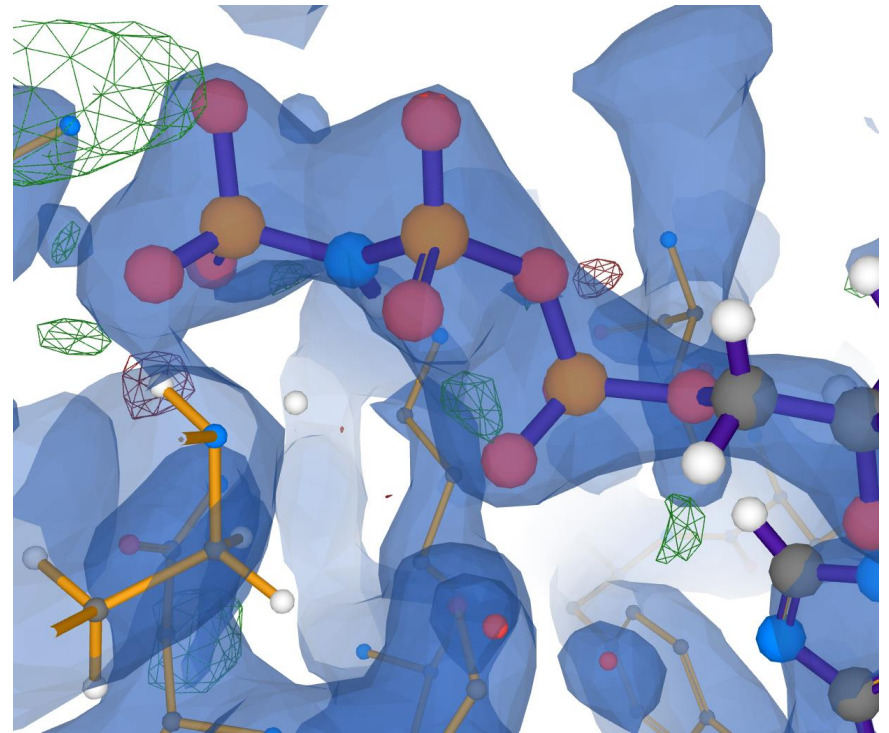


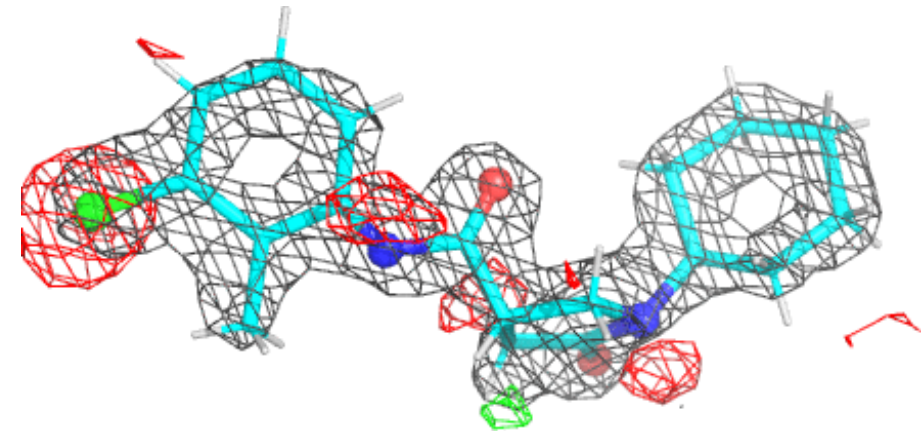
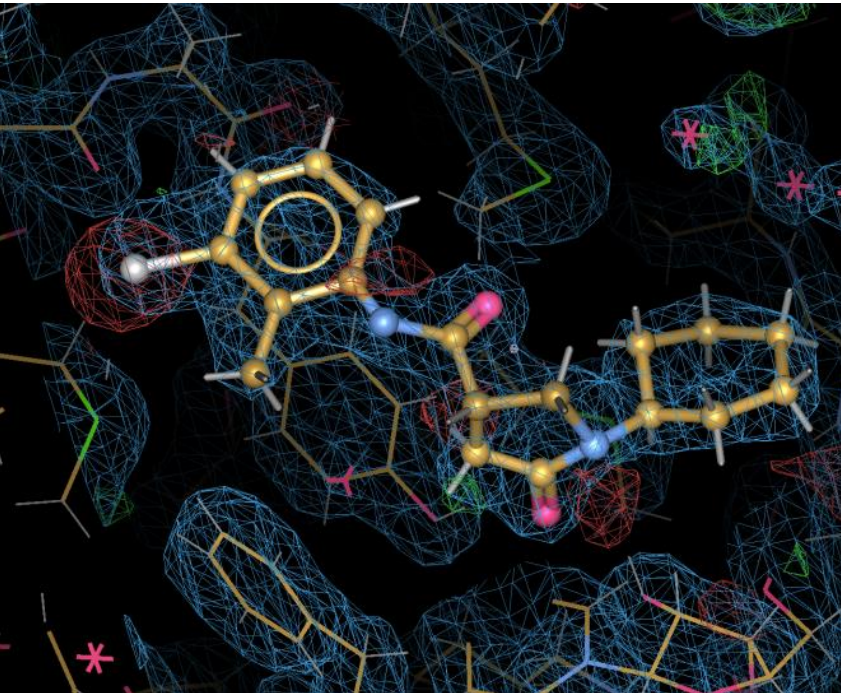
Electron density maps

- When looking at PDB structures
Electrons Density (ED) maps
are more/as important as the 3D
atomic model!
- ED is a 3D map of where the
scattering electron cloud is
according to the measured X-ray
data.
- 2Fo-Fc map indicates where
electrons are (according to SF
and model). Normally colour
blue or grey.
- Fo-Fc difference map:
 - green for positive difference: where
the current model fails to place
sufficient electrons
 - Red for negative difference: where
the current model places too many
electrons



www.ebi.ac.uk/pdbe/entry/pdb/4z9l/bound/ANP

Electron density for a ligand with poor fit



2h7p.pdb: ED around ligand, as visualized in buster-report

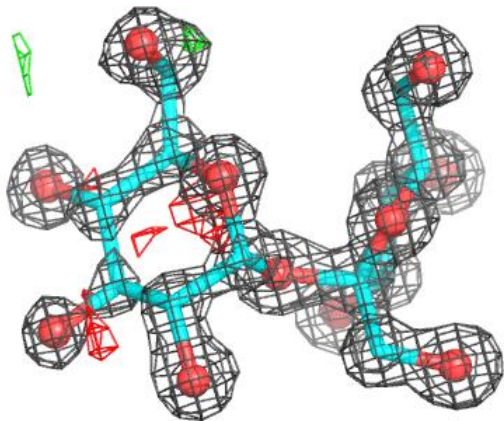
2h7p.pdb: ED visualized in coot

Notice difference density around ligand

2h7p has been obsoleted and replaced 4tzt with really superb ED and ligand fit

Data resolution affects electron density detail:

