

Field of Extremes:

If you refine them, they will come

Gregory Warren and Emanuele Perola



Field of Extremes:

If you re-refine them, they will come down

Gregory Warren and Emanuele Perola



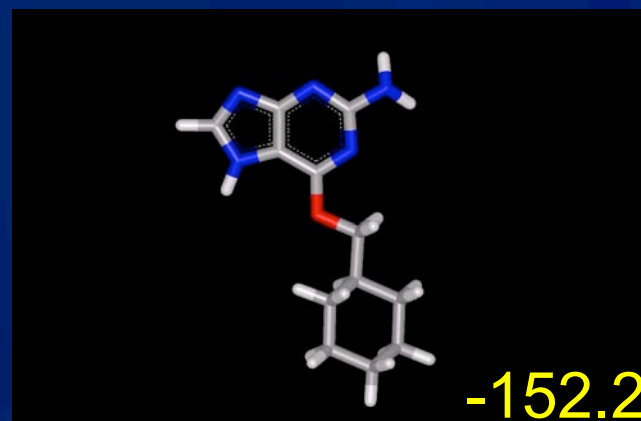
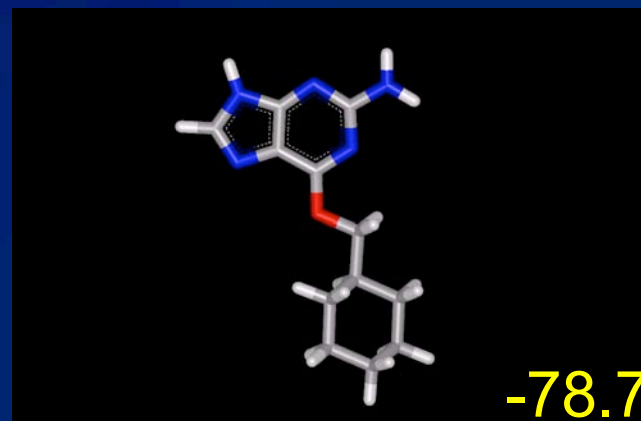
Introduction

- **Perola and Charifson**
 - J Med Chem, 2004, 47, 2499-2510
 - 60% don't bind in local minimum
 - 60% have strain energy <5 kcal/mol
 - 10% have strain energy > 9 kcal/mol
- **Number of explanations**
 - Force fields (Paul Labute)
 - Global minimum not important (Jim Snyder)
 - Comparing baseballs and footballs



Side Bar

- 1H1P



Introduction

- Perola and Charifson
 - High strain observed
- Number of explanations
 - Force fields (Paul Labute)
 - Solvent models
 - Can't find the global minimum (Jim Snyder)
 - Comparing baseballs and footballs
 - The Bears to the Black Sox

Your task – if you want to stop that annoying voice in your head...

- Re-refine all public structures with X-ray data
 - 100 -> 41
- Procedure
 - Refine deposited coordinates
 - Refine a low strain conformation
 - Compare
 - Use vacuum and Sheffield approximation



And like any good application scientist would do...

- Go find bugs in OpenEye software
 - Ah Roger, why did OEChem...
 - Ah Joe, why is Vida and Forge...
 - Ah Stan, please don't tell me Szybki did...
 - Ah Brian, I didn't want Afitt, Flynn or writedict to...
 - Ah Bruce, could you help me fix my bad python script?
 - Ah Ant, do I really have to give a talk at CUP – I gave two last year?
 - Ah Paul L, could I have a copy of MOE – please?

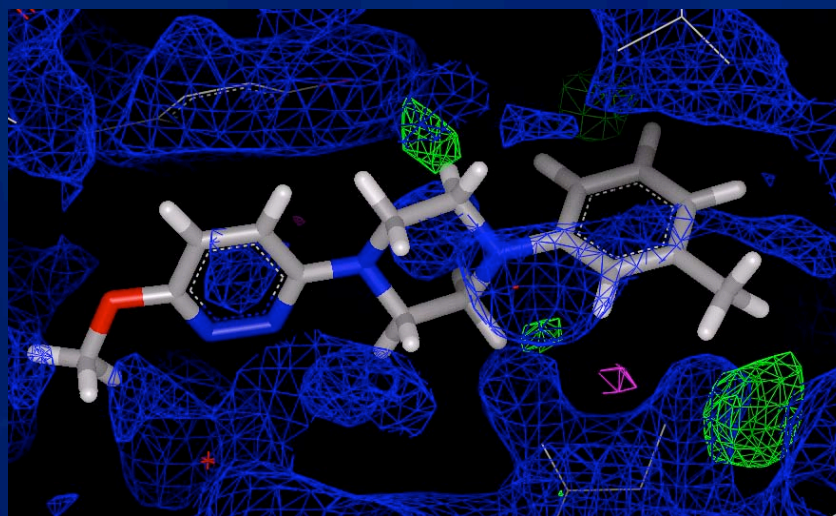
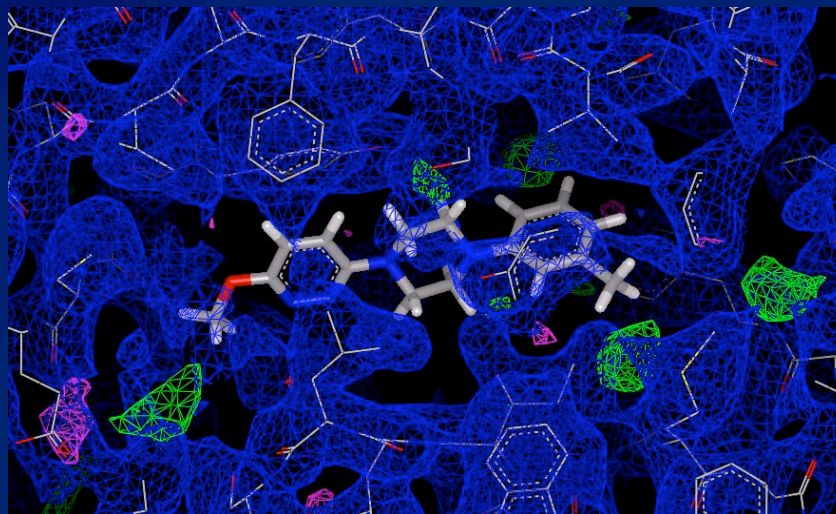


The hypothesis

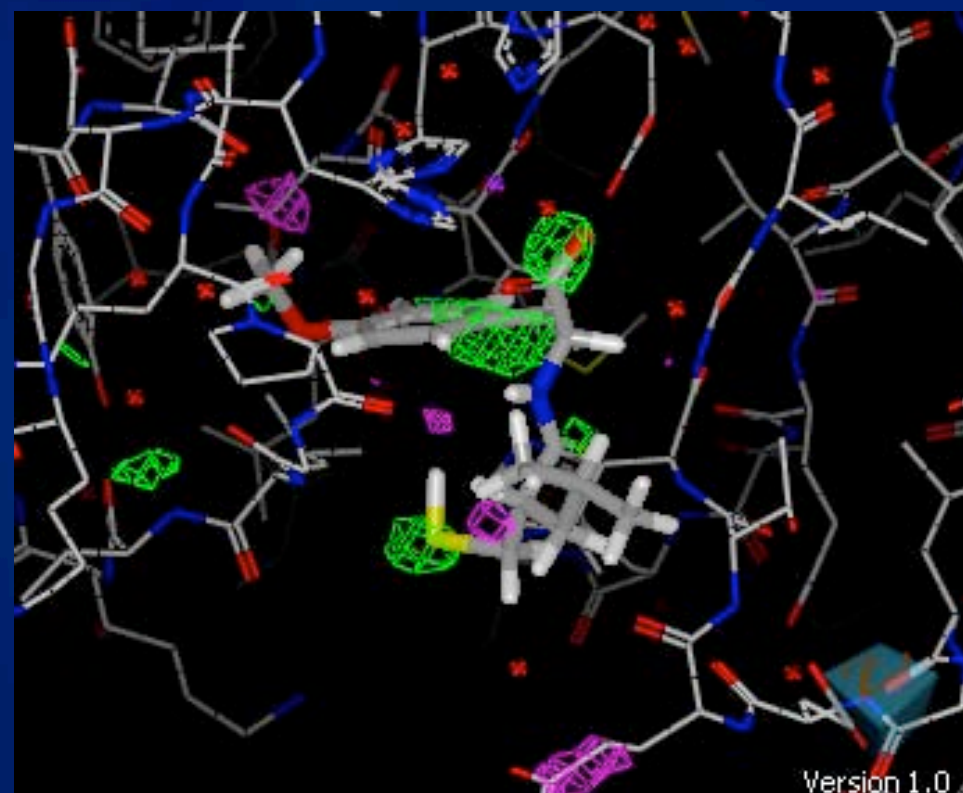
- For strained ligand conformation a low strain conformation can be found using either a force field derived refinement dictionary and/or a low strain starting conformation.

But first check the quality of the structures being used

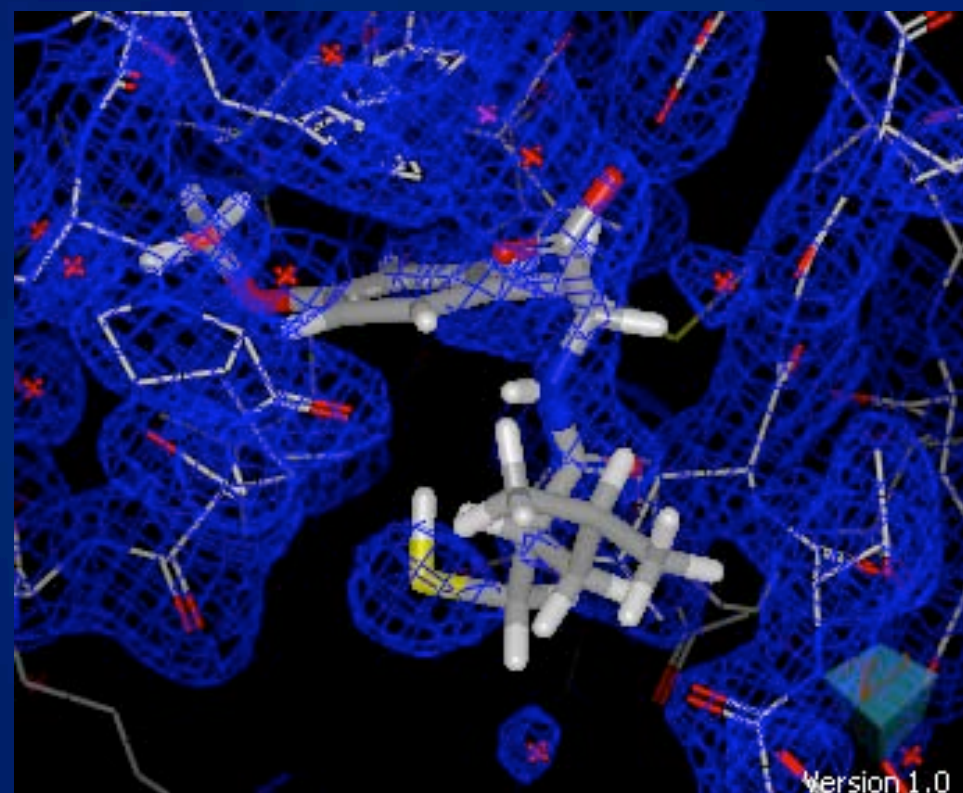
- 1r09
 - Virus structure
 - Resolution 2.70
 - DPI 0.79
 - R-free 0.46



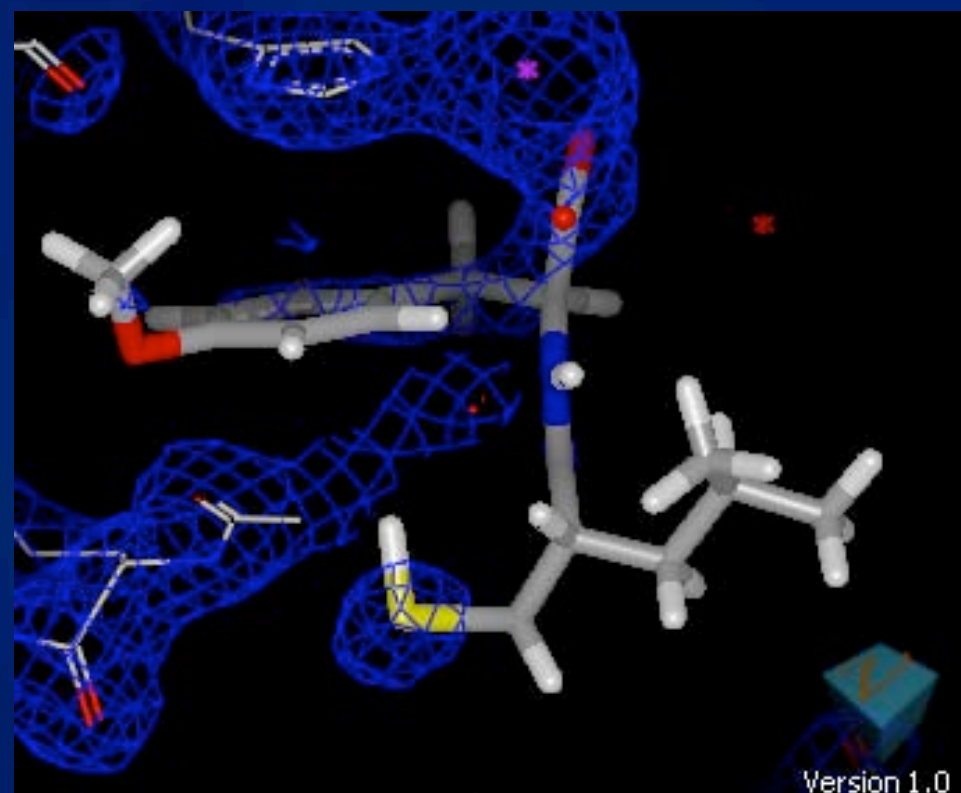
- 1atl
 - Metalloproteinase
 - Resolution 1.97
 - DPI 0.45
 - R-free 0.24
 - Oh where, oh where
 - Real Space Correlation Coefficient
 - 0.71, 0.72



- 1at1
 - Metalloproteinase
 - Resolution 1.97
 - DPI 0.45
 - R-free 0.24
 - Oh where, oh where
 - Real Space Correlation Coefficient
 - 0.71, 0.72

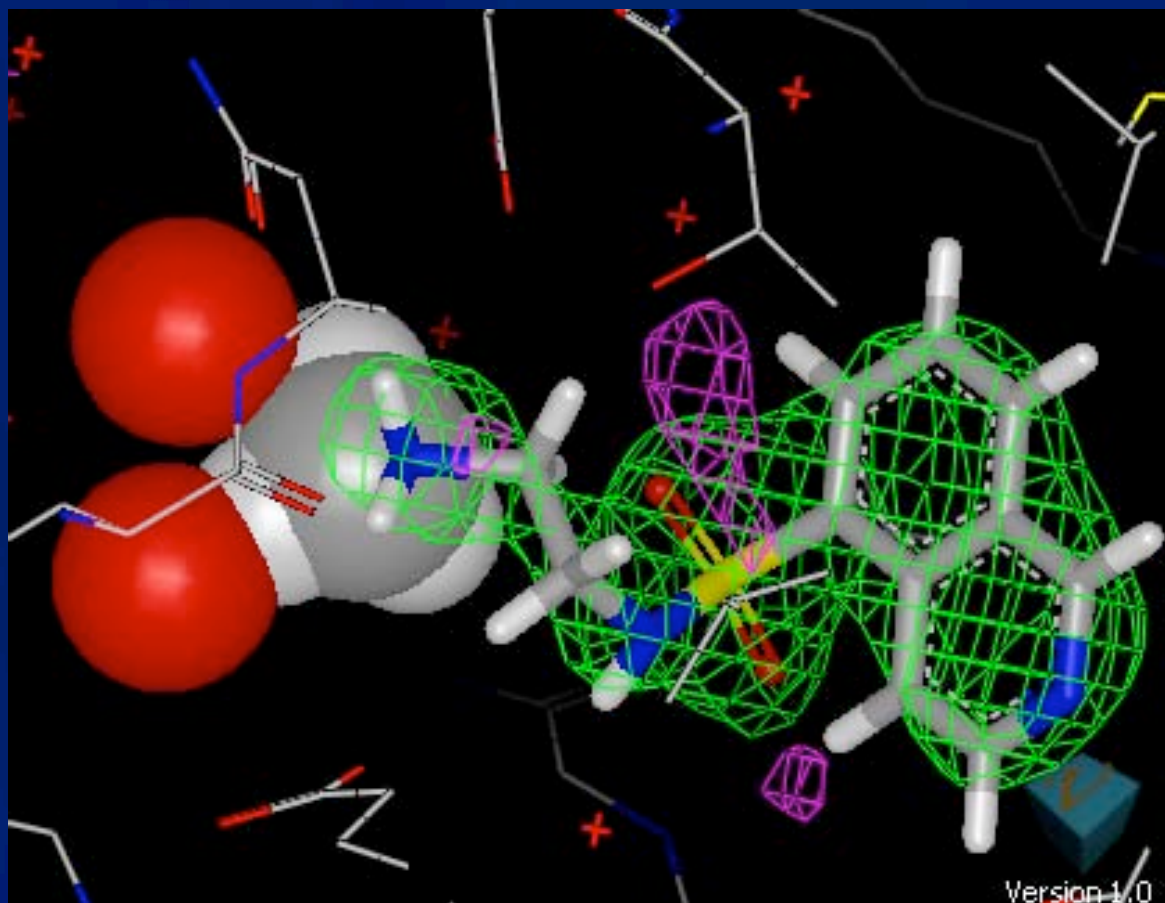


- 1atl
 - Metalloproteinase
 - Resolution 1.97
 - DPI 0.45
 - R-free 0.24
 - Oh where, oh where
 - Real Space Correlation Coefficient
 - 0.71, 0.72



