

# Handling ligands with PRODRG

Daan van Aalten

Division of Biological Chemistry and Molecular Microbiology

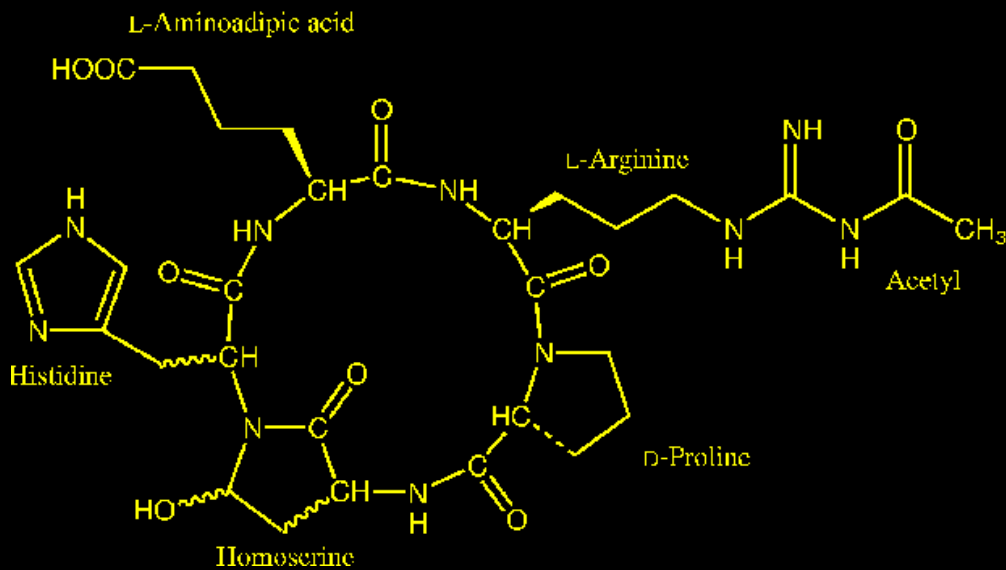


# Overview

- Definition of “the problem”
- Description of PRODRG
- What goes in and what comes out
- Building with PRODRG
- Example: from 0.5 mM to 50 nM



# What is the problem?



REFMAC/COOT?

CNS?

SHELX

O?



# What do we need?

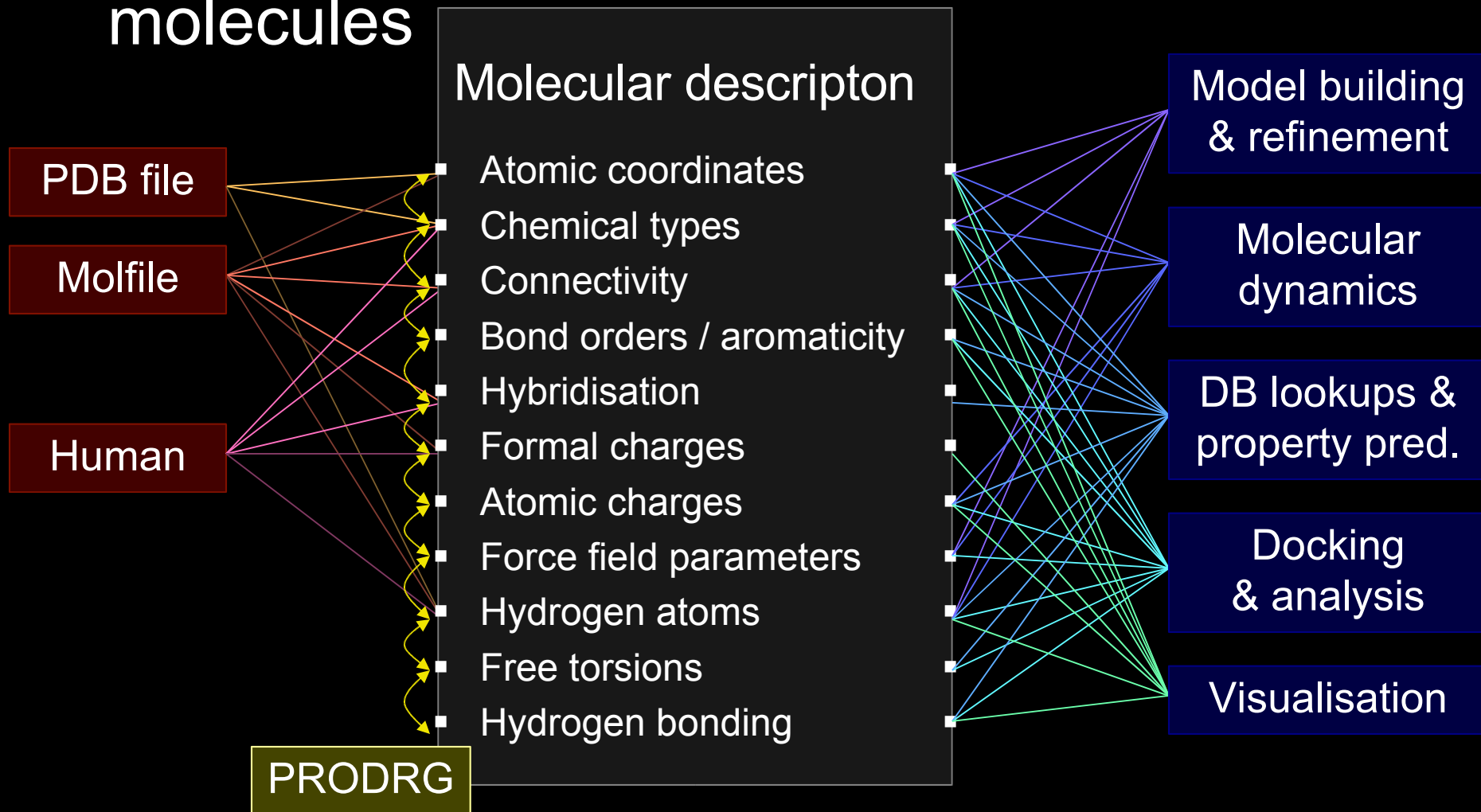
- Understand chemistry from different input formats (not just conversion)
- Chemical info -> coordinates
- Chemical info -> topologies
- Model building, refinement, docking, modelling, simulation and analysis
- Consistency – same chemical understanding of ligand throughout the various steps
- Automatic, fast & reliable
- Free and easily accessible

# What is PRODRG?

- Version 1 (1995)
  - Takes PDB file and generates ‘MOLDES’ and MD topologies
- Version 2 (2004)
  - Additional input formats
  - Additional output formats, including topologies for crystallographic software

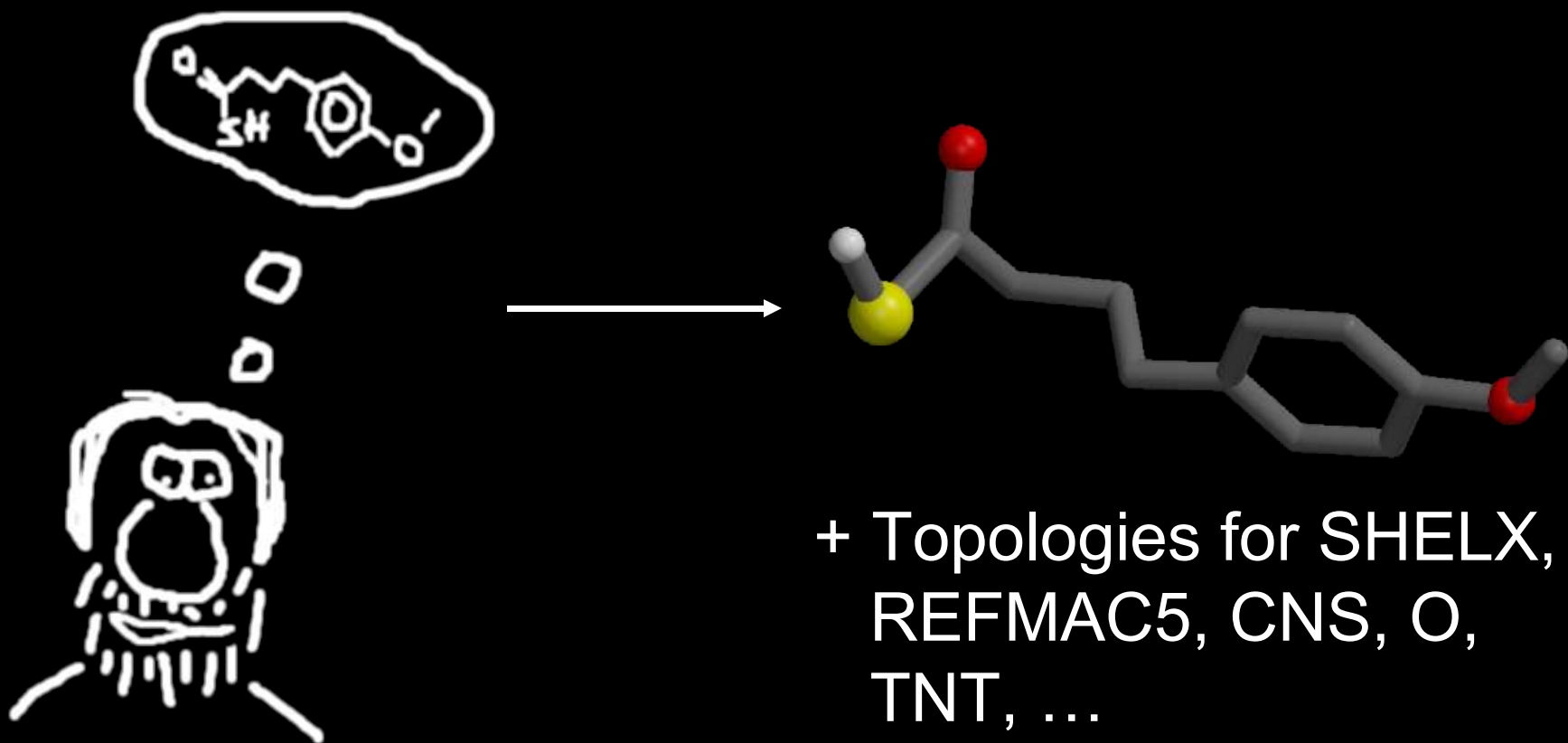
# What is PRODRG?

