Dictionary of ligands

all molecules

gBank: http://www.drugbank.ca/

C: http://zinc.docking.org/index.shtml

ODRUG: http://www.compbio.dundee.ac.uk/Web_Servers/prodrg_down.h

CTVS: http://www2.chemie.uni-erlangen.de/software/cactvs/

mbridge structural database - CSD: http://www.ccdc.cam.ac.uk/products/cs

cromolecules

B:

ropean EBI: http://www.ebi.ac.uk/msd/

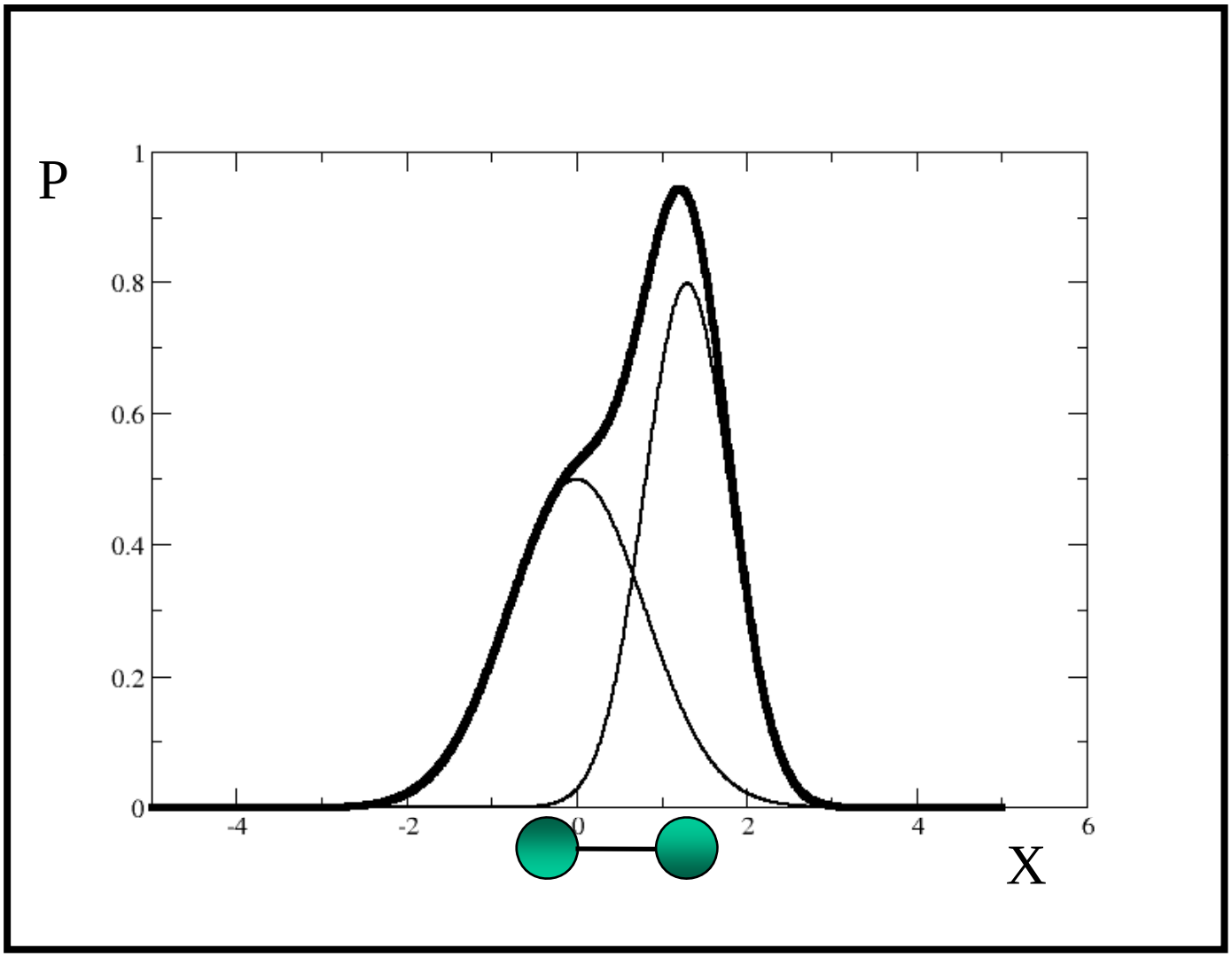
SA RSCB: http://www.rcsb.org/pdb/download/download.do