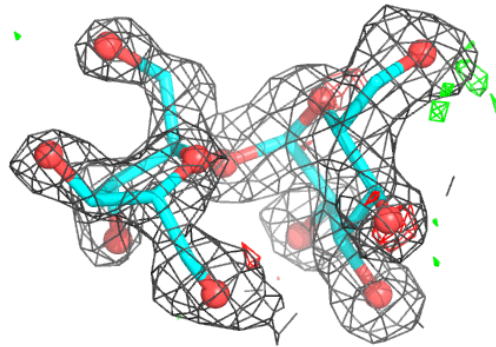
Well placed/refined sucrose ligand at different data resolutions:



1ylt 1.2Å resolution

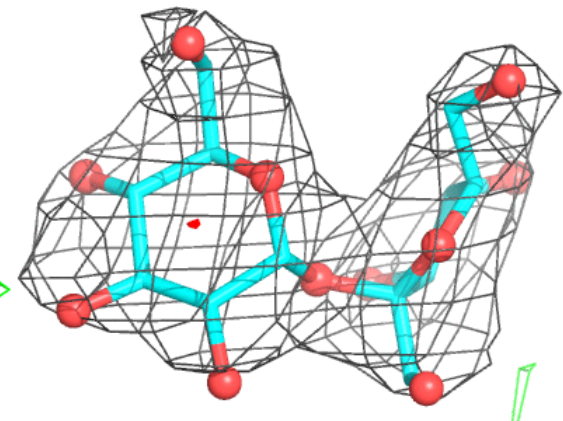
“atomic resolution”

2Fo-Fc ED shows atoms as individual blobs. Need higher resolution for hydrogen atoms



2pwe 2.0Å resolution

Typical medium resolution for ligand studies. Can see ring pucker

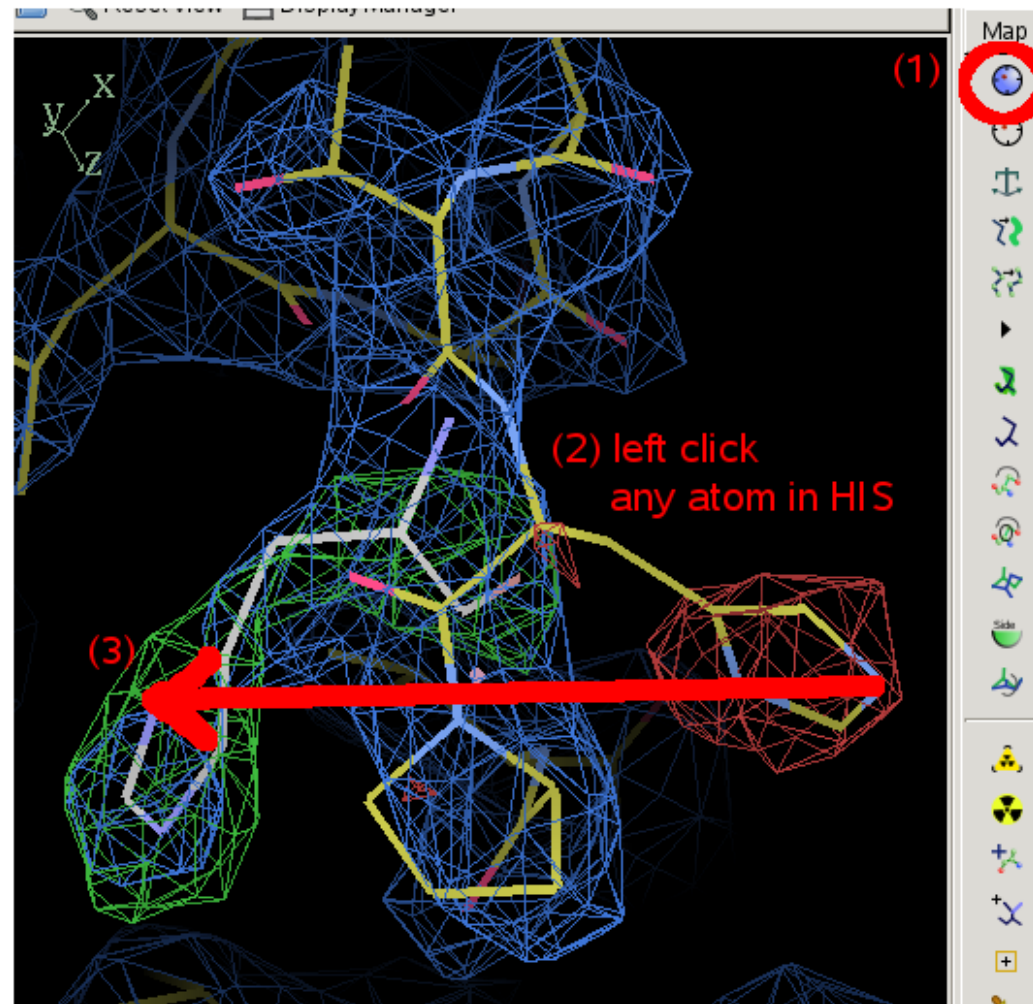


2qqv 3.0Å resolution

Low resolution. Ligand placement unambiguous but fine detail cannot be seen

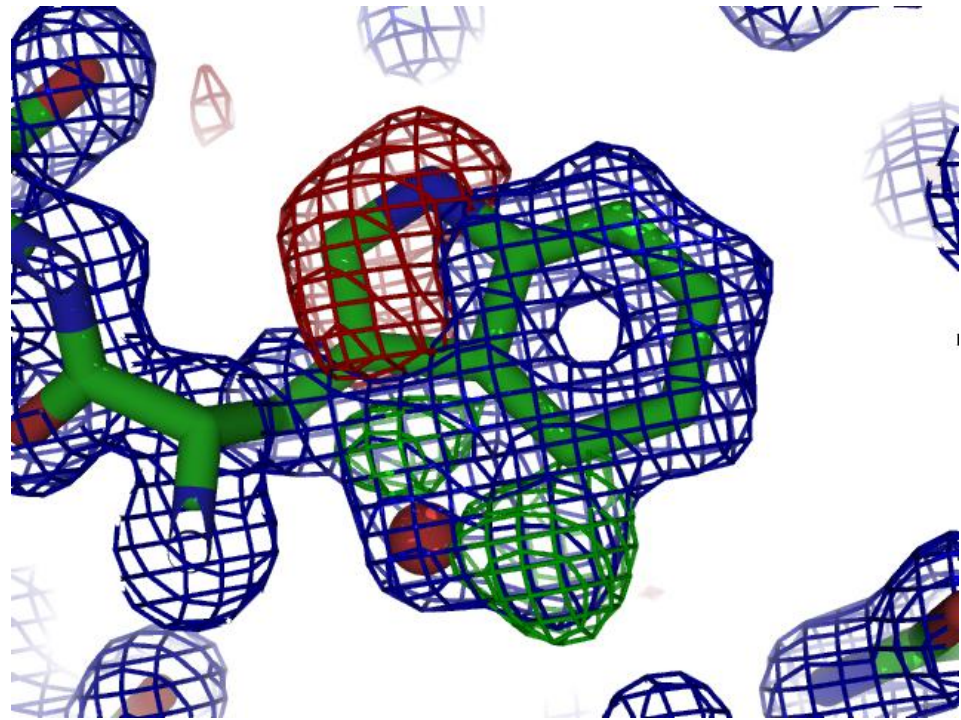
Model improvement

- Basically interpret electron density maps in real space to improve model
- Initially automated methods (warp/arpwarp) used
- But mostly manual corrections to the model are done using the Coot program
- Look at Fo-Fc difference map for both negative and positive features
- Build into 2Fo-Fc

