Practical course 1

1. Quality of experimental structures

A]

Experimental Method used to obtain 1N0S : X-RAY Diffraction
Experimental Method used to obtain 1T0V : Solution NMR

The **Resolution** of the X-ray structure is equal to 2 Å (good if above 2,5A) .

And the **R_values** are :

- R-Value Free: 0.243
- R-Value Work: 0.193

The R-value work (or R-value or R_{work} or R-factor) is a measure of the agreement between the crystallographic model and the experimental X-ray diffraction data. it mathematically describes the difference between the experimental observations and the ideal calculated values.

It is defined by the following equation:

$$R = \frac{\sum ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum |F_{\text{obs}}|}$$

where F is the structure factor, a mathematical description of how a material scatters incident radiation. The minimum possible value of the R-value work is zero, indicating perfect agreement between experimental observations and the structure factors predicted from the model. There is no theoretical maximum, but in practice, values are considerably less than one even for poor models, provided the model includes a suitable scale factor. Random experimental errors in the data contribute to even for a perfect model, and these have more leverage when the data are weak or few, such as for a low-resolution data set. Model inadequancies such as wrong or missing parts and unmodeled disorder are the other main contributors to , making it useful to assess the progress and final result of a crystallographic model refinement. For large molecules, the R-factor usually ranges between 0.6 and 0.2. Small molecules (up to *ca*. 1000 atoms) usually form better-ordered crystals than large molecules, for which it is possible to attain lower R-factors.

The R-Value Free is computed according to the same formula given above, but on a small, random sample of data that are set aside for the purpose and never included in the refinement. R-Value Free will always be greater than R-Value Work because the model is not fitted to the reflections that contribute to R-Value Free, but the two statistics should be similar because a correct model should predict *all* the data with uniform accuracy. If the two statistics differ significantly then that indicates the model has been over-parameterized, so that to some extent it predicts not the ideal error-free data for the correct model, but rather the error-afflicted data actually observed.

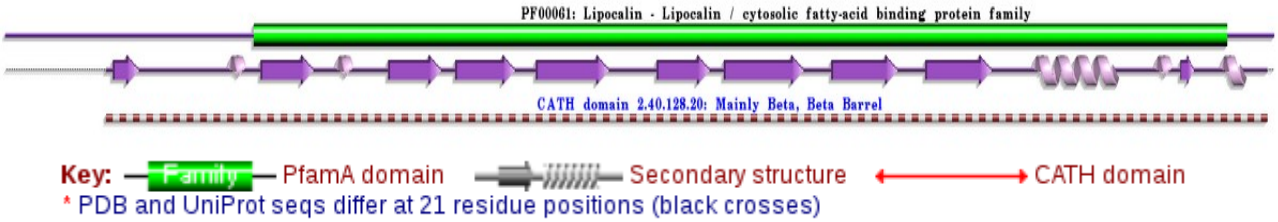
There are 100 calculated conformers and **20 submitted conformers**(= 20 «best» models) in the PDB file of the NMR structure.

B]

The tables below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for ≥ 3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and a grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with **a dot representing fractions ≤ 5 %**. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

1N0S :

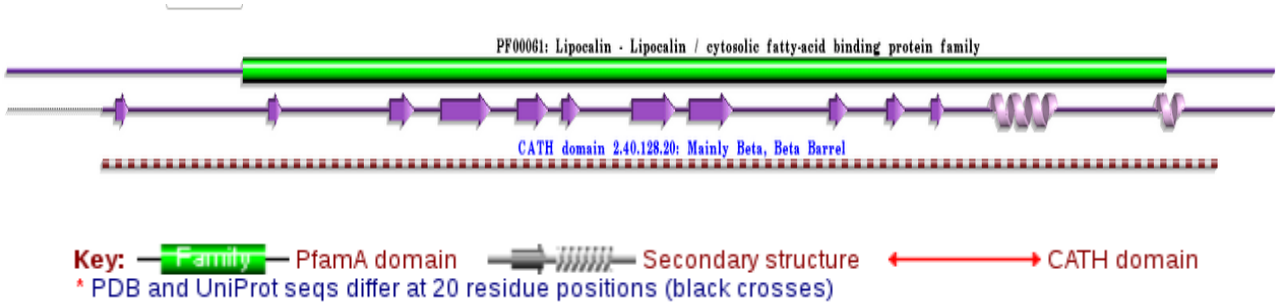
Mol	Chain	Length	Quality of chain
1	A	184	2% 72% 15% 7% • 6%
1	B	184	2% 74% 14% 5% • 6%



There are no Ramachandran outliers to report for 1N0S.

1T0V :

Mol	Chain	Length	Quality of chain
1	A	184	23% 56% 14% • 7%

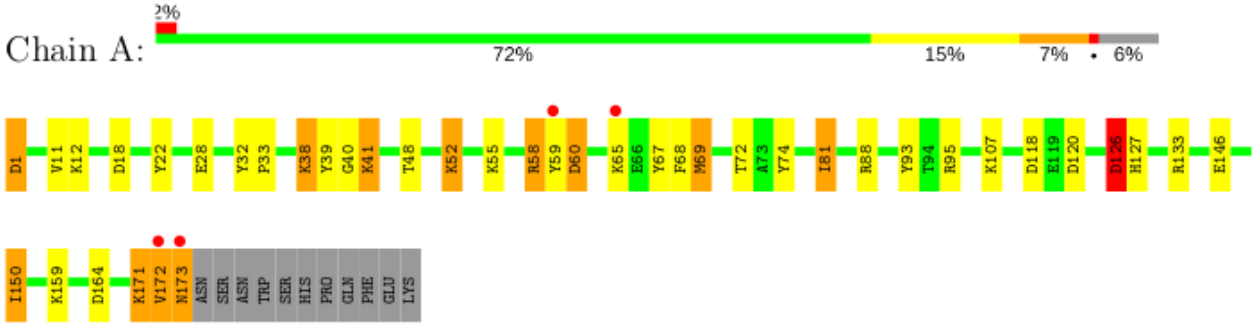


There are 34 unique Ramachandran outlier described for 1T0V. A residue is considered to be a Ramachandran plot outlier if the combination of its φ and ψ torsion angles is unusual, as assessed by MolProbity (Chen et al., 2010). The Ramachandran outlier score for an entry is calculated as the percentage of Ramachandran outliers with respect to the total number of residues in the entry for which the outlier assessment is available.

