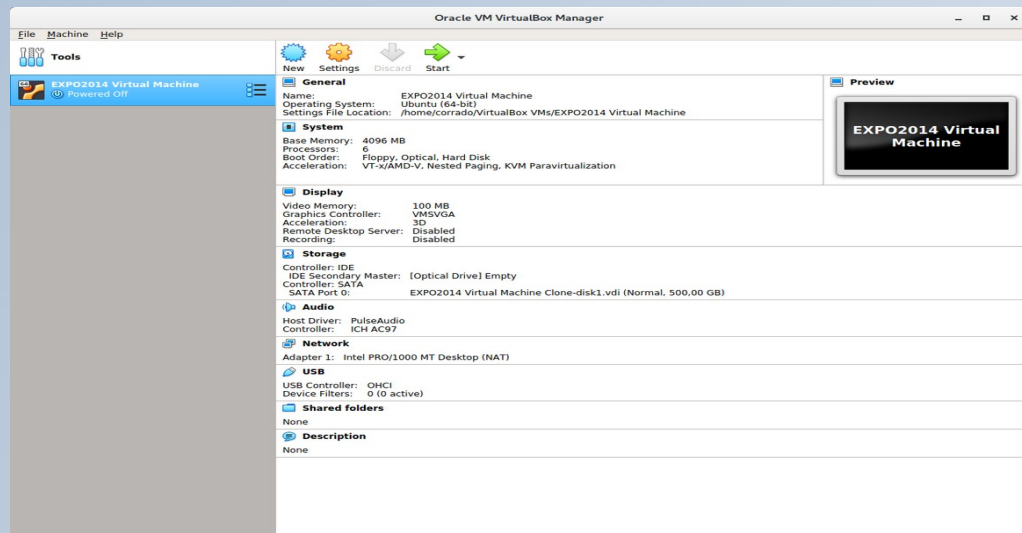
Running the Parallel Version of SIR

- Computer with multi-core CPUs and Linux environment.
- MPI installed.
- Compiling SIR from source and linking with MPI libraries
- Run SIR by using the launcher **mpirun** with the appropriate options.

```
mpirun -np 20 sir2024 input_file.exp
```



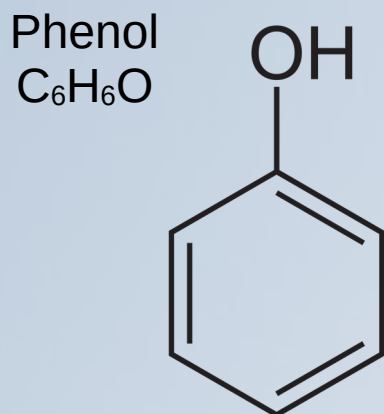
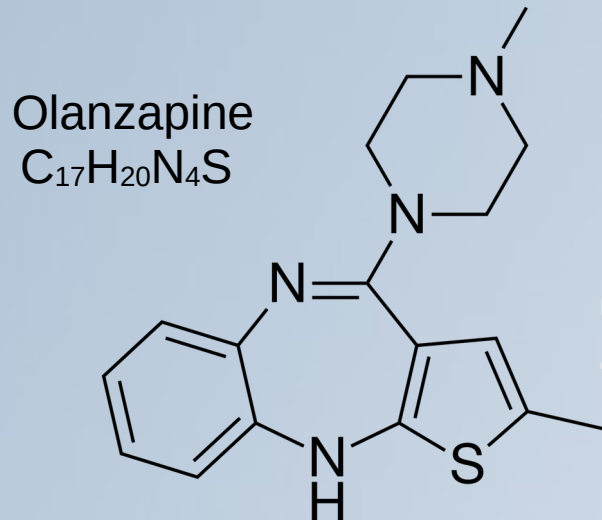
Running parallel version of SIR on Windows