- Generates information about small molecules



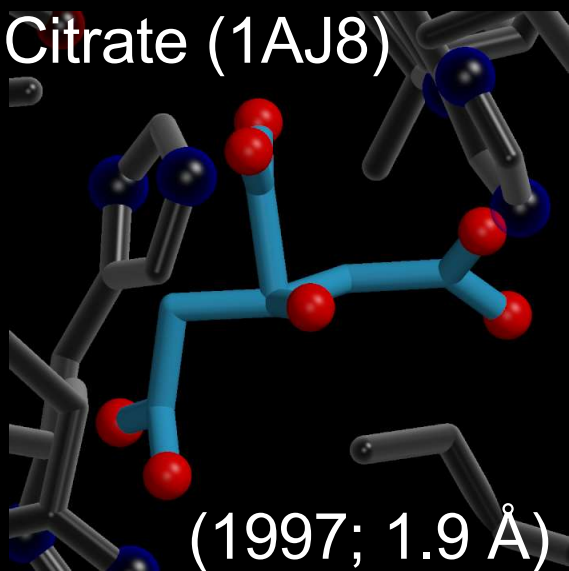
# How do I benefit from this?

- For many small molecules, we can go from imagination to a usable topology in minutes



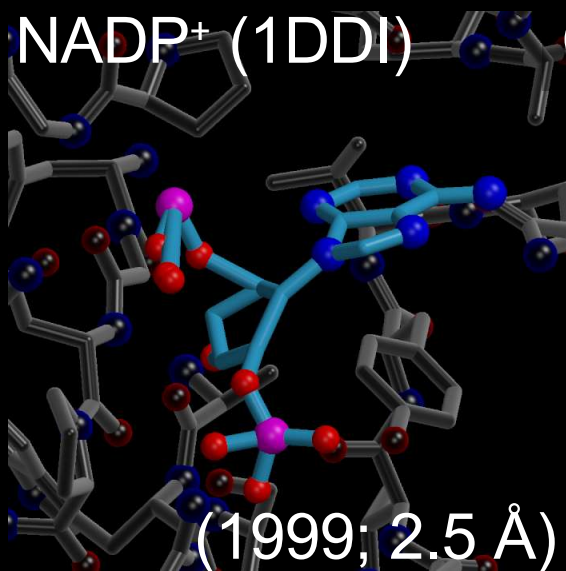
# Otherwise...

Citrate (1AJ8)



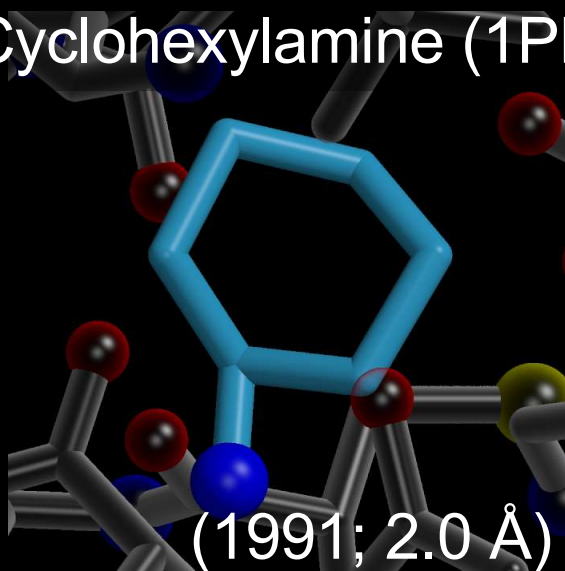
(1997; 1.9 Å)

NADP<sup>+</sup> (1DDI)



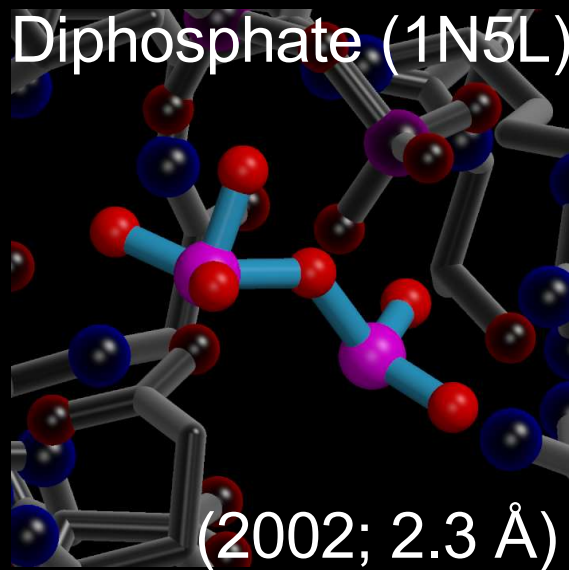
(1999; 2.5 Å)

Cyclohexylamine (1PPA)



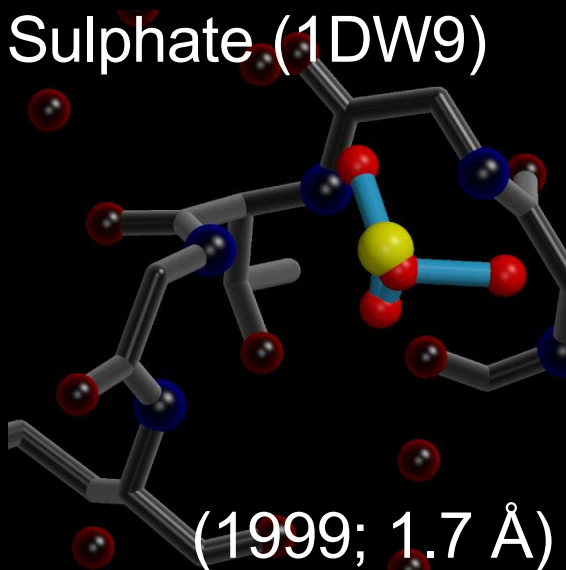
(1991; 2.0 Å)

Diphosphate (1N5L)



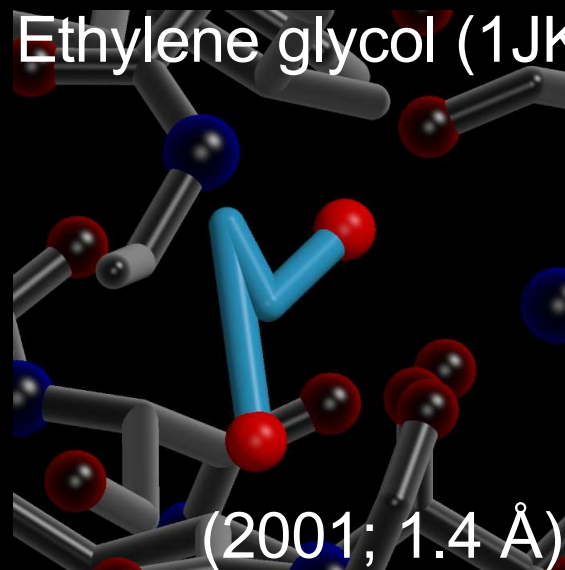
(2002; 2.3 Å)