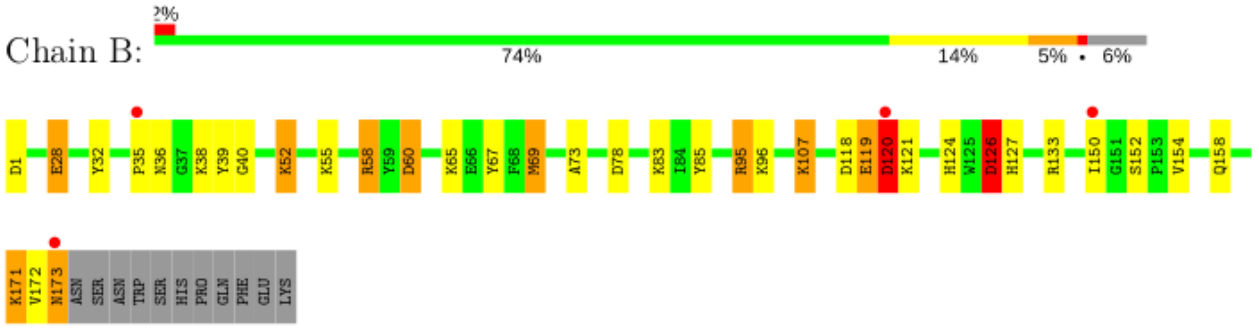
Conclusion : The quality of the structure 1N0S is higher than the structure of 1T0V. The numeric value of the yellow segment (fraction of residues that have perfect fit to the electron density, no outliers) is higher for 1N0S than for 1T0V. Regarding the protein backbone, there is no outliers for 1N0S and 34 for 1T0V. But we have to notice that RMN is in general less accurate than X-ray.

These plots below are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. **A red dot above a residue indicates a poor fit to the electron density (RSRZ>2)**. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.
The real-space R-value (RSR) is a measure of the quality of fit between a part of an atomic model (in this case, one residue) and the data in real space (Jones et al., 1991). The RSR Z-score (RSRZ) is a normalisation of RSR specific to a residue type and a resolution bin (Kleywegt et al., 2004). RSRZ is calculated only for standard amino acids and nucleotides in protein, DNA and RNA chains. A residue is considered an RSRZ outlier if its RSRZ value is greater than 2. The RSRZ outlier score as shown in the slider graph is calculated as the percentage RSRZ outliers with respect to the total number of residues for which RSRZ was computed. This is calculated by the EDS (Electron-Density Server) component of the validation pipeline which is a re-implementation of the software used by the Uppsala EDS server (Kleywegt et al., 2004).

- Molecule 1: Bilin-binding protein



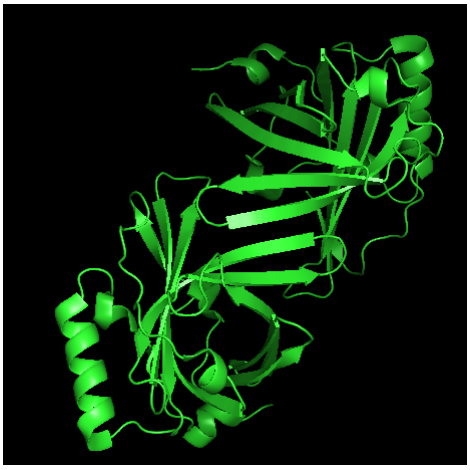
- Molecule 1: Bilin-binding protein

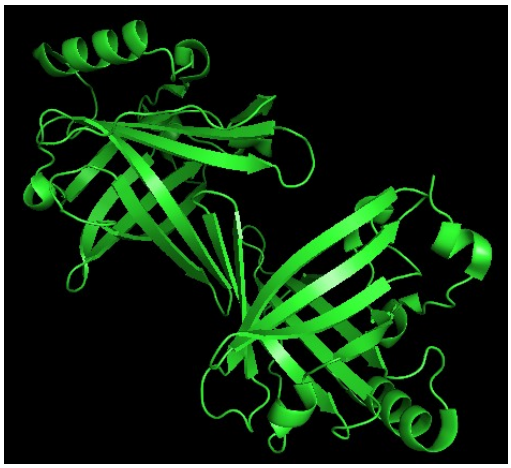


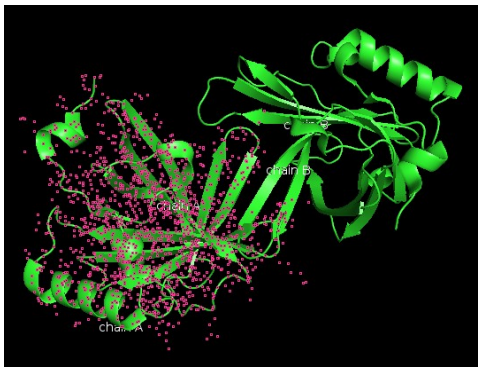
Y59 (tyrosine) , K65 (lysine), V172 (valine) and N173(asparagine) are marked with a red dot in the chain A. P35 (proline), D120(aspartate), I150(isoleucine) and N173(asparagine) are marked with a red dot in the chain B.

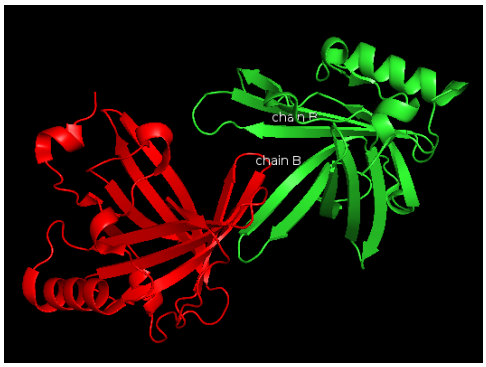
2. Visualization of structures

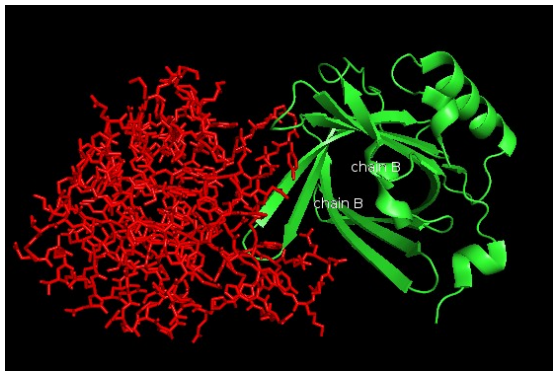
A]