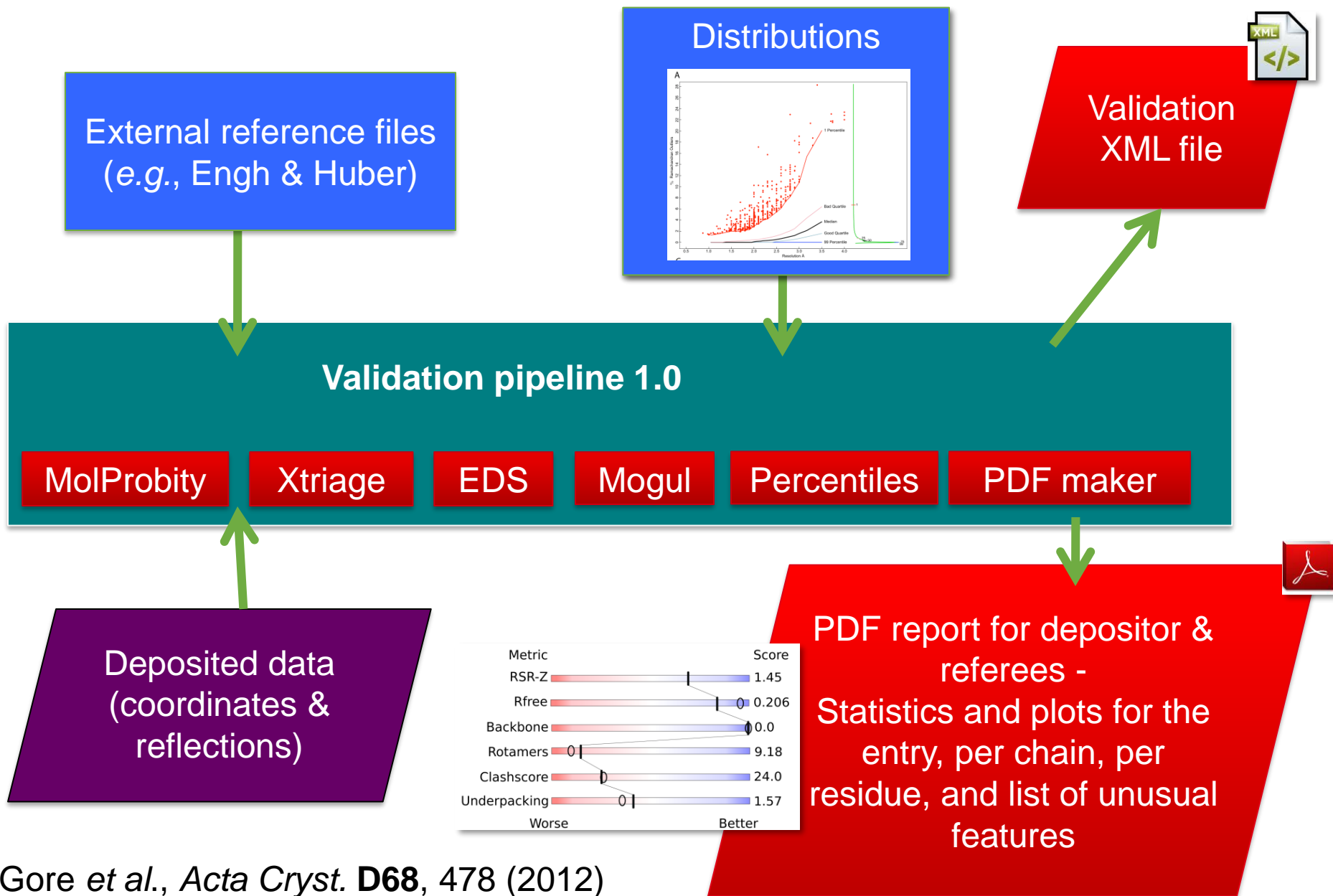


Why do crystallographers make mistakes?

- Limitations to the data
 - Incomplete
 - Weak
 - Limited resolution
 - Space and time averaged
 - Phase errors
- The human factor
 - Subjectivity and bias involved in map interpretation and refinement (even at atomic resolution!)
 - Inexperienced people do the work, use of black boxes, ...
 - Not everybody is a good chemist
 - Even experienced people make mistakes



wwPDB X-ray validation pipeline



wwPDB validation reports



Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2016 – 06:45 PM GMT

PDB ID : 1CBS
Title : CRYSTAL STRUCTURE OF CELLULAR RETINOIC-ACID-BINDING PROTEINS I AND II IN COMPLEX WITH ALL-TRANS-RETINOIC ACID AND A SYNTHETIC RETINOID
Authors : Kleywegt, G.J.; Bergfors, T.; Jones, T.A.
Deposited on : 1994-09-28
Resolution : 1.80 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry. We welcome your comments at validation@mail.wwpdb.org. A user guide is available at <http://wwpdb.org/validation/2016/XrayValidationReportHelp> with specific help available everywhere you see the ⓘ symbol.

Page 2 Full wwPDB X-ray Structure Validation Report 1CBS

1 Overall quality at a glance ⓘ

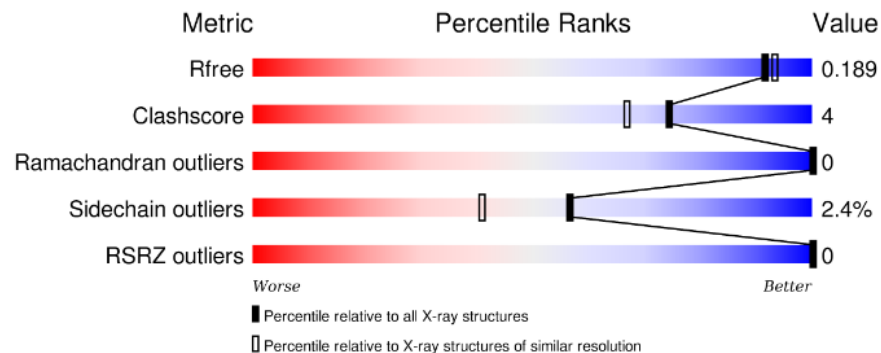
The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



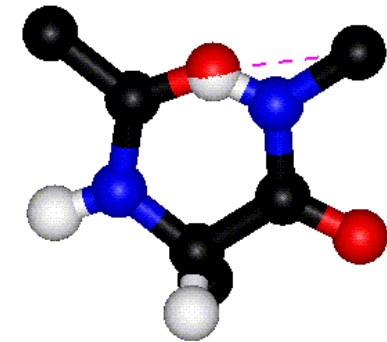
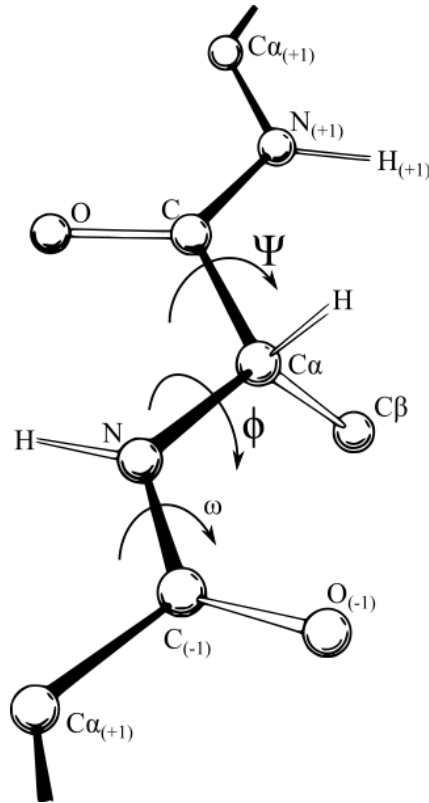
- Pipeline produces PDF report and XML output
- Slider graphic useful