SMOL (visualisation tool): http://rasmol.org/

OL (Java based visualisation tool): http://jmol.sourceforge.net/

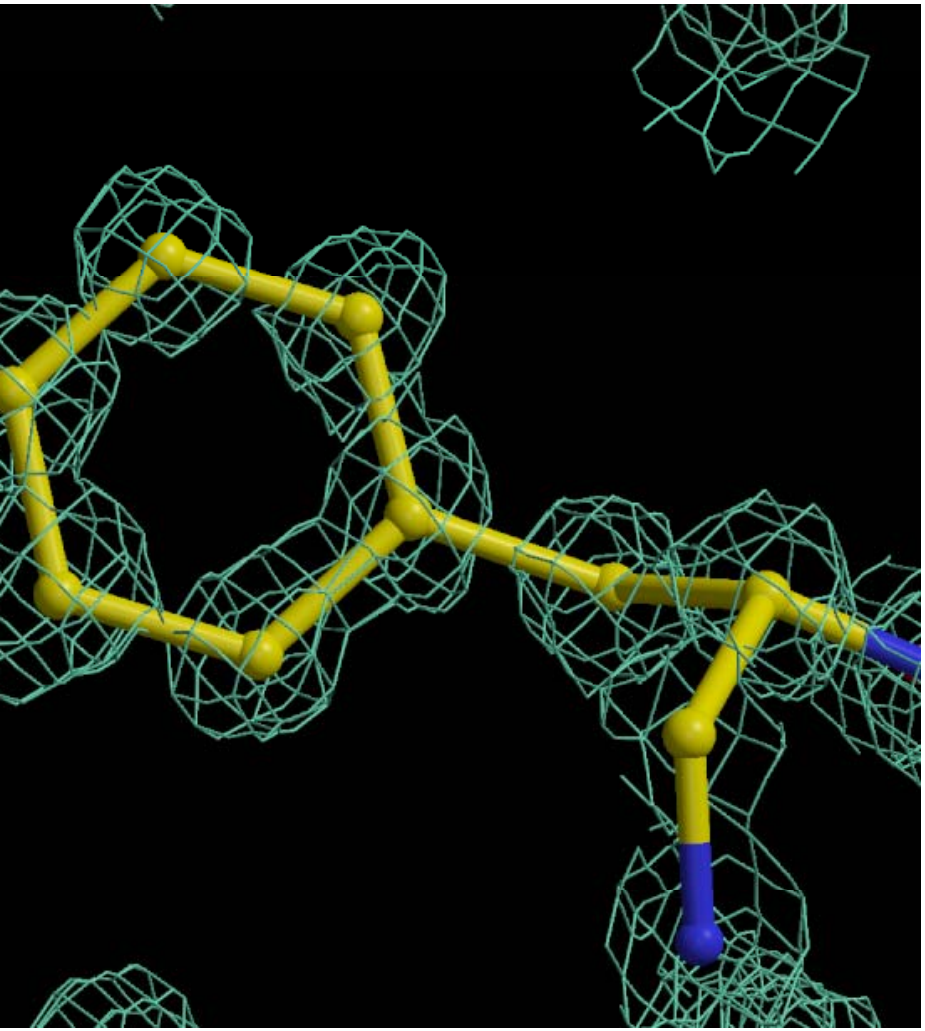
Why restraints:
Two atoms ideal case

Distance between atoms 1.3Å. B values 20 and 50

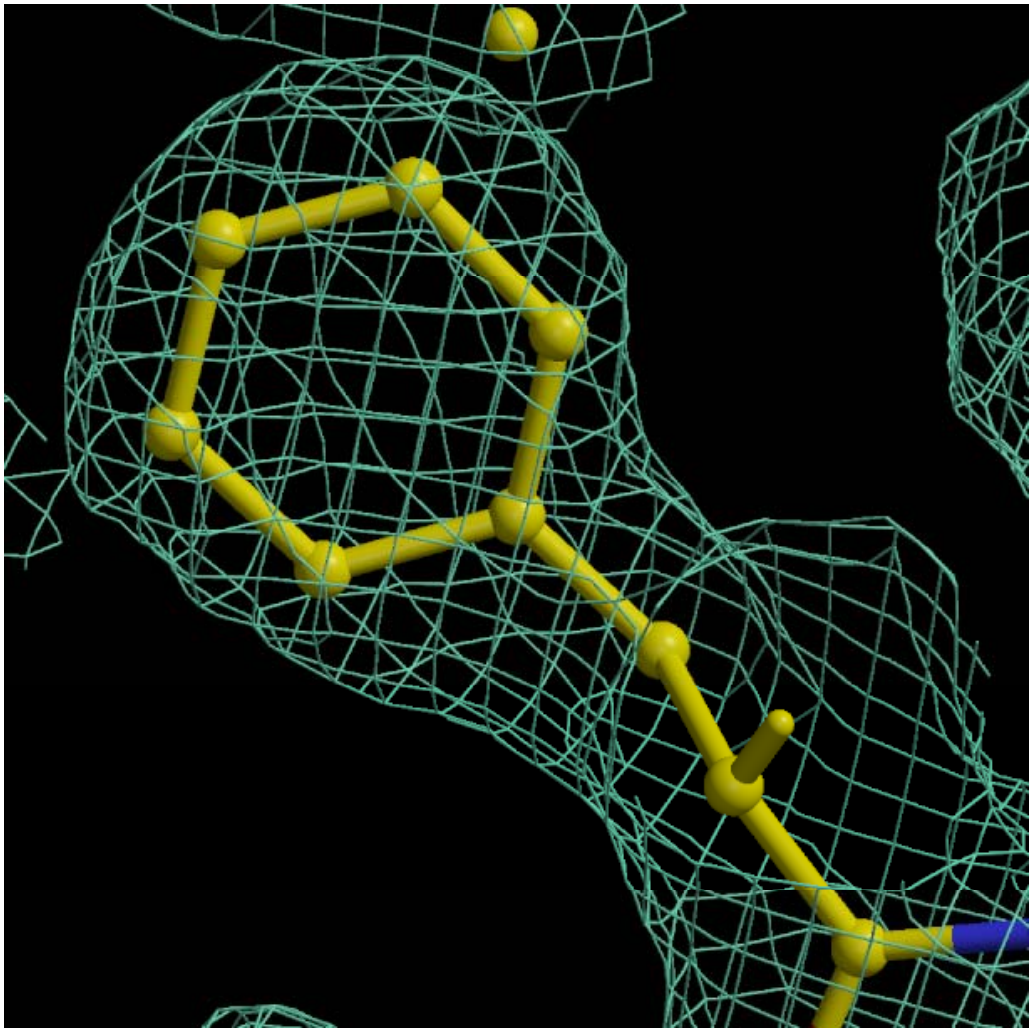


Chemical information: Phe at two different resolutions

0.88 Å



2 Å and High mobility



Role of restraints

hen atoms have high B values and/or data are at low solution then electron density may not show separate peak

restraints would not be used then chemistry of molecule ould be unreasonable.

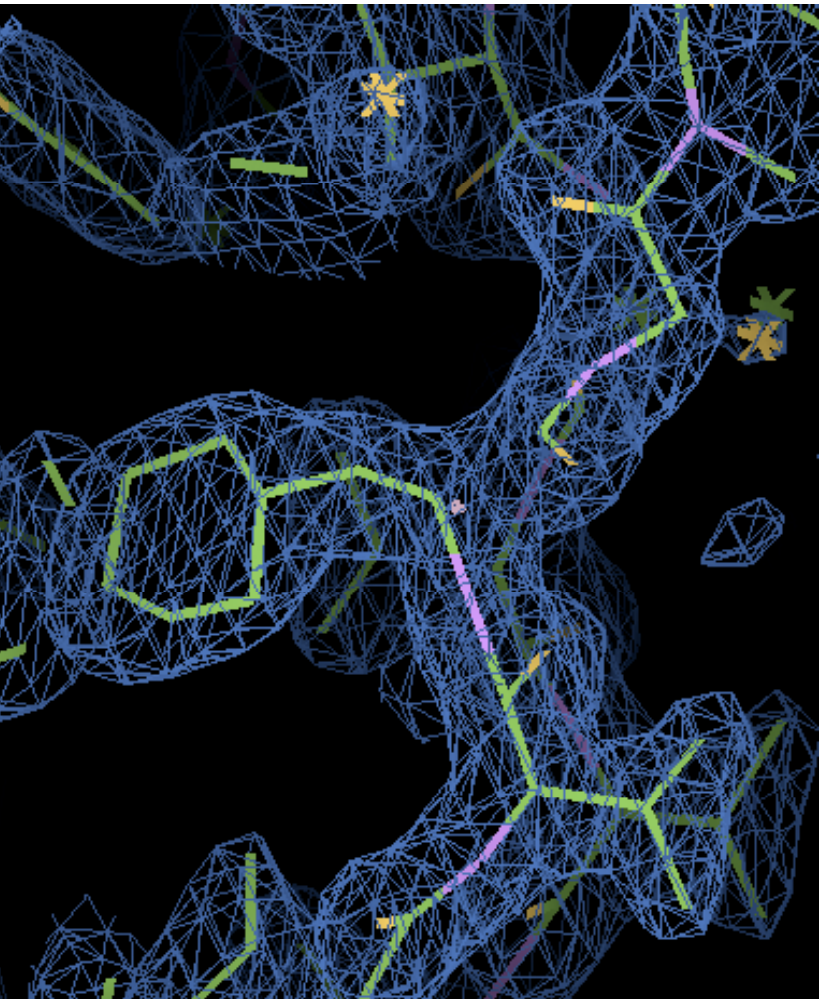
ole of restraints is that to retain chemistry of atoms and at e same time describe electron density optimally.

atoms are close to each other it is unlikely that they will ve hugely different B values

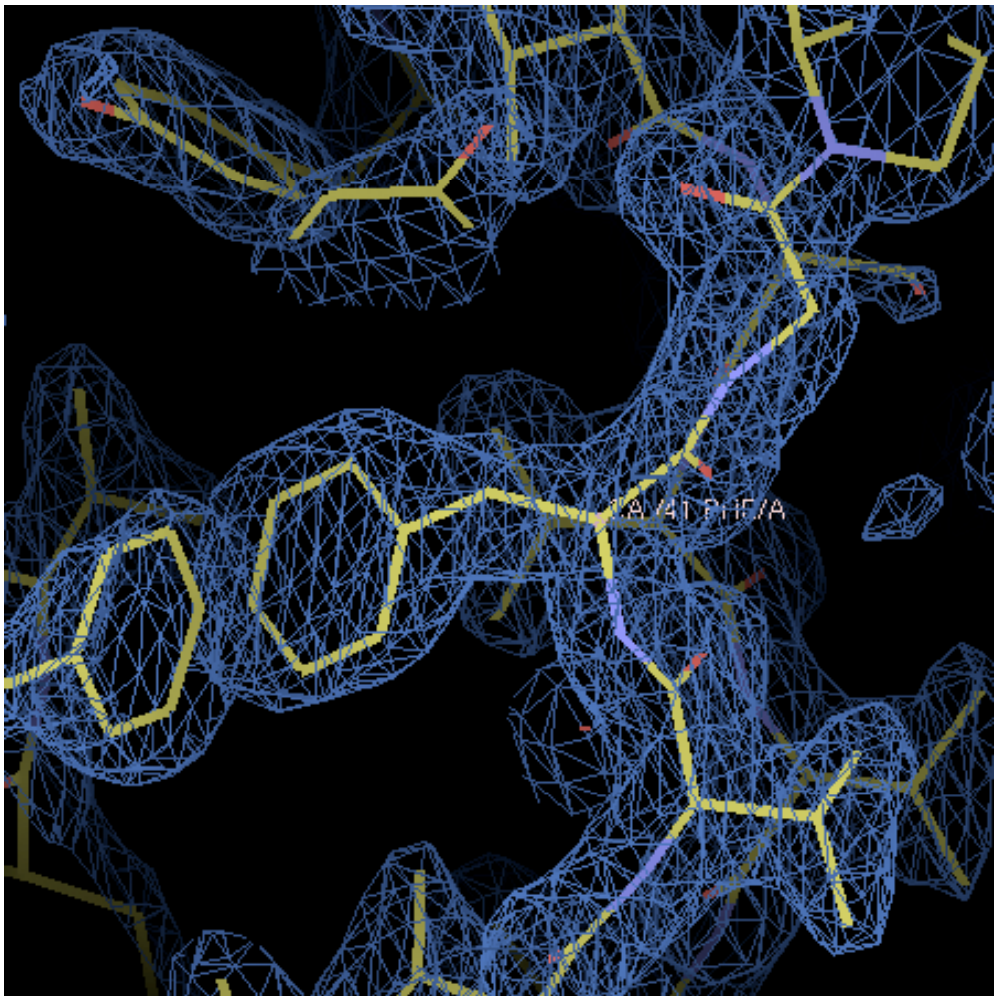
Example

a -
A

Unrestrained



Restrained



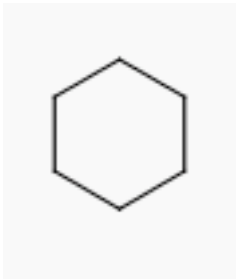
ard dictionary has description of around 1 500 small molecules. If one of the
crystal then the will be used automatically. In the new version there will be
000.