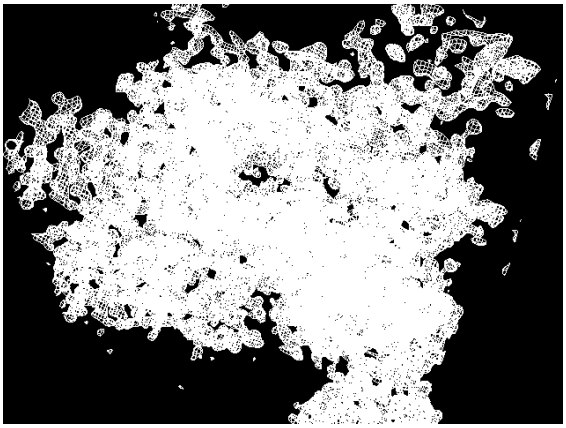


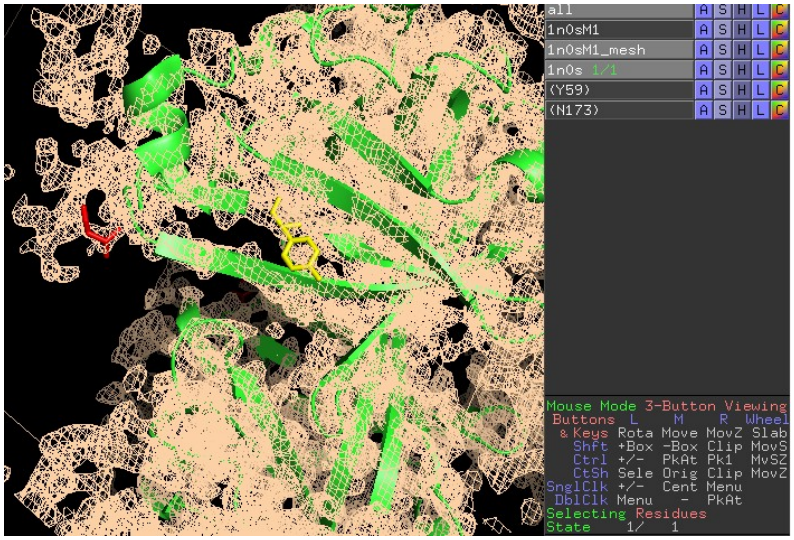


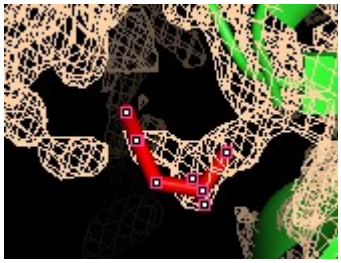


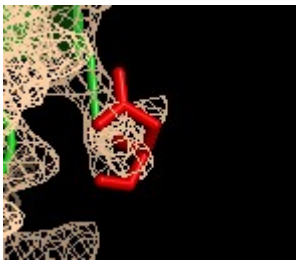


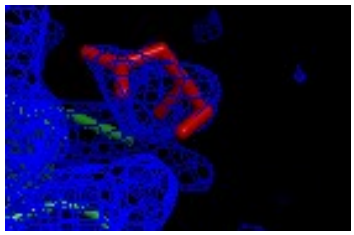
B]