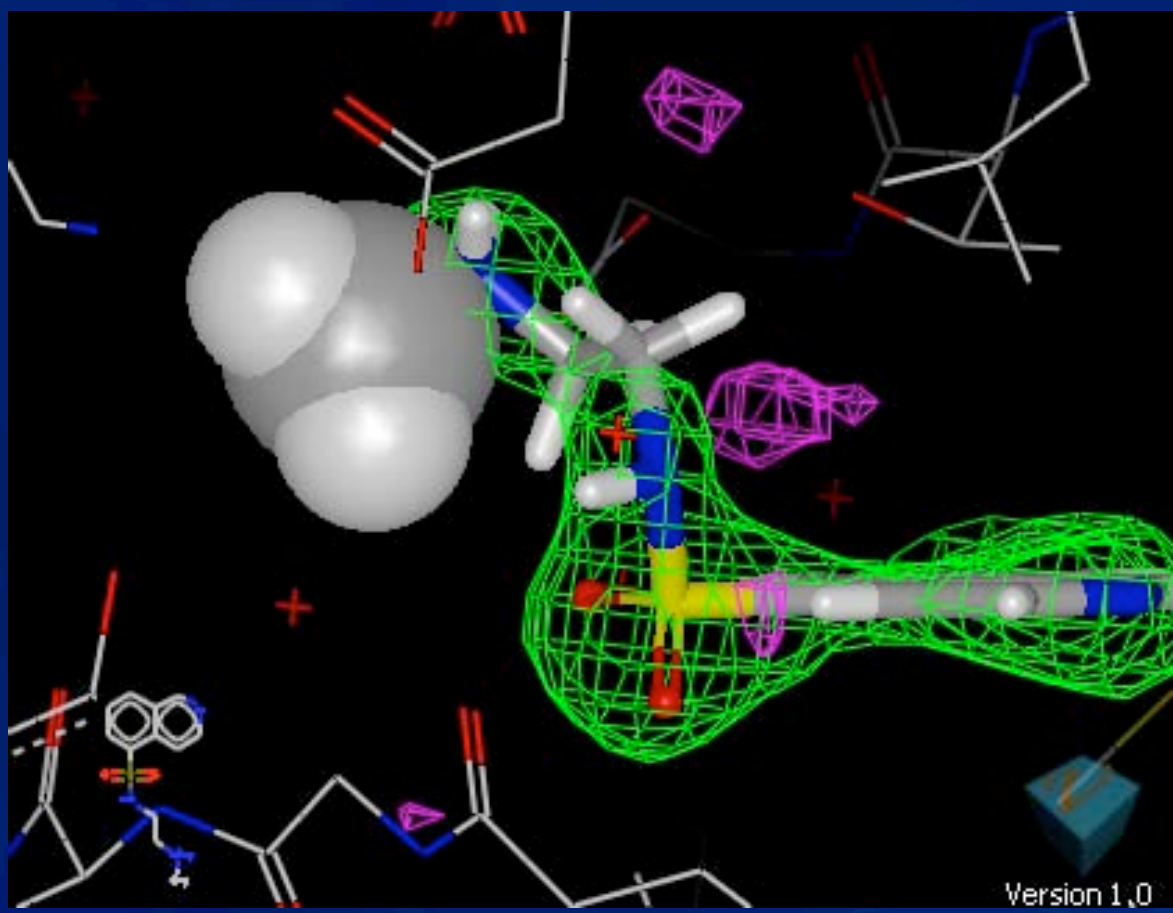
One last problem

- 1yds
 - Kinase
 - Resolution 2.1
 - DPI 0.37
 - R-free 0.26
 - Removed
 - HOH406
 - HOH411

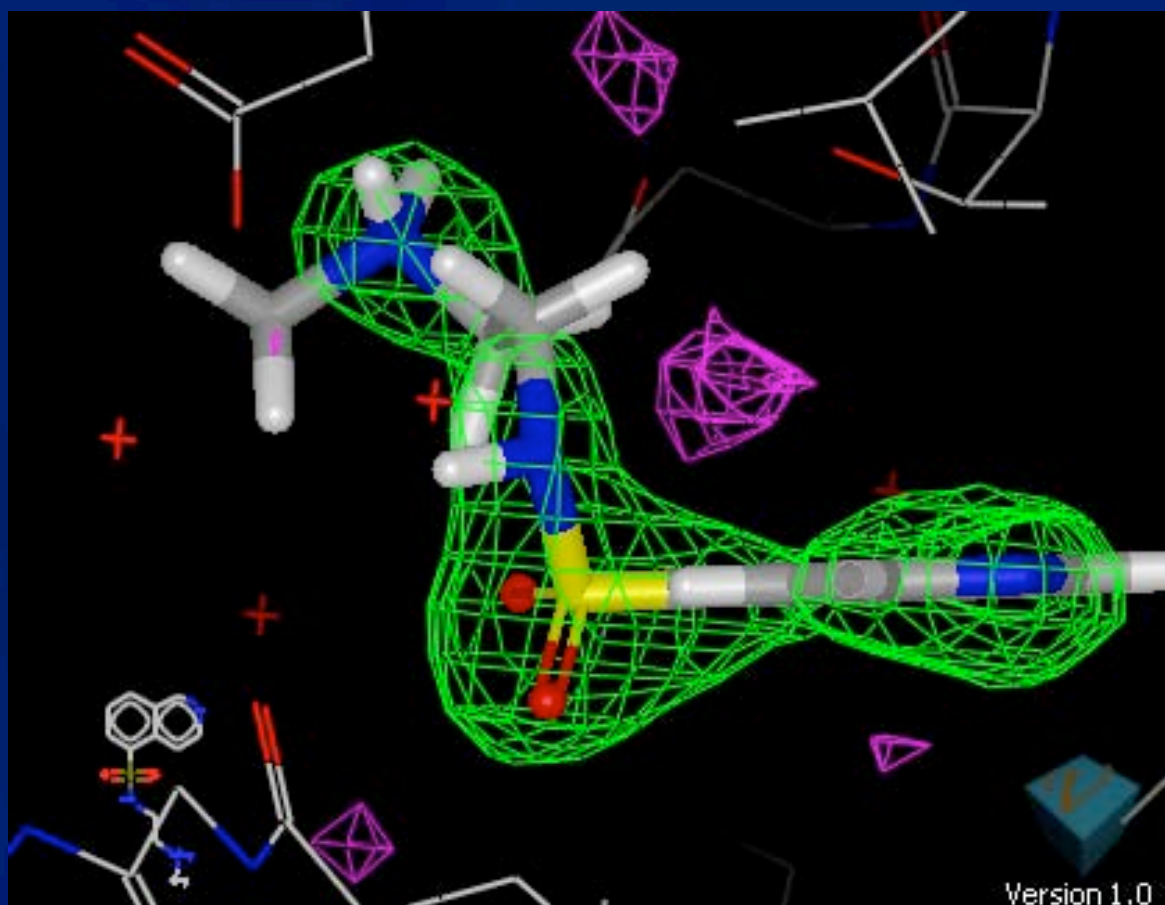


- 1yds

- Kinase
- Resolution 2.1
- DPI 0.37
- R-free 0.26
- Removed
 - HOH406
 - HOH411



- 1yds
 - Kinase
 - Resolution 2.1
 - DPI 0.37
 - R-free 0.26
 - Removed
 - HOH406
 - HOH411



So does it make a difference?

- Comparing raw structures optimizing only hydrogen
 - A smaller number is better

	Median	Mean
Refined Deposited - Deposited	-22.3	-36.0
Afitt -Deposited	-24.3	-36.8
Afitt – Refined Deposited	-0.3	-0.8

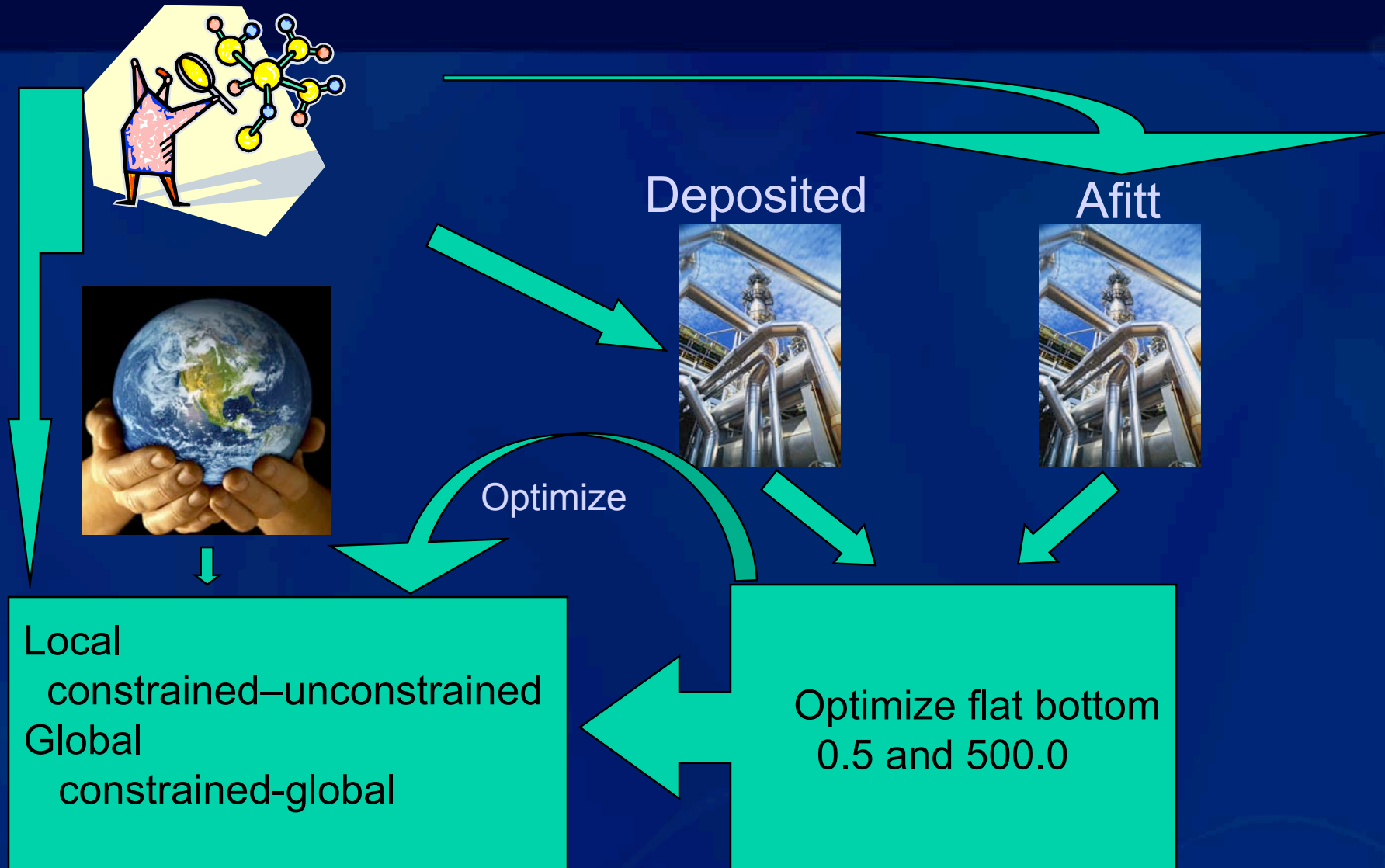


Now compare with a “global” low energy conformation

- Again compare raw structures with only hydrogen optimized

	Median	Mean
Deposited - Global	41.5	54.2
Refined Deposited - Global	18.2	19.8
Afitt – Global	17.3	19.0

How we measured strain

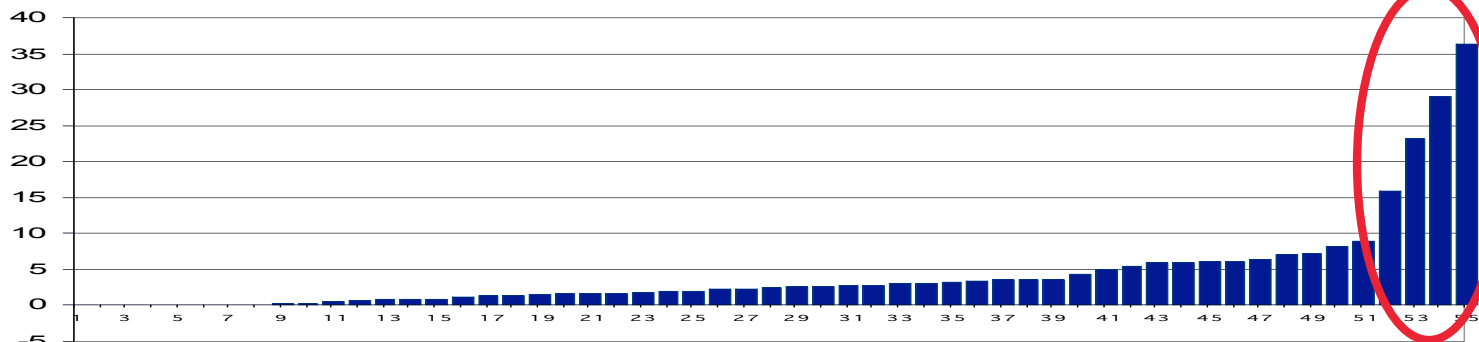


Ok?

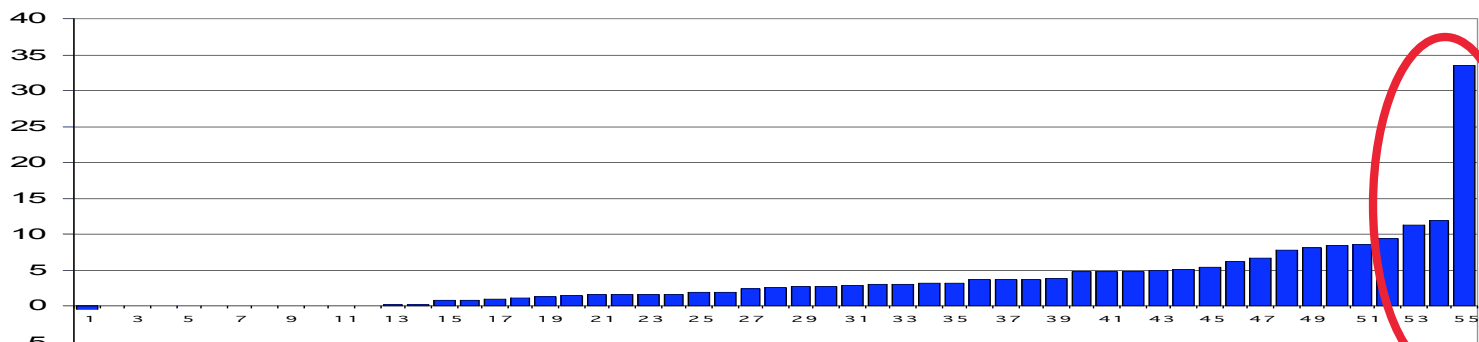
- **Local**
 - Constrained optimization of deposited structure (D), Refined deposited (RD), and Afitt (A) structures
 - The resulting structures were optimized with no constraints (local minima) to give DL, RDL and AL
 - Local strain is
 - D-DL, RD-RDL or A-AL
- **Global**
 - Use the D, R, or A structures with the global minimum structures
 - Global strain is
 - D-G, RD-G or A-G

Local Strain

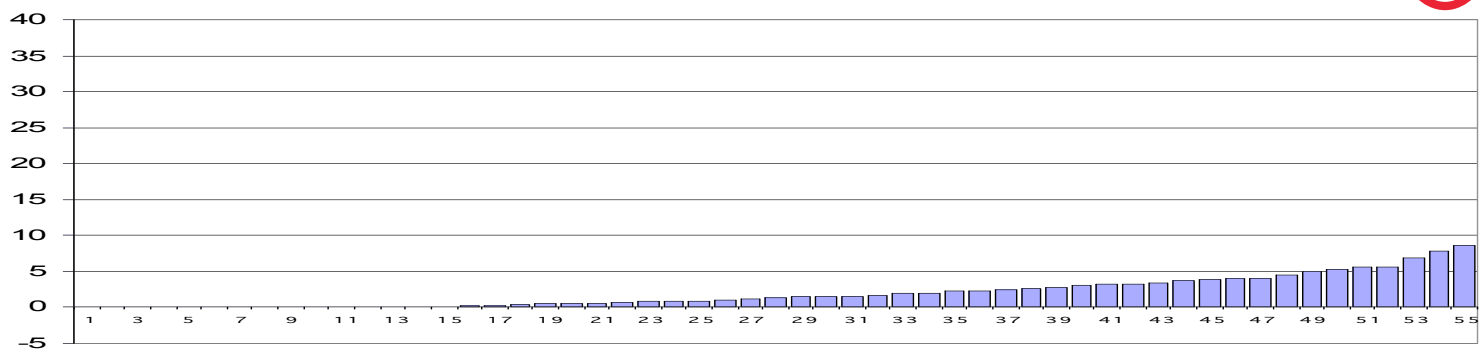
Strain Energy (kcal/mol)