Sulphate (1DW9)



(1999; 1.7 Å)

Ethylene glycol (1JKV)



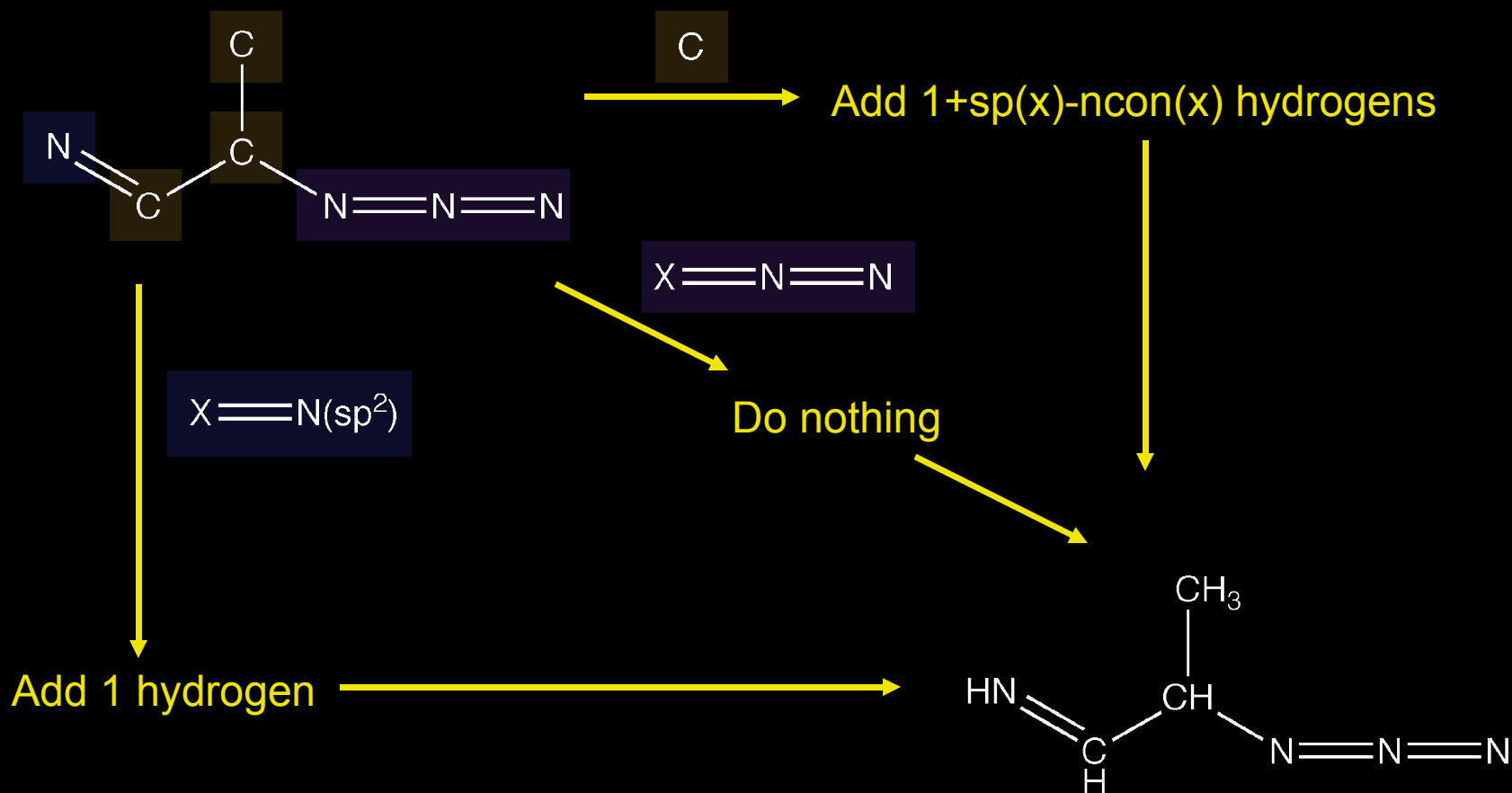
(2001; 1.4 Å)

# How does PRODRG work?

1. Analysis of input
2. Chemical typing  
(atoms/bonds)
3. Addition of hydrogens
4. Atom reordering
5. Topology generation
6. Formal and partial charges
7. Coordinates (minimisation)
8. Output

# How does PRODRG work?

- Most steps use 'chemical pattern matching'
- Example: hydrogen generation



# How does PRODRG work?

- Currently all Hs generated by 17 'rules'
- Chemical knowledge in data, not code
  - ⇒ More flexible
  - ⇒ Potentially user-configurable

# Limitations

- Few supported atom types
  - C,H,N,O,P,S,F,Cl,Br,I only
- Other chemical limitations
  - No more than 4 connections/atom
  - Standard version limited to  $\leq 300$  atoms
- Ignoring hydrogens and bond types may lead to unexpected results
- (Apolar hydrogens as second-class atoms)

# Basic usage: web server

- Four easy steps:

1. Go to <http://davapc1.bioch.dundee.ac.uk/programs/prodrg>

The screenshot shows a web browser window displaying the Dundee PRODRG2 Server interface. The browser title is "The Dundee PRODRG Server - Mozilla (Build ID: 2004031616)". The address bar shows the URL: <http://davapc1.bioch.dundee.ac.uk/programs/prodrg/prodrg.html>. The page has a blue background and features the title "The Dundee PRODRG2 Server" in yellow. Below the title, it says "Finally, a FAQ is available [here](#). READ it before using this server". There are three main sections: "Molecular topologies for ...", "... X-ray refinement/MD ...", and "... drug design/docking". Each section has a corresponding molecular structure image. A "Draw Molecule With JME" button is visible. Below this, there is a large white text area for input, with the instruction "Paste your input here (PDB coordinates, MDL MOLfile, text drawing). See below for instructions". At the bottom, there are checkboxes for "Chirality", "Full charges", and "Energy minimisation", each with a "Yes" or "No" option. There are also "Run PRODRG" and "Clear" buttons. A red message at the bottom says "Please be patient, this can take up to 2 minutes".

# Basic usage: web server

- Four easy steps:

1. Go to <http://davapc1.bioch.dundee.ac.uk/programs/prodrg>
2. Paste input

The Dundee PRODRG2 Server

Finally, a FAQ is available [here](#). READ it before using this server

Molecular topologies for ... X-ray refinement/MD ... drug design/docking

Funded by The Wellcome Trust

Draw Molecule With JME

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

```
C=C N
| |
-C-C=O
|
C---O
| |
C-C-C
| | |
O O C-O
```

Chirality Full charges Energy minimization  
Yes No Yes Run PRODRG Clear

Please be patient, this can take up to 2 minutes

# Basic usage: web server

- Four easy steps:

1. Go to <http://davapc1.bioch.dundee.ac.uk/programs/prodrg>

2. Paste input

3. Edit settings

- Chirality restraints?
- Reduced charges?
- Coordinates?

The Dundee PRODRG2 Server

Finally, a FAQ is available [here](#). READ it before using this server

Molecular topologies for ... X-ray refinement/MD ... drug design/docking

Funded by The Wellcome Trust

Draw Molecule With JME

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

```
C-C-C N
|
N-C-C-C=O
|
C---O
|
C-C-C
|
O O C-O
```

Chirality Full charges Energy minimization  
Yes No Yes

Run PRODRG Clear

Please be patient, this can take up to 2 minutes

Chirality Full charges Energy minimization

Yes

No

Yes

# Basic usage: web server

- Four easy steps:

1. Go to <http://davapc1.bioch.dundee.ac.uk/programs/prodrg>
2. Paste input
3. Edit settings
4. Run it

The Dundee PRODRG2 Server

Finally, a FAQ is available [here](#). READ it before using this server

Molecular topologies for ... X-ray refinement/MD ... drug design/docking

Draw Molecule With JME

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

```
C-C-C N
*   | |
N-C-C-C=O
|
C---O
| |
C-C-C
| | |
O O C-O
```