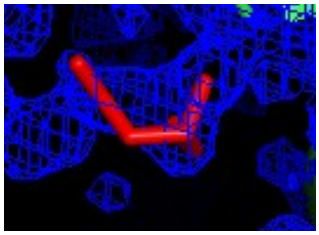






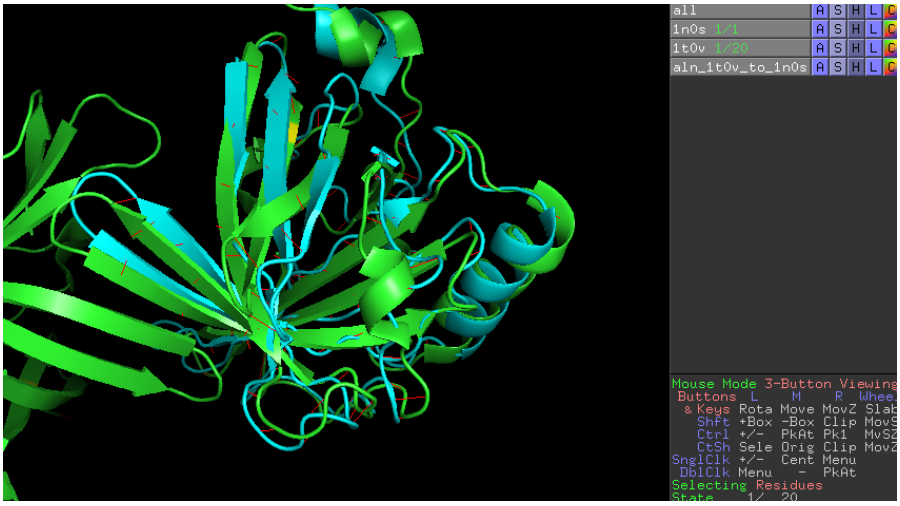


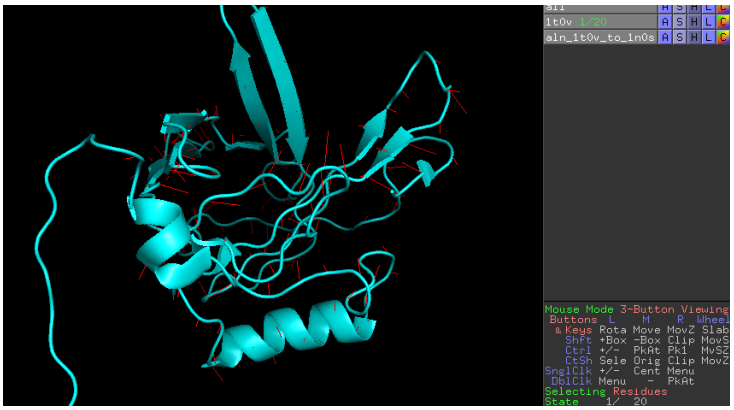


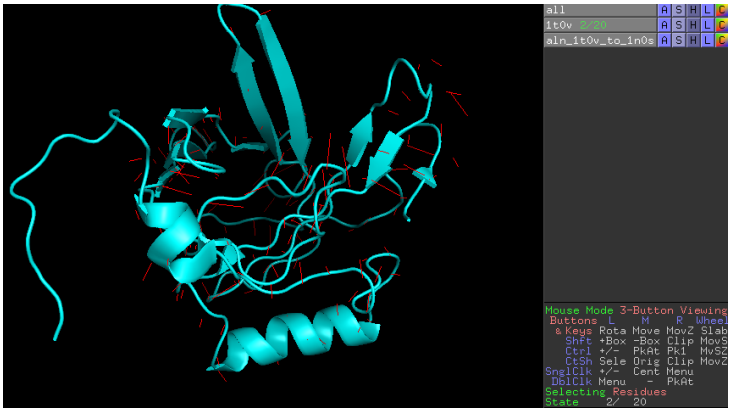


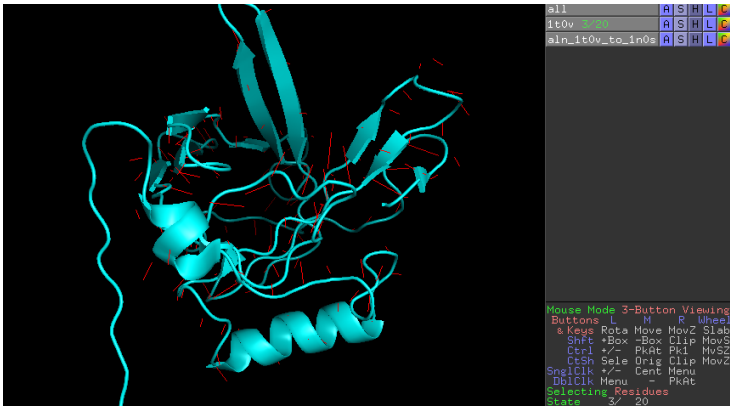
There is a good conformity between the electron density of this residue and the position of its atoms. We can see it on the pictures because the red stick residue is encompassed in the blue representation of the electronic density.

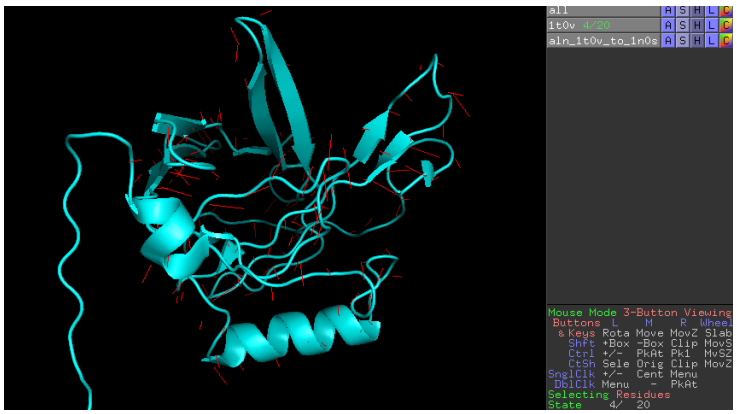
C]

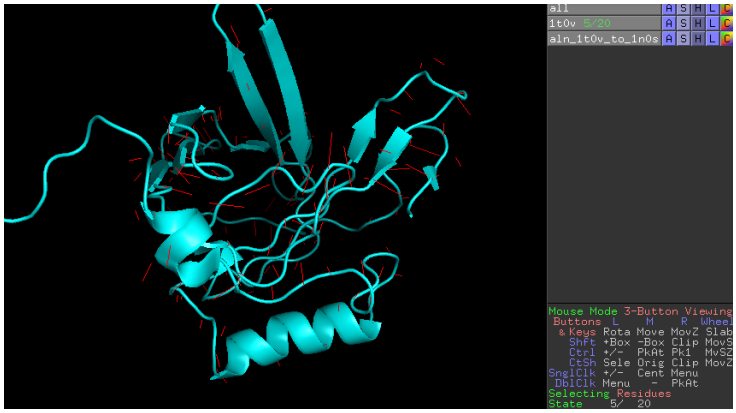


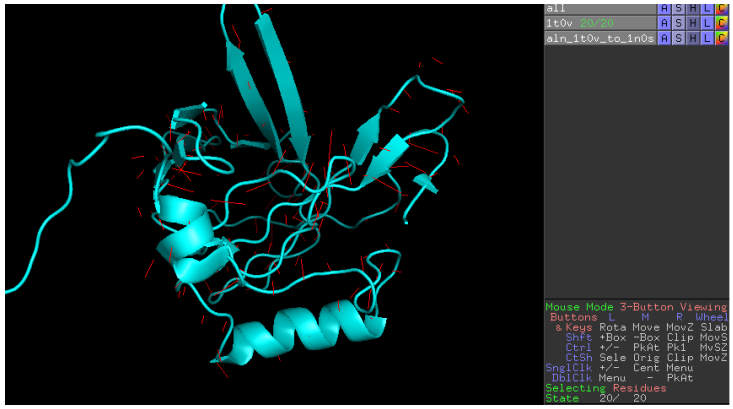












The secondary structure of 20 conformers of this NMR structure are globally conserved. But the random coils are not conserved and because of this, they change some angles and thus some orientations of chains.

3. Enzyme-ligand complex

A]