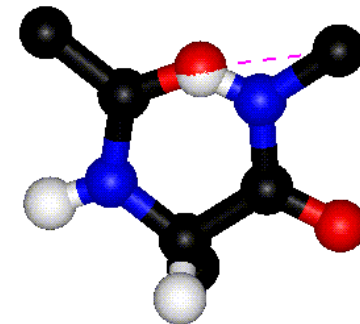
- Current PDF is “rather verbose”

Ramachandran plot

- Look at main chain dihedral angles phi and psi
- Ramachandran et al. (1963) worked out only certain combinations of phi/psi cause clashes

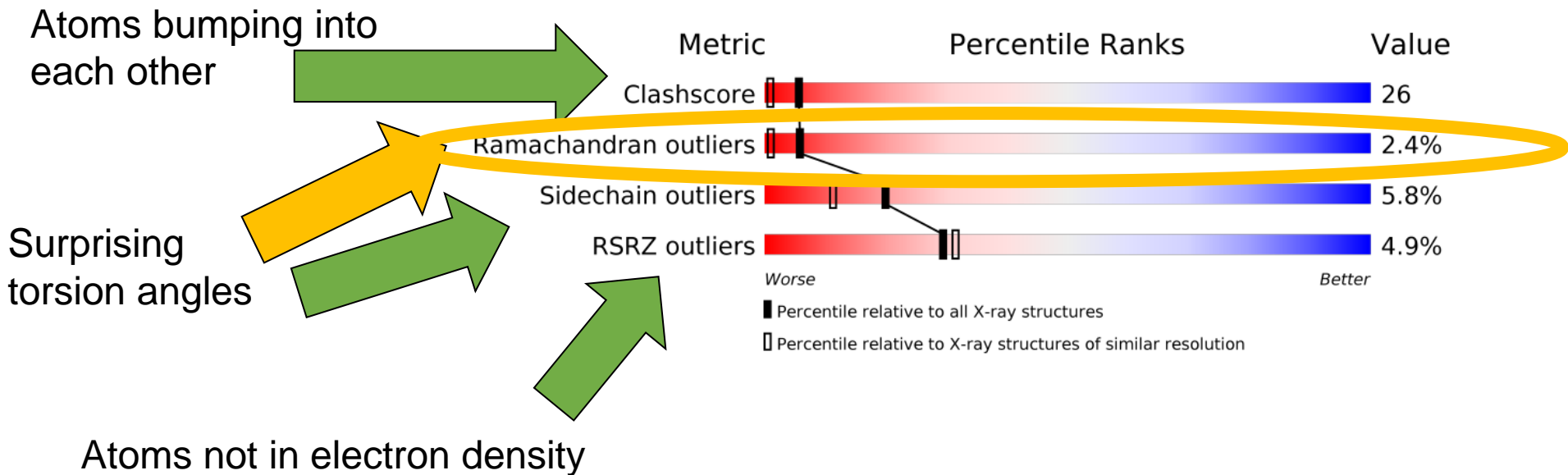


Rotation around ϕ with $\psi=0^\circ$



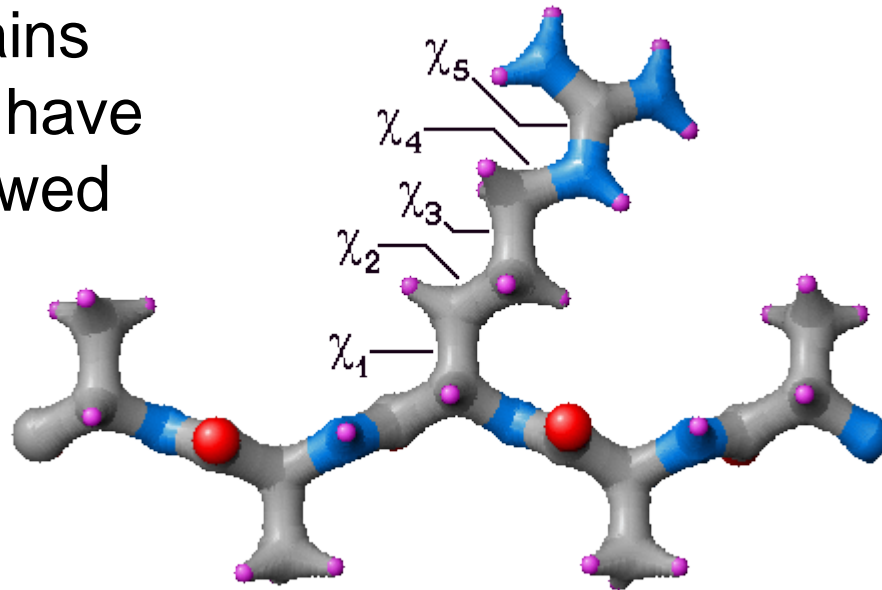
Rotation around ψ with $\phi=0^\circ$
(Images kindly provided by David Sanders, University of Saskatchewan.)

Summary 'Sliders' Validation information for users



Sidechain outliers

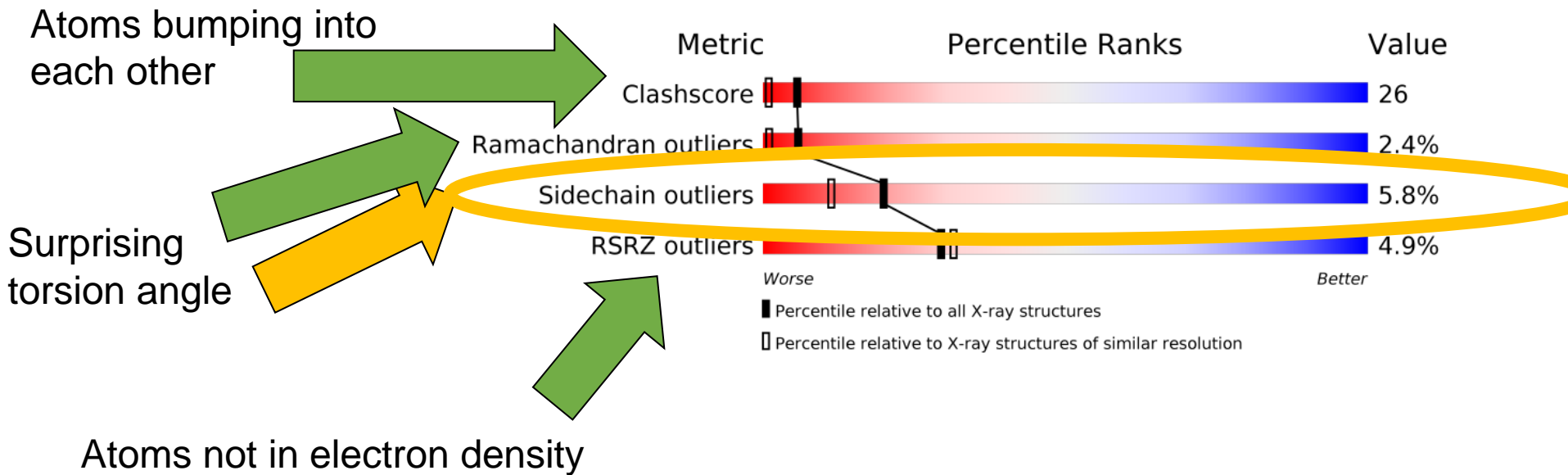
- Just like the main chain phi and psi dihedral angles amino acid side chains have chi angles with preferred and disallowed regions