Chirality Full charges Energy minimisation  
Yes No Yes

Run PRODRG Clear

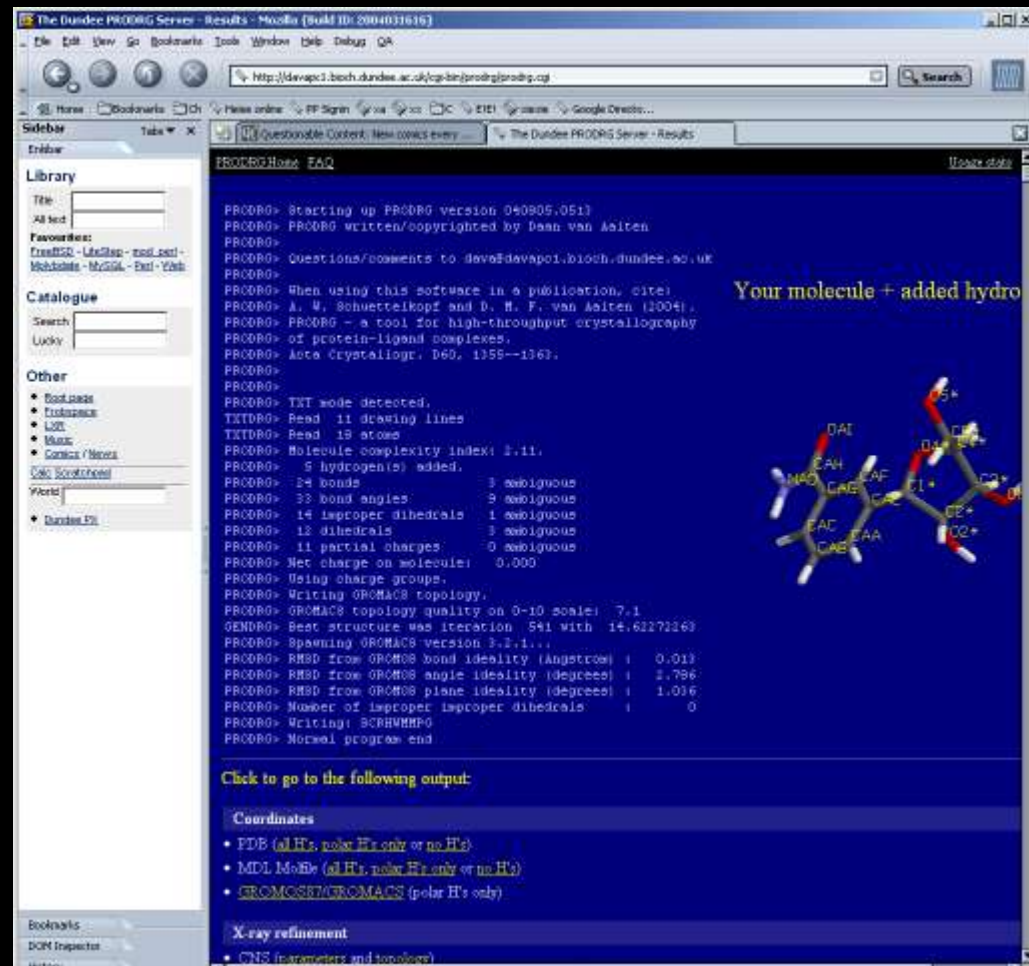
Please be patient, this can take up to 2 minutes

# Basic usage: web server

- Four easy steps:

1. Go to <http://davapc1.bioch.dundee.ac.uk/programs/prodrg>
2. Paste input
3. Edit settings
4. Run it

Success!



The screenshot shows a web browser window displaying the Dundee PRODrg Server interface. The browser address bar shows the URL <http://davapc1.bioch.dundee.ac.uk/programs/prodrg.cgi>. The page content is a blue background with white text and a 3D ball-and-stick model of a molecule. The text output is as follows:

```
PRODrg> Starting up PRODrg version 040905.051
PRODrg> PRODrg written/copyrighted by Daan van Asiten
PRODrg>
PRODrg> Questions/comments to dave@davapc1.bioch.dundee.ac.uk
PRODrg>
PRODrg> When using this software in a publication, cite:
PRODrg> A. W. Schuettelkopf and D. M. F. van Asiten (2004).
PRODrg> PRODrg - a tool for high-throughput crystallography
PRODrg> of protein-ligand complexes.
PRODrg> Acta Crystallogr. D60, 1355-1363.
PRODrg>
PRODrg>
PRODrg> TXT mode detected.
TXTDRG> Read 11 drawing lines
TXTDRG> Read 13 atoms
PRODrg> Molecule complexity index: 3.11.
PRODrg> 5 hydrogen(s) added.
PRODrg> 24 bonds
PRODrg> 33 bond angles
PRODrg> 14 isproper dihedrals
PRODrg> 12 dihedrals
PRODrg> 11 partial charges
PRODrg> Net charge on molecule: 0.000
PRODrg> Using charge groups.
PRODrg> Writing ORMAC8 topology.
PRODrg> ORMAC8 topology quality on 0-10 scale: 7.1
GENDRG> Best structure was iteration 541 with 14.62273269
PRODrg> Spawning ORMAC8 version 3.2.1...
PRODrg> RMSD from ORMAC8 bond ideality (Angstroms) | 0.013
PRODrg> RMSD from ORMAC8 angle ideality (degrees) | 1.796
PRODrg> RMSD from ORMAC8 plane ideality (degrees) | 1.036
PRODrg> Number of isproper isproper dihedrals | 0
PRODrg> Writing: BCRHVMF0
PRODrg> Normal program end
```

Below the text output, there is a section titled "Click to go to the following output:" with three bullet points:

- [Coordinates](#)
- [PDB \(all H's, polar H's only or no H's\)](#)
- [MDL Molfile \(all H's, polar H's only or no H's\)](#)
- [ORMAC8SET/ORMAC8 \(polar H's only\)](#)

At the bottom, there is a section titled "X-ray refinement" with one bullet point:

- [CNS \(parameters and topology\)](#)

On the right side of the page, there is a 3D ball-and-stick model of a molecule with the text "Your molecule + added hydro" above it. The model shows a central carbon atom bonded to several other atoms, including oxygen and nitrogen, with hydrogen atoms added to complete the structure.

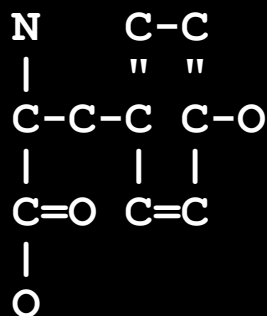
# Input: PDB

- PDB coordinates (discouraged!)

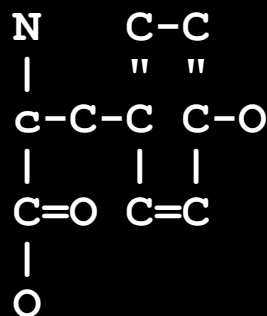
ATOM	2217	CD	GLU	C	539	-10.874	0.063	44.507	1.00	0.00	C
ATOM	2218	OE1	GLU	C	539	-11.105	-0.322	43.337	1.00	0.00	O
ATOM	2219	OE2	GLU	C	539	-11.357	-0.504	45.510	1.00	0.00	O
ATOM	2220	N	ALA	C	540	-10.551	4.531	41.396	1.00	0.00	N
ATOM	2221	CA	ALA	C	540	-11.155	5.596	40.651	1.00	0.00	C
ATOM	2222	C	ALA	C	540	-11.368	4.997	39.260	1.00	0.00	C
ATOM	2223	O	ALA	C	540	-11.560	5.686	38.255	1.00	0.00	O
ATOM	2224	CB	ALA	C	540	-10.274	6.827	40.570	1.00	0.00	C
ATOM	2225	N	THR	C	541	-11.318	3.665	39.263	1.00	0.00	N
ATOM	2226	CA	THR	C	541	-11.498	2.806	38.097	1.00	0.00	C
ATOM	2227	C	THR	C	541	-10.515	3.041	36.956	1.00	0.00	C
ATOM	2228	O	THR	C	541	-10.444	4.134	36.389	1.00	0.00	O
ATOM	2229	CB	THR	C	541	-12.950	2.881	37.567	1.00	0.00	C
ATOM	2230	OG1	THR	C	541	-13.446	4.220	37.688	1.00	0.00	O

# Input: Text Drawing

- Atoms represented by their element symbols
- Connected by bonds
  - Single: – or |
  - Double: = or "
  - Triple: #
- Change case of symbol to invert chirality



D-Tyr



L-Tyr

# Input: JME Editor

The Dundee PRODRG Server

http://davapc1.bioch.dundee.ac.uk/programs/prodrgr/

Energy and ...products UK Apple .Mac Amazon eBay Yahoo! News

PRODRG Home [FAQ](#) [PRODRG Beta](#) [How to obtain](#)

JME Molecular Editor

The screenshot shows the JME Molecular Editor window. The main canvas displays a chemical structure of a bicyclic amine. It consists of a cyclopentane ring fused to a cyclobutane ring. The cyclopentane ring has a bromine atom (Br) at the 2-position and is substituted at the 1-position with a 2-amino-1-cyclobutyl group. The cyclobutane ring has an amino group (NH2) at the 1-position. The editor includes a toolbar with icons for clearing (CLR), deleting (DEL), undo (D-R), redo (+/-), loading (LDO), and saving (JME). A vertical element menu on the left lists atoms: C, N, O, S, F, Cl, Br, I, P, X. Below the canvas are buttons for 'Transfer To PRODRG Window', 'Close', and 'Help'. At the bottom of the editor window, it says 'JME Editor courtesy of Peter Ertl, Novartis'. The background is a blue website page with 'Molecular' and 'G2 Se' visible.