List of protein-ligand interactions

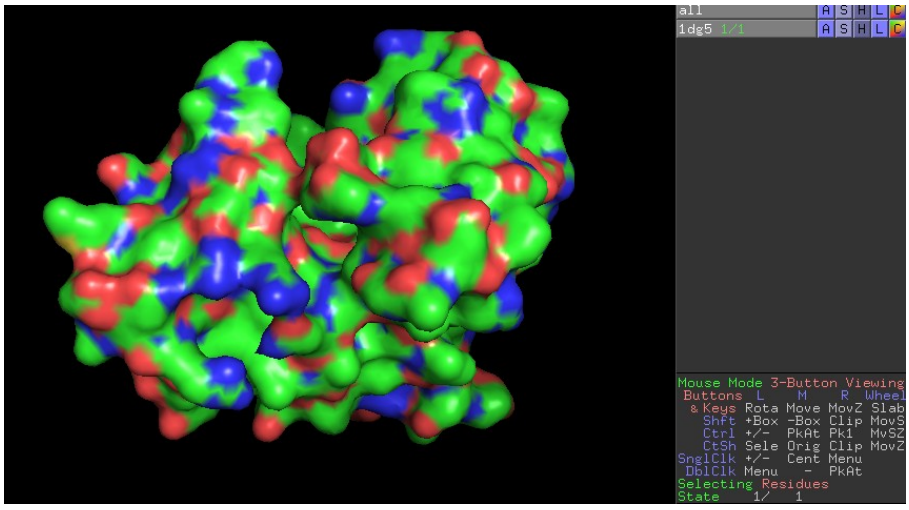
PDB code: 1dg5 Ligand TOP

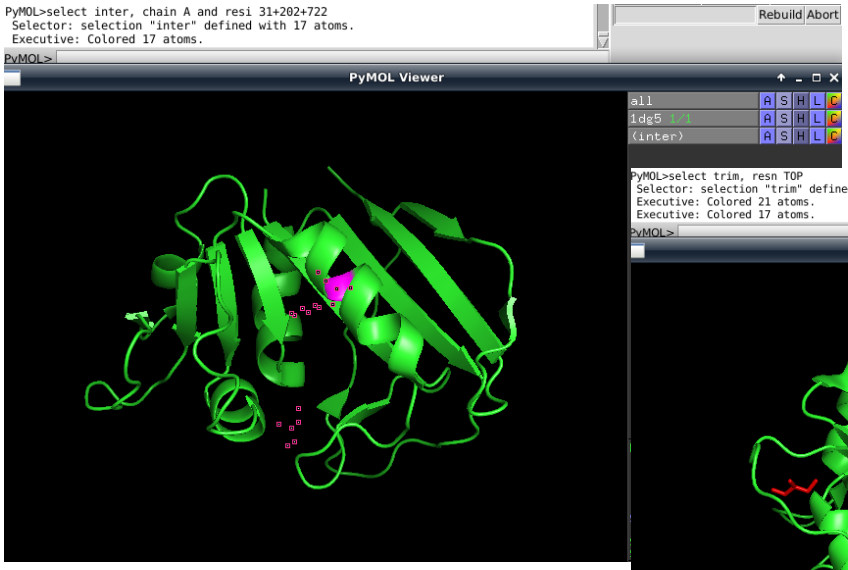
Hydrogen bonds

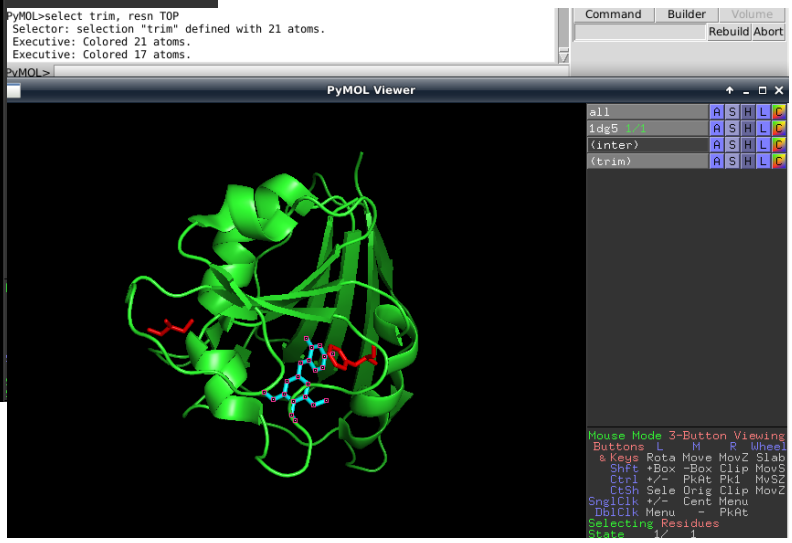
	<----- A T O M 1 ----->					<----- A T O M 2 ----->					
	Atom no.	Atom name	Res name	Res no.	Chain		Atom no.	Atom name	Res name	Res no.	Chain	Distance
1.	31	O	ILE	5	A	<--	1300	N7	TOP	201	A	2.76
2.	202	OD1	ASP	27	A	<--	1297	N4	TOP	201	A	2.81
3.	722	O	ILE	94	A	<--	1300	N7	TOP	201	A	2.85

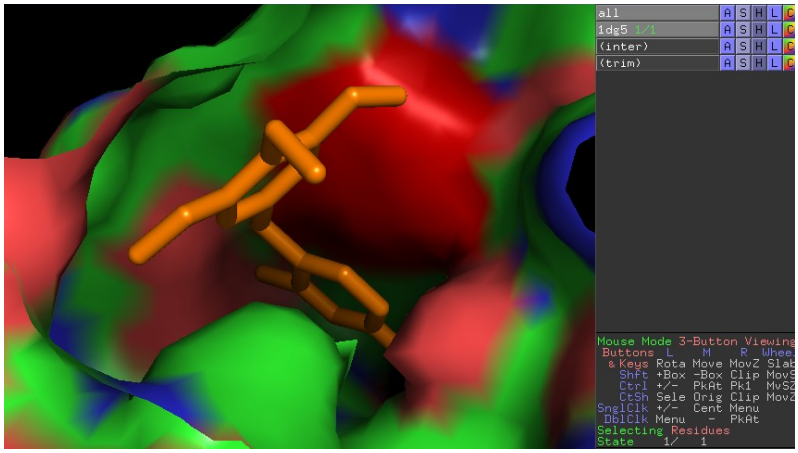
There are 3 amino acids which are in interaction with trimethoprim : I31, D202 et I722.

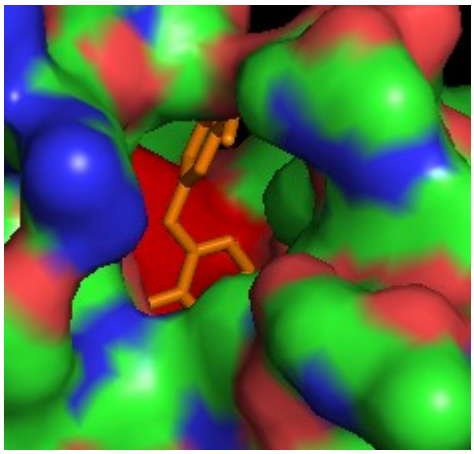
B]