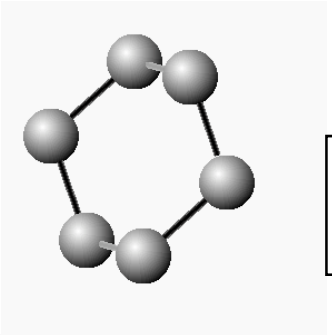
happens if you have a ligand that is not in the dictionary. Then it is your
sibility to create chemically sensible description.

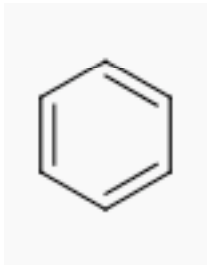
starting to create a description you need to study bonding structure of your

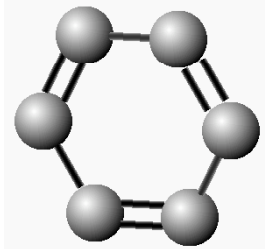
2D

3D









These two molecules will refine very differently (oviously)

various
e

d”

DrugBank

Home | Browse | Search | About | Downloads | Contact Us

Search: Search

The DrugBank database is a unique bioinformatics and cheminformatics resource that combines detailed drug (i.e. chemical, pharmacological and pharmaceutical) data with comprehensive drug target (i.e. sequence, structure, and pathway) information. The database contains nearly 4800 drug entries including >1,350 FDA-approved small molecule drugs, 123 FDA-approved biotech (protein/peptide) drugs, 71 nutraceuticals and >3,243 experimental drugs. Additionally, more than 2,500 non-redundant protein (i.e. drug target) sequences are linked to these FDA approved drug entries. Each DrugCard entry contains more than 100 data fields with half of the information being devoted to drug/chemical data and the other half devoted to drug target or protein data.

DrugBank is supported by David Wishart, Departments of Computing Science & Biological Sciences, University of Alberta.

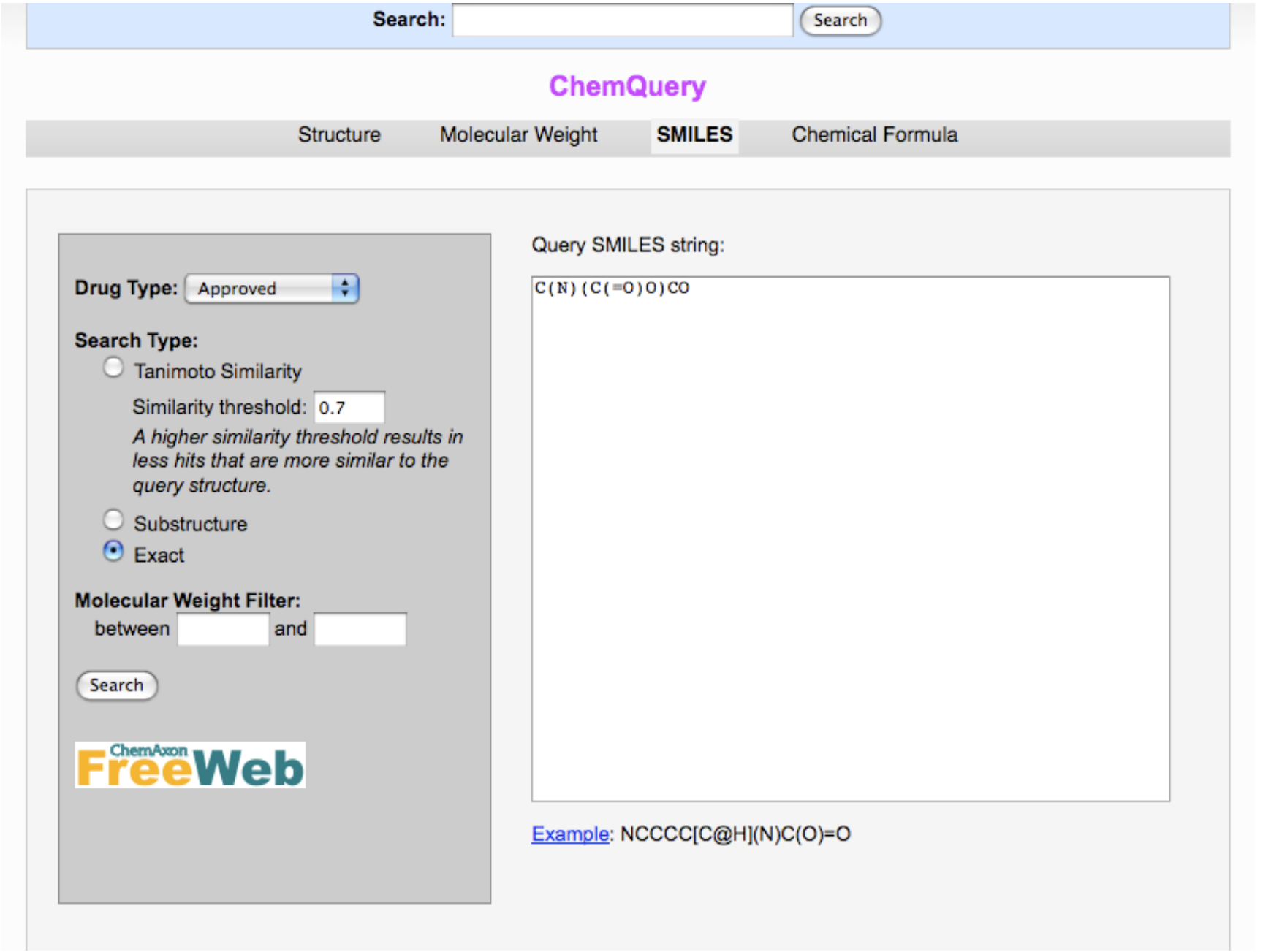
More about DrugBank

What's New?

- We have implemented the ChemAxon solution for structure searches. You can now perform similarity (tanimoto), substructure, and exact searches via the ChemQuery function. This system replaces an outdated structure search and should be faster and more accurate. We have only added the most basic features for this release, so if you would like to see more/different features added, please let us know.
- We have added a new page containing links to other useful drug and small molecule databases. The other databases page

an be
d using
tools. One
s smile
arch can be

ure



SMILES

S notation is the most popular notation and almost all computational
cal websites, programs use this notation. They can read and write
S.

sed on several simple rules. Full description of SMILES can be find from
nt websites.
www.daylight.com/dayhtml/doc/theory/theory.smiles.html

S stands for Simplified Molecular Input Line Entry System.

ncise and widely spread. It is very easy to learn. It was originally
ed for manual input using text only editors. SMILES has become as a
rd and it is a useful thing to know about.

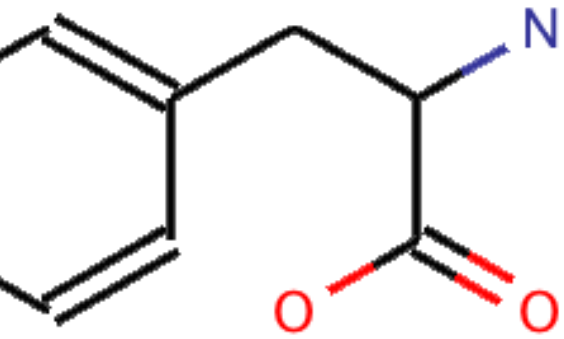
SMILES

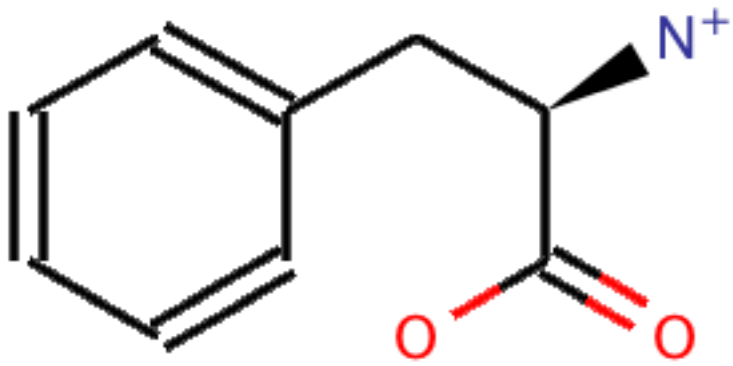
S uses several very simple rules (these rules are sufficient to generate
S from structure and structure from SMILES).

c symbols used for atoms
gen atoms as a rule are implicit. They are deduced using valence
ation about atoms
bouring atoms stand one after another
e, double, triple and aromatic bonds are denoted using “-”, “=”, “#” and “:”
tively. Single and aromatic bonds are usually not shown.
nes represented by parentheses
are added by using matching digits on connecting atoms
atic atoms are denoted using lower cases.

rules are sufficient to describe most of the cases. Let us consider some
les

SMILES





O)C(N)Cc1ccccc1

ens are implicit, aromatic and single are not shown. Stereochemistry is ar.

presentation is not unique. Here is ILES produced by MarvinSketch: 1=CC=CC=C1)C(O)=O

[NH3+][C@H](CC1=CC=CC=C1)C(O)=O

Explicit charges are shown as an attributer of atom. Stereochemistry is shown using @ or @@. Chirality and other stereochemical information in SMILES are local and can be understood using immediate neighbourhood of the atom for which @ or @@ symbols are defined.