Transfer To PRODRG Window Close Help

[JME Editor](#) courtesy of Peter Ertl, Novartis

Chirality Full charges Energy minimization

Yes No Yes

Run PRODRG Clear

# Input: MDL Molfile

-ISIS- 11030314422D

```
13 13 0 0 0 0 0 0 0 0999 V2000
  2,4089 -1,8503 0,0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  1,9977 -2,5631 0,0000 N 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  1,1713 -2,5603 0,0000 N 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  0,7642 -1,8485 0,0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  1,9982 -1,1391 0,0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  1,1782 -1,1427 0,0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  0,7669 -0,4356 0,0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  1,1744 0,2757 0,0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  1,9974 0,2754 0,0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  2,4051 -0,4324 0,0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  3,2339 -1,8501 0,0000 N 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  3,6466 -2,5645 0,0000 N 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  4,2792 -1,2792 0,0000 C1 0 0 0 0 0 0 0 0 0 0 0 0 0 0
  5 6 1 0 0 0 0
  6 7 2 0 0 0 0
  2 3 1 0 0 0 0
  7 8 1 0 0 0 0
  5 1 1 0 0 0 0
  8 9 2 0 0 0 0
  3 4 2 0 0 0 0
  9 10 1 0 0 0 0
  10 5 2 0 0 0 0
  4 6 1 0 0 0 0
  1 11 1 0 0 0 0
  11 12 1 0 0 0 0
  1 2 2 0 0 0 0
M END
```

# Controlling PRODRG

- Normally things work plug and pray
- Additional commands/hints in input file:
  - PATCH (hybridisation)
  - INSHYD and DELHYD
  - PATCH (chirality)
  - PATCH (torsions)
  - BUILD

# Hybridisation hints

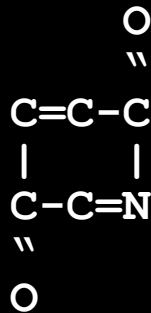
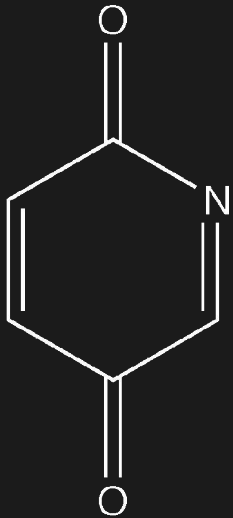
PATCH <atom> <number>

Number	Hybridisation
1	sp (triple bond)
2	sp <sup>2</sup>
3	sp <sup>3</sup>
10	sp (allene)
20	sp <sup>2</sup> (amide N)
21	sp <sup>2</sup> (not amide N)

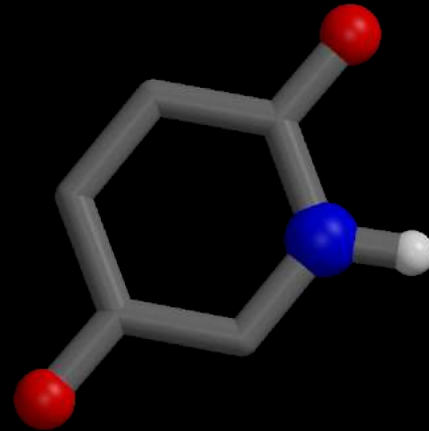
# Hybridisation hints

PATCH <atom> <number>

- Useful if PDB analysis did not quite work
- Allows to nudge PRODRG in right direction:



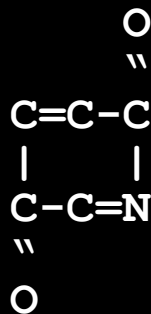
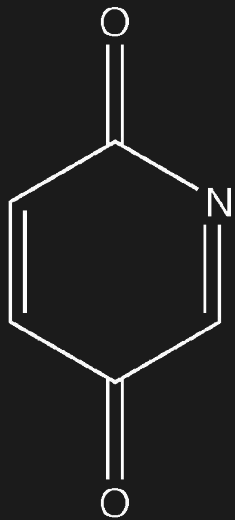
PRODRG> WARNING: multiplicity of generated molecule is not 1.  
PRODRG> WARNING: bond type assignment failed at CAF .



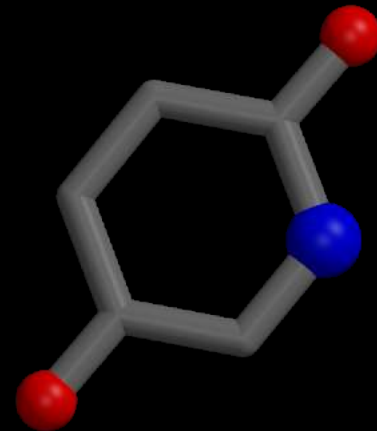
# Hybridisation hints

PATCH <atom> <number>

- Useful if PDB analysis did not quite work
- Allows to nudge PRODRG in right direction:



PATCH NAG 21

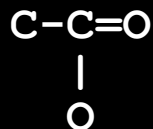


# Adding/removing hydrogens

INSHYD <atom>

DELHYD <atom>

- Allows to override default protonation
- Often not actually what you want



PRODRG> Cannot assign type to atom 'OAD'.  
ERRDRG> Error in GROMOS atom names/types.  
PRODRG> Drug topology not made, sorry!

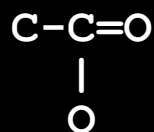
INSHYD OAD

# Adding/removing hydrogens

INSHYD <atom>

DELHYD <atom>

- Allows to override default protonation
- Often not actually what you want



PATCH OAD 3

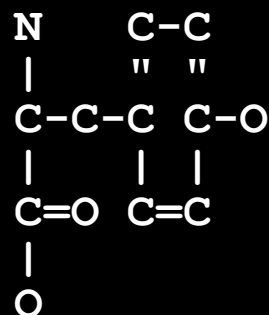
# Modifying chirality

PATCH <atom> -1

- Inverts stereocenter <atom>, useful for PDB input

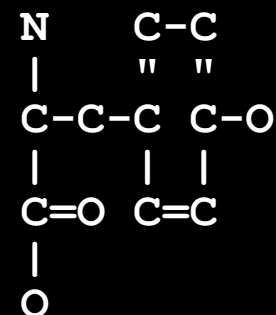
PATCH <atom> <pattern>

- 'Absolute' chirality for certain classes of molecules



PATCH CA L

L-Tyr



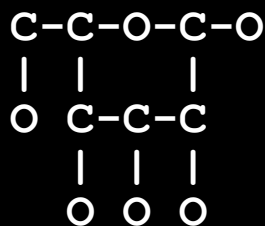
PATCH CA D

D-Tyr

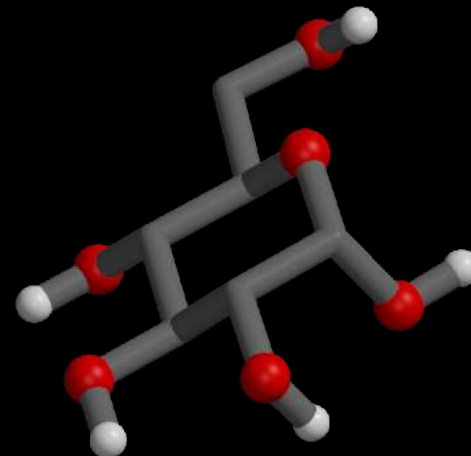
# Adding dihedral restraints

PATCH <atom> ><pattern>

- After EM pyranose rings often found in undesirable conformations
- PATCH statement introduces additional dihedral restraints to fix conformation