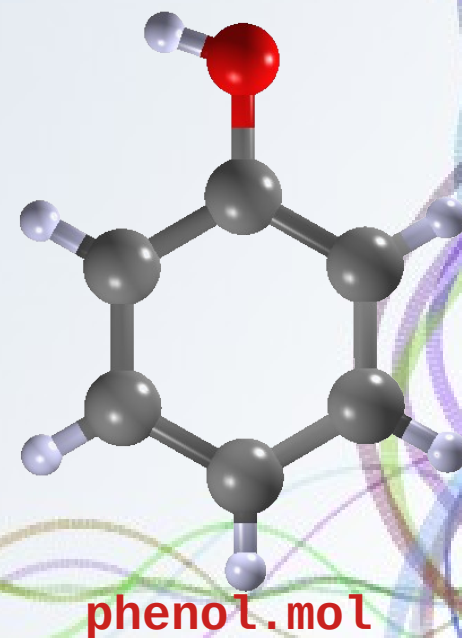
Windows Subsystem for Linux (WSL)

<https://learn.microsoft.com/en-us/windows/wsl/install>

Example 9: crystal structure determination of olanzapine cocrystal *



PubChem



* Acta Cryst. (2020). B76, 1036–1044

Load crystal structures from COD

“File” menu → “Import Fragment” menu → “From COD”

The screenshot displays the expo-2.3.7 software interface. The 'Import Fragment from COD' dialog box is open, showing search results for the formula C17H20N4S. The 'JAV Molecular Viewer' window shows a 3D ball-and-stick model of the molecule.

Import Fragment from COD

Basic Search | Cell | Space Group | Bibliography | Filters

Text:

Formula:

Elements:

NOT these elements:

No. of elements min and max:

COD id:

Downloads | **6 results found** | Search

<input type="checkbox"/>	COD number	Spacegroup	Formula	Cell Volume	Cell Parameters
<input checked="" type="checkbox"/>	1572205	P 1 21/c 1	C17 H20 N4 S	1664	10.708 16.476 10.065 90 110.43 90
<input type="checkbox"/>	2018274	P 1 21/c 1	C17 H20 N4 S	1622.7	9.913 16.5329 9.9992 90 98.023 90
<input type="checkbox"/>	2203128	P 1 21/c 1	C17 H20 N4 S	1600.9	10.388 14.839 10.567 90 100.64 90
<input type="checkbox"/>	2237783	P -1	C17 H20 N4 S	1640.2	9.172 12.224 15.563 86.933 86.99 70.382
<input type="checkbox"/>	4507586	P 1 21/c 1	C17 H20 N4 S	1556	10.3411 14.521 10.5314 90 100.291 90
<input type="checkbox"/>	4507587	P 1 21/c 1	C17 H20 N4 S	1586.9	9.8544 16.314 9.9754 90 98.304 90

Details

Chemical Names: olanzapine

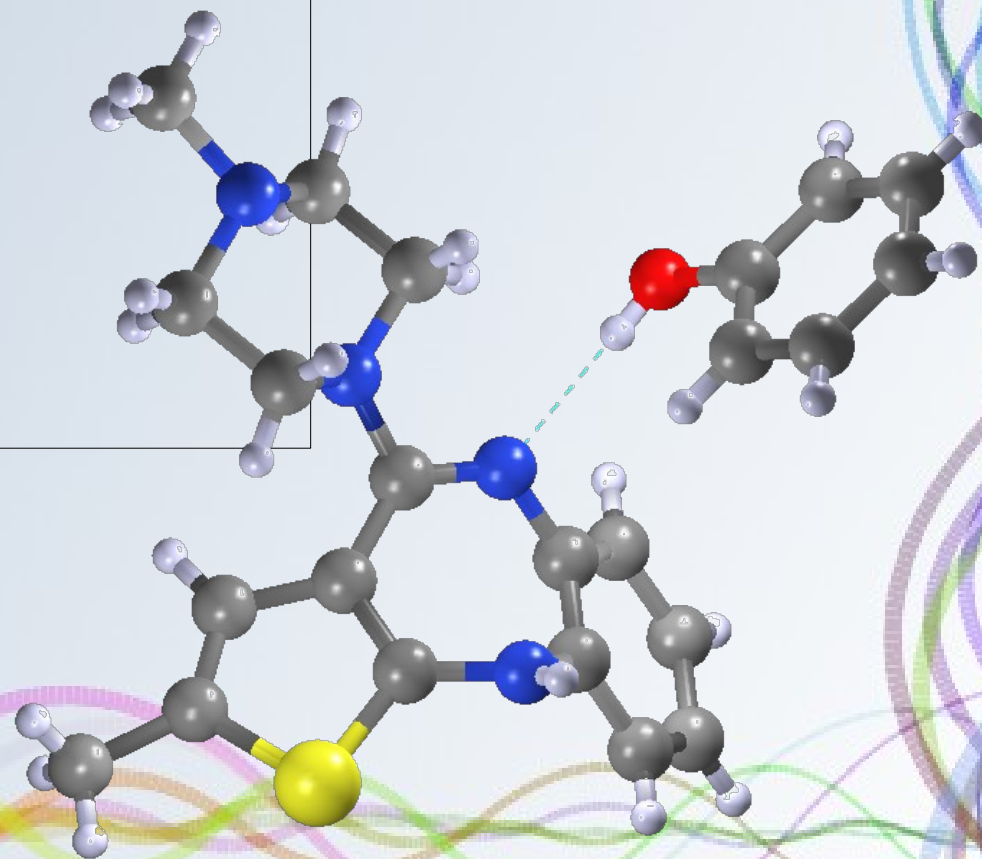
Bibliography: Anyfanti, Goulielmina; Husanu, Elena; Andrusenko, Iryna; Marchetti, Danilo; Gemmi, Mauro
The crystal structure of olanzapine form III.
IUCr, **2024**, 11
<https://doi.org/10.1107/S2052252524007383>