SMILES: atom specification

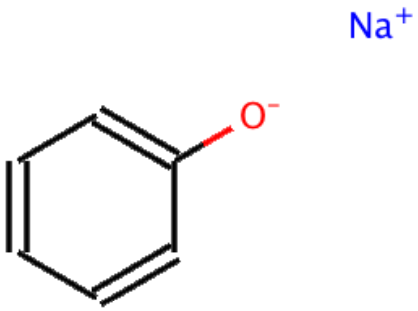
s are specified using their atomic symbol. Atoms that are not in the organic
et (C, N, O, S, P, F,Cl) or valency is different from “normal” or are isotopes o
ged then they are shown inside square brackets. General notation for atoms

ht atom chirality charge]
xample
- shows deuterium with atomic charge +1.
H+] - shows positively charged nitrogen atom that is a chiral centre
+2] - iron atom with atomic charge +2

ogen atoms are not usually specified. In this case the number of hydrogens
would satisfy “normal” valency of atoms is assumed. However hydrogens ca
ecified explicitly. The number of hydrogens is shown using a digit
diately after hydrogen. It should not cause problem with ring closure since
ogens can make only one bond (except in reaction intermediates).

SMILES: disconnected atoms

nsecutive atoms are not connected then between them “.” is added. For
nple:
].c1ccccc1[O-] id fc



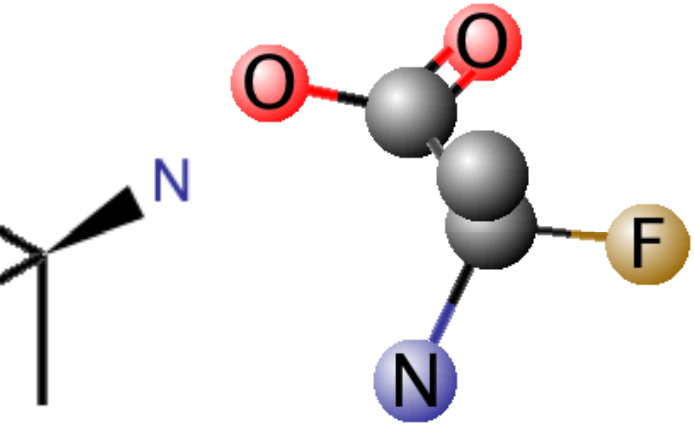
that if a SMILES string has “.” it does not mean that it is a disconnected
ture. For example: C1.C1 is same as CC (note that matching digits show
hese atoms must be connected). This notation is used in extension of
LES to represent reaction (e.g. reagents are separated by “.”)

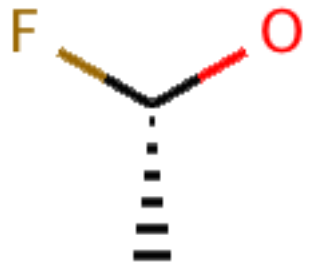
SMILES: stereochemistry

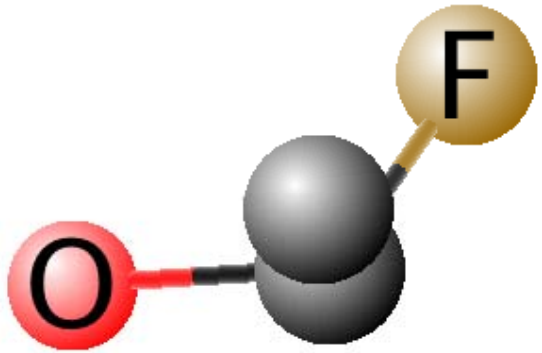
om makes bond with four atoms (e.g. sp^3 carbon) and all these atoms a
then we can arrange these atoms in general in two different ways. C
cannot be generated from another using rotations and translations only.
sh these two structures SMILES uses chirality notations - @ or @@. If
r an atom is specified then it means that if we take the first atom attached
om and the three remaining atoms appear anti-clockwise. For example
N[C@](C)(F)C(O)=O we look down to the chiral carbon from the first atom
thyl group, F and carboxyl group appear anti-clockwise. If we write @@ th
ear clockwise.
atom is not the first atom and it has an implicit hydrogen then it is taken as t
n. For example in case N[C@](C)C(O)=O neighbours H (implicit hydrogen)
nd carboxyl groups appear anti-clockwise. To avoid confusion in these case
n can be written explicitly.
atom is the first atom then implicit hydrogen is an atom “from” where we loo
iral atom.

ot find specification what happens if a chiral atom is the first atom and it has
cit hydrogens.

MarvinSketch







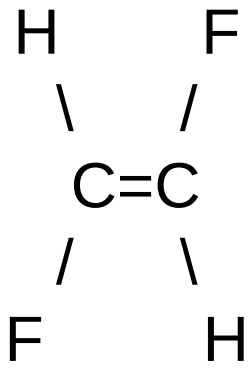
@](N)(F)C(O)=O

C[C@@H](O)F

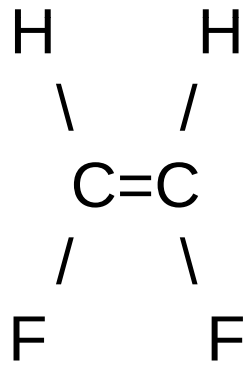
SMILES: stereochemistry

ration around double bonds are denoted using matching “\” and/or “/”. For
e

F (or F/C=C/F) denotes



F (or F/C=C\F) denotes



at SMILES chirality and are local chirality.

The list of compounds

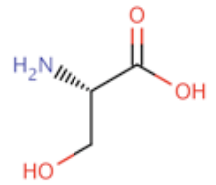
rious info
h
They can
sed using

on tools

es

ChemQuery Search Results

ChemQuery search returned 1 results.

DrugBank ID	Name / CAS Number	Structure	Formula	Average Weight / Monoisotopic Weight
DB00133 DRUGCARD	L-Serine / 56-45-1		$C_3H_7NO_3$	105.0926 / 105.0426
	Canonical SMILES: NC(CO)C(O)=O Isomeric SMILES: N[C@@H](CO)C(O)=O			

3D view

Showing drug card for L-Serine (DB00133)

target field | enzyme field | Show Similar Structures for Approved drugs

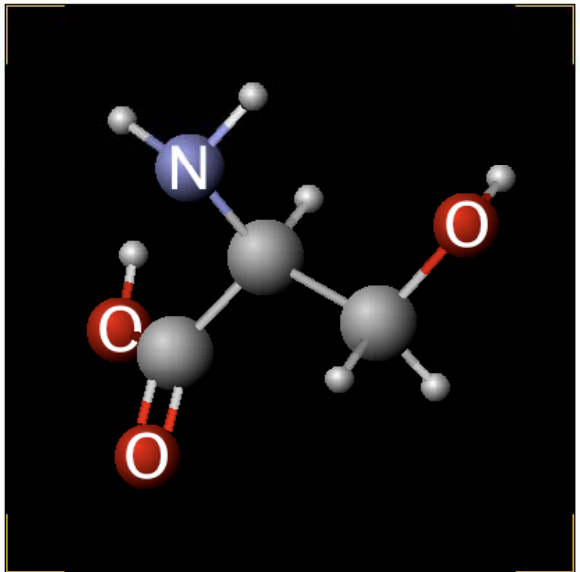
2.5
2005-06-13 13:24:05
2008-08-26 14:00:56
DB00133
• NUTR00053
L-Serine
• Approved • Nutraceutical • Small Molecule
A non-essential amino acid occurring in natural form as the L-isomer. It is synthesized from glycine or threonine. It is involved in the biosynthesis of purines; pyrimidines; and other amino acids. [PubChem]
1. (-)-Serine 2. (S)-2-Amino-3-hydroxypropanoic acid 3. (S)-Serine

Structure Viewer

Displaying structure for L-Serine

Double click on the structure to save and manipulate the structure and

Large molecules may take a while to load



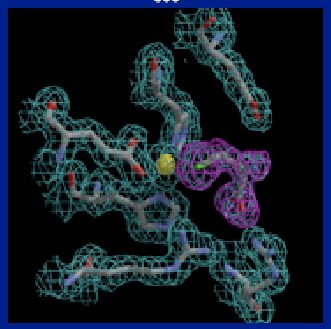
PRODRG server

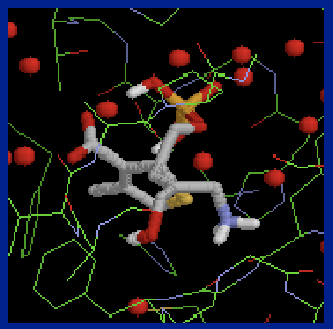
RODRG Beta How to obtain

The Dundee PRODRG2 Server

Finally, a FAQ is available here, READ it before using this server

gies for ... X-ray refinement/MD ... . drug design/docking





Funded by:

The Wellcome Trust

Draw Molecule With JME

... or ...