The 5 chi angles of an arginine side chain

<http://www.ccp14.ac.uk/ccp/web-mirrors/garlic/garlic-1.5/commands/dihedrals.html>

Summary 'Sliders' Validation information for users



MolProbity – clash score

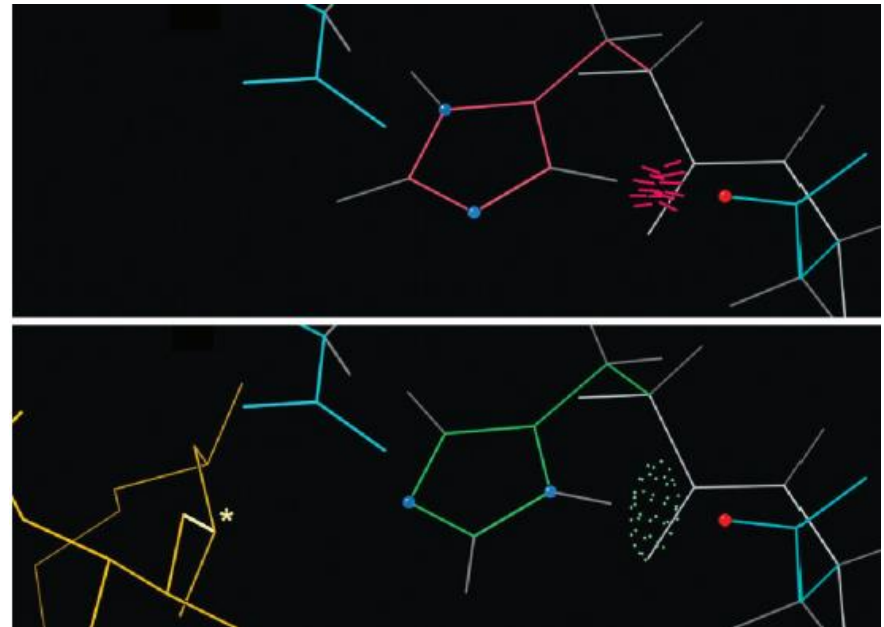
Nucleic Acids Research, 2007, Vol. 35, Web Server issue W375–W383
doi:10.1093/nar/gkm216

- Idea is to look for bad non-bonded contacts after hydrogen atoms have been added to the model
- Very powerful method
- Suggests NQH flips
- Included in wwPDB validation reports
- Or Use from:
 - Molprobity web site
 - Or within coot

MolProbity: all-atom contacts and structure validation for proteins and nucleic acids

Ian W. Davis¹, Andrew Leaver-Fay², Vincent B. Chen¹, Jeremy N. Block¹, Gary J. Kapral¹, Xueyi Wang², Laura W. Murray¹, W. Bryan Arendall III¹, Jack Snoeyink², Jane S. Richardson¹ and David C. Richardson^{1,*}

¹Department of Biochemistry, Duke University, Durham, NC, USA and ²Department of Computer Science, UNC Chapel Hill, Chapel Hill, NC, USA



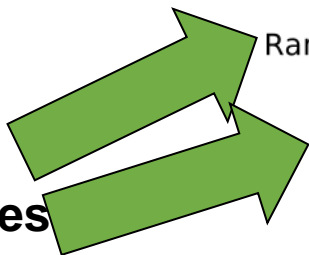
Validation information for users

Summary 'Sliders'