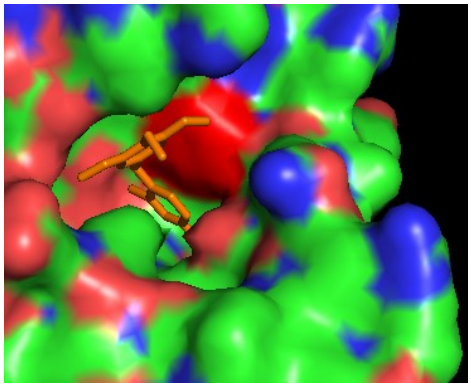


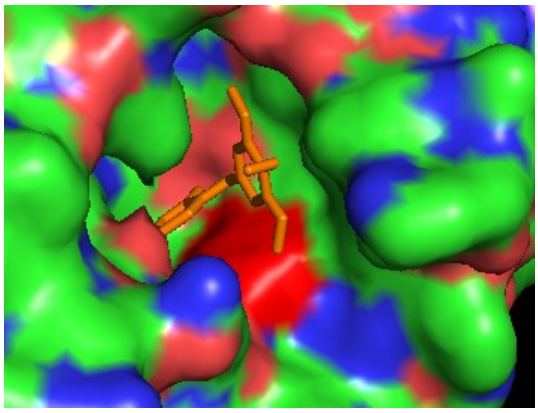

PyMOL>select inter, chain A and resi 31+202+722
Selector: selection "inter" defined with 17 atoms.
Executive: Colored 17 atoms.
PyMOL>
Rebuild
Abort
PyMOL Viewer
all
1dg5 1/1
(inter)

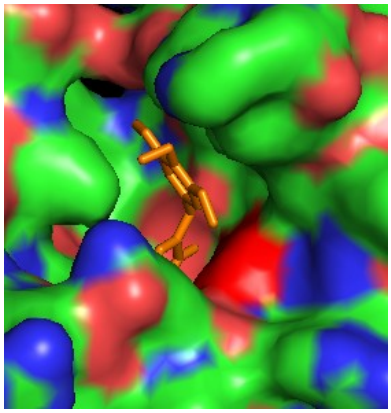

PyMOL>select trim, resn TOP
Selector: selection "trim" defined with 21 atoms.
Executive: Colored 21 atoms.
Executive: Colored 17 atoms.
PyMOL>
Command
Builder
Volume
Rebuild
Abort
PyMOL Viewer
all
1dg5 1/1
(inter)
(trim)
Mouse Mode 3-Button Viewing
Selecting Residues


all
1dg5 1/1
(inter)
(trim)
Mouse Mode 3-Button Viewing
Selecting Residues



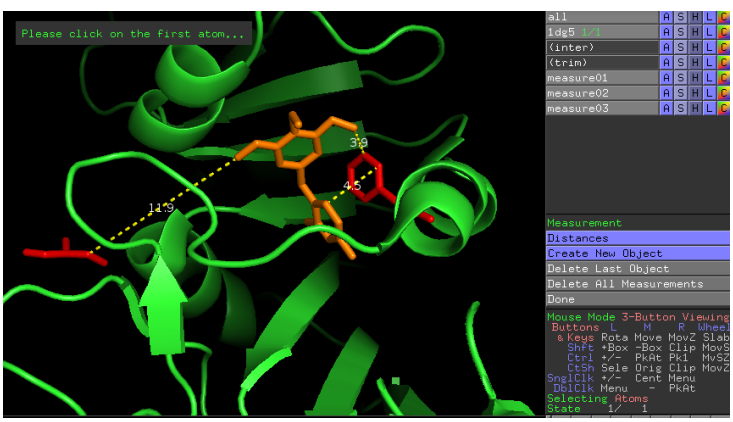


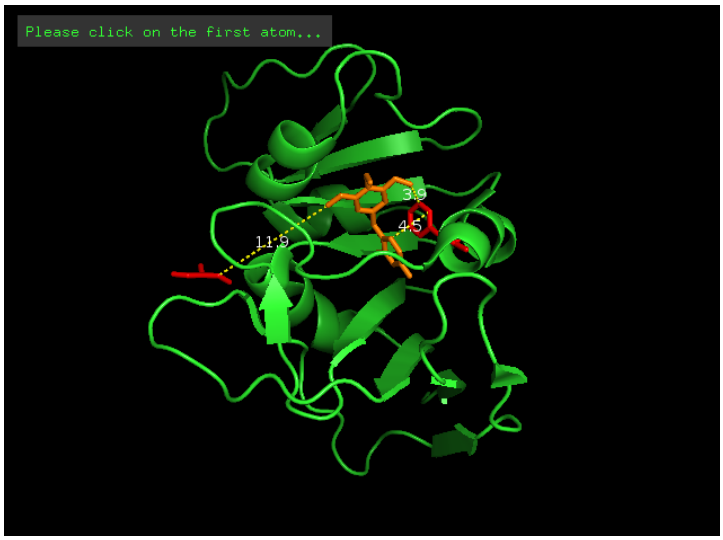




C]