Paste your input here (PDB coordinates, MDL MOLfile, text drawing). See below for instructions.

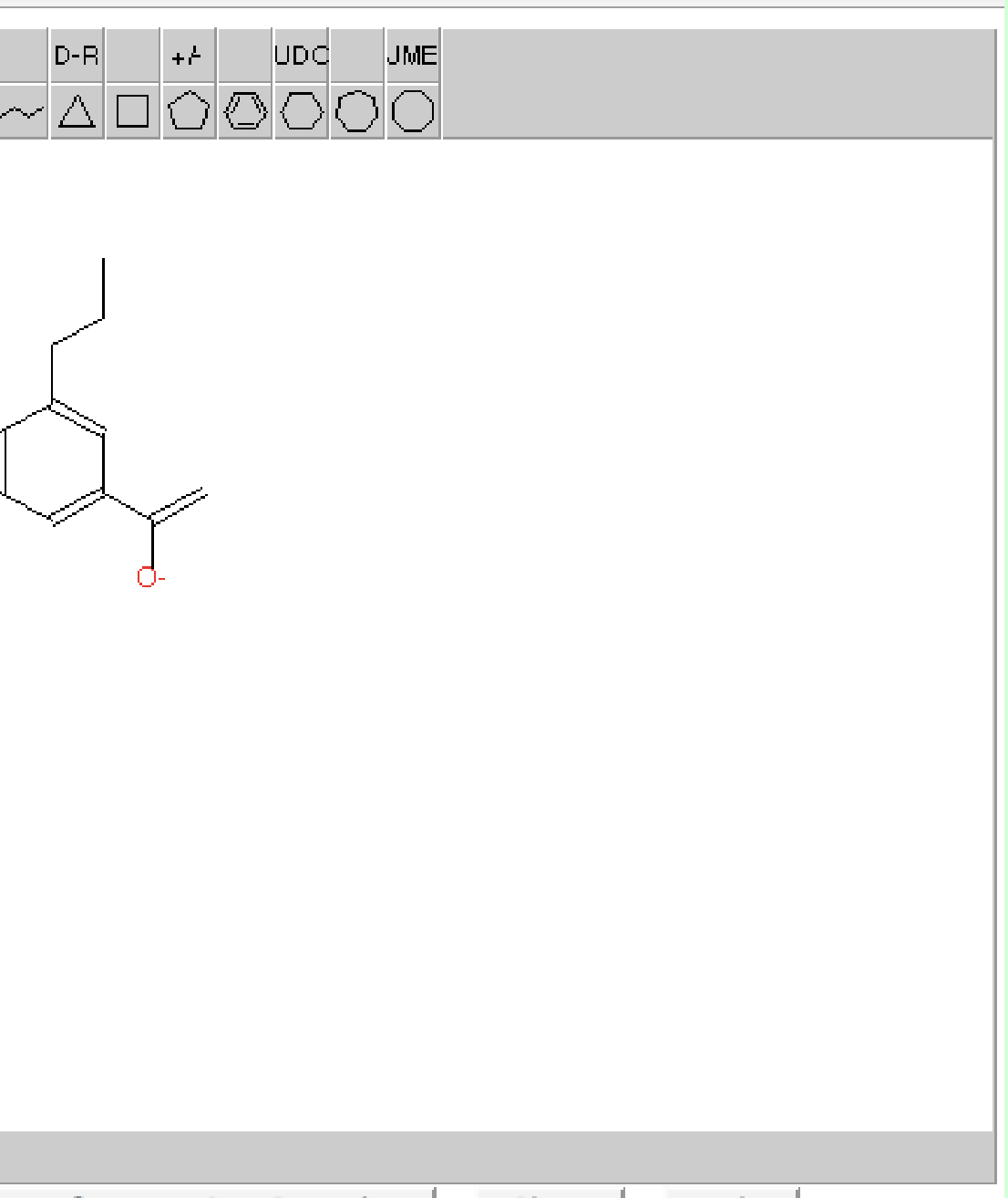
Chirality: Yes | Full charges: No | Energy minimization: Yes | Run PRODRG | Clear

Please be patient, this can take up to 2 minutes

JME

Load your

JME is java based program for drawing of small compounds. It used in PRODRG2, MSDchem



Draw your ligand, transfer to PRODRG window and run

RODRG Beta How to obtain

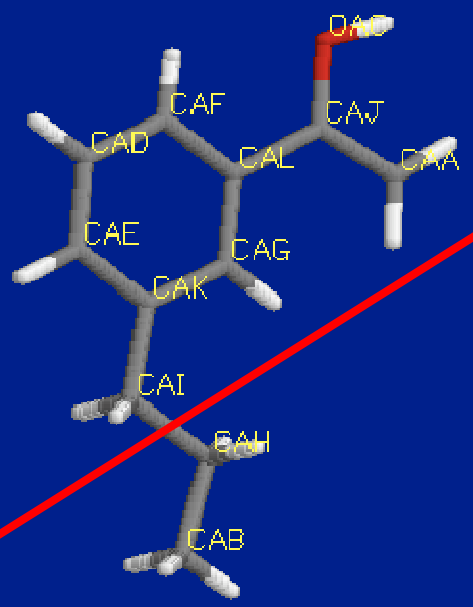
```
 up PRODRG version 061128.0522
ritten/copyrighted by Daan van Aalten
ander Schuettelkopf

s/comments to dava@davapc1.bioch.dundee.ac.uk

ng this software in a publication, cite:
huettelkopf and D. M. F. van Aalten (2004).
 a tool for high-throughput crystallography
in-ligand complexes.
stallogr. D60, 1355--1363.


 detected.
o information found in input file.
 complexity index: 2.00.
ogen(s) added.
s                  1 ambiguous
 angles            3 ambiguous
oper dihedrals     1 ambiguous
drals              0 ambiguous
ial charges        0 ambiguous
ge on molecule:    0.000
arge groups.
GROMACS topology.
topology quality on 0-10 scale:  7.7
ucture was iteration  841 with   0.70210928
 GROMACS version 3.2.1...
m GROMOS bond ideality (Angstrom) :   0.017
m GROMOS angle ideality (degrees) :   2.257
m GROMOS plane ideality (degrees) :   0.432
f improper improper dihedrals     :       0
 SCRHWMMPG
rogram end
```

Your molecule + added hydrogens



following output:

r H's only or no H's)
H's, polar H's only or no H's)
OMACS (polar H's only)

and topology)

entry, pre-9.x refi dictionary and 9.x dictionary)

It can write out repres
in various formats suit
various popular softw

B is Protein Data Bank. It has all macromolecular structures determined
erimentally as well as theoretically. There are more than 56000 macromole
ctures available in the PDB.

many cases protein structures are determined with some ligands (small mol
npounds). These small molecular structures are available from PDB. There
0 such small molecules in the PDB.

re are websites that allows people to view macromolecular structures as w
all molecular compounds. These sites are located in USA, Europe and Japa

nd Chemistry ? Energy types ?

sistent and enriched library of ligands, small
ecules and monomers that are referred as residues
hetgroups in any PDB entry (8702 currently in the
base - Release:03-2008_02_22)

Dchem : Molecule | Search | Reset

tter code	=		
e	=		
ecule name	like		
mula	formula range		edit
stereo smile	has substructure		edit
eo smile	exact stereo structure		edit
gments	fragment expression		edit
gerprint	common segments		edit