☒ Import ☐ Close

Example 2: crystal structure determination of olanzapine cocrystal

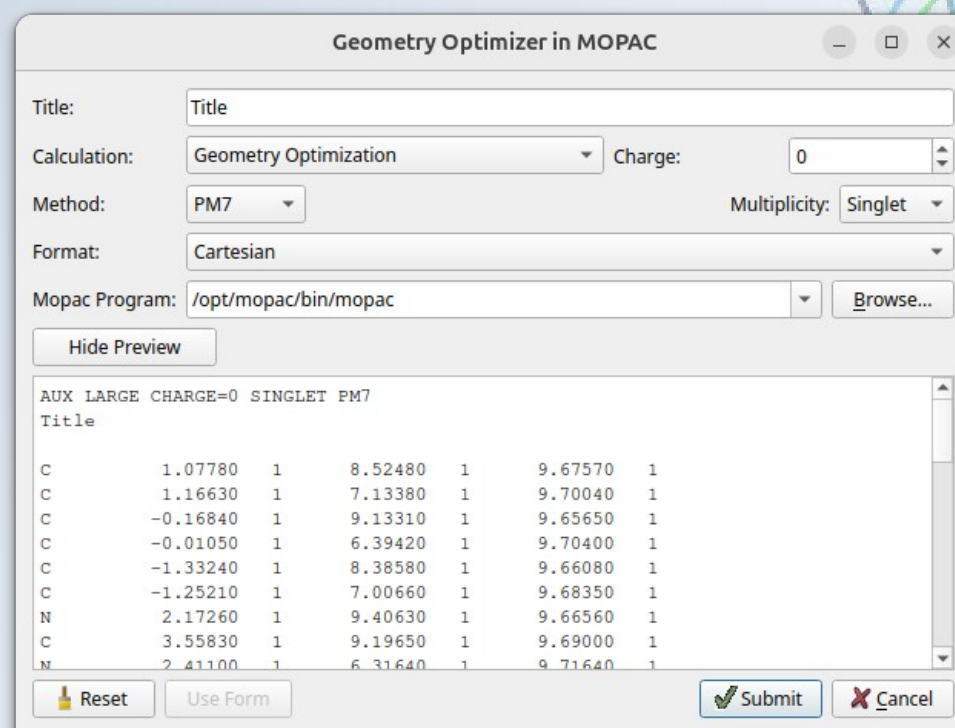
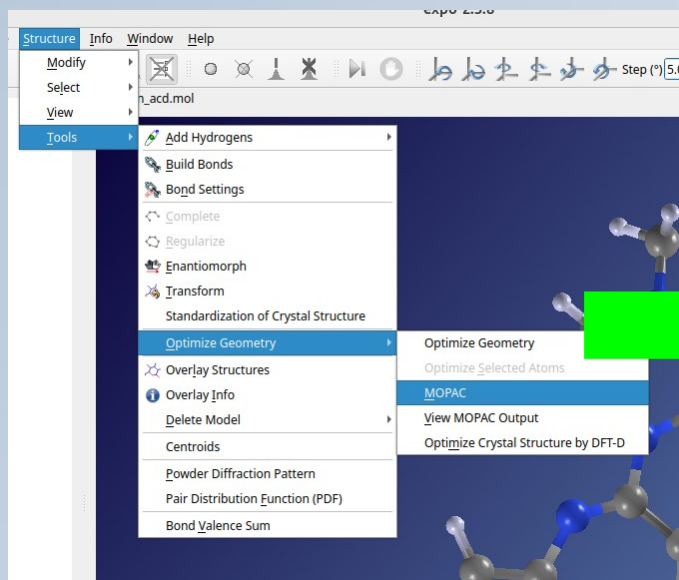
Example 2/test 1/olanphen.sir

```
%window
%structure olanphen
%job olanzapine+phenol
%data
  cell  9.00 10.6 12.0 92.0 97.0 103.0
  formula ((C17 H20 N4 S) (C6 H5 O H))2
  space P -1
  reflections olanphen.hkl
  format (3i4,2f10.2)
  electrons
!   wave 0.035
!   known olanphen.cif
%fragment 1572205.cif
%fragment phenol.mol
%anneal
%end
```