all
1dg5 1/1
(inter)
(trim)
Mouse Mode 3-Button Viewing
Selecting Residues

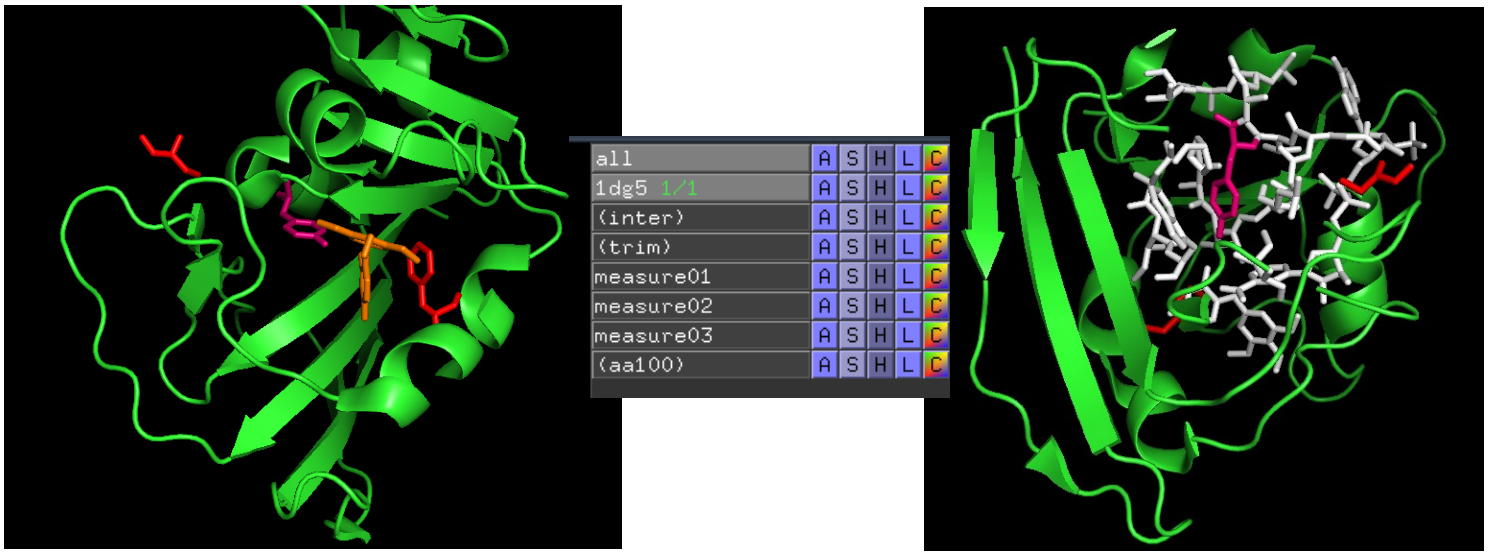

Please click on the first atom...
all
1dg5 1/1
(inter)
(trim)
measure01
measure02
measure03
Measurement
Distances
Create New Object
Delete Last Object
Delete All Measurements
Done
Mouse Mode 3-Button Viewing
Selecting Atoms

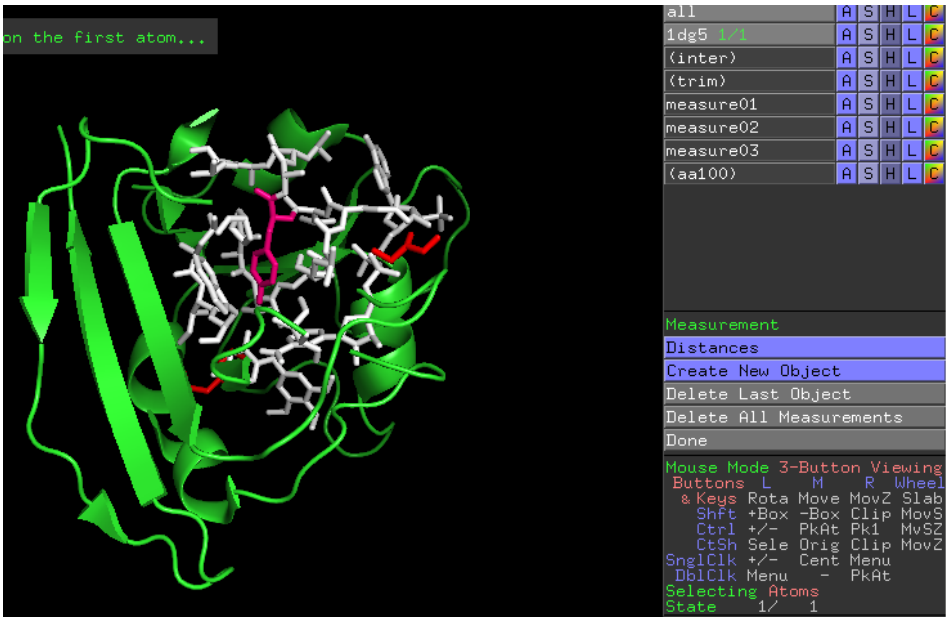


4. Compactness of the protein core

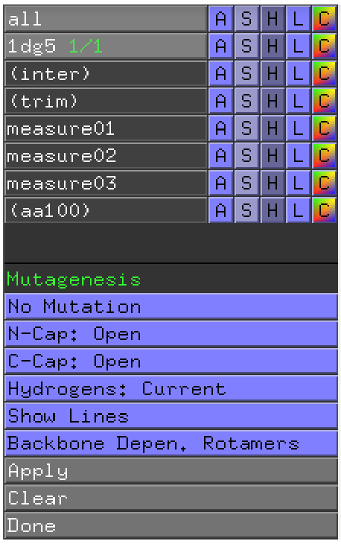
A]

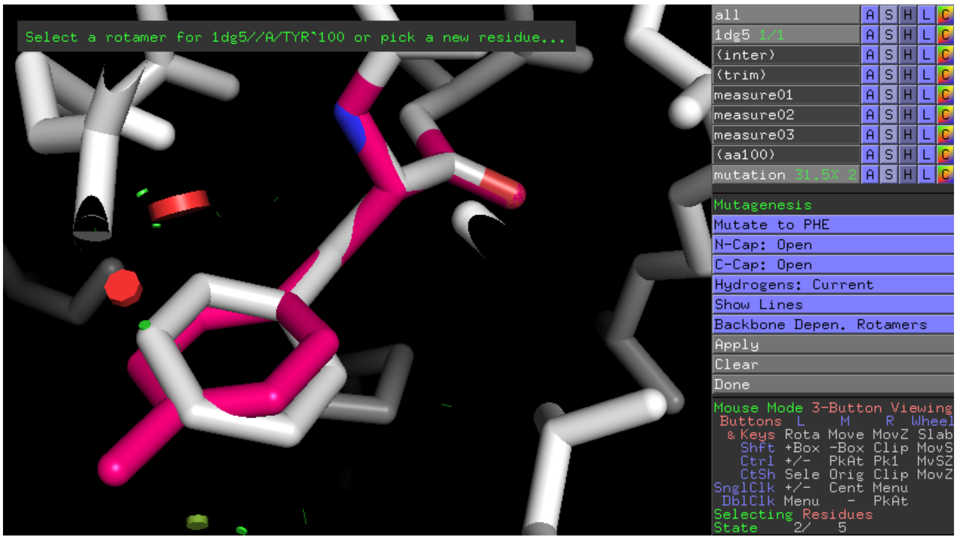
The pink residue is the amino acid number 100 (aa100) of 1DG5. The surrounding of this residue is in white.





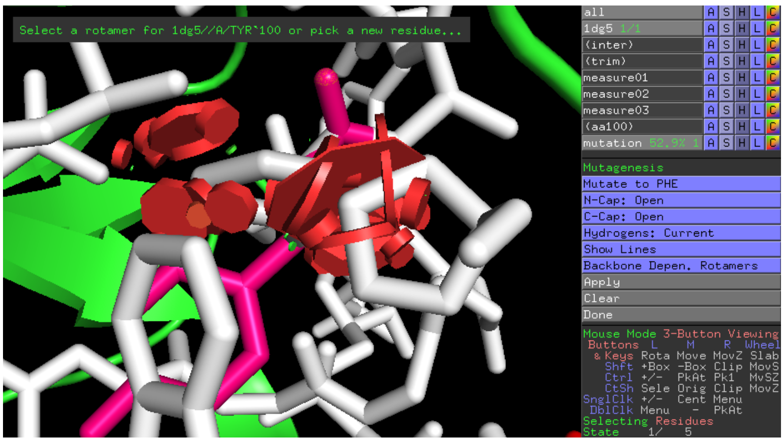
We mutate this core amino acid using the menu "Wizard –Mutagenesis" to illustrate the protein compactness.