And
Or
ieve: Html

Substructure, common segment, exact stereo

CLR DEL D-R UDC JME

C N O S F Cl Br I P X

JME 2004.10

Load OK Cancel

or Load File Browse... **(SDF,MO**

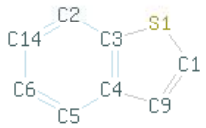
or give Smile String **(i.e. c1ccccc1)**

or give Code of Existing Molecule **(i.e. ATP)**

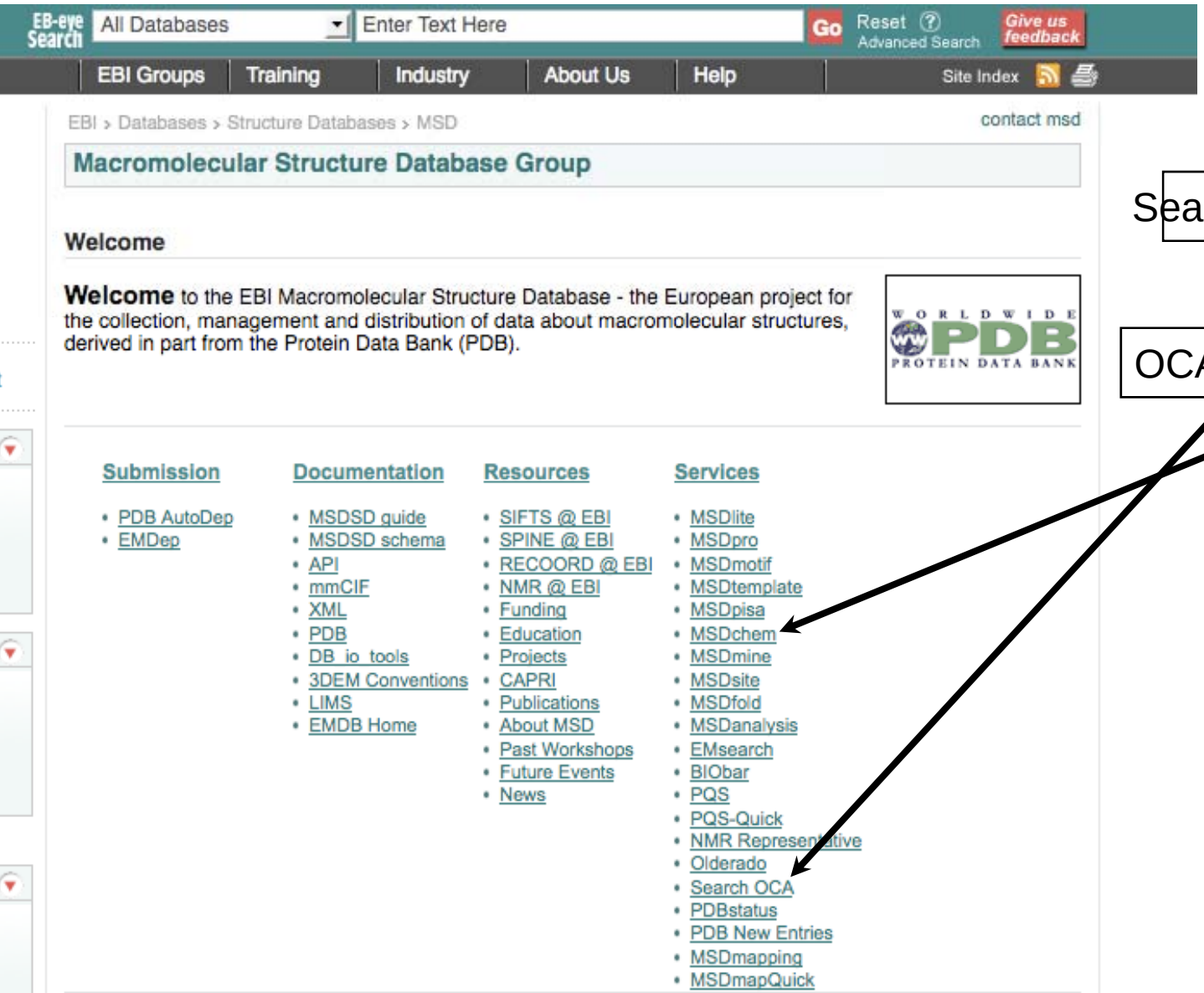
(Press the Load Button to Load the Molecule with that smile or 3 letter

Select chemical fragment pattern

acetylurea	acridine	acridone			
actinophenoxazine	adenine	alkaloid			
barbit	barbiturates	barbiturgroup			
benzimidazole	benzodiazepine	benzofuran			
benzoisoquinoline	benzothiadiazide	benzothiazole			
benzothiophen	benzoxazole	bilirubin			
biotin	carbazole	cephalosporin			
chromen	cinnoline	coumarine			
cyclobutane	cyclohexane	cyclopentane			
cyclopropane	cytosine	deoxyribose			
dibenzofuran	dibenzothiophen	dithiolane			
flavin	furan	furanose			
glycerophos	guanine	imidazole			
indole	inosine	isoquinoline			
isoxadiazole	isoxazole	naphtyridine			
napthalene	oxadiazole	oxazole			
oxazolidinedione	oxepin	peptide			
penicillin	phenanthrene	phenanthridine			
phenanthroline	phenazine	phenothiazine	phenyl	phthalazine	piperazine
porphin	prosto	pteridine	pteroyl	purine	pyran
pyranose	pyrazine	pyrazole	pyridazine	pyridine	pyrimidine
pyrole	quinazoline	quinoline	quinoxaline	rauwolfia	ribose
steroid	succinimide	thiadiazole	thiazole	thiepin	thiophen
tocol	vitaminAcore	xanthen			



Fragment: min: max: none: any: Add

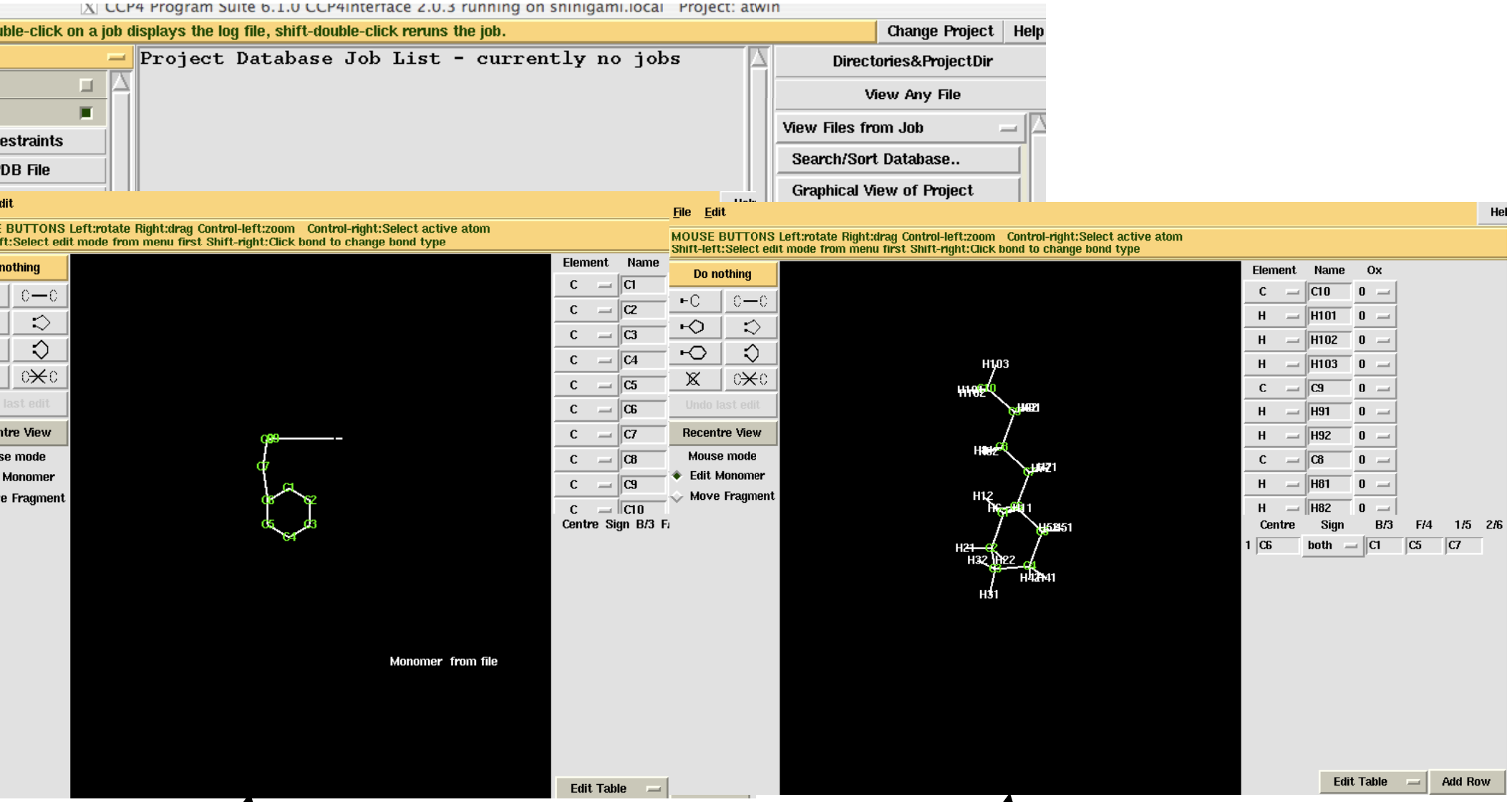


Search for macromolecules and sma

OCA

MSDchem

etcher is under Refinement/Restraint Preparation/Monomer library etcher.