Deposited



Refined

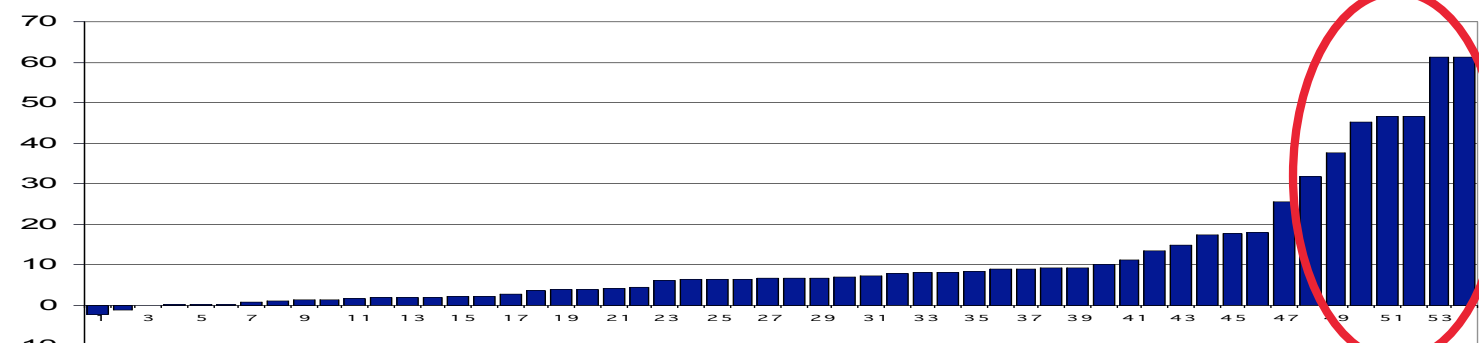


Afit

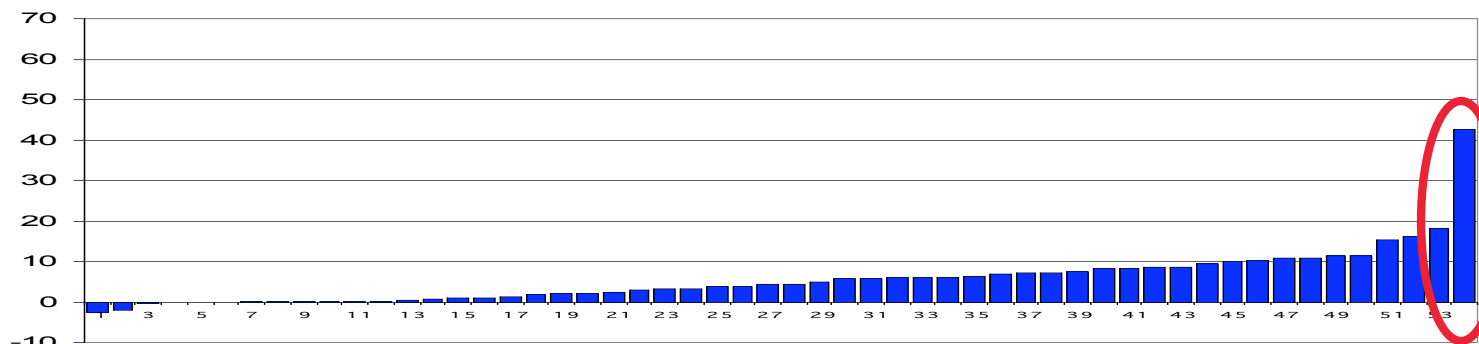


Global Strain

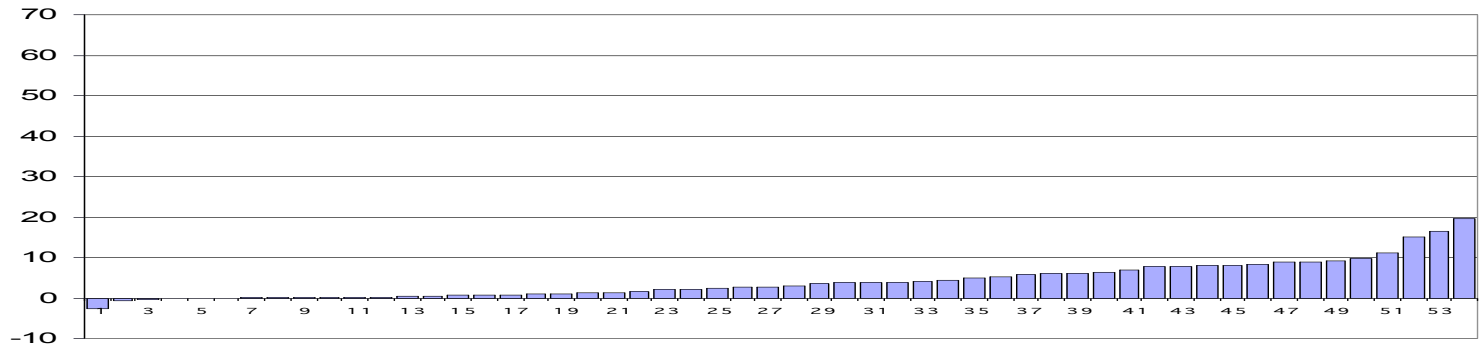
Strain Energy (kcal/mol)



Deposited



Refined

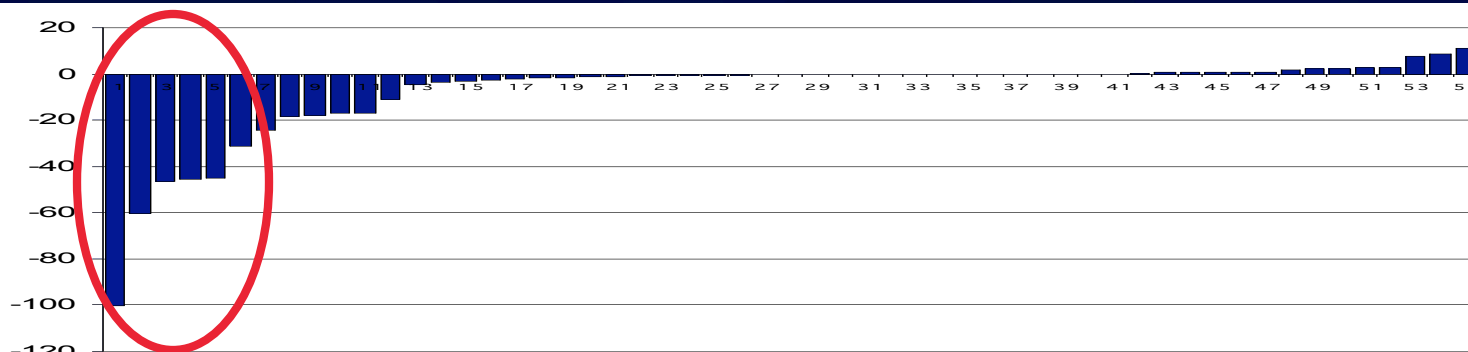


Afitt

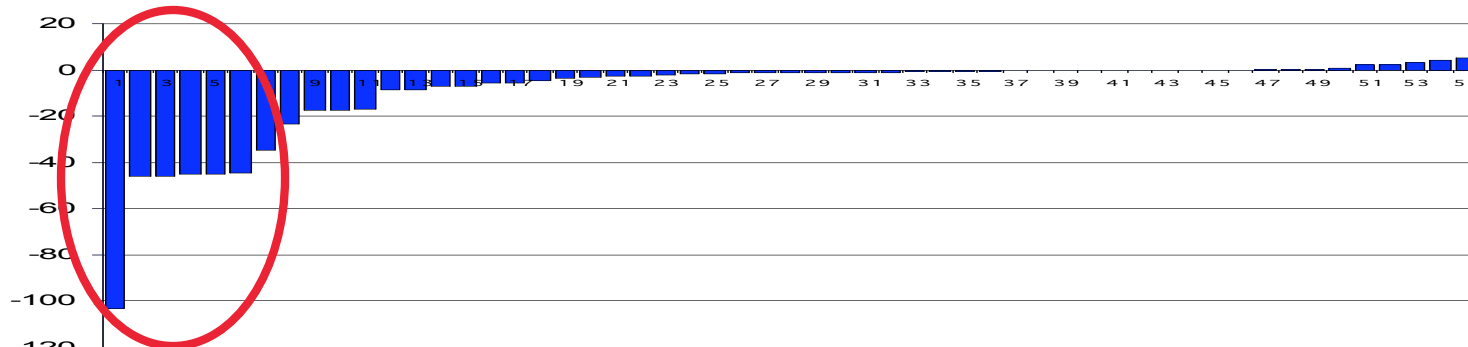


Comparison of optimized structures

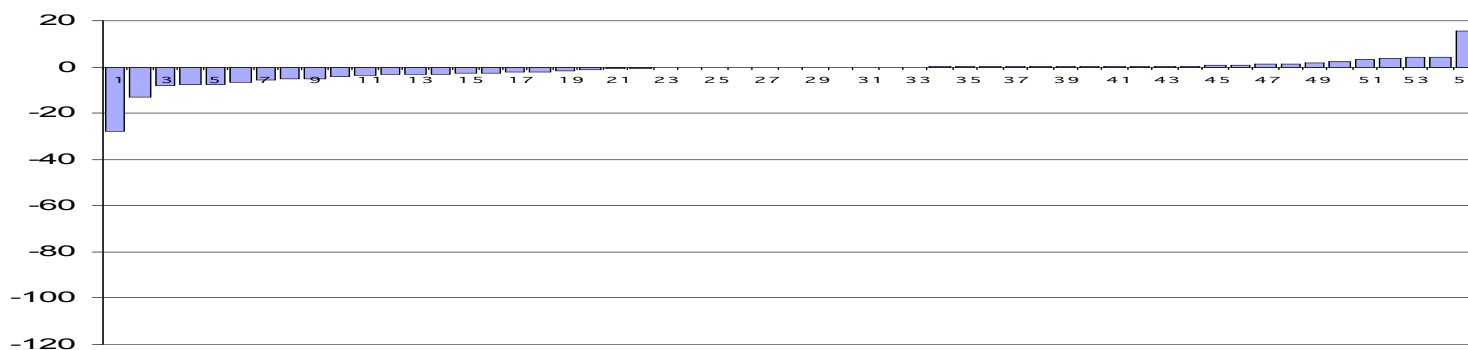
Strain Energy (kcal/mol)