Geometry optimization by MOPAC

MOPAC is a semiempirical quantum chemistry software package available FREE.
Download link: <http://openmopac.net/downloads.html>



The 'Geometry Optimizer in MOPAC' window is shown. It contains the following fields and options:

- Title: Title
- Calculation: Geometry Optimization
- Charge: 0
- Method: PM7
- Multiplicity: Singlet
- Format: Cartesian
- Mopac Program: /opt/mopac/bin/mopac
- Buttons: Hide Preview, Reset, Use Form, Submit, Cancel

The output window displays the following data:

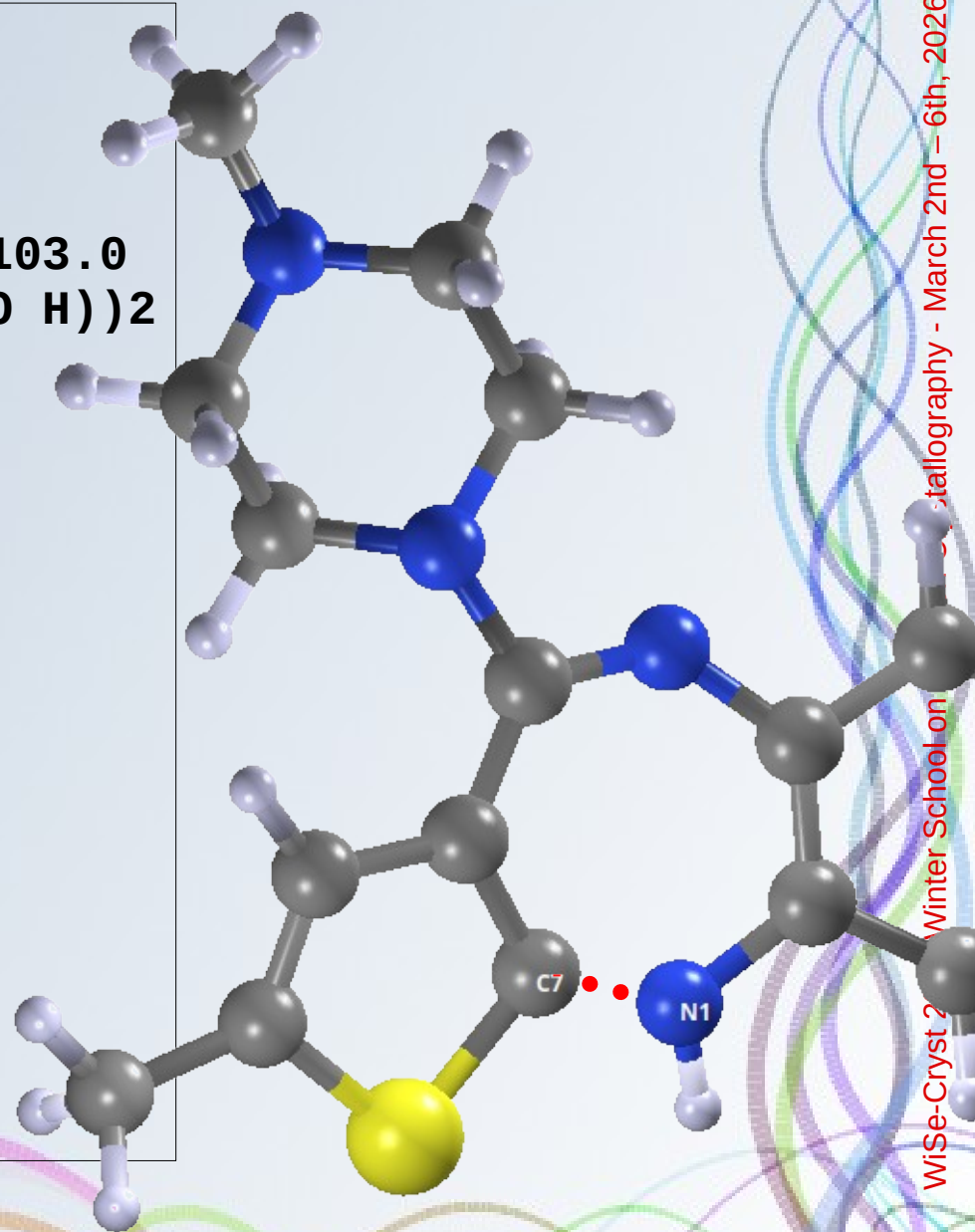
AUX LARGE CHARGE=0 SINGLET PM7						
Title						
C	1.07780	1	8.52480	1	9.67570	1
C	1.16630	1	7.13380	1	9.70040	1
C	-0.16840	1	9.13310	1	9.65650	1
C	-0.01050	1	6.39420	1	9.70400	1
C	-1.33240	1	8.38580	1	9.66080	1
C	-1.25210	1	7.00660	1	9.68350	1
N	2.17260	1	9.40630	1	9.66560	1
C	3.55830	1	9.19650	1	9.69000	1
N	2.41100	1	6.31640	1	9.71640	1

Graphical User Interface for MOPAC

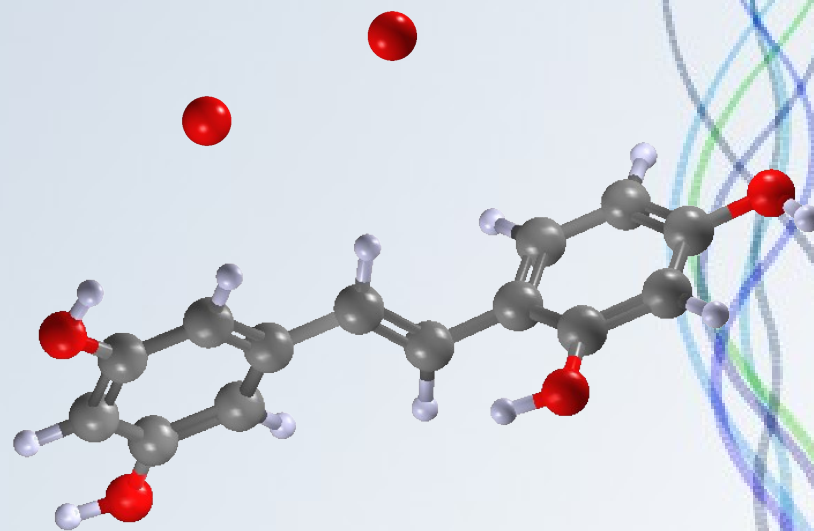
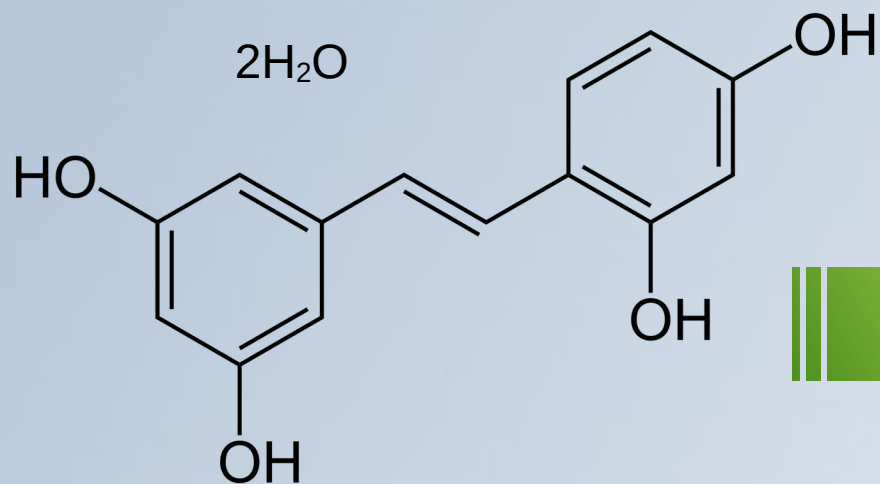
Example 2: crystal structure determination of olanzapine cocrystal