etch your ligand

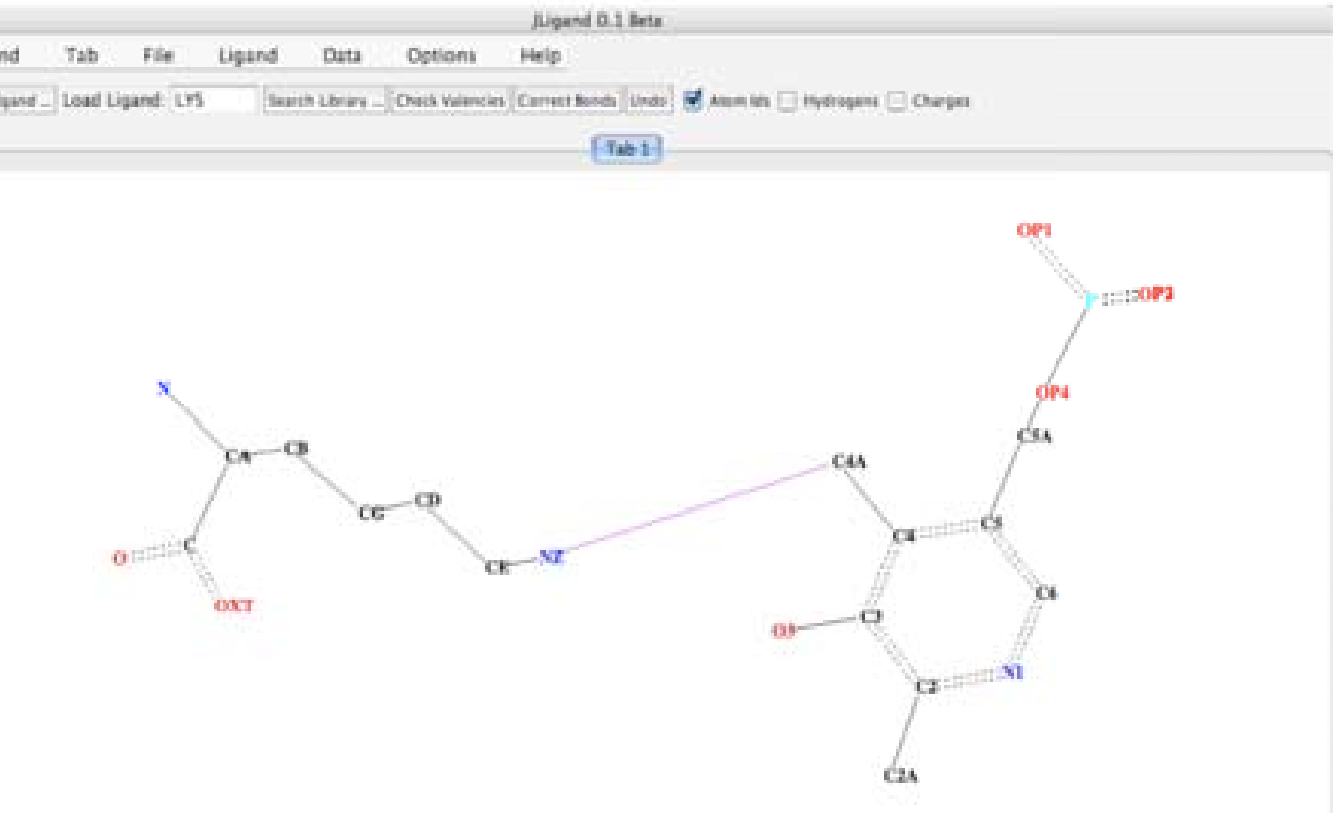
After regularisation

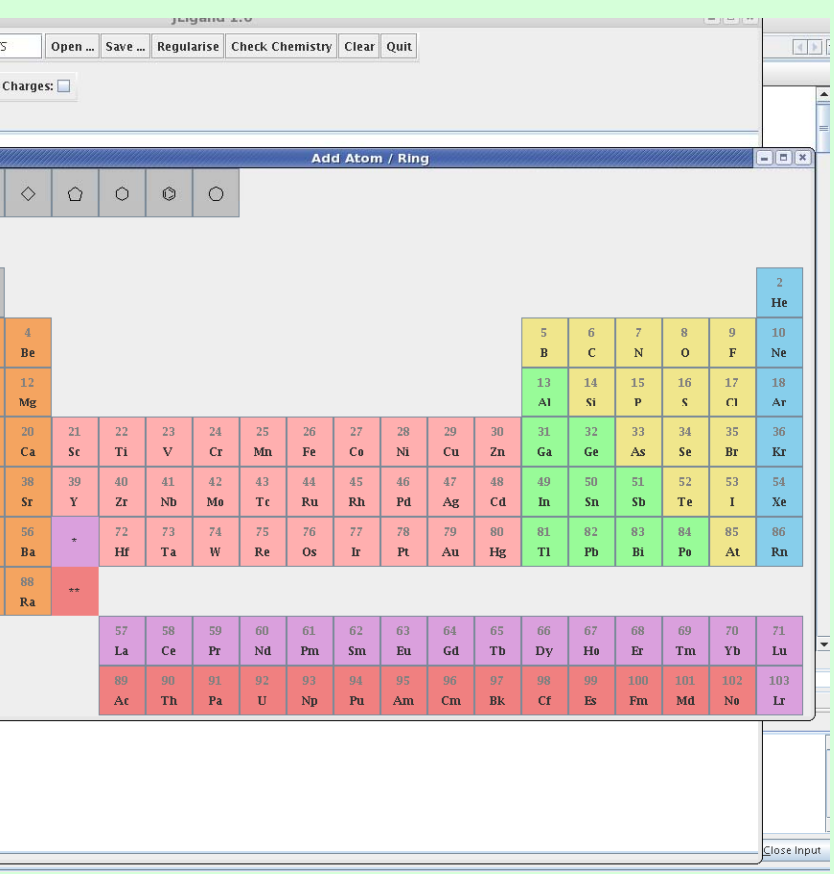
nd

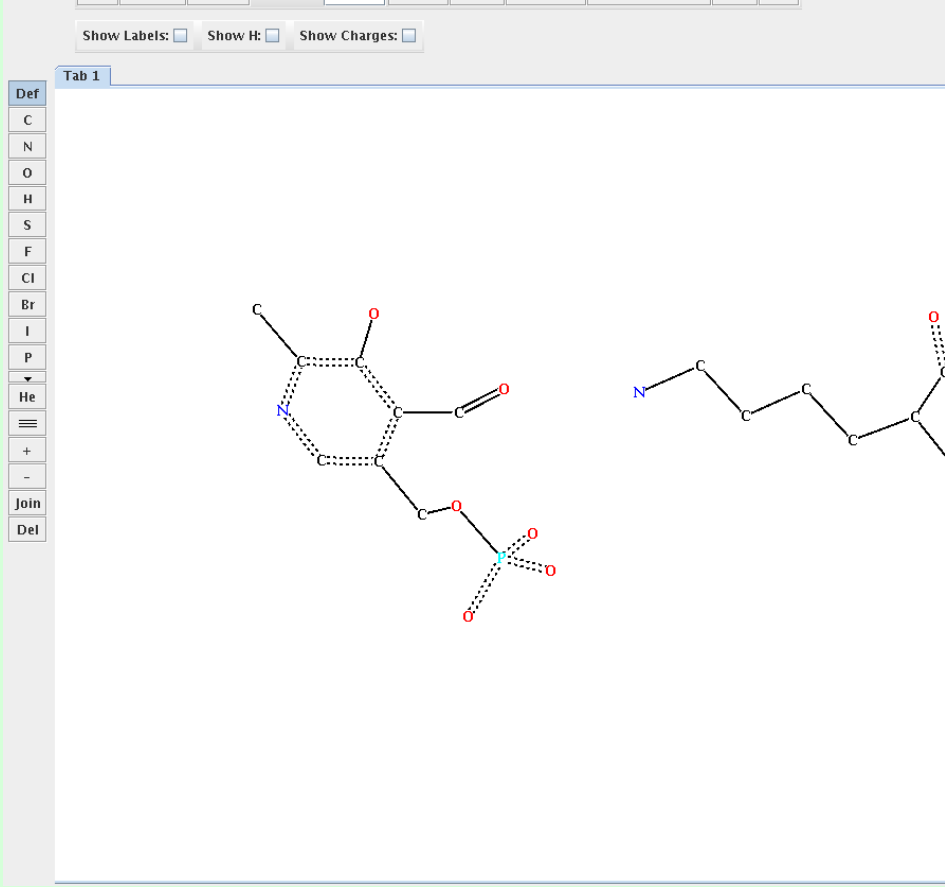
JLigand 0.1 Beta	Download
urce (Netbeans Project - Java 1.5)	java source
ecutable	java executable
file	README
Info	JLigand Info