RD-D



A-D

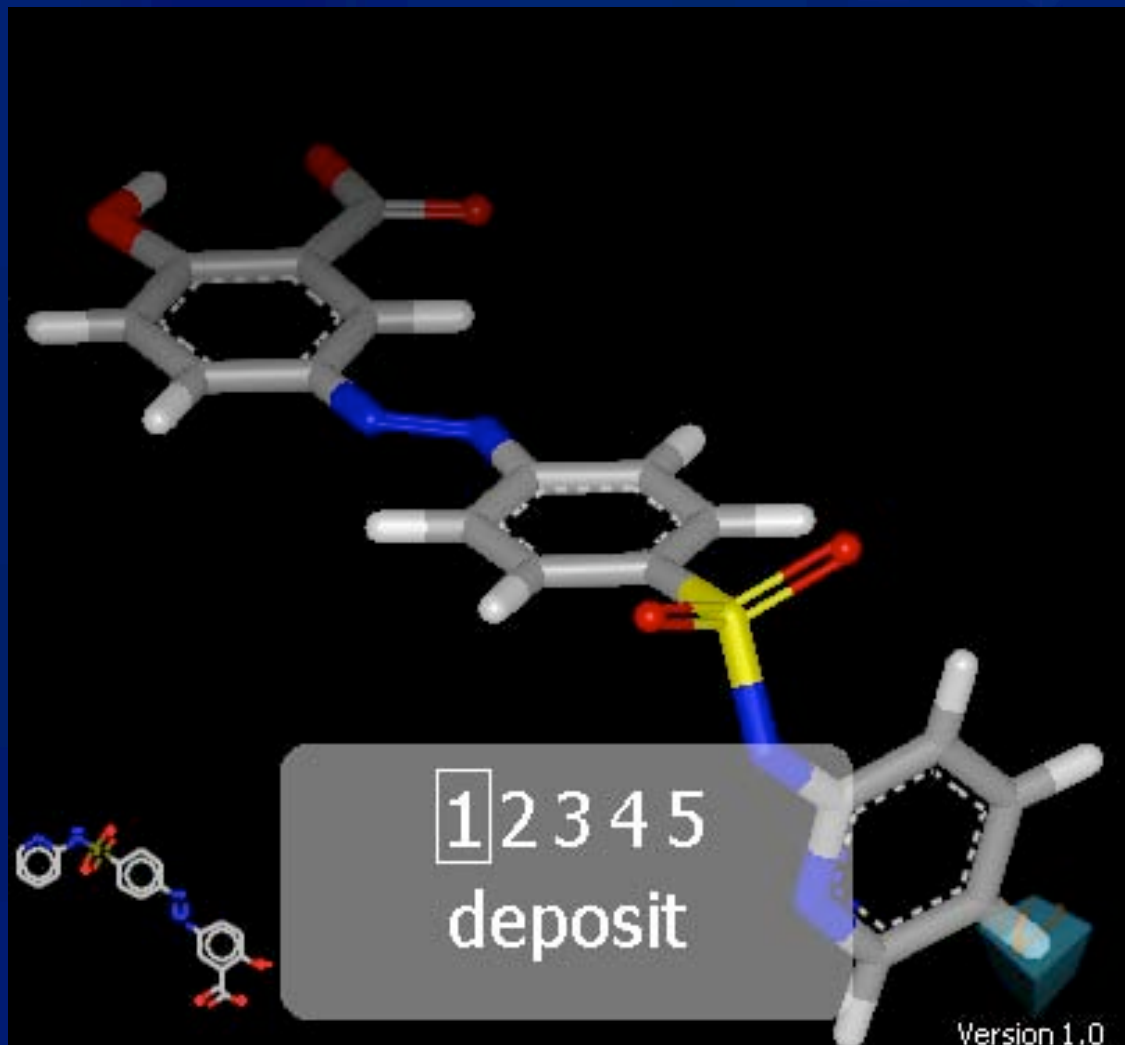


A-RD

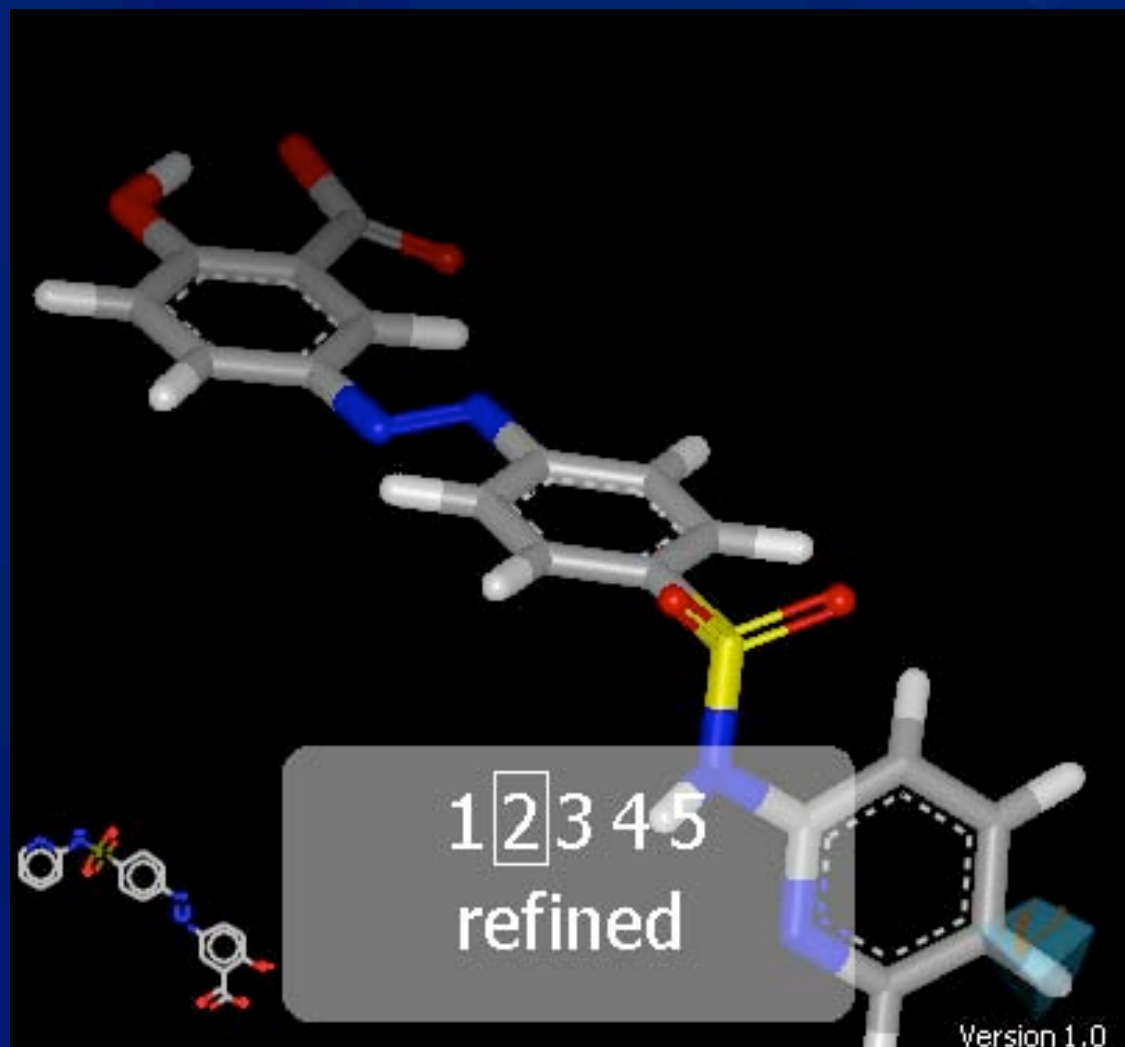


13GS

- Deposited global=61.2



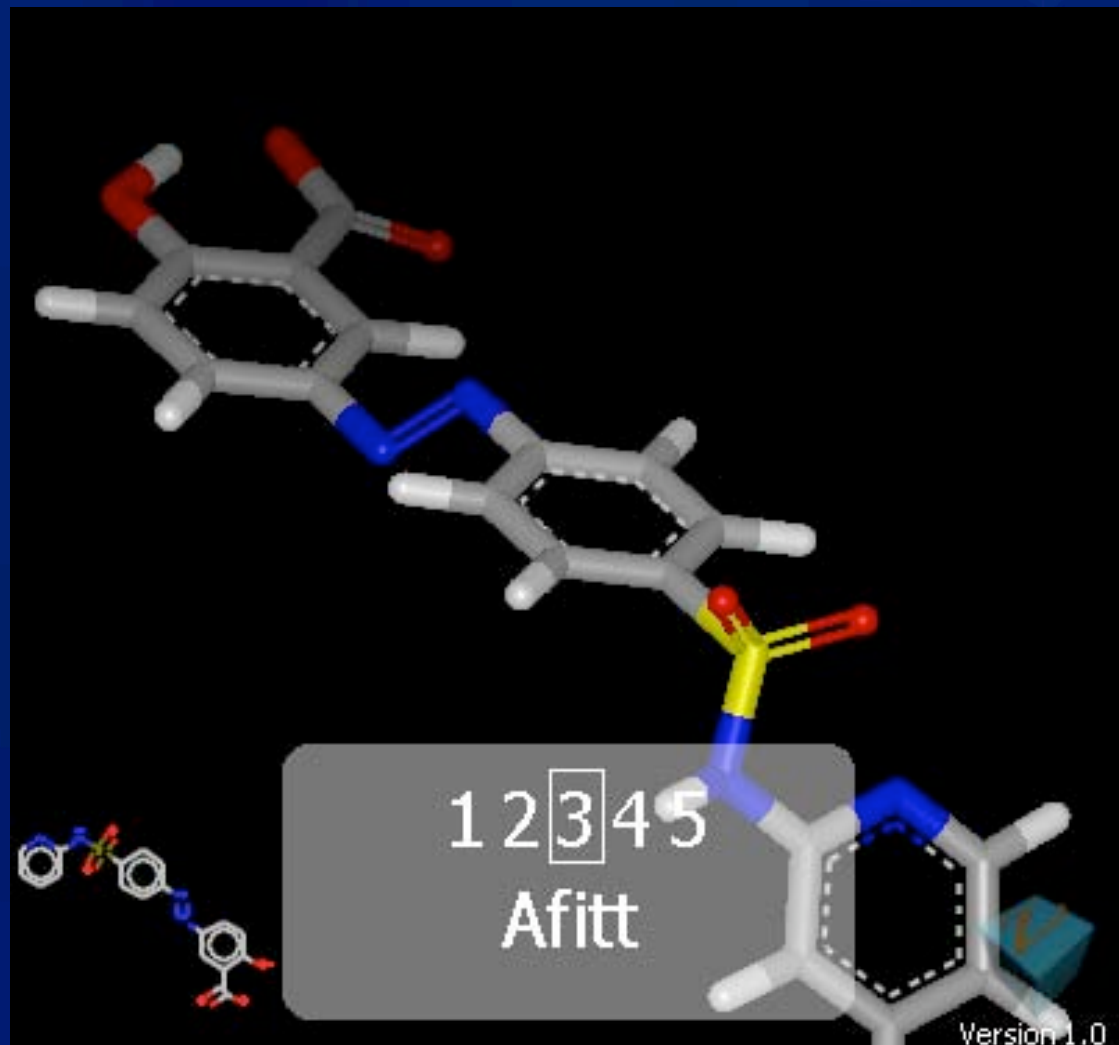
- Deposited
global=61.2



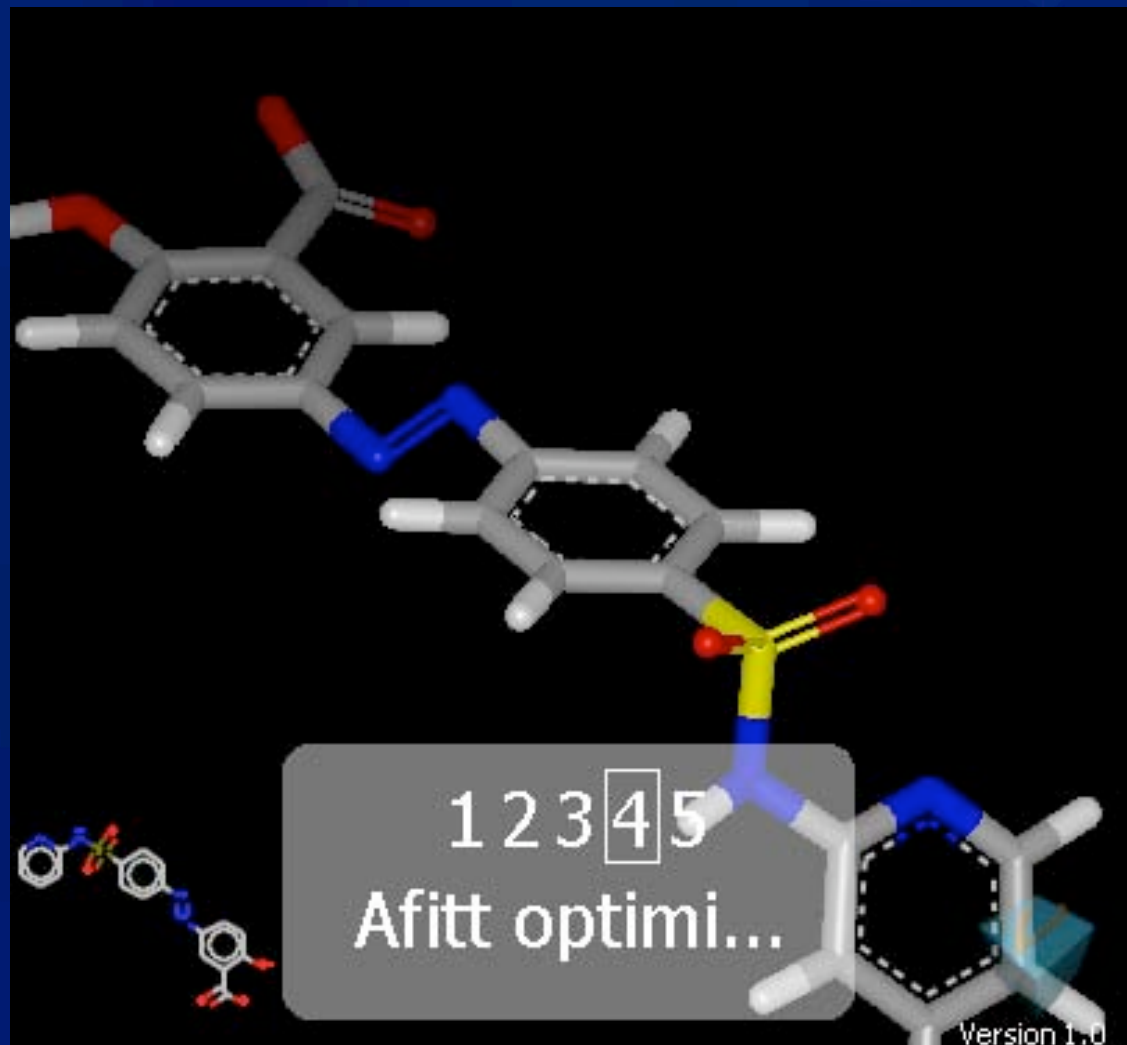
Version 1.0



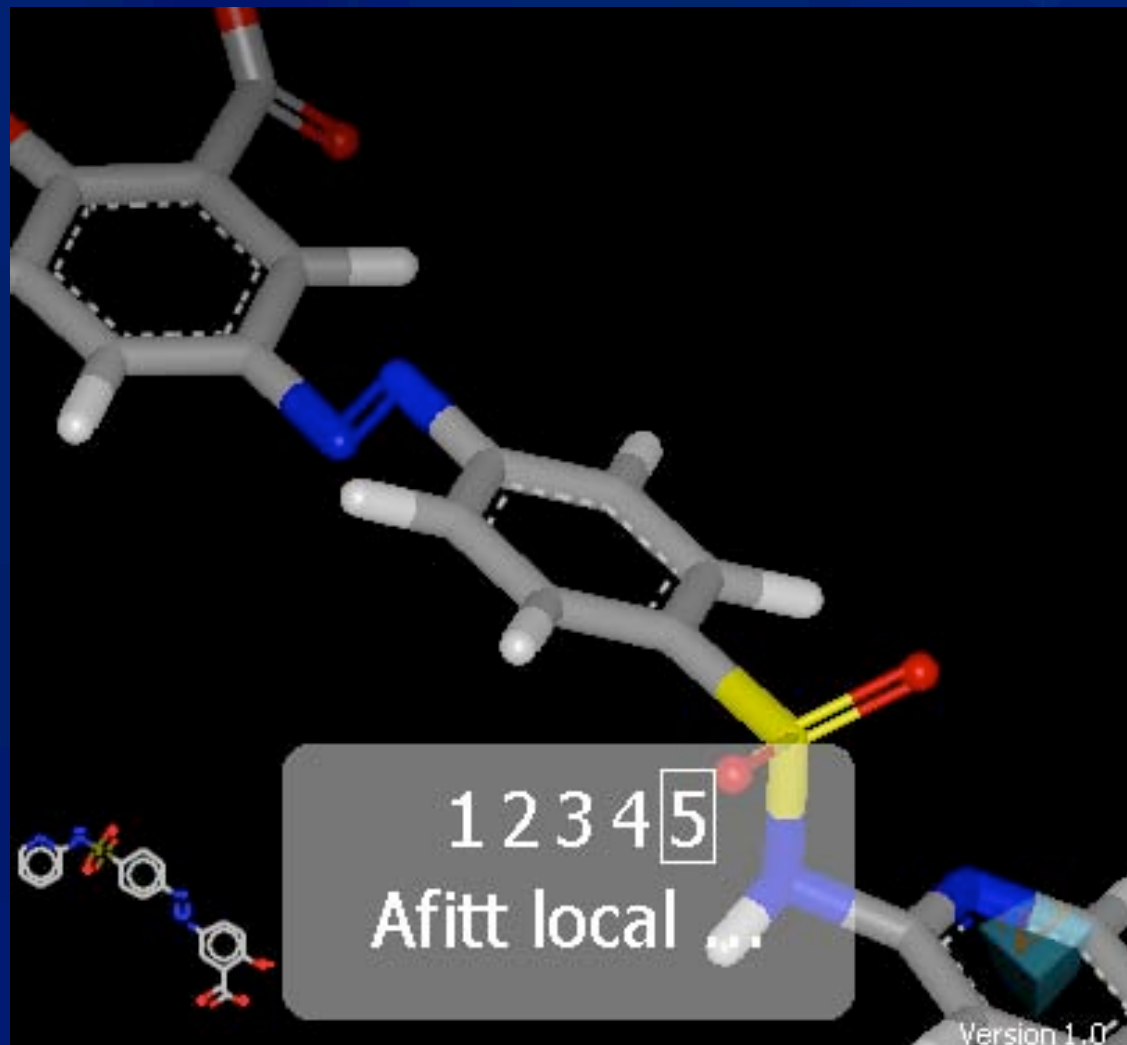
- Deposited
global=61.2



- Deposited
global=61.2

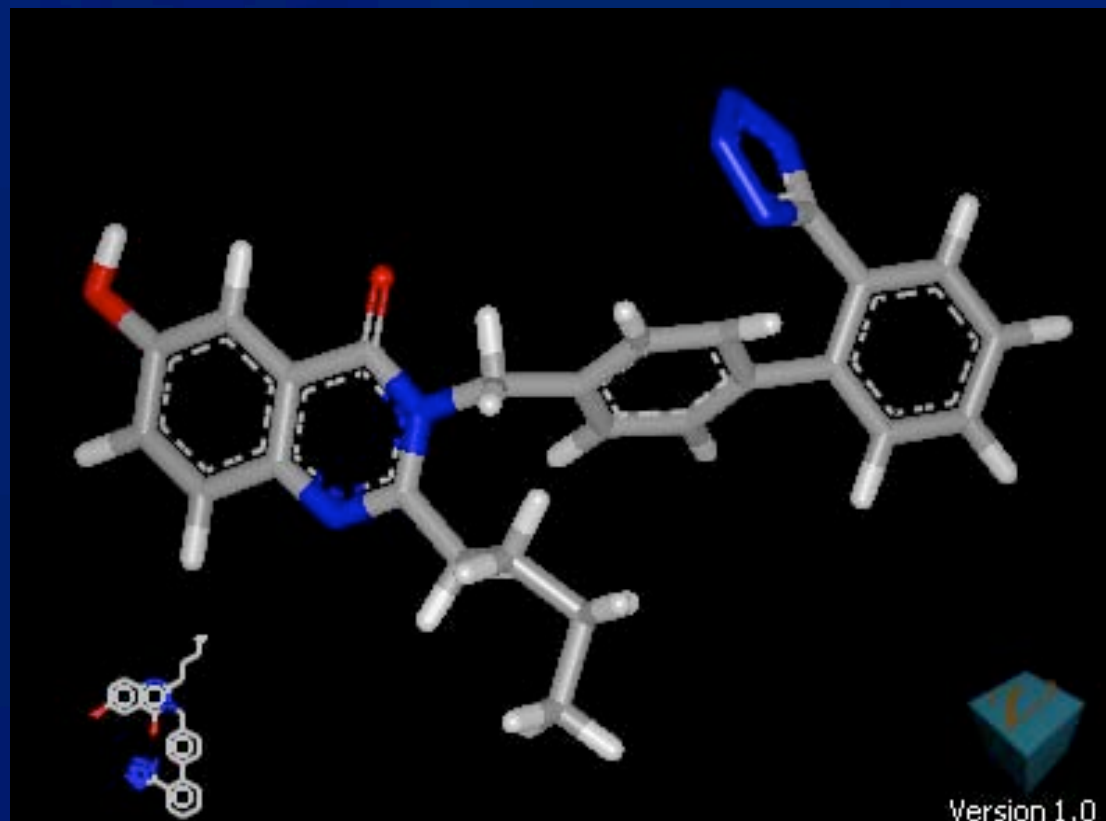


- Deposited
global=61.2

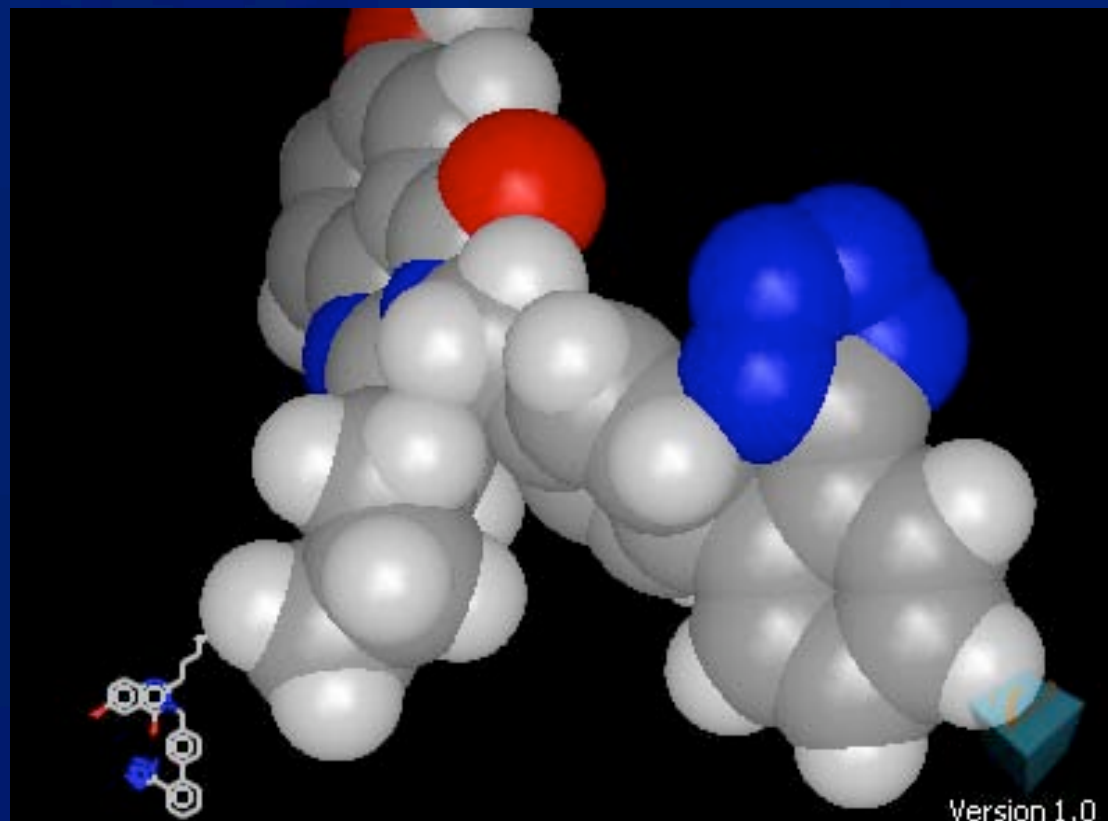


1A8T

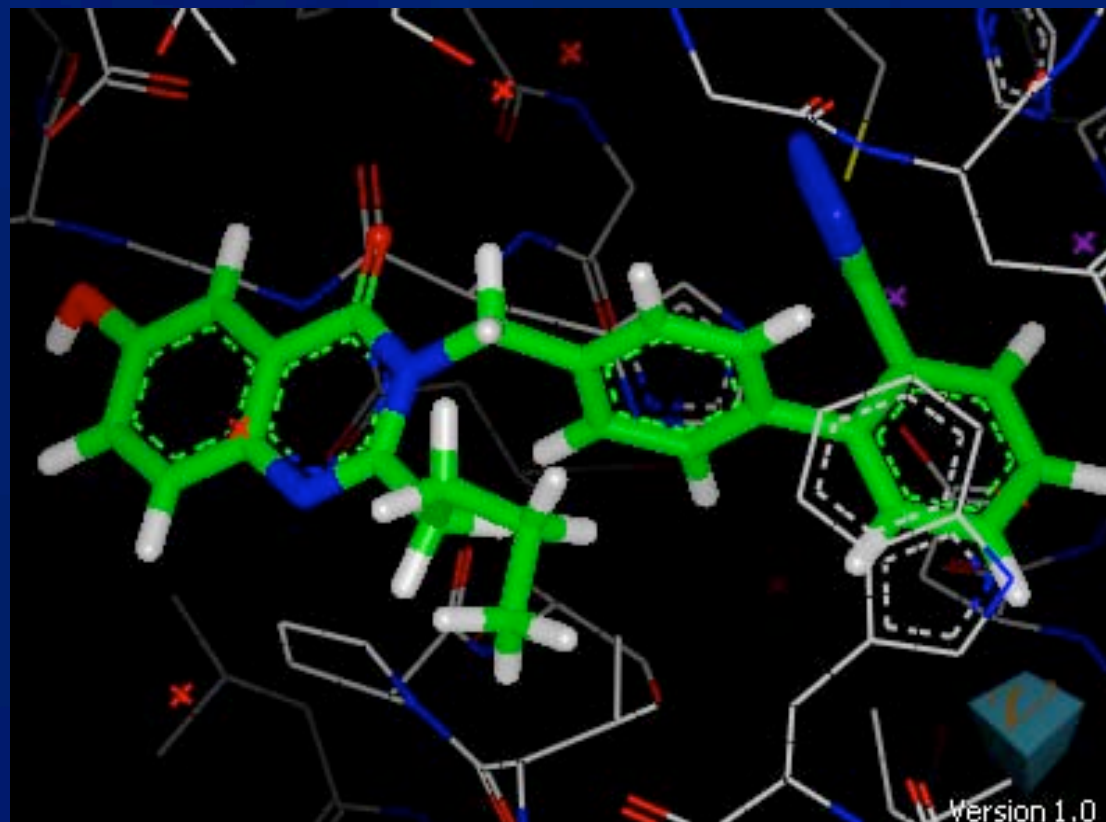
- Deposited to global=37.67
- Refined to global=6.2
- Afitt to global=3.1



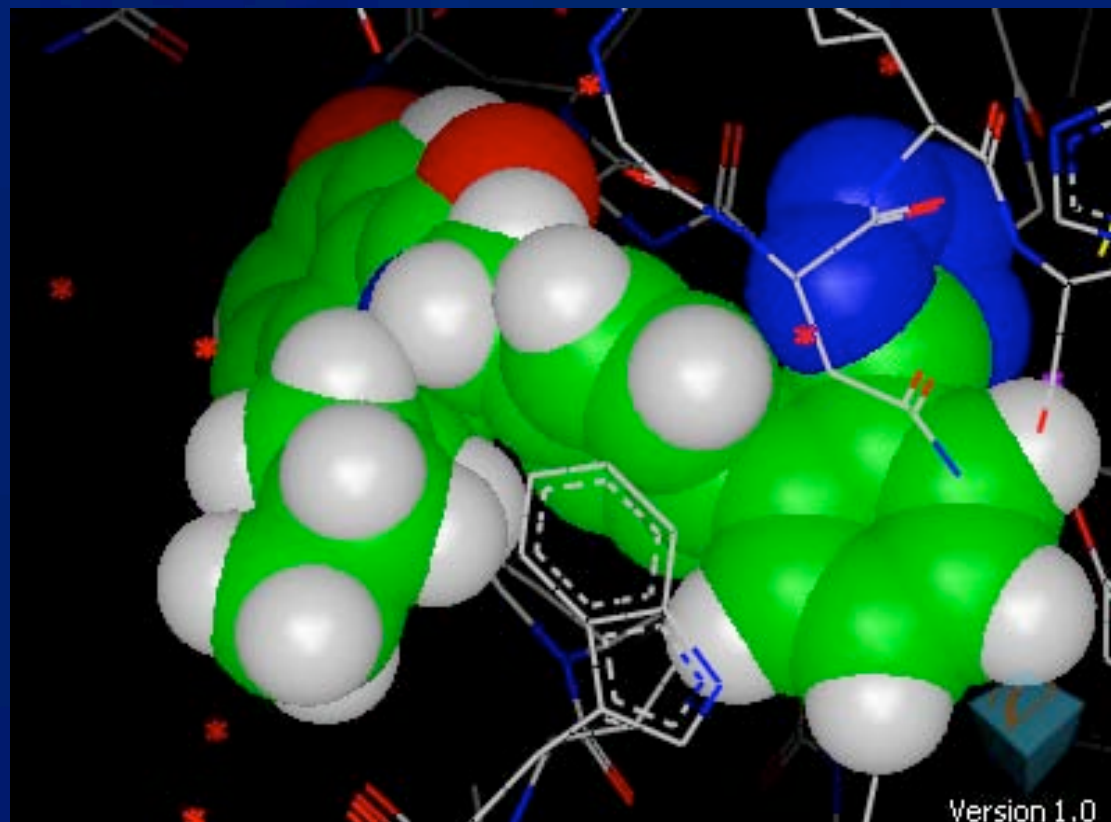
- Deposited to global=37.67
- Refined to global=6.2
- Afitt to global=3.1



