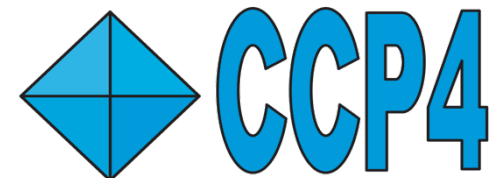


Low Resolution Refinement

DLS/CCP4 2017

Oleg Kovalevskiy

MRC Laboratory of Molecular Biology, Cambridge, UK
okovalev@mrc-lmb.cam.ac.uk



Regularisation

Use of available knowledge (prior information):

High-low resolution:

- Geometry restraints (chemical information)

Medium-low resolution:

- Local NCS restraints
- Jelly body restraints

Low resolution (and medium-low resolution model building):

- Jelly body restraints
- External restraints

Jelly-body restraints

Regularisers without a target:

$$f = \sum_{i: \text{close atom pairs}} \frac{1}{\sigma^2} (|d_i| - |d_{i,\text{current}}|)^2$$

$|d_i|$: interatomic distance

$|d_{i,\text{current}}|$: current interatomic distance

σ : restraint standard deviation

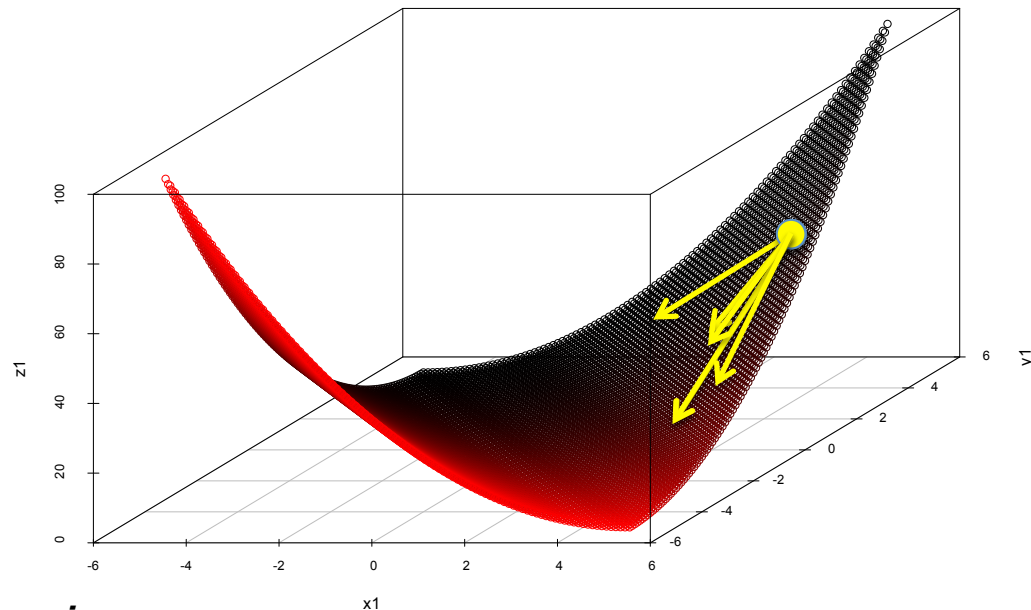
Does not change likelihood function.

Does not change derivative.

Does change 2nd derivative – curvature.

Model should be less prone to fitting into noise

Should only work if parameters are near the minima (model is good)