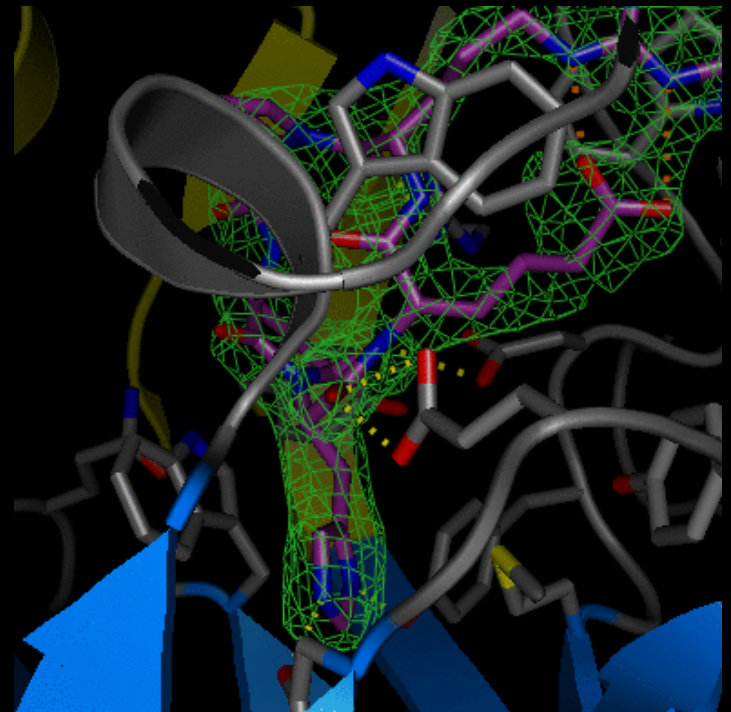
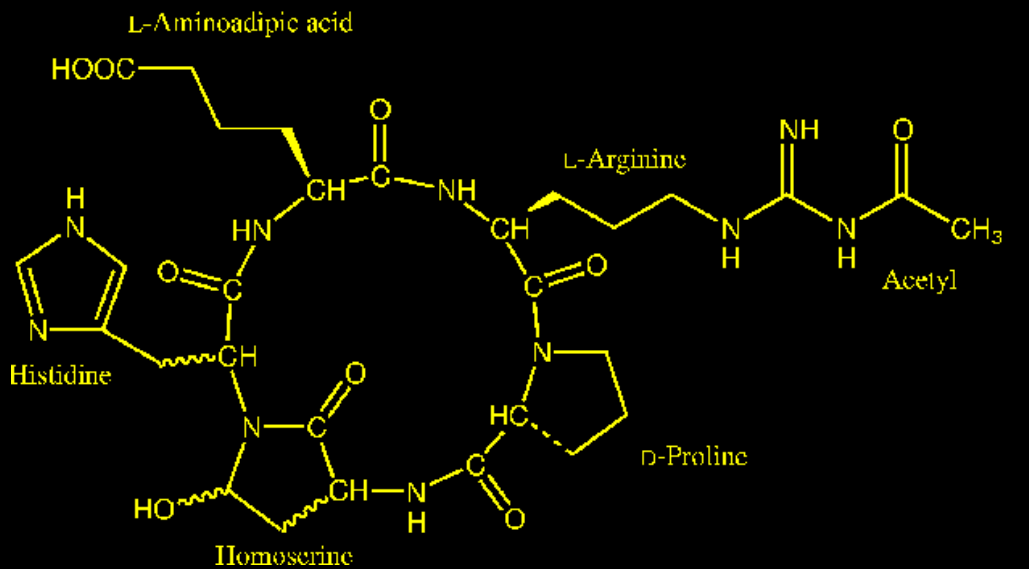
```
PATCH C1 ALPHA
PATCH C2 D
PATCH C3 L
PATCH C4 D
PATCH C5 D
PATCH C1 >4C1
```



$\alpha$ -D-Glucose

# Output

- Coordinates: PDB, GRO (etc.)
- Topologies: SDF, MOL2, O, REFMAC, Coot, CNS, SHELX, GROMOS, GROMACS, WHATIF, AUTODOCK (etc.)



# Protein-ligand H-bonds

- PRODRG plug-in for WHAT IF
- Automatic protein-ligand H-bond tables

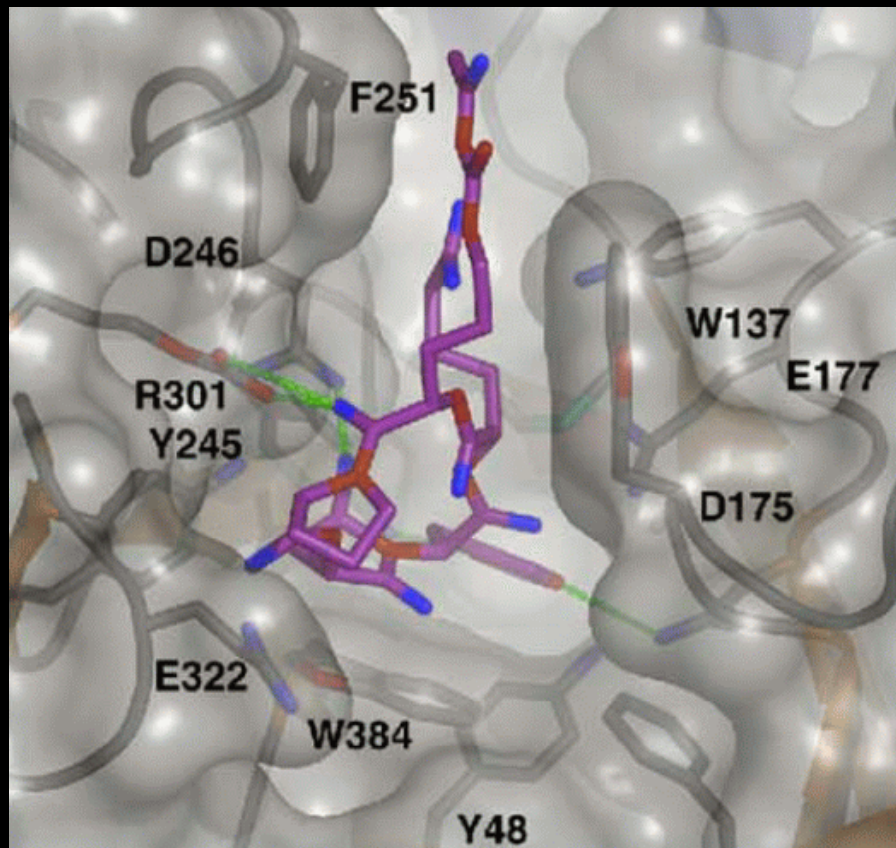


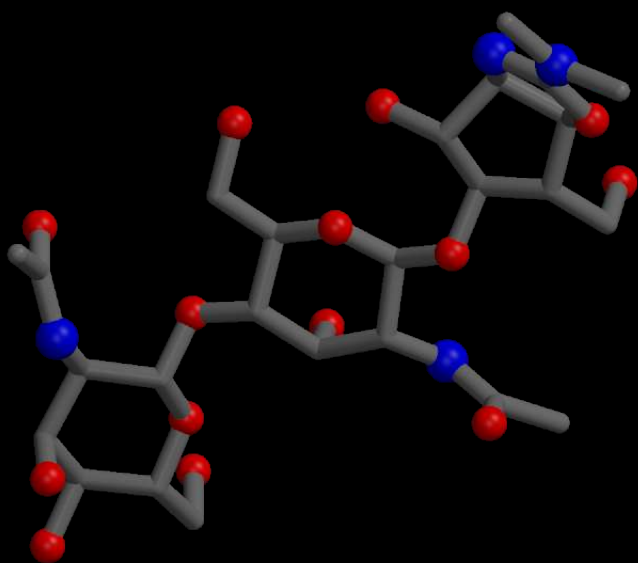
Table 1. Details of Argifin Binding to Family 18 Chitinases

Argifin	AfChiB1
Binding/inhibition ( $\mu\text{M}$ )	$K_d = 0.46/\text{IC}_{50}=1.1$
Buried surface ( $\text{\AA}^2$ )	112
Internal energy (kJ/mol)	-212
H bonds (protein)	
3,O $\delta$ 1	W137,N $\epsilon$ 1 (0.80)
1,O $\iota$ 1	Y245,O $\eta$ (0.66)
1,O	R301,N $\eta$ 1 (0.16)
1,O	R301,N $\eta$ 2 (0.44)
1,N $\iota$ 2	D175,O $\delta$ 2 (0.55)
1,N $\iota$ 2	E177,O $\epsilon$ 2 (0.37)
1,N $\eta$ 1	E177,O $\epsilon$ 2 (0.70)
1,N $\eta$ 2	D246,O $\delta$ 2 (0.53)
1,N $\epsilon$	E177,O $\epsilon$ 1 (0.71)