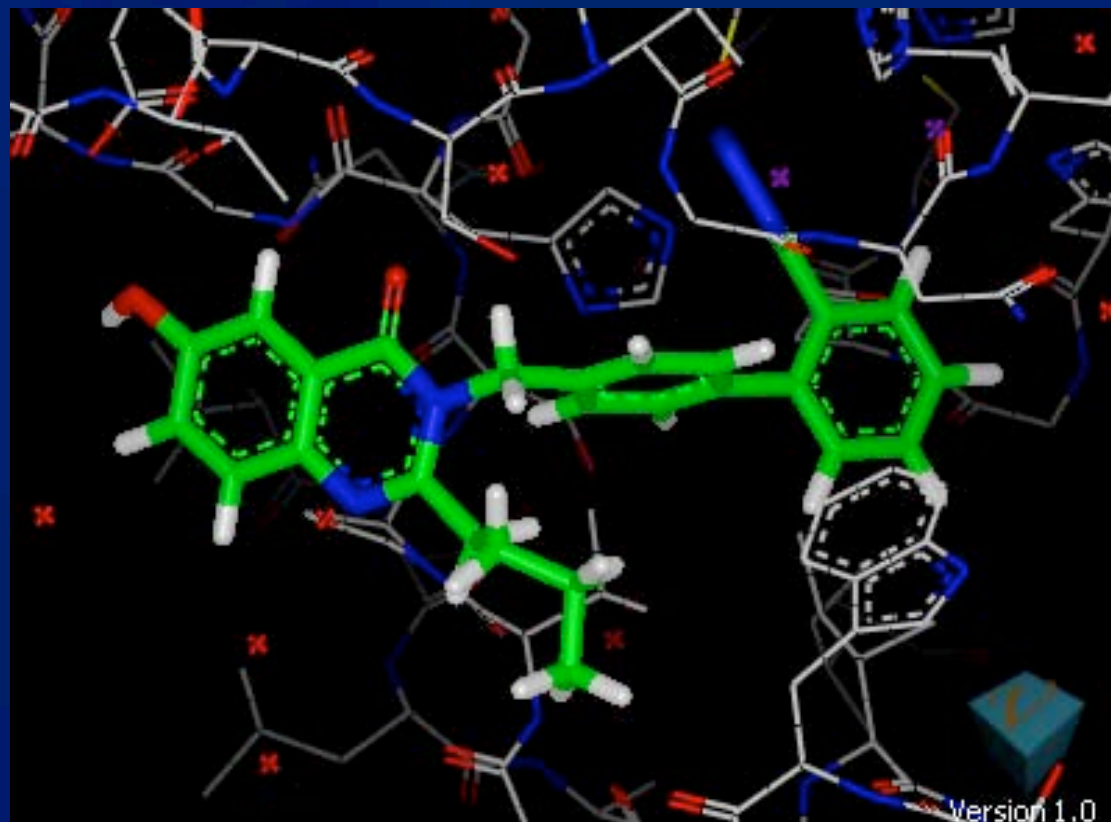
- Deposited to global=37.67
- Refined to global=6.2
- Afitt to global=3.1



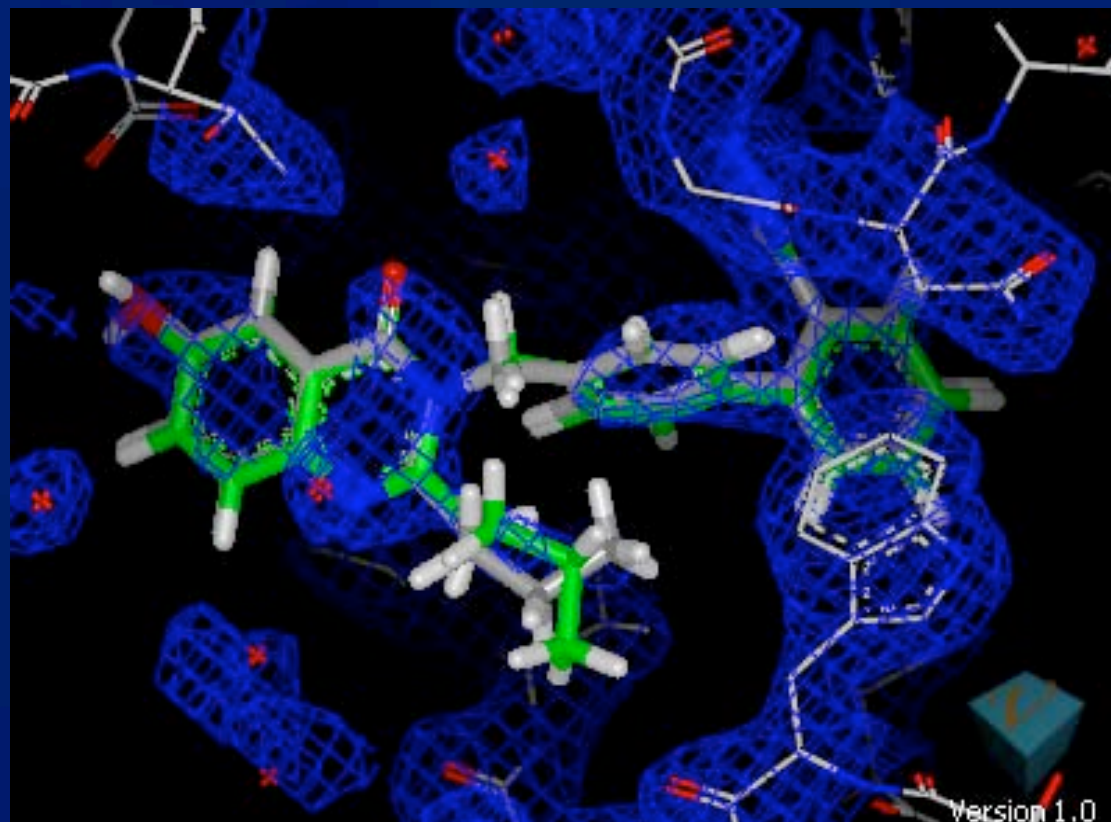
- Deposited to global=37.67
- Refined to global=6.2
- Afitt to global=3.1



- Deposited to global=37.67
- Refined to global=6.2
- Afitt to global=3.1

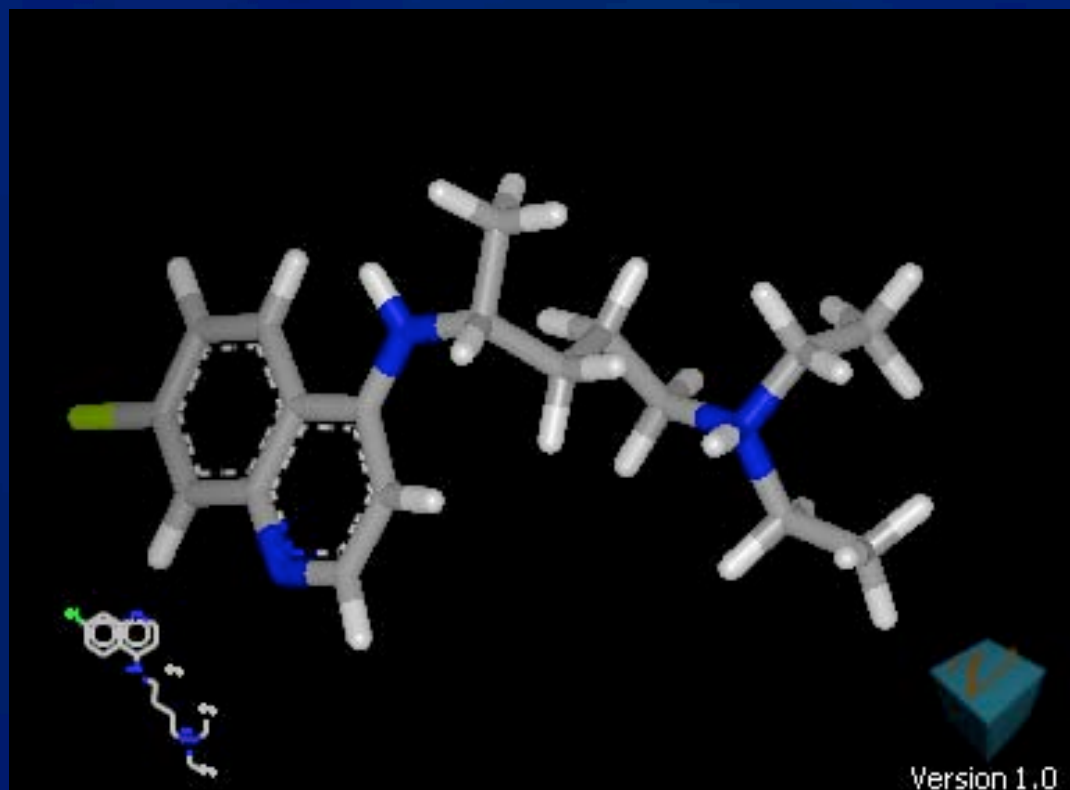


- Deposited to global=37.67
- Refined to global=6.2
- Afitt to global=3.1

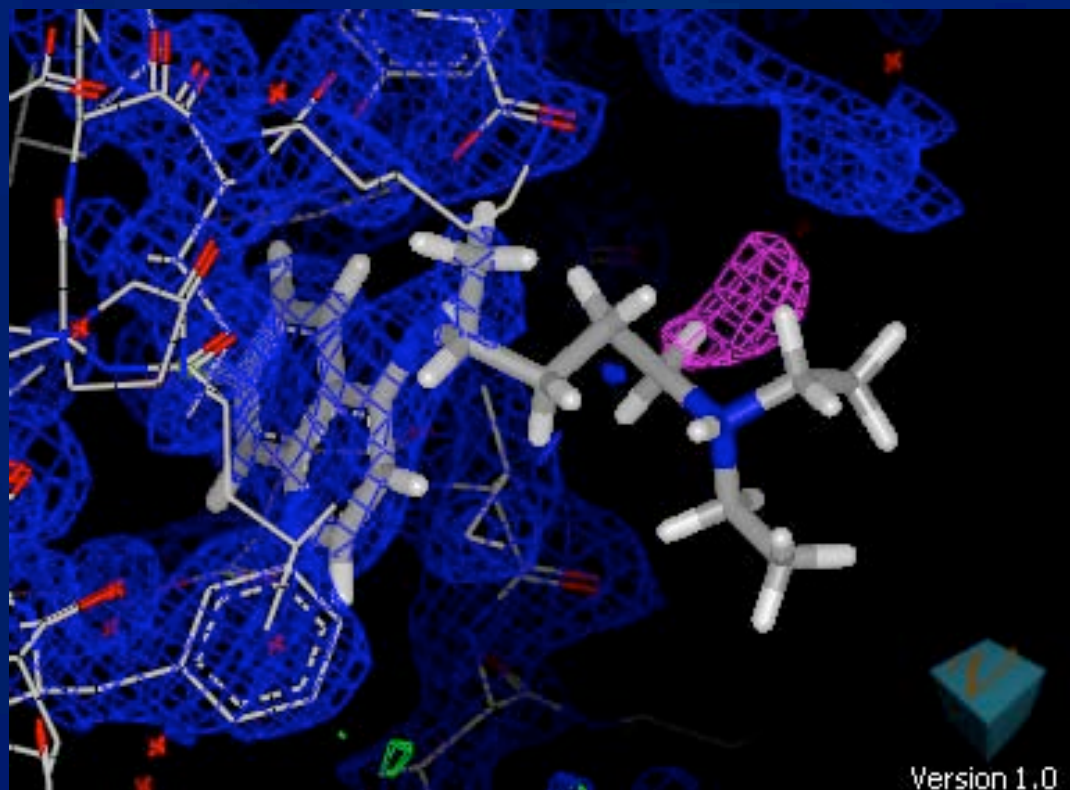


1CET

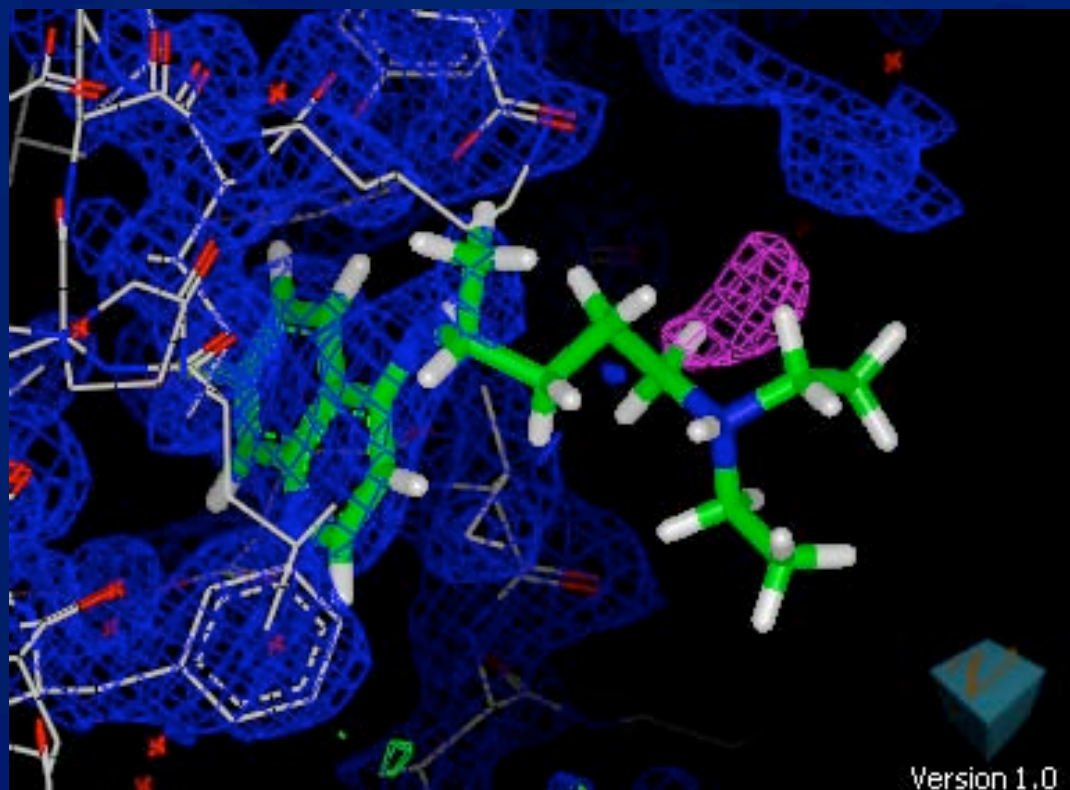
- Refined to global=16.3
- Refined to local=11.3
- Afitt to global=8.1



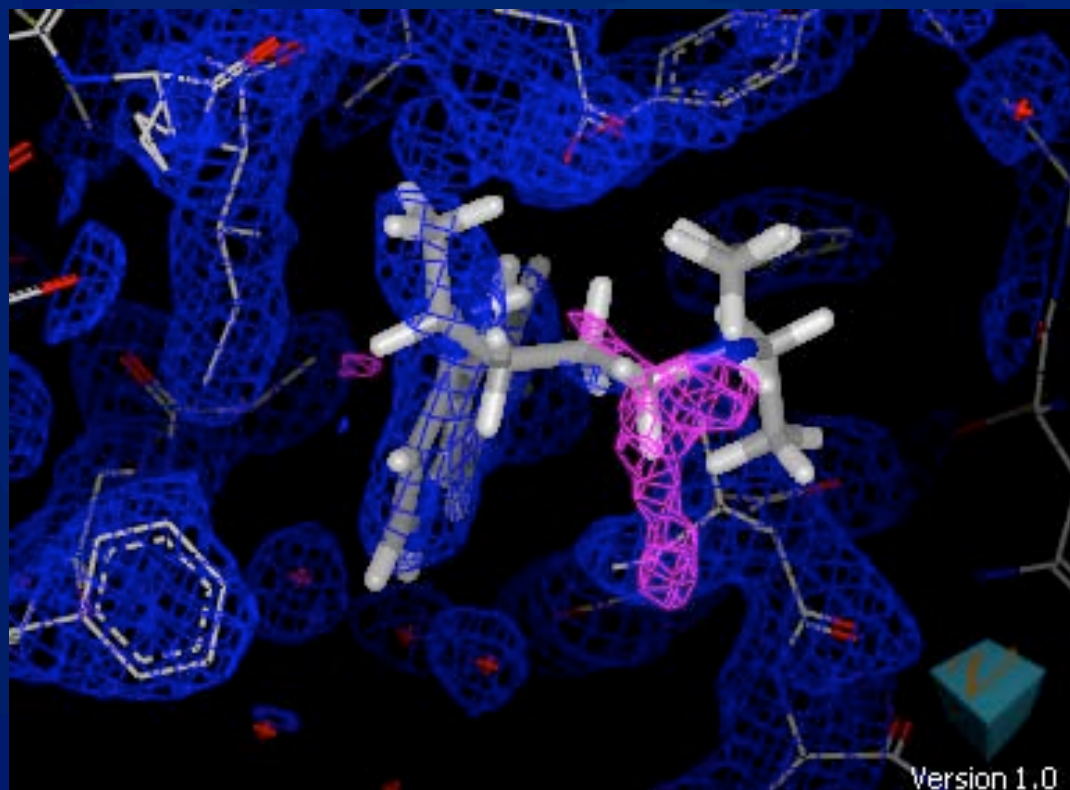
- Refined to global=16.3
- Refined to local=11.3
- Afitt to global=8.1