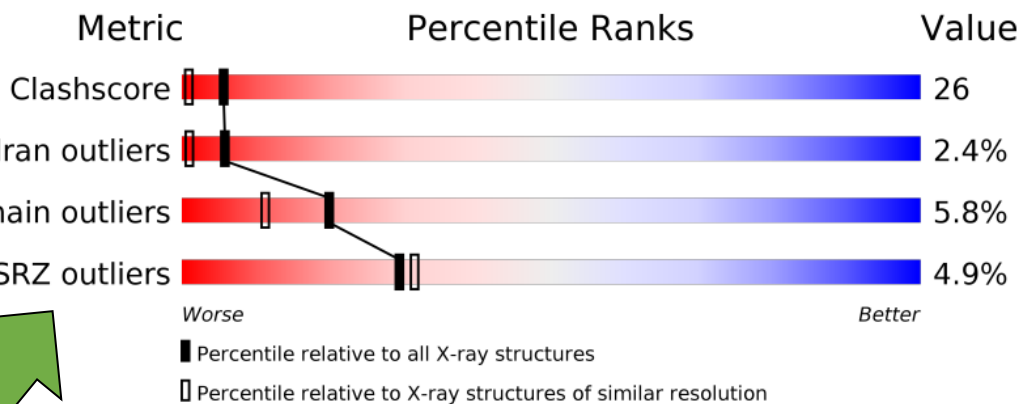
Atoms bumping into each other



Surprising torsion angles



Atoms not in electron density

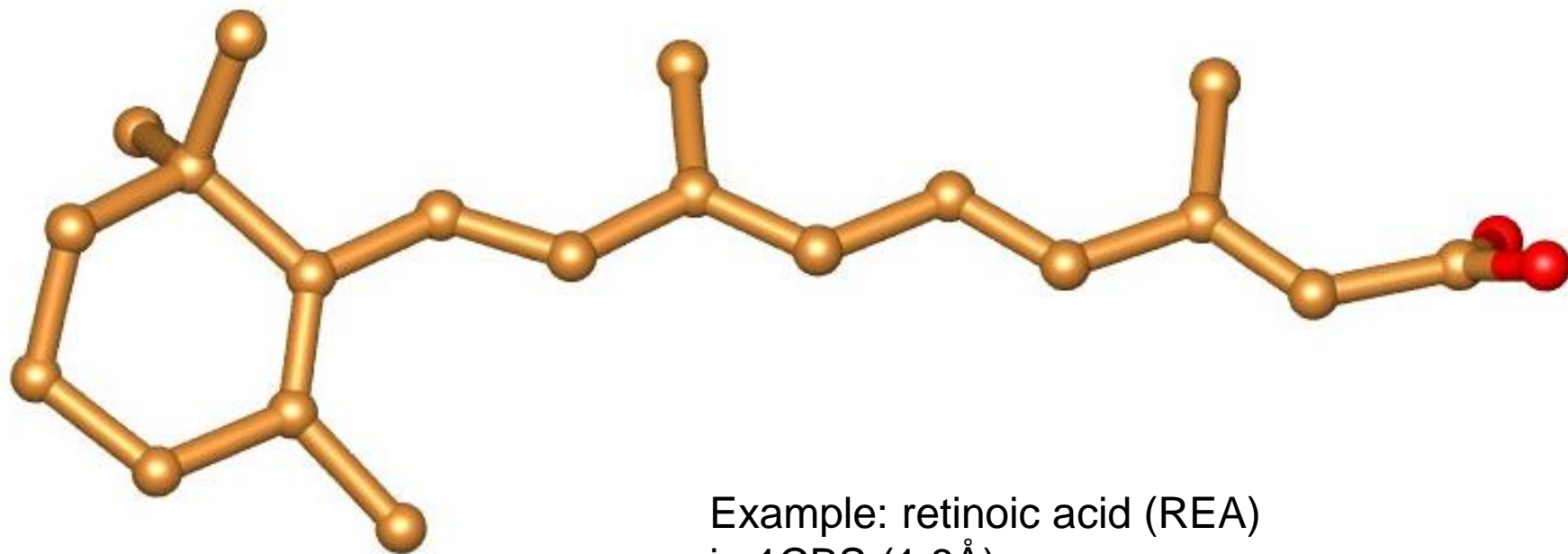


Real-space fit

- Quantitative, real-space measure of how well a residue fits its local density (Jones *et al.*, 1991)
- Express as R-value (RSR) or correlation coefficient (RSCC)
- $$\text{RSR} = \frac{\sum |\rho_{\text{obs}} - \rho_{\text{calc}}|}{\sum |\rho_{\text{obs}} + \rho_{\text{calc}}|}$$
- Sums extend over all grid points inside a mask around the residue

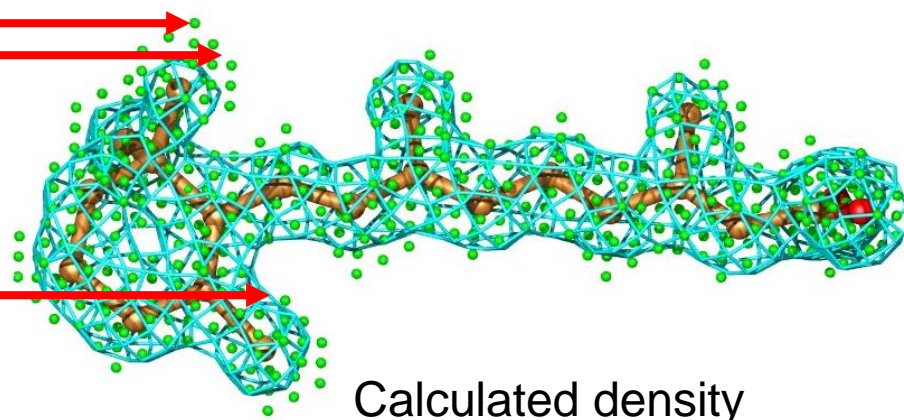
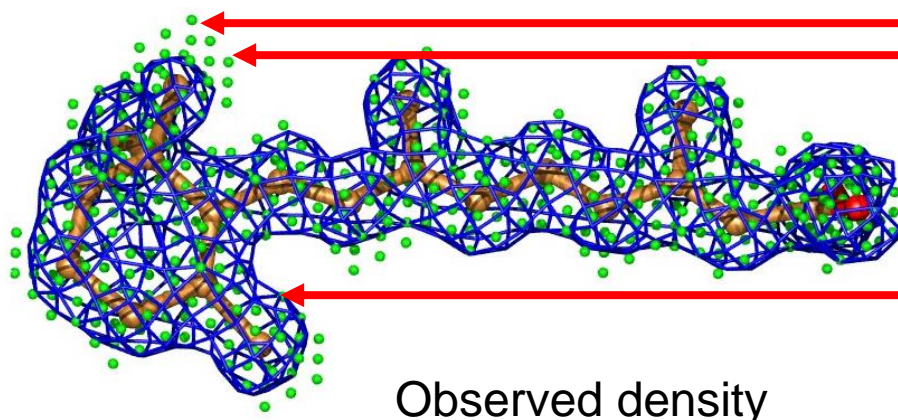
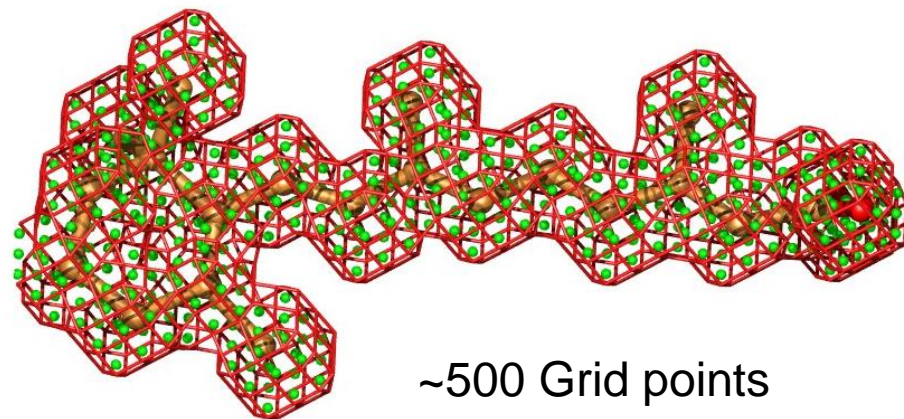
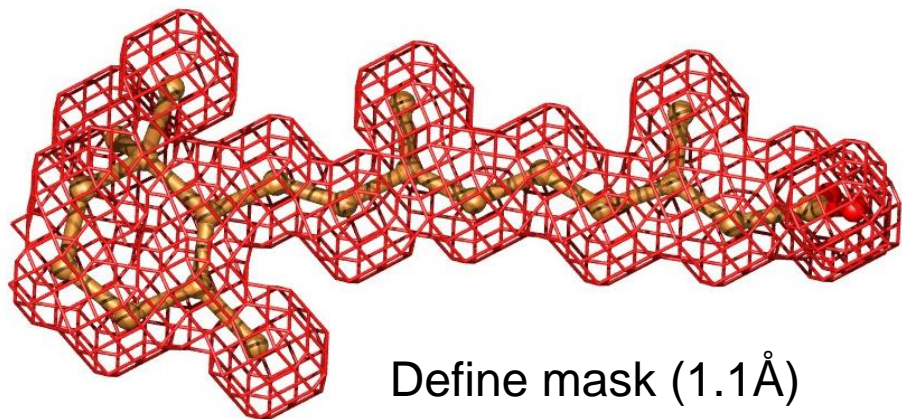
RSR - real-space R-value

- $$\text{RSR} = \frac{\sum |\rho_{\text{obs}} - \rho_{\text{calc}}|}{\sum |\rho_{\text{obs}} + \rho_{\text{calc}}|}$$



Example: retinoic acid (REA)
in 1CBS (1.8Å)

RSR - real-space R-value



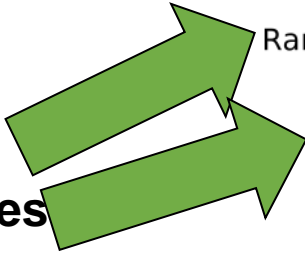
$$RSR = \frac{\sum |\rho_{obs} - \rho_{calc}|}{\sum |\rho_{obs} + \rho_{calc}|}$$

RSRZ is reported in Summary 'Sliders'