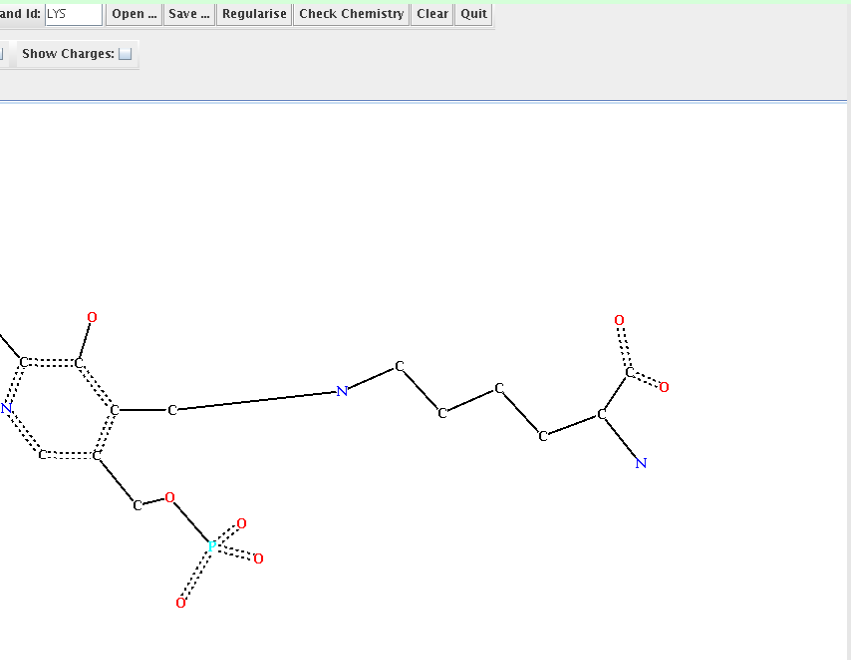
download the latest refmac, libcheck and monomers ligand dictionary from Garib's Refmac pages (Note: libcheck is part of the r

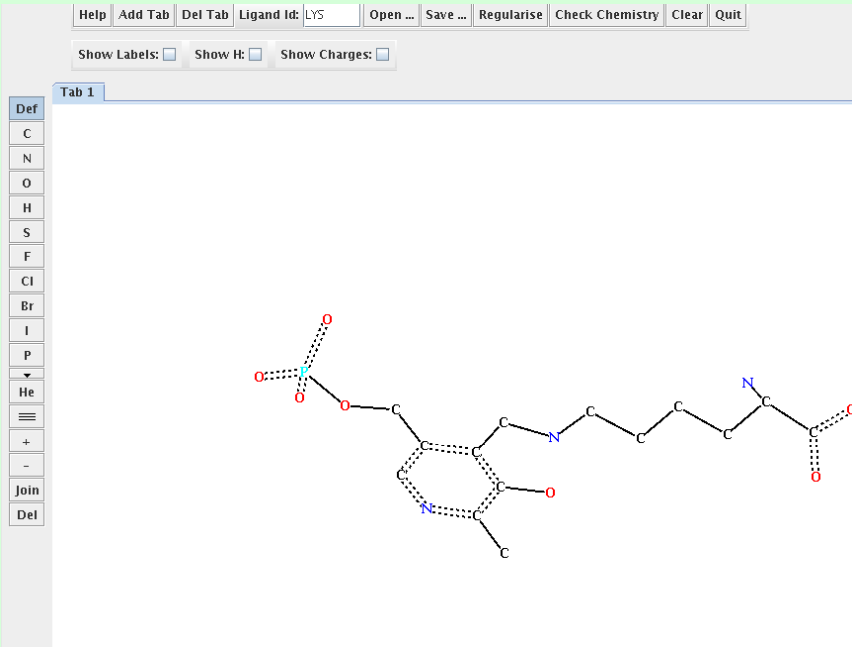
not:




Open ...
Save ...
Regularise
Check Chemistry
Clear
Quit
Charges:
Add Atom / Ring
Close Input


Show Labels:
Show H:
Show Charges:
Tab 1
Def
C
N
O
H
S
F
Cl
Br
I
P
He
+
-
Join
Del


Open ...
Save ...
Regularise
Check Chemistry
Clear
Quit
Show Charges:


Help
Add Tab
Del Tab
Ligand Id:
Open ...
Save ...
Regularise
Check Chemistry
Clear
Quit
Show Labels:
Show H:
Show Charges:
Tab 1
Def
C
N
O
H
S
F
Cl
Br
I
P
He
+
-
Join
Del

: JLigand 1.0
ion: ()
rsion: ()
Young, Andrey Lebedev, Alexei Vagin, Garib Murshidov

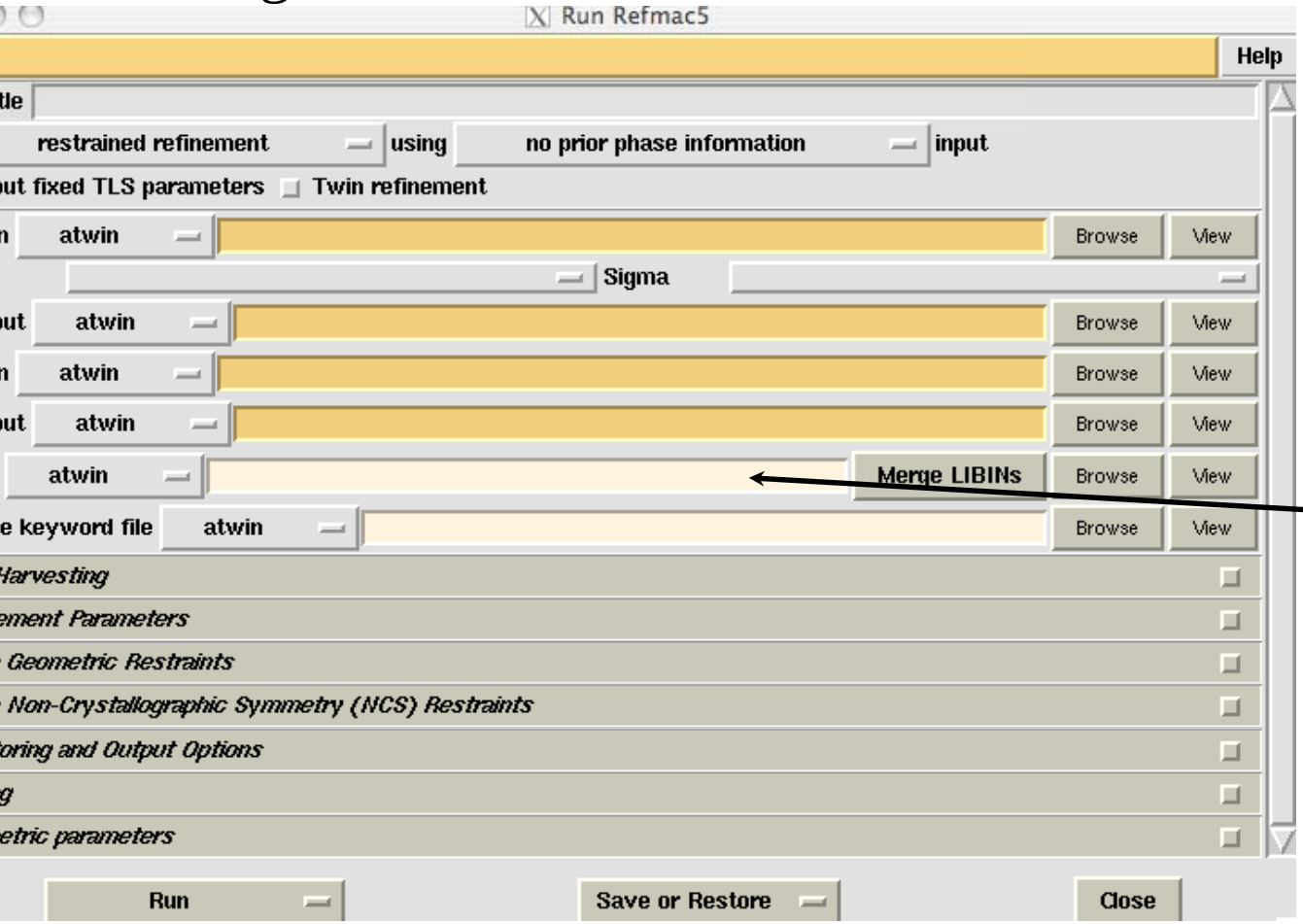
mp_id_1
d_id_1
up_comp_1
mp_id_2
d_id_2
up_comp_id_2
me
E PHEmod . PHE1 PHE1mod . 'PHEPHE1'

me
mp_id
oup_id
Emod' PHE .
HE1mod' PHE1 .

HE1

nd.link_id

ce satisfactory dictionary has been created they can be used in refinement and
del building. In refinement



Your dictionary should go here.