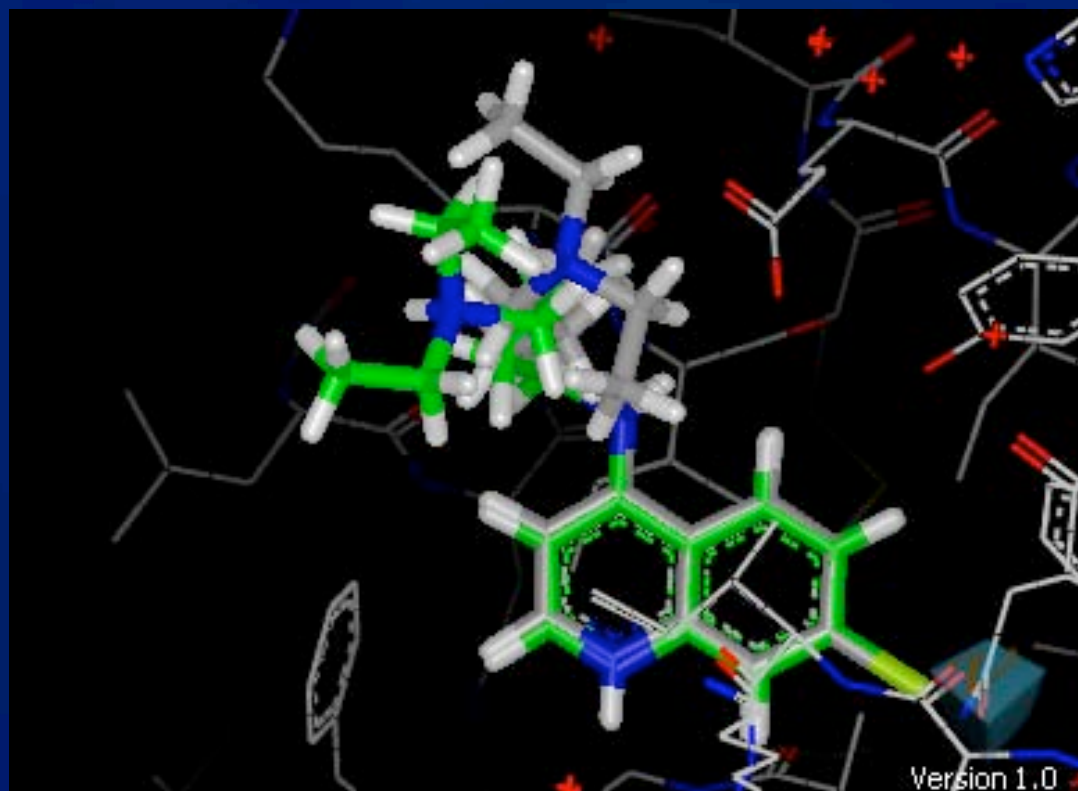
- Refined to global=16.3
- Refined to local=11.3
- Afitt to global=8.1



- Refined to global=16.3
- Refined to local=11.3
- Afitt to global=8.1

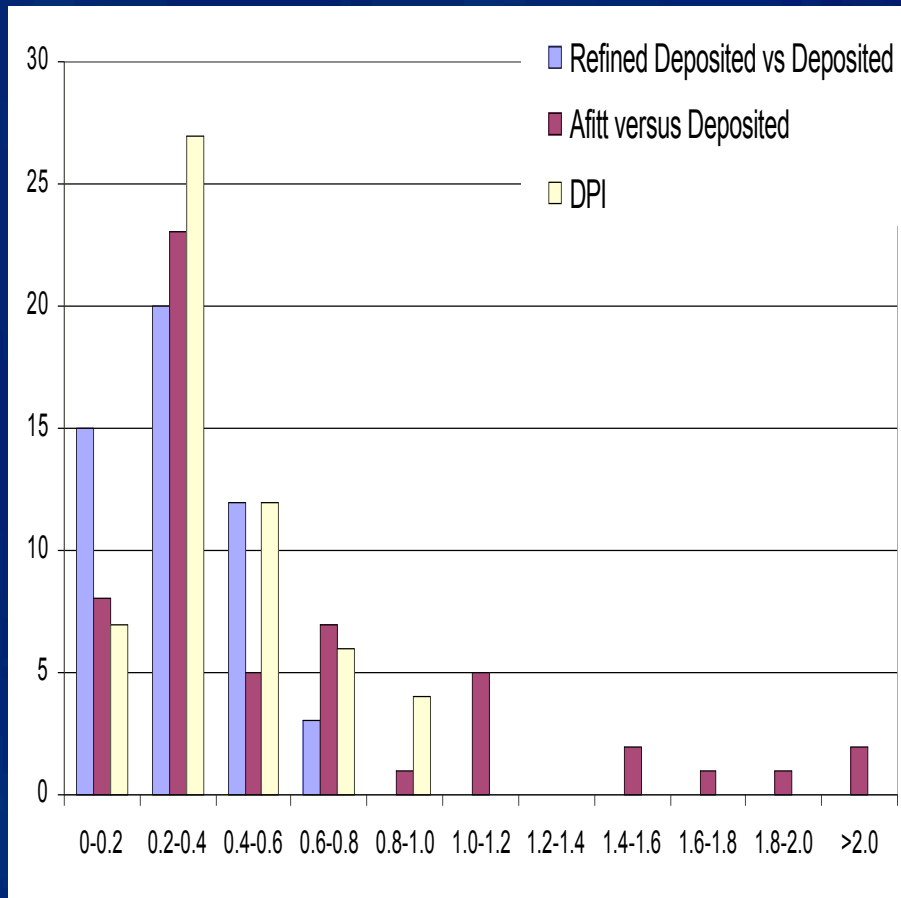


- Refined to global=16.3
- Refined to local=11.3
- Afitt to global=8.1

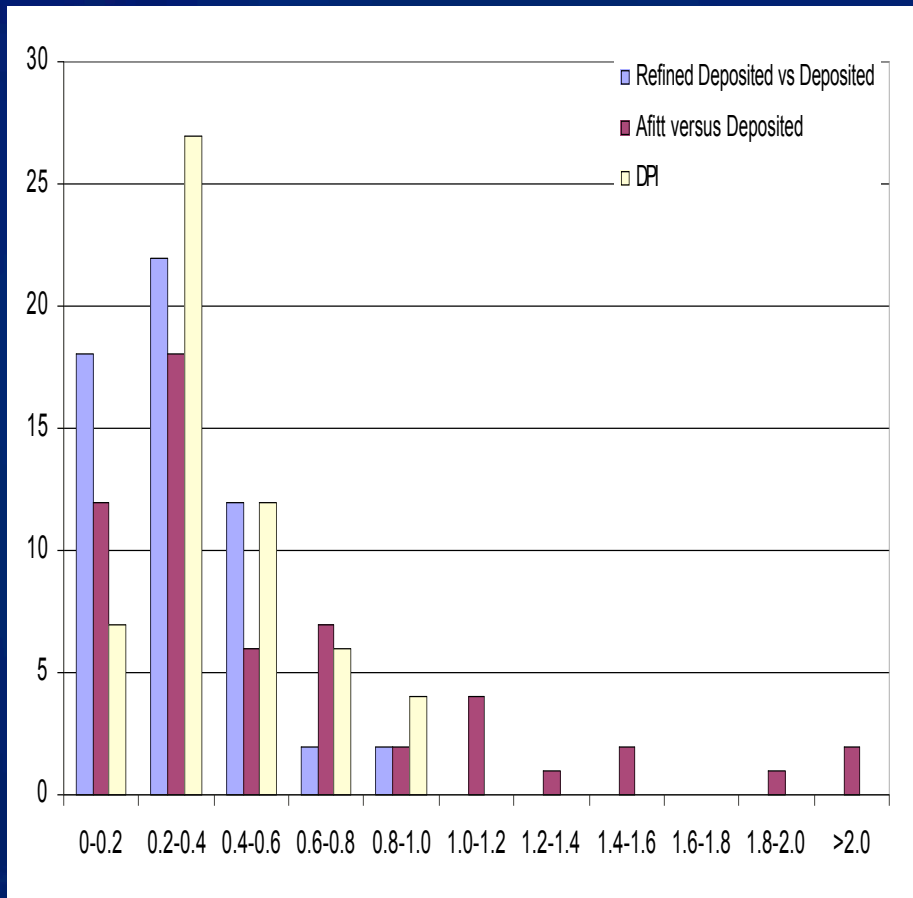


So are the structures different?

Non-Optimized

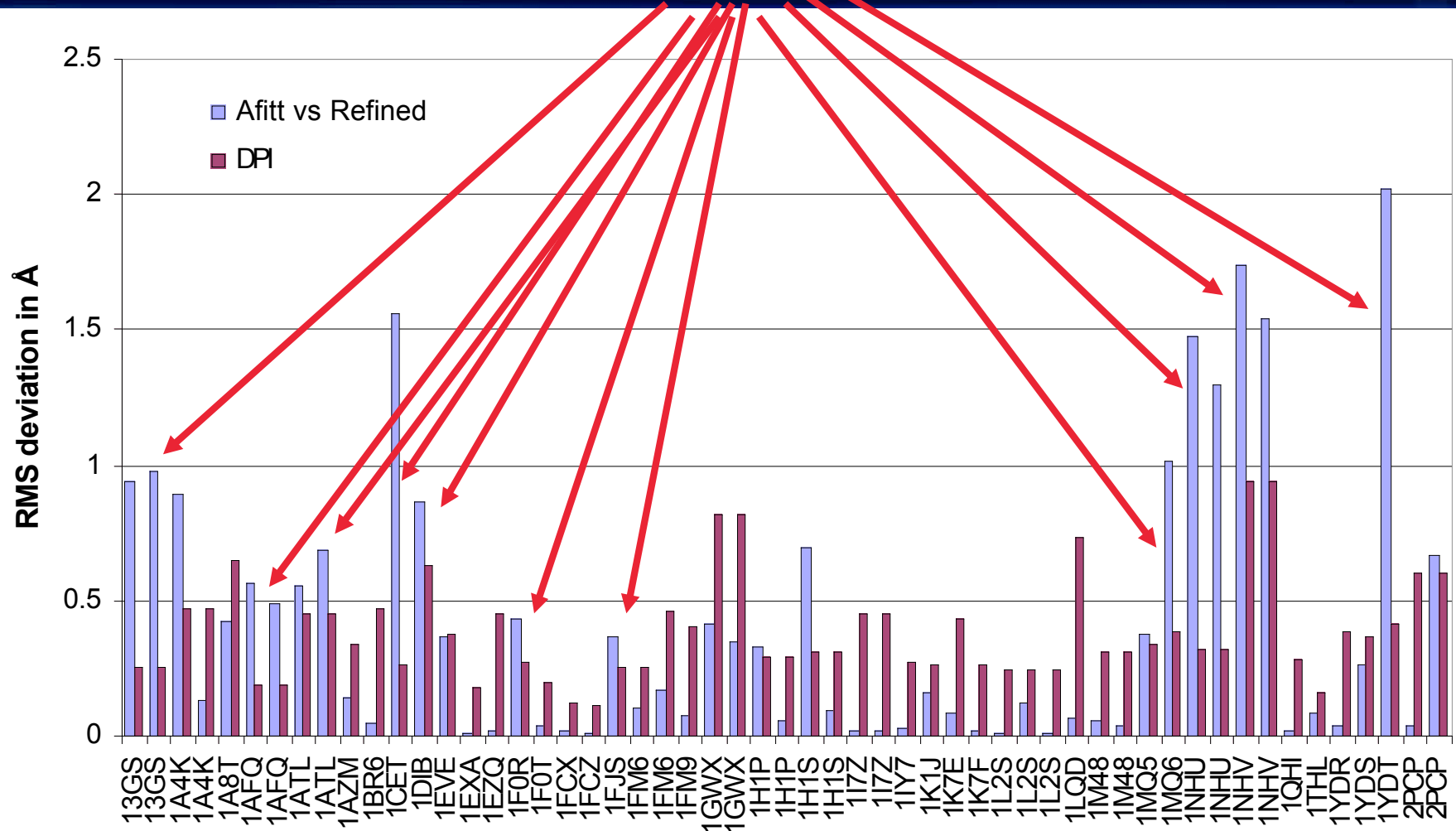


Optimized



RMSD Afitt from refined deposited

Outliers

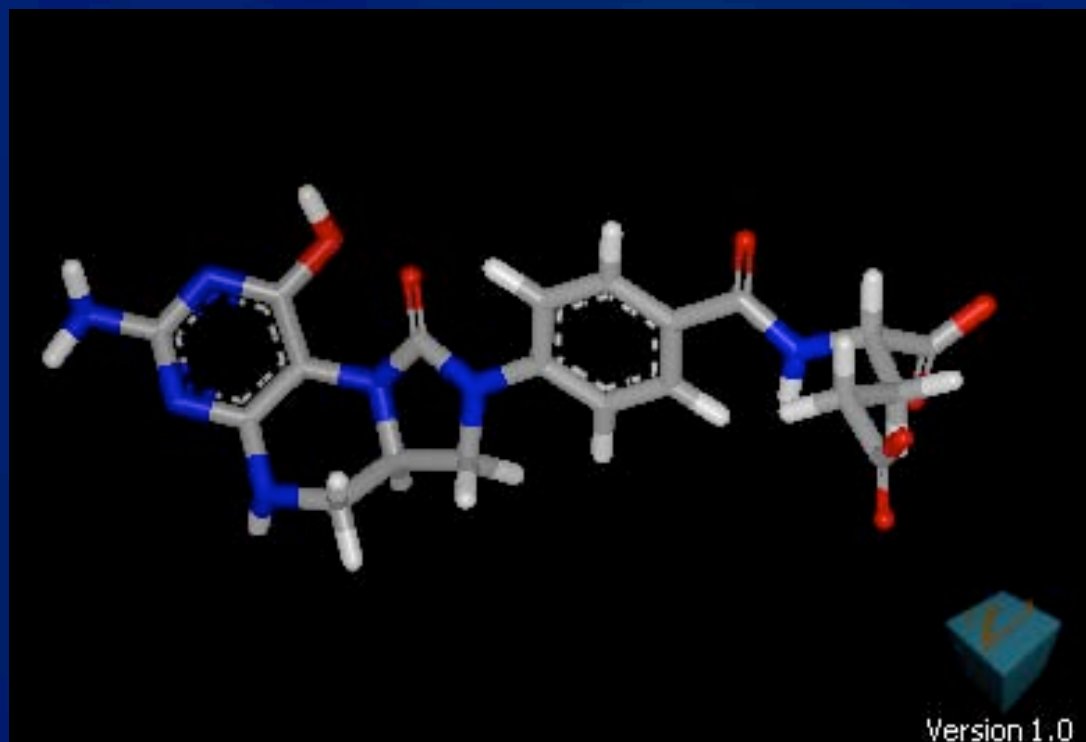


Some of the Outliers

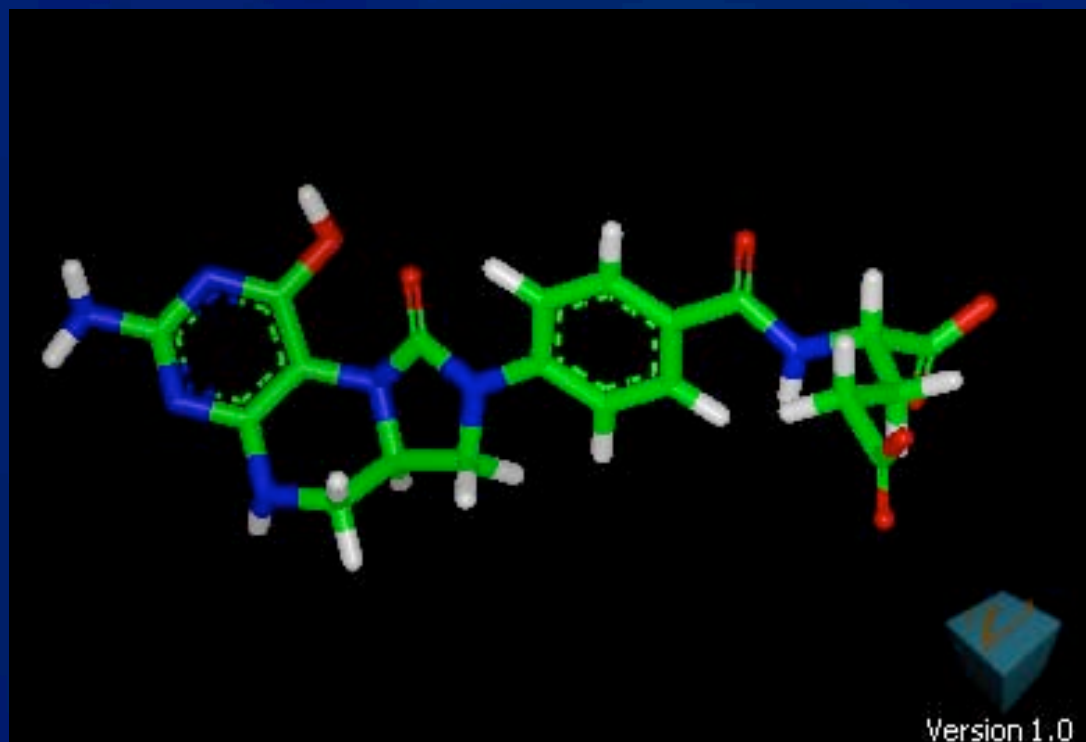
- 13GS
- 1ATL
- 1CET
- 1F0R
- 1MQ6
- 1NHV
- Force field
- No density
- Strain-local,global
- Strain-global
- Strain-local
- Strain-global