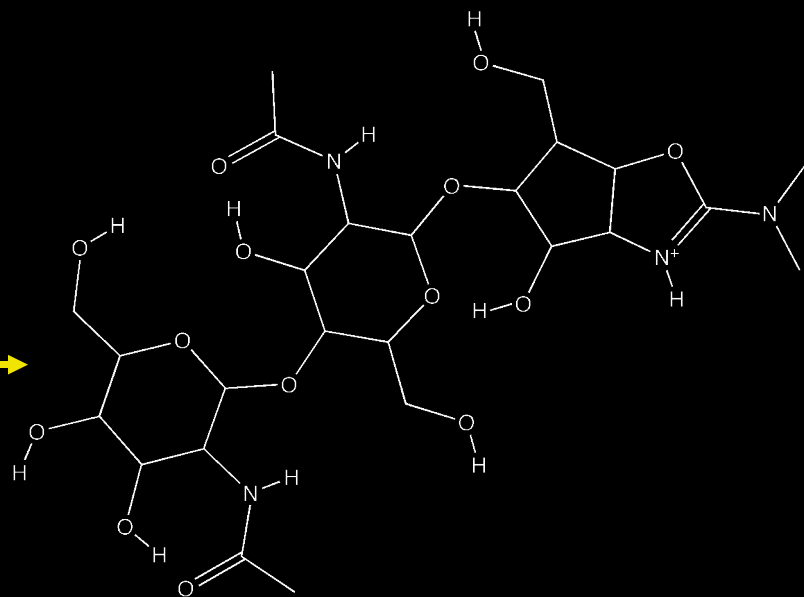
# Generating 2D structures

- Modified (internal) coordinate generator



Allosamidin

PRODRG

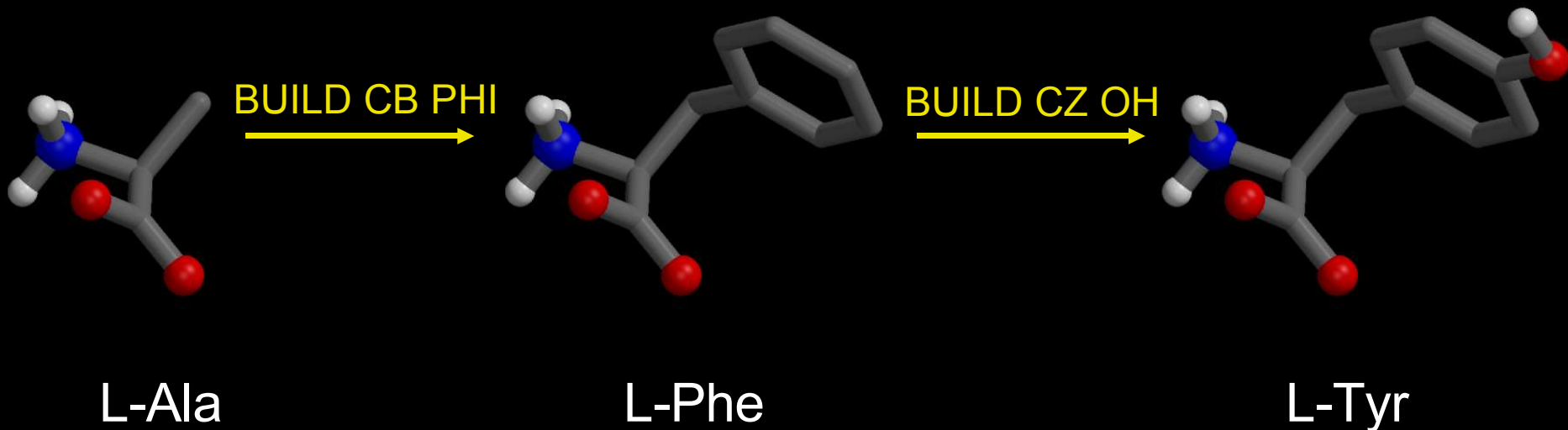


Still allosamidin

# Building

- PRODRG can add molecular fragments to existing molecules:

BUILD <atom> <fragment>



# Building

- Allows quick alterations to existing molecules
- Preserves coordinates of root structure
- Fragment libraries contain text drawings – easy to define:

```
FRAG OH
```

```
X-O
```

```
FRAG PHI
```

```
X-C-C=C
```

```
"   |
```

```
C-C=C
```

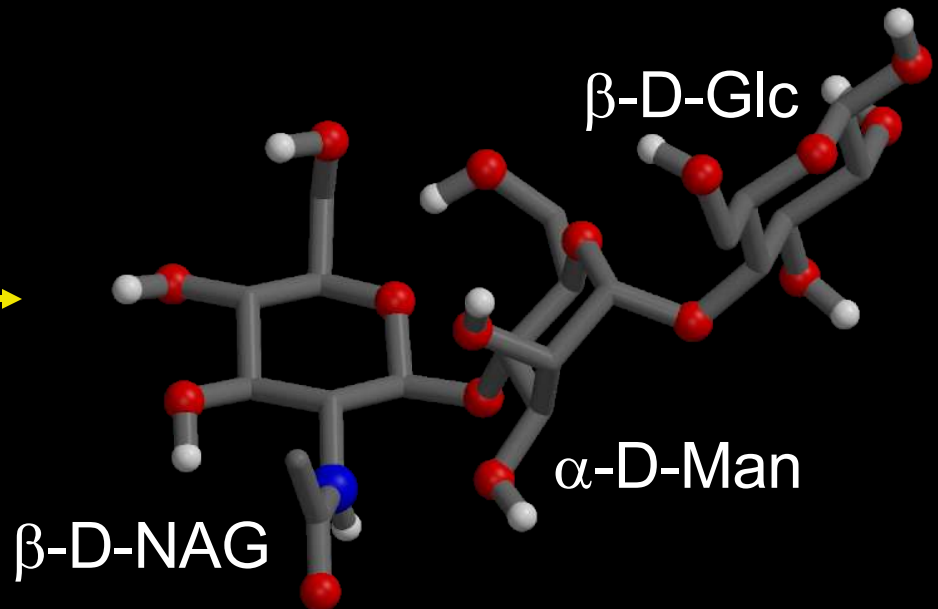
```
FRAG ...
```

# Building

- Can also be used to generate oligopeptides and oligosaccharides, using BUILD and START <fragment>

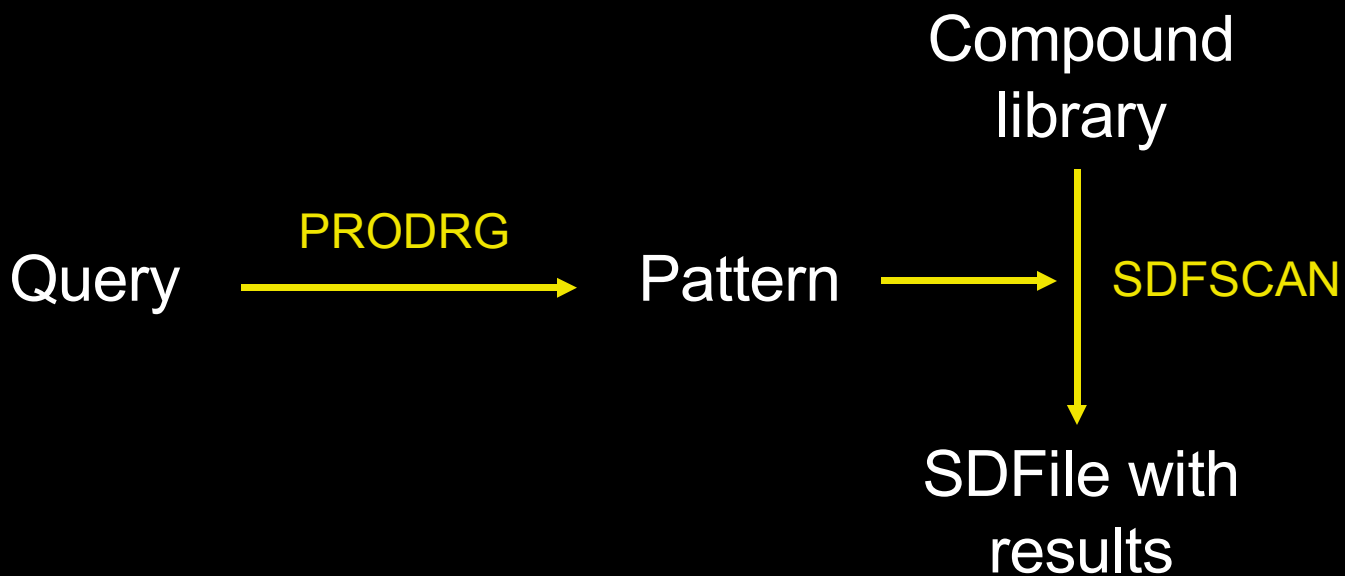
```
START      bdGLC
BUILD O4   adMAN1
BUILD O0F  bdNAG1

PATCH C1  >4C1
PATCH C0B >4C1
PATCH C1B >4C1
```