Example 2/test 2/olanphen.sir

```
%window
%structure olanphen
%job olanzapine+phenol
%data
    cell  9.00 10.6 12.0 92.0 97.0 103.0
    formula ((C17 H20 N4 S) (C6 H5 O H))2
    space P -1
    reflections olanphen.hkl
    format (3i4,2f10.2)
    electrons
!    wave 0.035
!    known olanphen.cif
%fragment olanzapine_mopac.mol
cut N1 C7
%fragment phenol.mol
%anneal
rest N1 C7 1.36 0.01
niter 1000
nrun 20
%end
```



Example 10: crystal structure determination of oxyresveratrol dihydrate*



oxyresveratrol.mol

*ACS Omega 2024, 9, 41555–41564

Imposing anti-bumping restraints

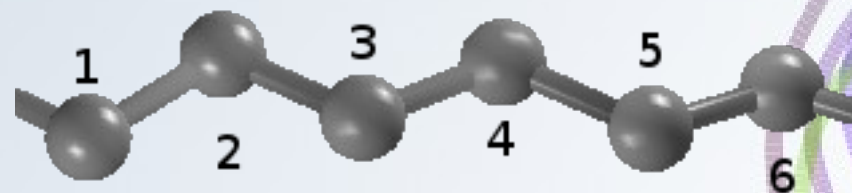
$$CF_{bump} = \sum_{ij}^n w_{ij} (d_{ij}^{min} - d_{ij}^{model})^{2k}$$

$$k = 2$$

$$d_{ij}^{model} < d_{ij}^{min}$$

$$d_{ij}^{min} = \epsilon(R_i^{vdW} + R_j^{vdW})$$

All nonbonded interactions between atoms that are separated by a path of bonds containing 4 rotatable bonds or less are excluded