1DIB

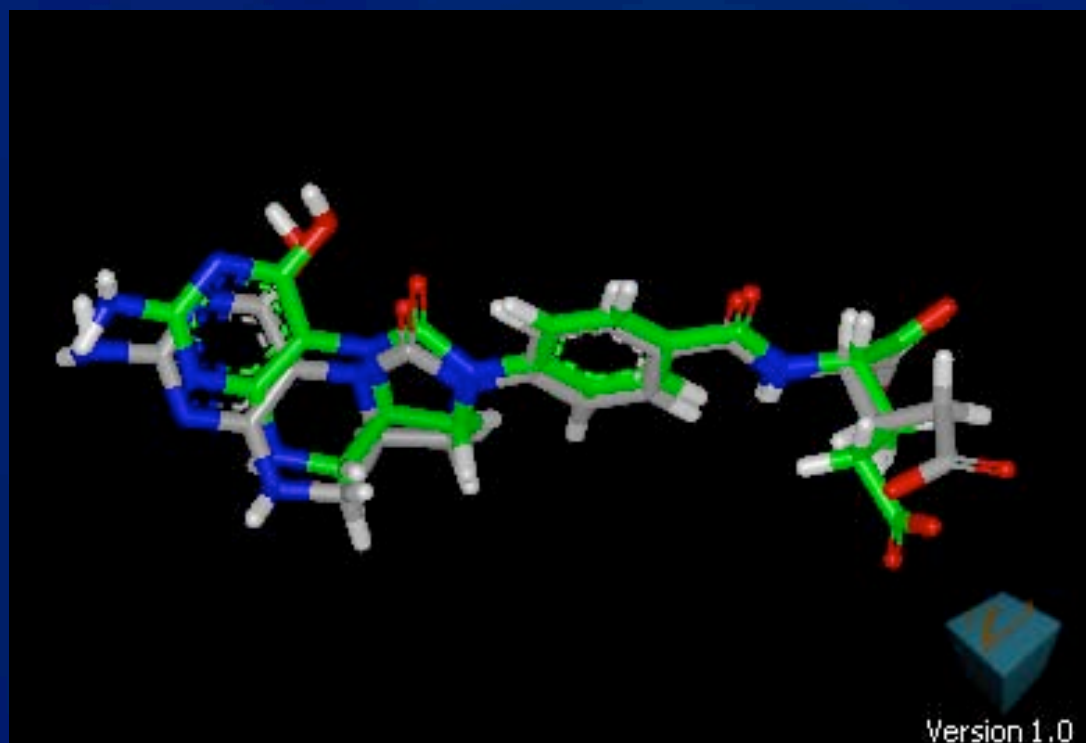
- Refined to global=15.4
- Afitt to global=2.5



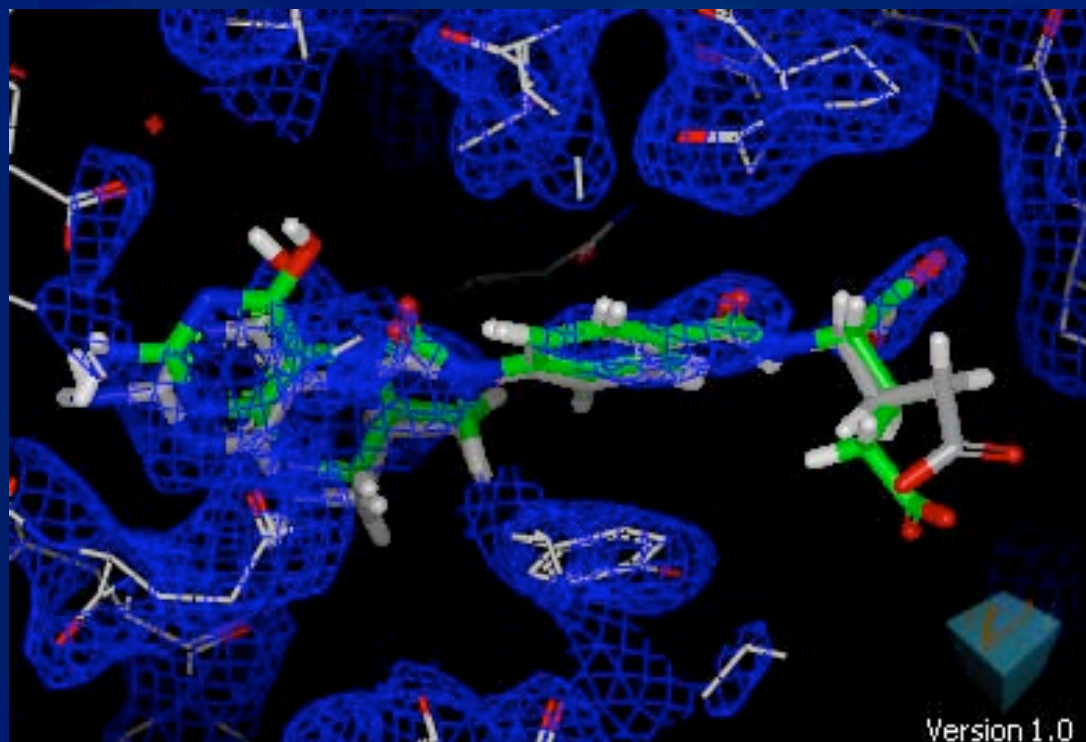
- Refined to global=15.4
- Afitt to global=2.5



- Refined to global=15.4
- Afitt to global=2.5

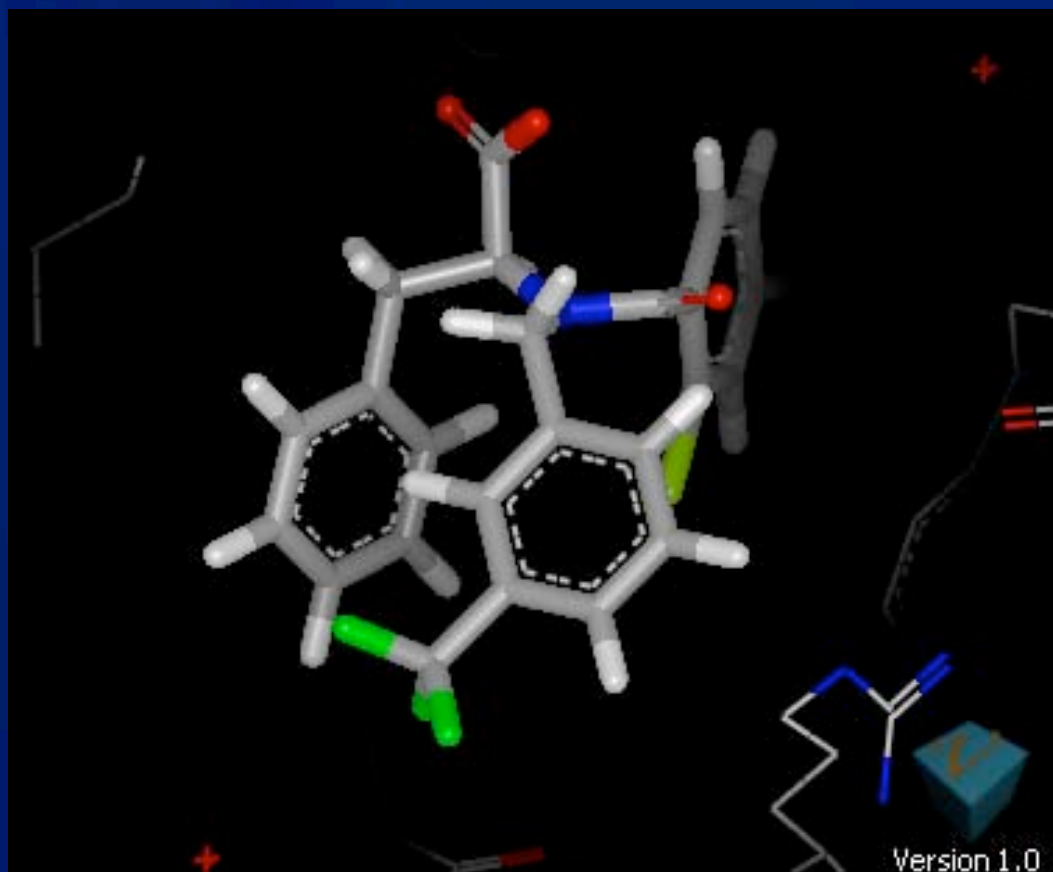


- Refined to global=15.4
- Afitt to global=2.5

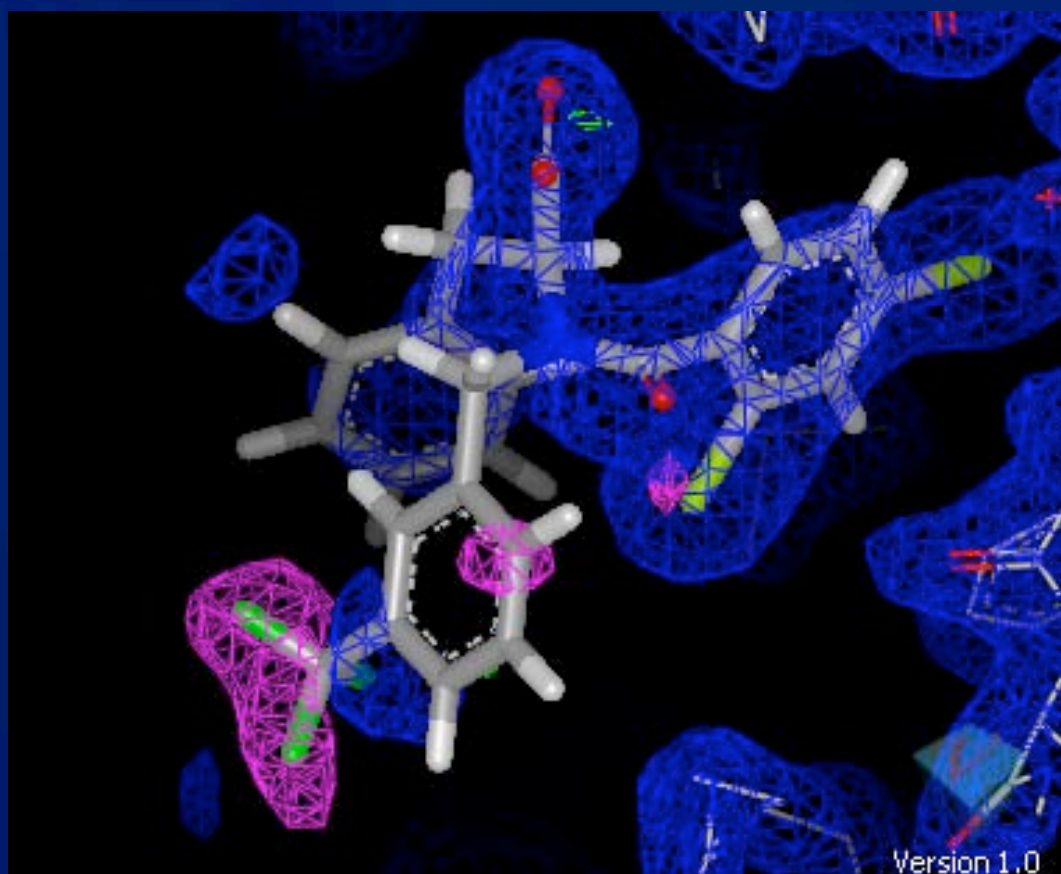


1NHU

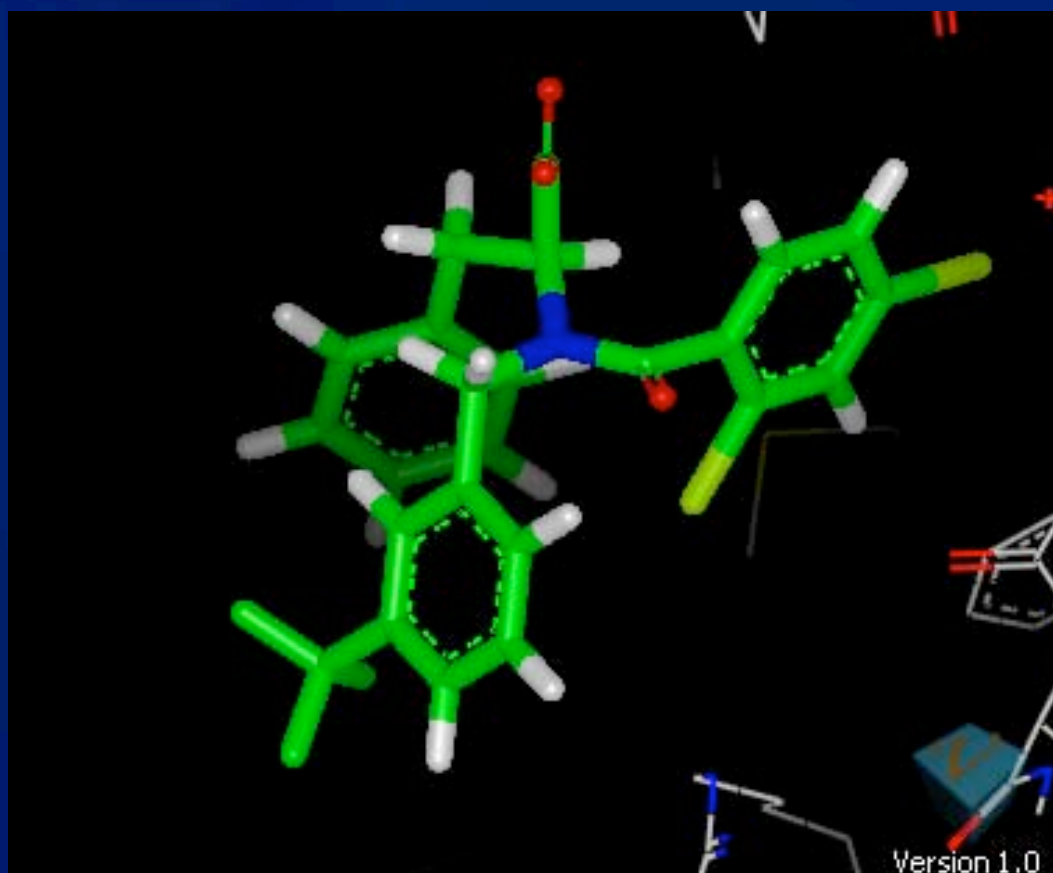
- Refined to global=4.3
- Afitt to global=-0.6



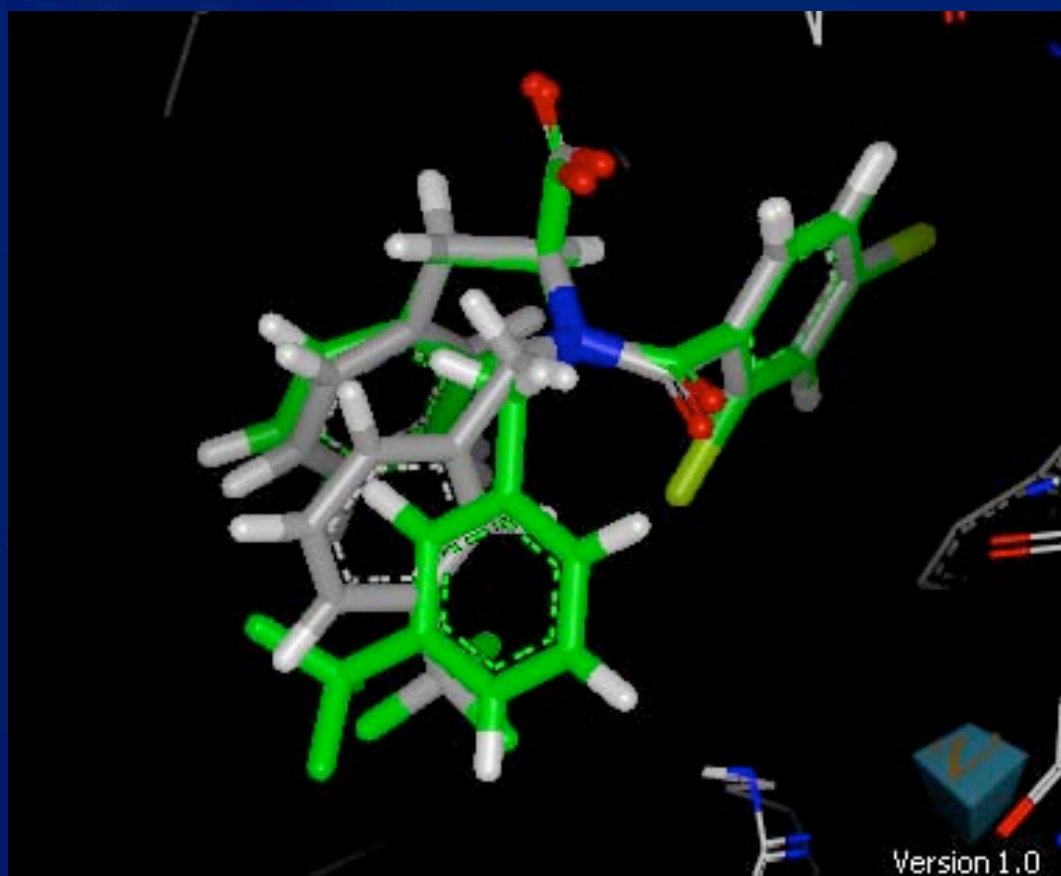
- Refined to global=4.3
- Afitt to global=-0.6



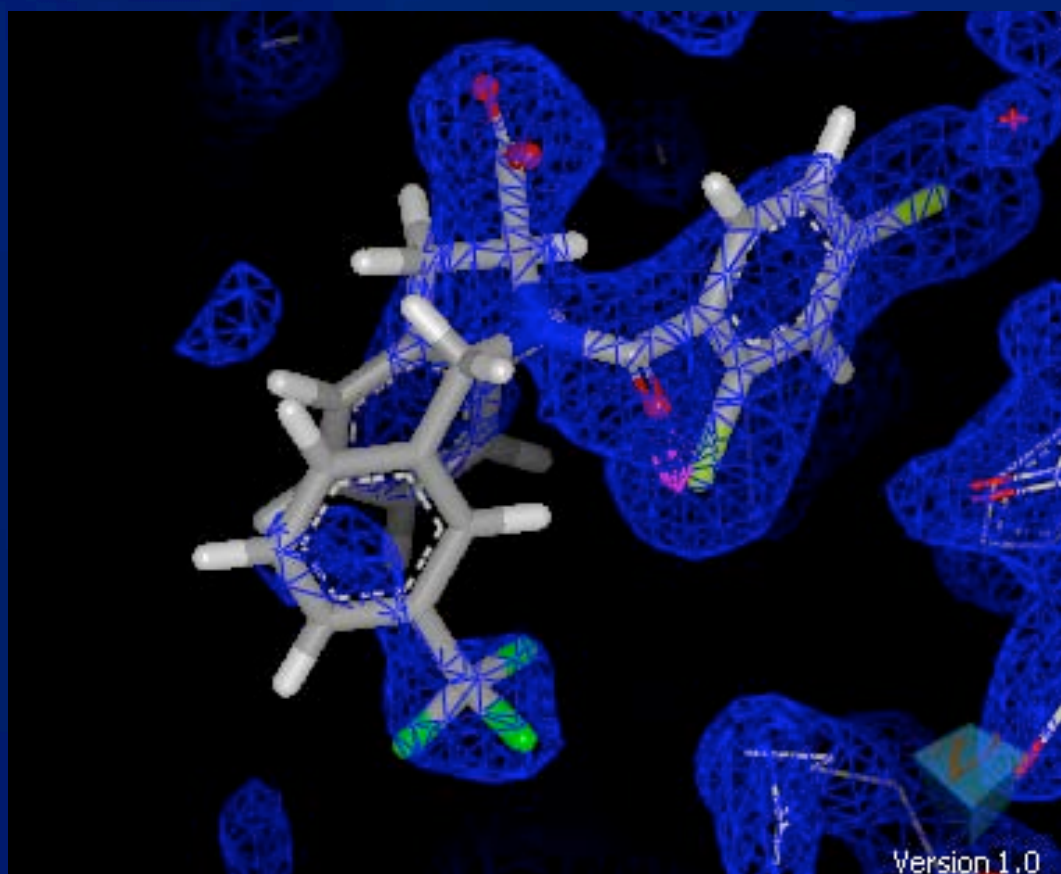
- Refined to global=4.3
- Afitt to global=-0.6