# Searching SDFiles

- Pattern matching to select compounds from 'compound database' (SDF / internal format)
- Powerful, reasonably fast, multipattern queries



# Docking with PRODRG

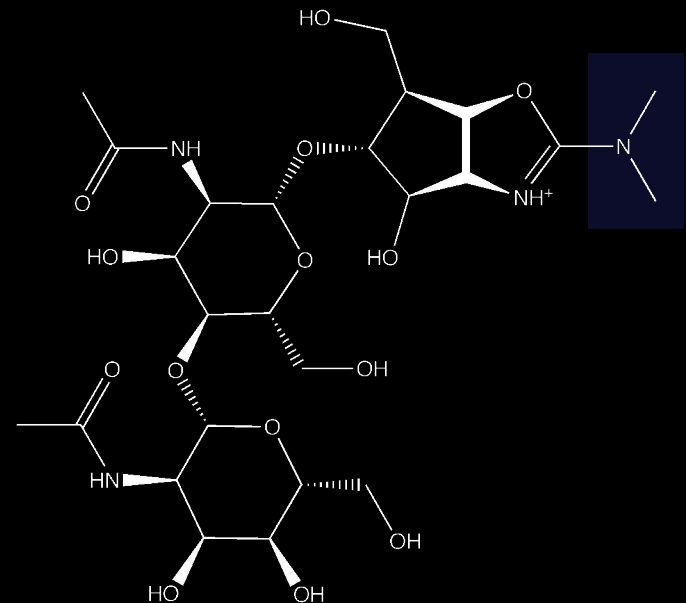
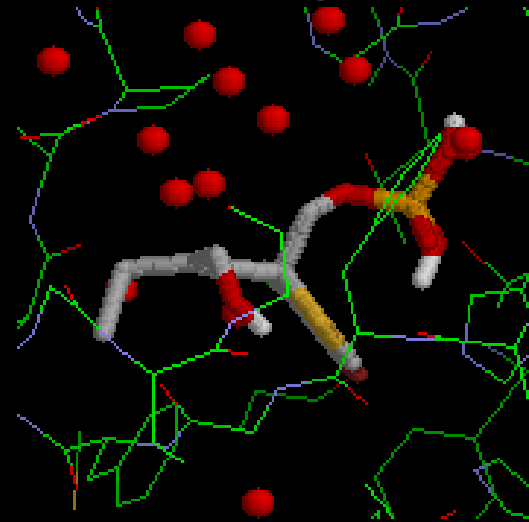
- Building/matching useful for ligand design
- We require energy minimisation in the context of the receptor
- Integration into PRODRG?

⇒ LIGTOR

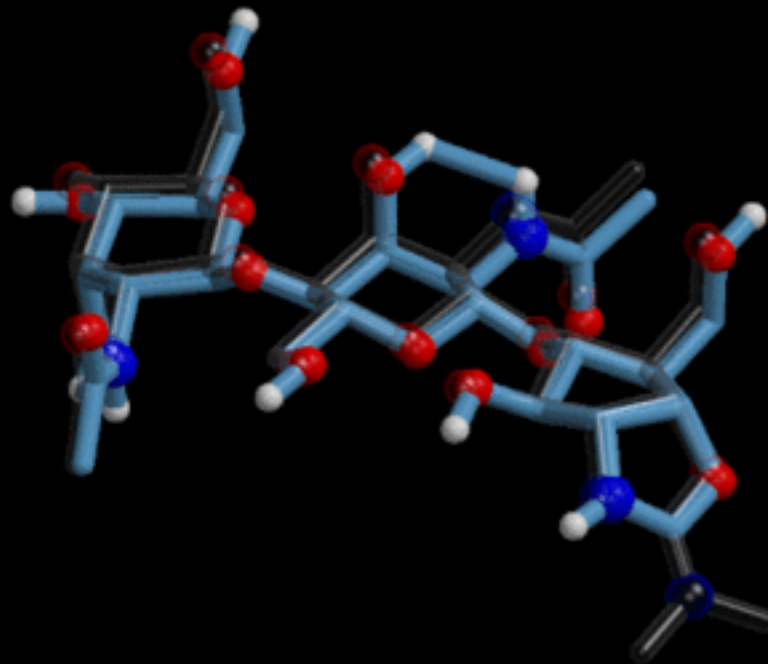
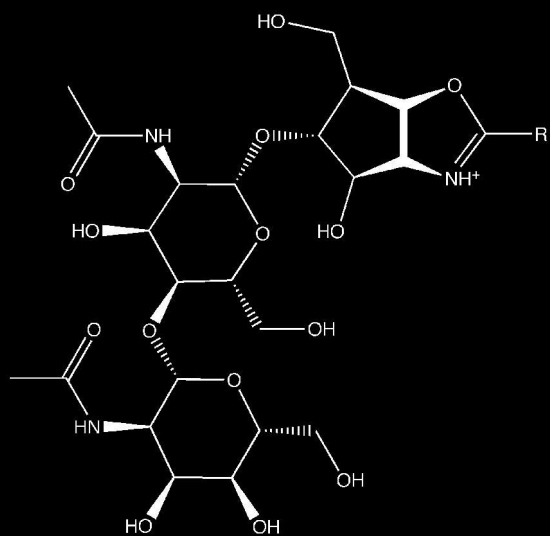
- Extremely simple docking program
- Grid-based, uses Autodock 3 force field
- Fast, but not very powerful

# LIGTOR in building

- BUILD adds fragments
- Growing protocols
- Example: allosamidin



# LIGTOR in building



R	#tor	$\Delta G$ / kcal/mol	$\Delta G_{\text{RBR}}$ / kcal/mol	$\text{RMSD}_{\text{Root}}$ / Å
-H	0	-11.5	-12.1	0.37
-NH <sub>2</sub>	1	-11.8	-12.5	0.38
-NHCH <sub>3</sub>	1	-12.5	-13.2	0.36
-N(CH <sub>3</sub> ) <sub>2</sub>	1	-13.4	-14.1	0.38
-NHC <sub>2</sub> H <sub>5</sub>	2	-12.8	-13.5	0.38
-N(CH <sub>3</sub> )C <sub>2</sub> H <sub>5</sub>	2	-13.7	-14.4	0.44
-N(C <sub>2</sub> H <sub>5</sub> ) <sub>2</sub>	3	-13.7	-14.4	0.42

# What's next?

- Release of PRODRG as part of CCP4
- Coot plug-in for on-the-fly topologies
- SMILES in/out
- Better (proper) treatment of stereochemistry
- Release of LIGTOR

# Acknowledgements

- Alex Schuettelkopf, PRODRG users

Structure, Vol. 11, 1051–1059, September, 2003, ©2003 Elsevier Science Ltd. All rights reserved. DOI 10.1016/S0969-2126(03)00186-2

## Pound-Wise but Penny-Foolish: How Well Do Micromolecules Fare in Macromolecular Refinement?

## Ways & Means

Gerard J. Kleywegt,<sup>1,\*</sup> Kim Henrick,<sup>2</sup>  
Eleanor J. Dodson,<sup>3</sup> and Daan M.F. van Aalten<sup>4</sup>

today. However, fewer than half of these entries are  
unique at the level of the sequence. The remaining en-

Acta Crystallographica Section D  
Biological  
Crystallography  
ISSN 0907-4449

## *PRODRG*: a tool for high-throughput crystallography of protein–ligand complexes

Alexander W. Schüttelkopf and  
Daan M. F. van Aalten\*

The small-molecule topology generator *PRODRG* is  
described, which takes input from existing coordinates or  
various two-dimensional formats and automatically generates

Received 25 February 2004  
Accepted 13 May 2004