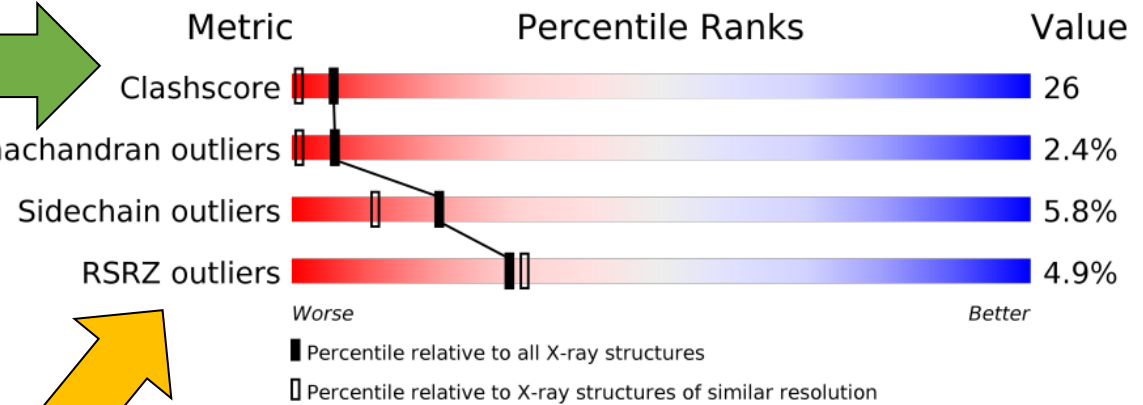
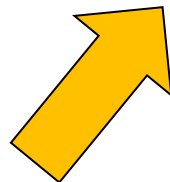
Atoms bumping into each other



Surprising torsion angles



Atoms not in electron density



PDBe simplification of validation sliders

(2) < 1 > Entry 1 to 2 of 2 Quality (desc) ▾

[1det](#) RIBONUCLEASE T1 CARBOXYMETHYLATED AT GLU 58 IN COMPLEX WITH 2'GMP

(2) Ishikawa K, Suzuki E, Tanokura M, Takahashi K
Biochemistry (1996) [PMID: 8679590]

Source organism: [Aspergillus oryzae](#)

Assembly composition: protein only structure

Interacting compounds: [2GP](#) [NA](#)

(1) Add to basket Download files

(1)

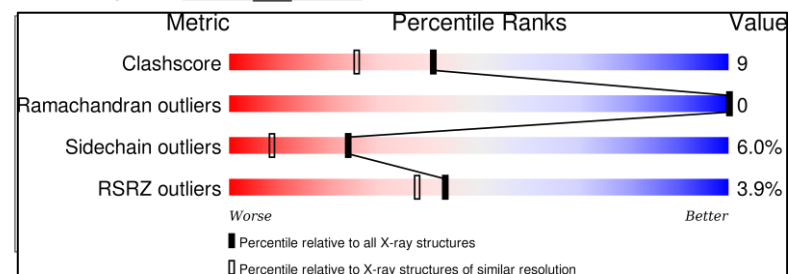
X-ray diffraction

1.8Å resolution

Released: 11 Jul 1996

Model geom...

Fit model/da...



[3syu](#) Re-refined coordinates for pdb entry 1det - ribonuclease T1 carboxymethylated at GLU 58 in complex with 2'GMP

(2) Smart OS, Womack TO, Bricogne G
Acta Crystallogr. D Biol. Crystallogr. (2012) [PMID: 22505257]

Source organism: [Aspergillus oryzae RIB40](#)

Assembly composition: protein only structure

Interacting compounds: [NA](#) [2GP](#)

(1) Add to basket Download files

(1)

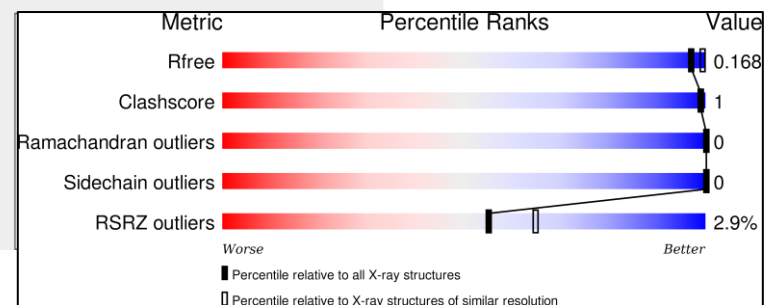
X-ray diffraction

1.95Å resolution

Released: 28 Mar 2012

Model geom...

Fit model/da...



“Best molecule” – integration of validation information in PDBe query system

Search results

Refine query:

[Reset](#) ["Protein-serine/threonine kinases"](#)

[View basket \(0\)](#)

[resolution:\[1.5 TO 2\]](#)

[Save search](#) [Download](#) Per page: 10

[Entries](#) [Macromolecules](#) [Compounds](#) [Protein families](#)

(200) [1](#) [2](#) [3](#) ... [19](#) > Macromolecule 1 to 10 of 183

Protein: [3-phosphoinositide-dependent protein kinase 1](#)

Best example found in:

[4rqk](#) Crystal structure of PDK1 in complex with ATP and the PIF-pocket ligand RS1

Rettenmaier TJ, Wells JA
Proc. Natl. Acad. Sci. U.S.A. (2014) [PMID: [25518860](#)]

Source organism: *Homo sapiens*

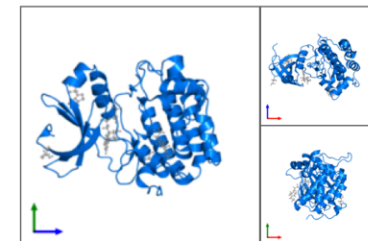
Assembly composition: protein only structure

[Add to basket](#) [Download files](#)

X-ray diffraction
1.55Å resolution

Released: 17 Dec 2014

Model geometry 
Fit model/data 



Macromolecules (100+)

- Cyclin-dependent kinase 2 (200)
- Mitogen-activated protein kinase 14 (82)
- cAMP-dependent protein kinase catalytic subunit alpha (80)
- Casein kinase II subunit alpha (52)
- Serine/threonine-protein kinase Chk1 (44)
- Mitogen-activated protein kinase 1 (37)
- Death-associated protein kinase 1 (33)
- Serine/threonine-protein kinase pim-1 (28)
- 3-phosphoinositide-dependent protein kinase 1 (24)
- Serine/threonine-protein kinase PDK1 (22)

Molecule type (1)

- Protein (1031)

Interacting macromolecules (98)

Interacting compounds (100+)

Species name (38)

- Zea mays (27)
- Rattus norvegicus (24)
- Arabidopsis thaliana (16)
- Saccharomyces cerevisiae (10)
- Synechococcus elongatus PCC 7942 (8)
- Toxoplasma gondii (8)
- Xenopus laevis (8)