- Refined to global=4.3
- Afitt to global=-0.6

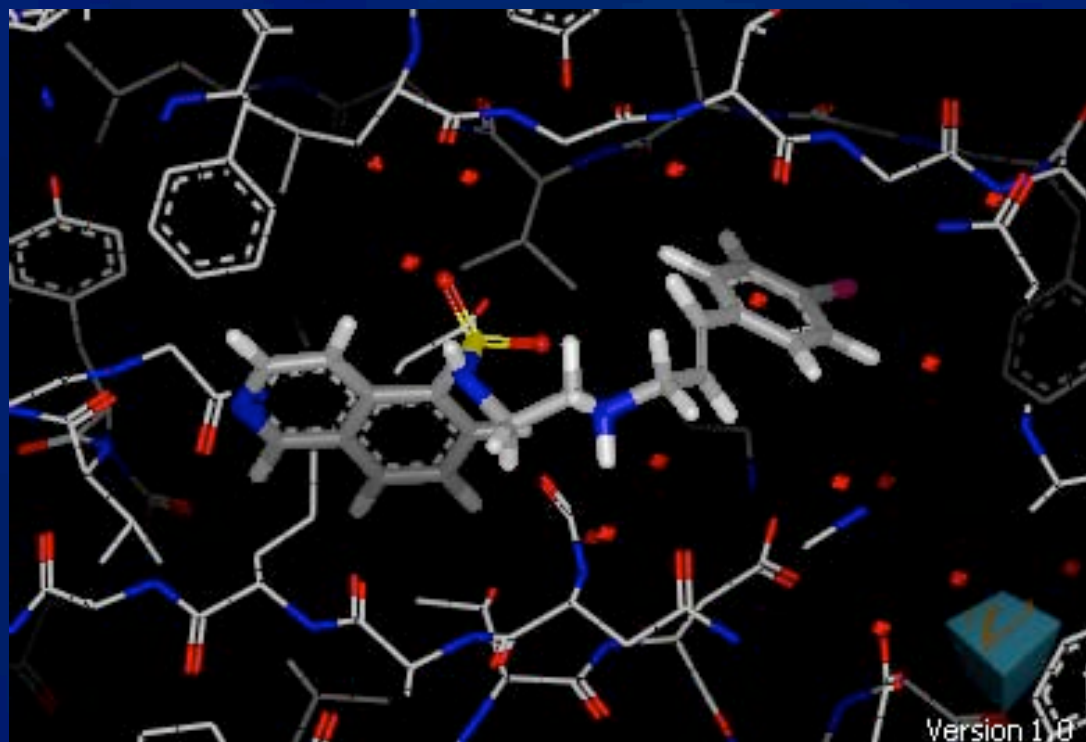


- Refined to global=4.3
- Afitt to global=-0.6

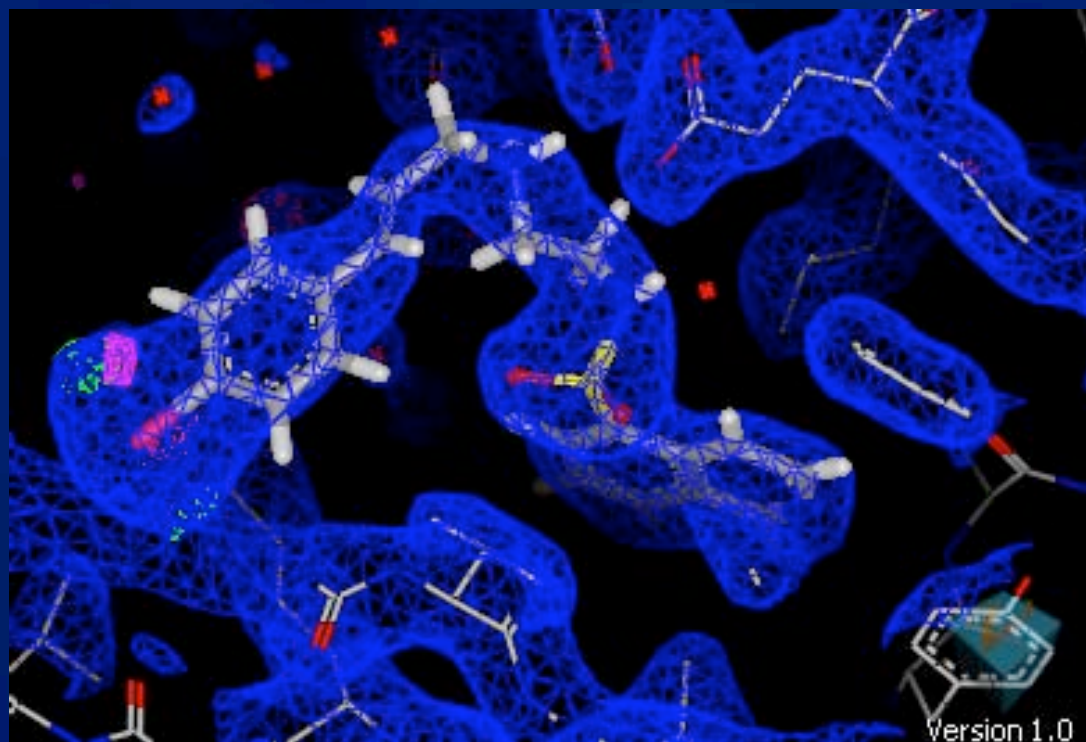


1YDT

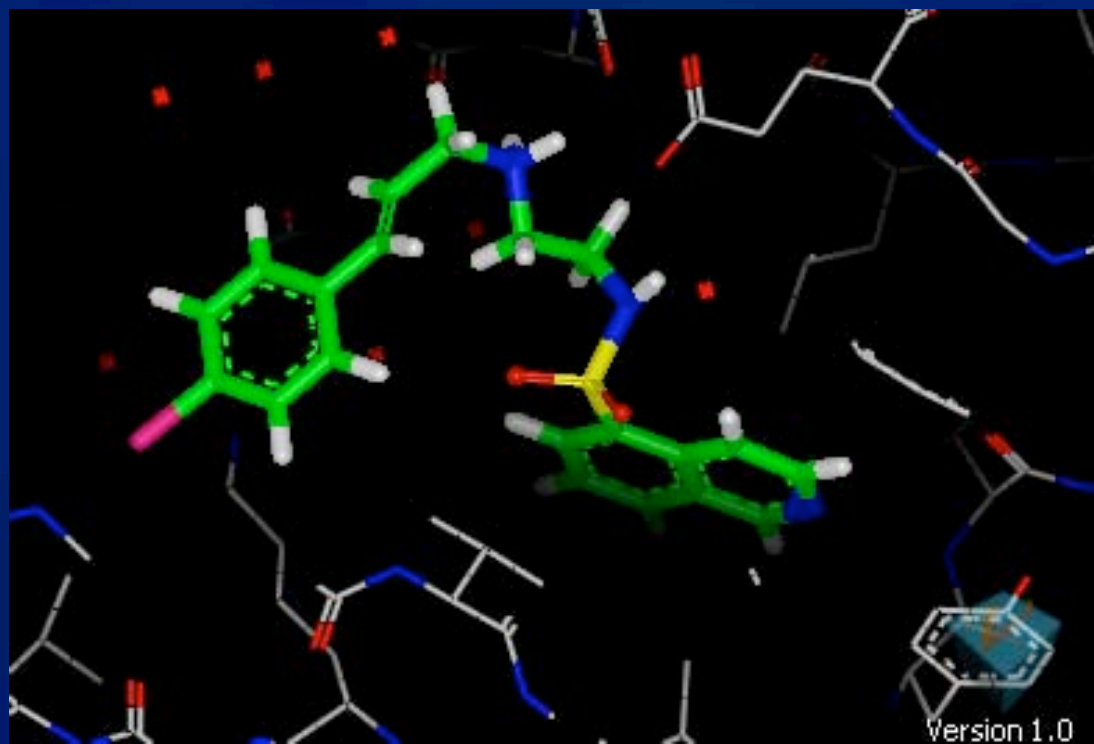
- Refined to global=11.3
- Afitt to global=8.0
 - But it's probably wrong



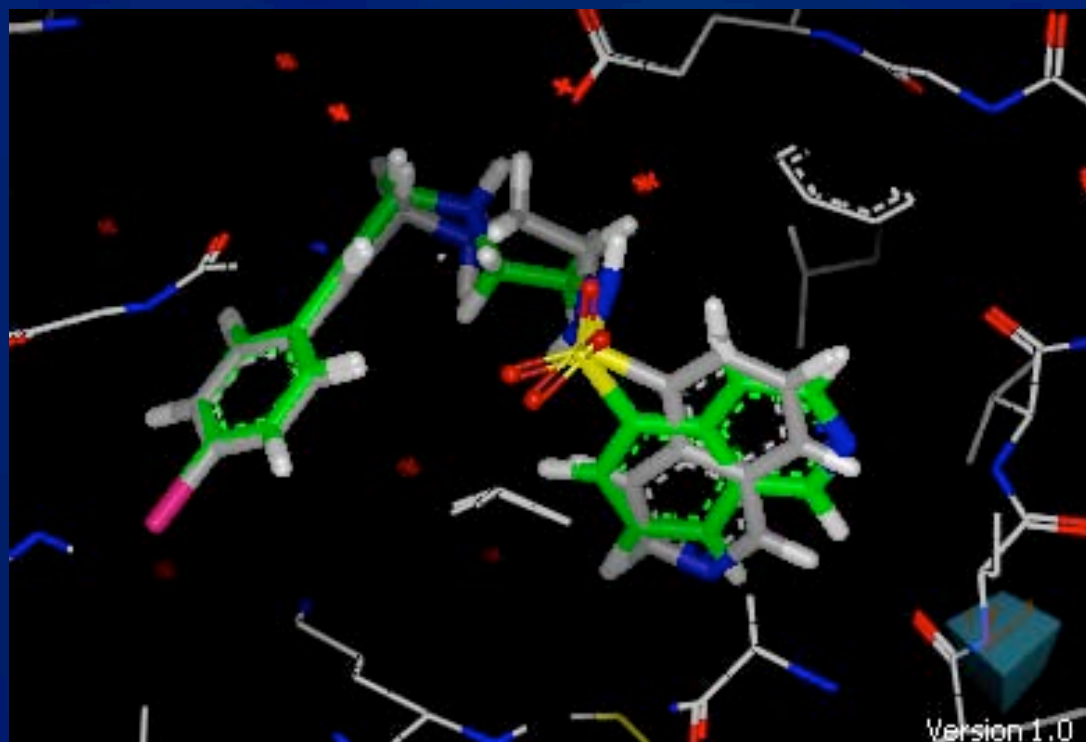
- Refined to global=11.3
- Afitt to global=8.0
 - But it's probably wrong



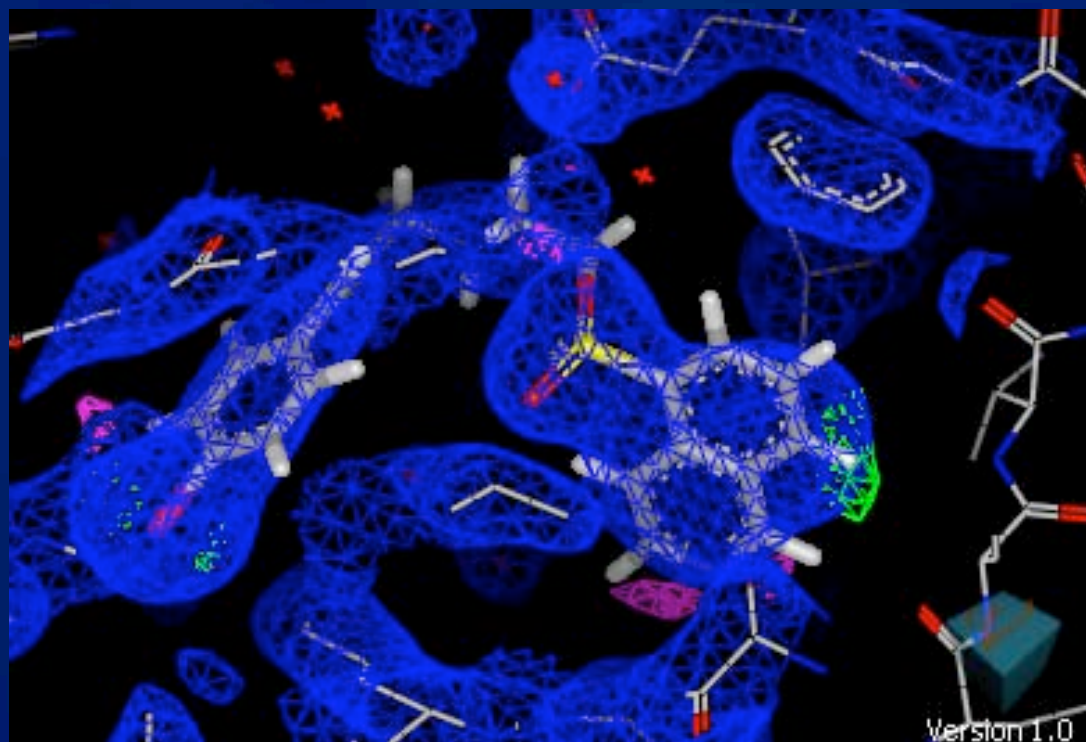
- Refined to global=11.3
- Afitt to global=8.0
 - But it's probably wrong



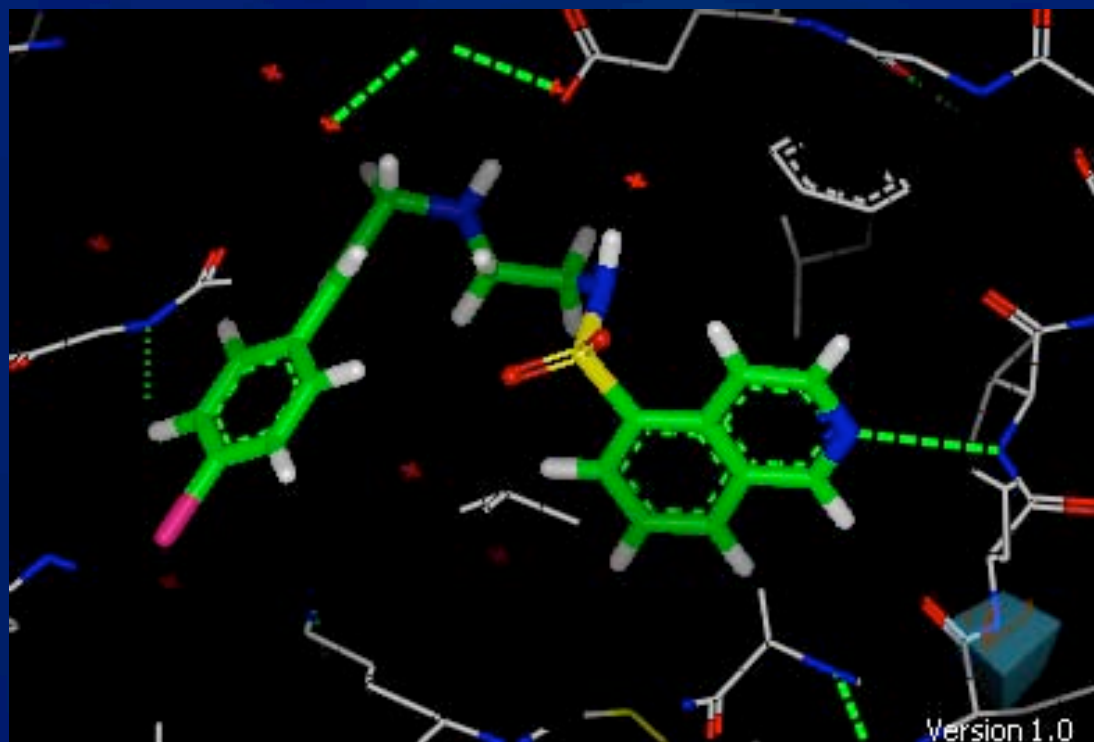
- Refined to global=11.3
- Afitt to global=8.0
 - But it's probably wrong



- Refined to global=11.3
- Afitt to global=8.0
 - But it's probably wrong



- Refined to global=11.3
- Afitt to global=8.0
 - But it's probably wrong



Conclusion

- Low strain conformations can be found
 - 20% reduction in global strain for Afitt over deposited structures
- Sometimes very different conformations fit the data with equivalent strain
- Caveat – small data set
- Implications
 - Smaller energy windows for conformer searching.
 - Maybe there is no contradiction between binding affinity and ligand strain.



Arthur Schopenhauer

- In the sphere of thought, absurdity and perversity remain the masters of the world, and their dominion is suspended only for brief periods.



Acknowledgement

- Roger Sayle
- Joe Corkery and Kevin Schmidt
- Stan Wlodek
- Brian Kelley
- Brian Cole (Bruce)
- Anthony Nicholls
- Paul Labute
- Paul Hawkins and Geoff Skillman



PDB codes

- 13GS.pdb, 1A4K.pdb, 1A8T.pdb, 1AFQ.pdb, 1ATL.pdb, 1AZM.pdb, 1BR6.pdb, 1CET.pdb, 1DIB.pdb, 1EVE.pdb, 1EXA.pdb, 1EZQ.pdb, 1F0R.pdb, 1F0T.pdb, 1FCX.pdb, 1FCZ.pdb, 1FJS.pdb, 1FM6.pdb, 1FM9.pdb, 1GWX.pdb, 1H1P.pdb, 1H1S.pdb, 1I7Z.pdb, 1IY7.pdb, 1K1J.pdb, 1K7E.pdb, 1K7F.pdb, 1L2S.pdb, 1LQD.pdb, 1M48.pdb, 1MQ5.pdb, 1MQ6.pdb, 1NHU.pdb, 1NHV.pdb, 1QHI.pdb, 1R09.pdb, 1THL.pdb, 1YDR.pdb, 1YDS.pdb, 1YDT.pdb, 2PCP.pdb