Only 1-5 interactions are considered

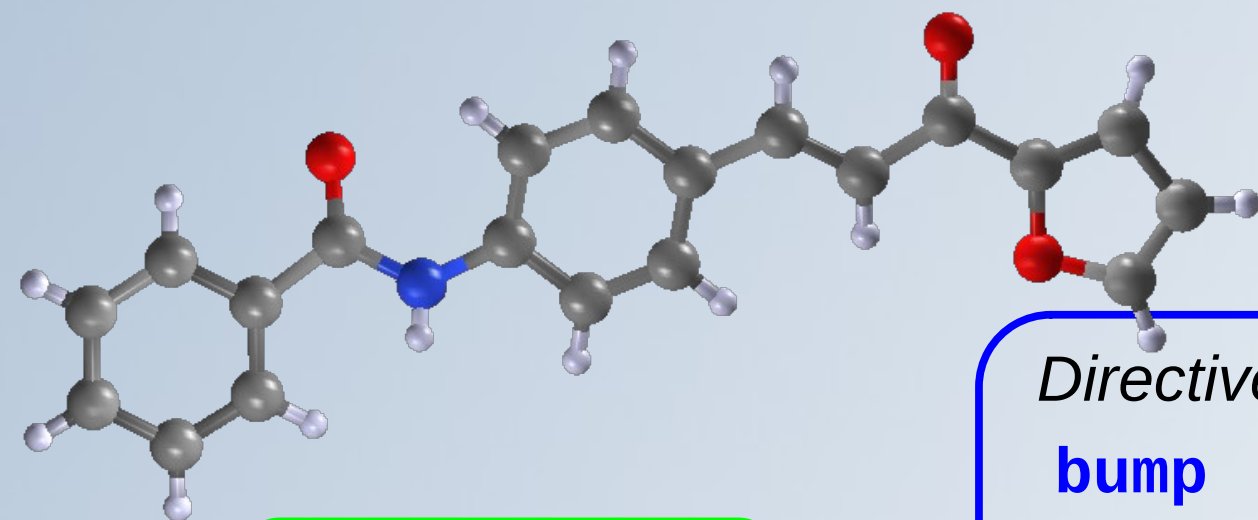
Imposing anti-bumping restraints

$$CF_{bump} = \sum_{ij}^n w_{ij} (d_{ij}^{min} - d_{ij}^{model})^{2k}$$

$$k = 2$$

$$d_{ij}^{model} < d_{ij}^{min}$$

$$d_{ij}^{min} = \epsilon(R_i^{vdW} + R_j^{vdW})$$



```
%sannel  
bscale 0.8  
bump
```

Directive:

bump

or

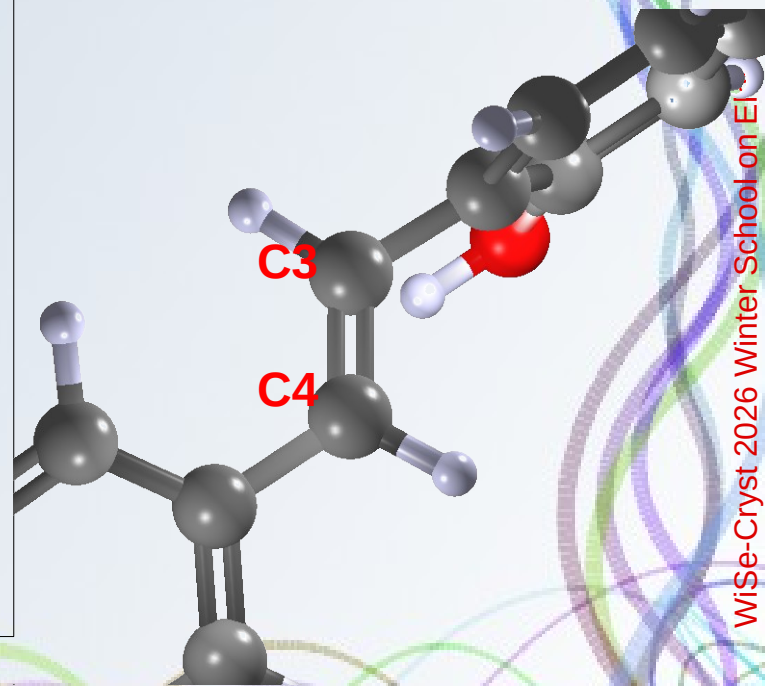
bump atoms1 atoms2 [dist]

Warning: time-consuming procedure, use only if the diffraction data are not of sufficient quality

Example 10: crystal structure determination of oxyresveratrol dihydrate*

Example 10/Dihydrate_Oxyresveratrol.sir

```
%window
%structure Dihydrate_Oxyresveratrol
%job Dihydrate Oxyresveratrol
%data
  cell 8.872 8.112 9.292 90 90.929 90
  space Pc
  formula (C14 H12 O4 (H2O)2)4
  refl Dihydrate_Oxyresveratrol.hkl
  format (3i4,2f10.2)
  electrons
  !known Dihydrate_Oxyresveratrol.cif
  !wave 0.0335
%fragment Oxyresveratrol.mol
%fragment atoms 02
!deletehydro
%sanneal
intdof C3 C4 0
bump
%end
```



*ACS Omega 2024, 9, 41555–41564