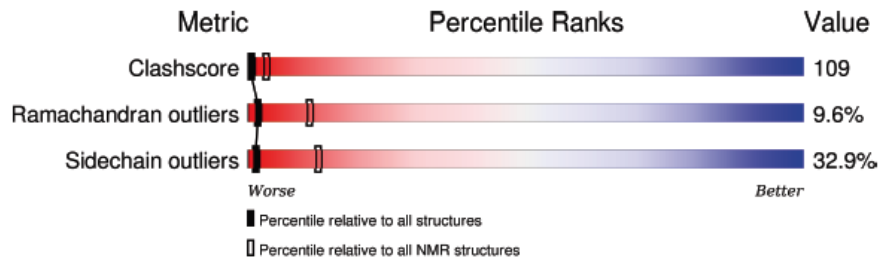
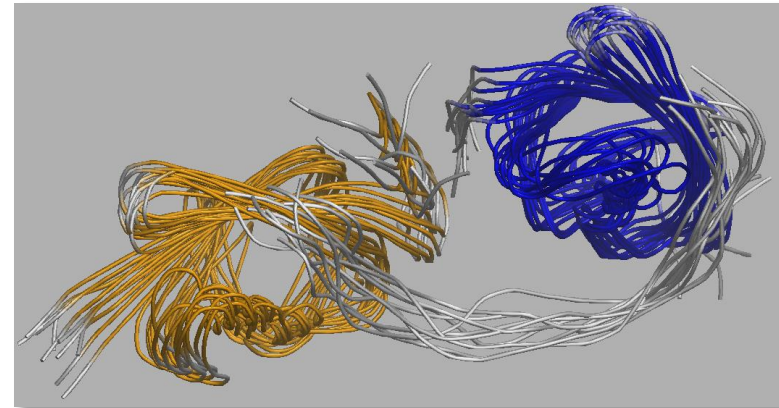
Other entries (23)

[Add to basket](#)

Entry ID	Model geometry	Fit model/data	Quality Bar	Bookmarks	Human	X-RAY	Upload	Hexagon
3hrf	1.9Å	Model geometry						
4rqv	1.502Å	Model geometry						
4ct1	1.85Å	Model geometry						
4a07	1.85Å	Model geometry						
3rwp	1.92Å	Model geometry						

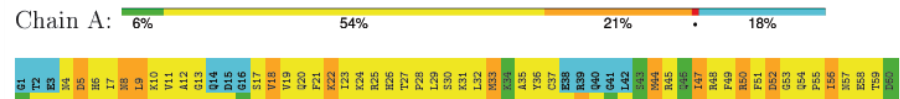
NMR validation

- NMR VTF recommendations published
- Global quality scores reported for “well-defined residues” only
 - As averages over the ensemble
 - Medoid model only

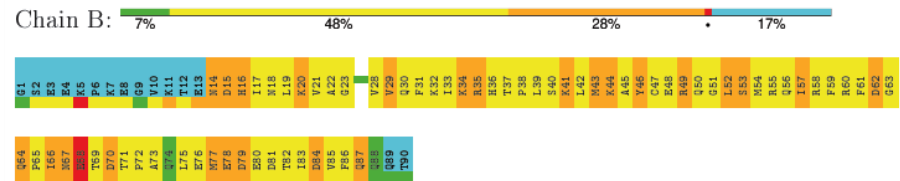


Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	114402	11133
Ramachandran outliers	111179	9975
Sidechain outliers	111093	9958

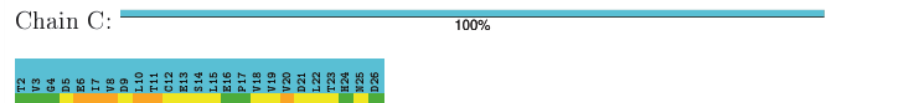
- Molecule 1: Small ubiquitin-related modifier 3



- Molecule 2: Small ubiquitin-related modifier 3



- Molecule 3: E3 ubiquitin-protein ligase RNF4



EM validation reports



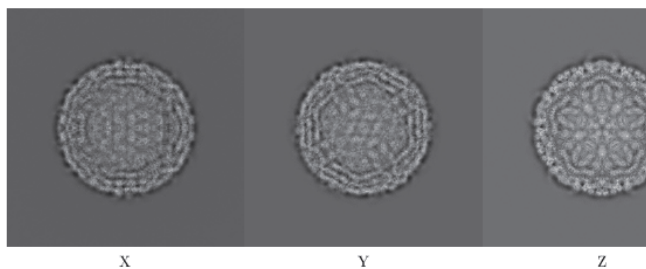
- Prototype EM map-validation reports
 - Most of the PDBe “Visual analysis” functionality implemented

4 Map analysis ⓘ

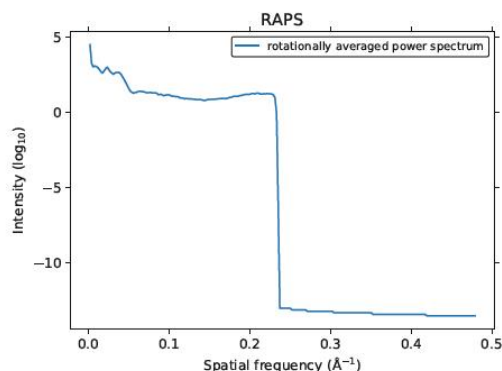
4.1 Map parameters ⓘ

Property	Value
Endianness	little-endian
Pixel size	1.04
Axis order	XYZ
Number of pixels in X	768
Number of pixels in Y	768
Number of pixels in Z	768
Minimum density	-19.308
Maximum density	26.264
Average density	0.0267
Standard deviation of densities	1.129
Range of densities	45.572
Recommended contour	3.5

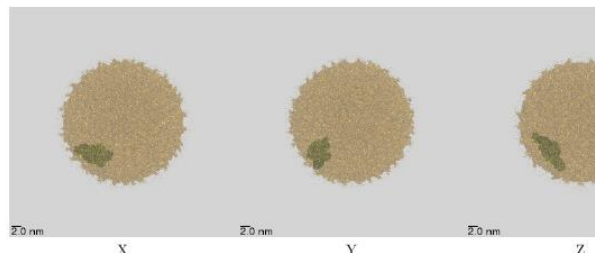
4.2 Orthogonal projections ⓘ



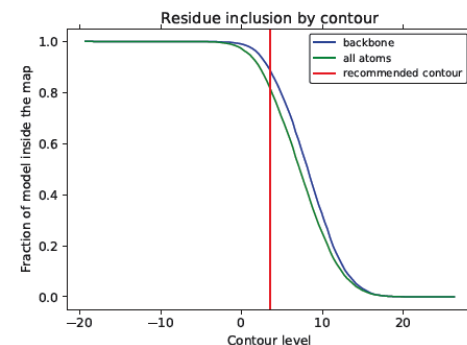
4.7 Rotationally averaged power spectrum ⓘ



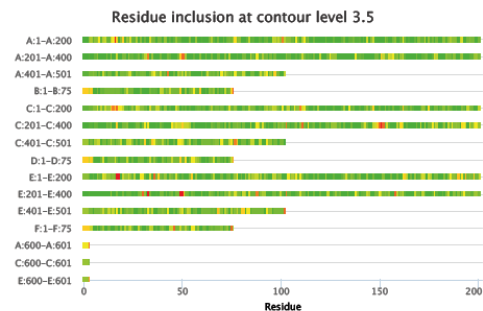
4.8 Model to map fitting ⓘ



4.9 Residue inclusion by contour ⓘ



4.10 Residue inclusion at recommended contour ⓘ



Ligands in proteins

- So you have successfully navigated all the hazards so far have great data, well integrated, successful MR, refinement model building, Ramachandran analysis
- You have density in the active site and the whole point of the structure is to find how the interesting drug candidate ligand binds
- Here be dragons!