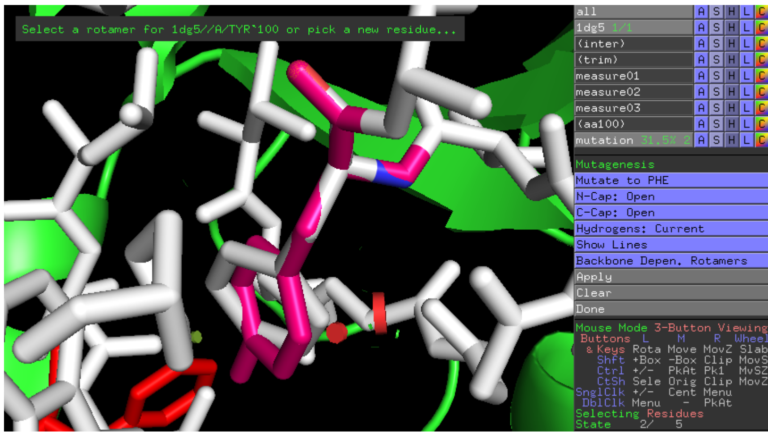


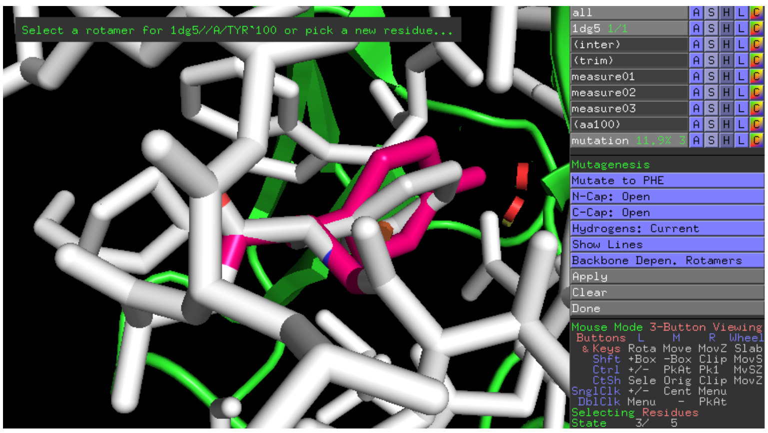


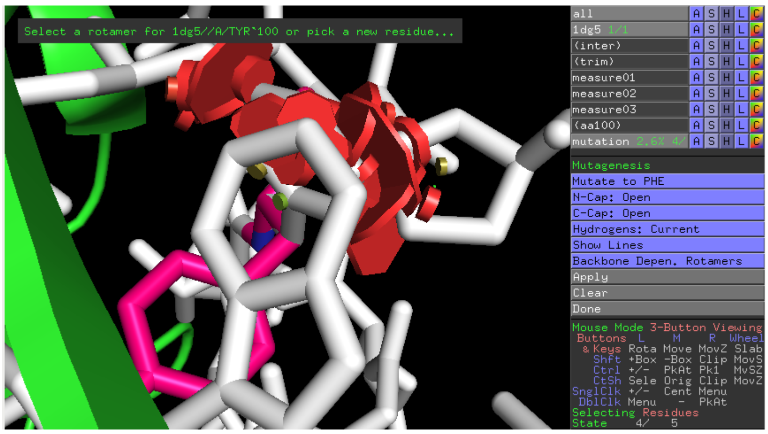
We can test the different possible conformers of the side chain with the arrows on the bottom right of the Pymol window. The red disks show the steric clashes.

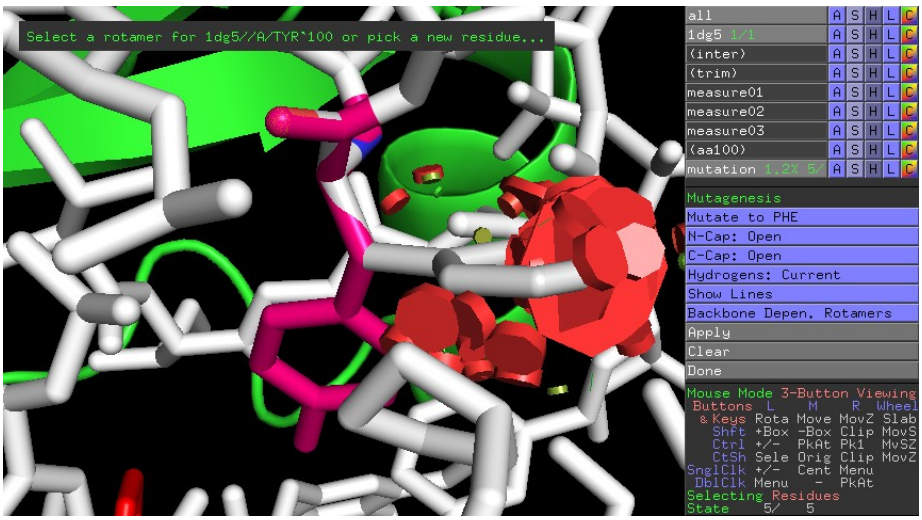
Here are the 5 different conformers:










Select a rotamer for 1dg5//A/TYR`100 or pick a new residue...
all A S H L C
1dg5 1/1 A S H L C
(inter) A S H L C
(trim) A S H L C
measure01 A S H L C
measure02 A S H L C
measure03 A S H L C
(aa100) A S H L C
mutation 1.2% 5/ A S H L C
Mutagenesis
Mutate to PHE
N-Cap: Open
C-Cap: Open
Hydrogens: Current
Show Lines
Backbone Depen. Rotamers
Apply
Clear
Done
Mouse Mode 3-Button Viewing
Buttons L M R Wheel
&Keys Rota Move MovZ Slab
Shft +Box -Box Clip MovS
Ctrl +/- PkAt Pk1 MvSZ
CtSh Sele Orig Clip MovZ
SnglClk +/- Cent Menu
DblClk Menu - PkAt
Selecting Residues
State 5/ 5