Typical: $\sigma = 0.01-0.02$

Distance threshold: $4.2\text{\AA}$

Jelly-body restraints

Transferrin



1ryx

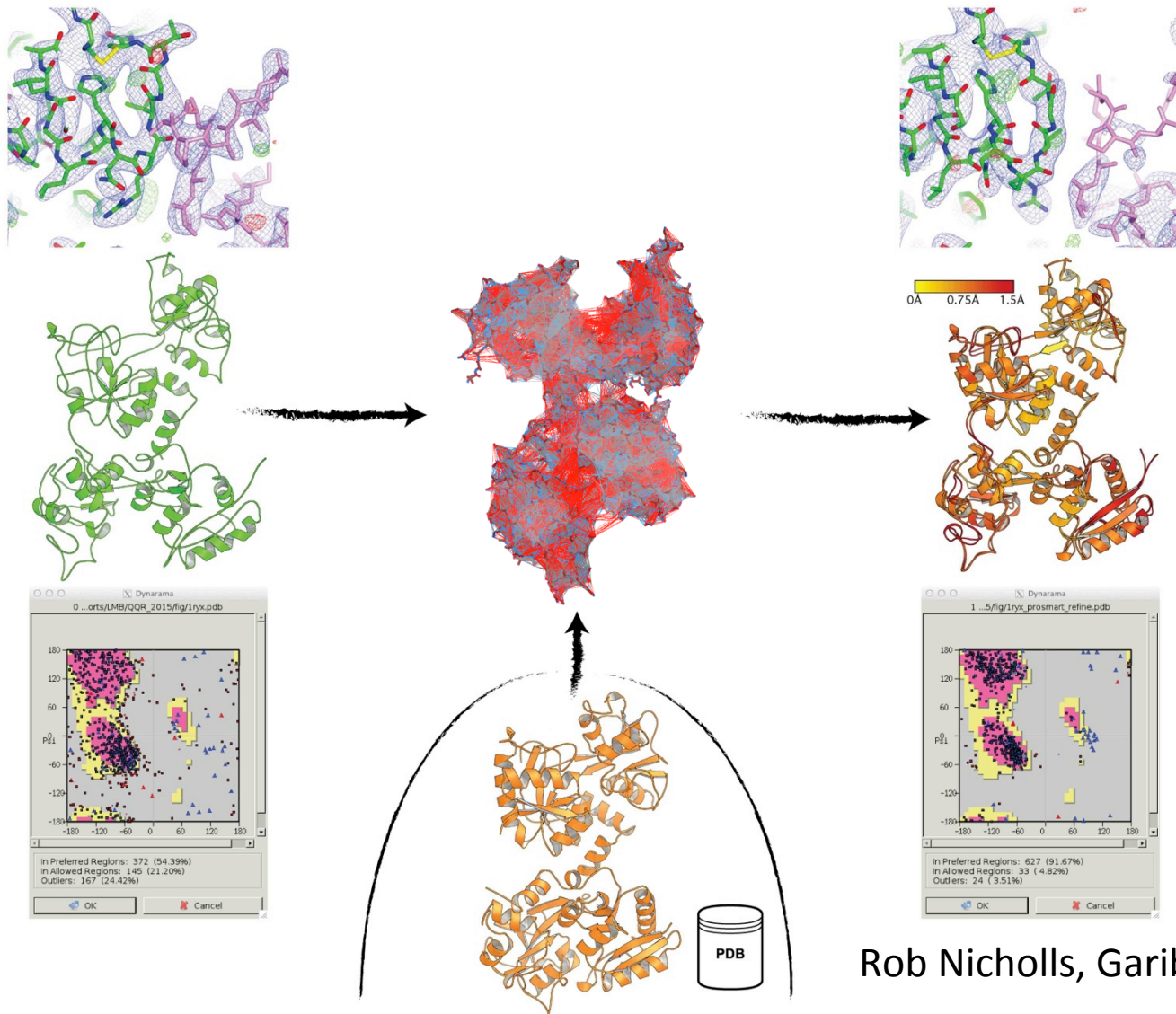
2003, 3.5 Å

Starting: $R_{\text{work}}/R_{\text{free}}$: 28.7/32.8 %

Std. refinement: $R_{\text{work}}/R_{\text{free}}$: 27.1/34.2 %

Jelly-body: $R_{\text{work}}/R_{\text{free}}$: 28.3/32.4 %

ProSMART – external restraints



Rob Nicholls, Garib's group

External restraint generation:

Injection of prior knowledge to aid new structure determination:

- External Restraints from homologous structures
- Hydrogen bond restraints

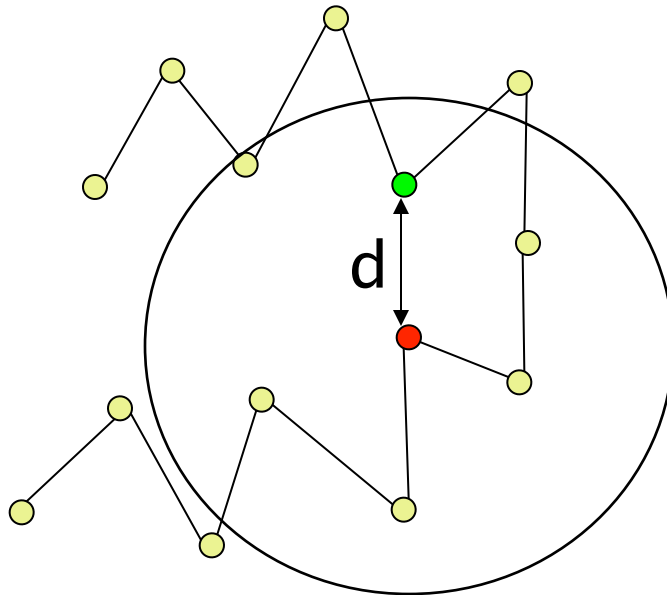
Restraints utilised by: REFMAC5, Coot

Executed from: CCP4i, CCP4mg, Coot, command line

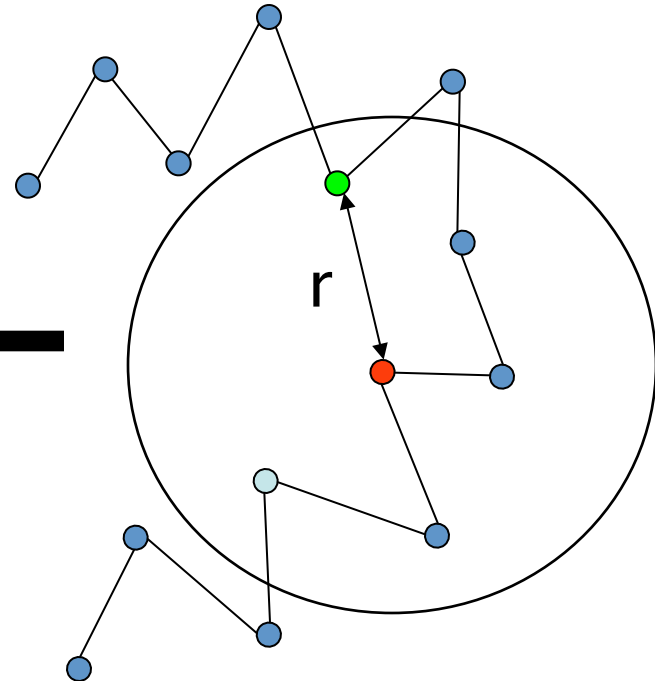
Independent of global conformation

ProSMART – external restraints

structure to be refined



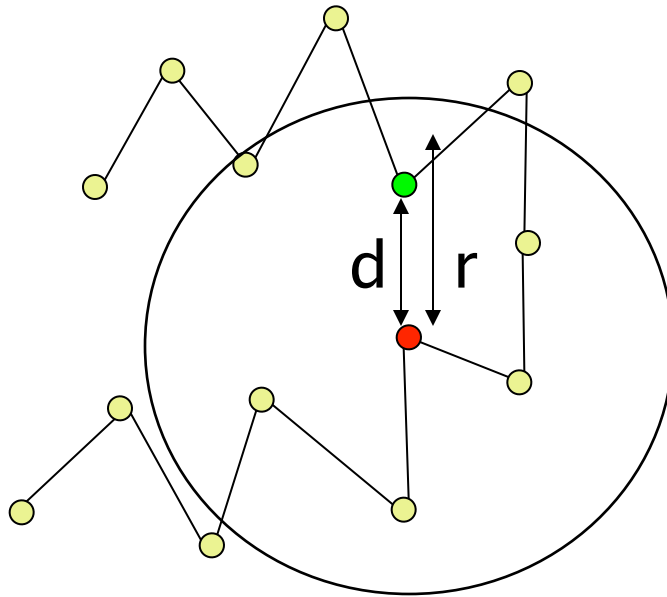
known similar structure (prior)



(abstract representation of an atomic model; circles = atoms)

ProSMART – external restraints

structure to be refined



$$d \sim N(r, \sigma^2)$$

(abstract representation of an atomic model; circles = atoms)

All Restraints Great and Small

ProSMART: restraints from homologous structures

Input Data | Options | Advanced Options

Job title

 Use data from job as input below..

Main inputs

	Atomic model	1 Atomic model imported from 1ryx.pdb by job 1		
	Reflections	1 1ryx: unknown170412 imported by job 1		
	Free R set	1 1ryx: unknown170412 imported by job 1		

Optional additional inputs

	Phases	..is not used		
	TLS coefficients	..is not used		
	Ligand description	..is not used		
	Reference model	3 Reference model imported from 1n04.pdb by job 3		

Rob Nicholls, Garib's group

All Restraints Great and Small

ProSMART: restraints from homologous structures

Run Refmac5

Job title

Do **restrained refinement** using **no prior phase information** input

☐ Input fixed TLS parameters

no twin refinement

Use Prosmart: **no** (low resolution refinement)

☐ Run libg to g no

☐ Run Coot:fin to generate external restraints to reference structure (protein)

MTZ in Tutor to generate external restraints to reference structure (dna/ma)

FP to generate h-bond restraints (e.g. secondary structure restraints)

MTZ out Tutor to generate fragment restraints (e.g. secondary structure element conformations)

PDB in Tutorial input existing restraints file

Browse View

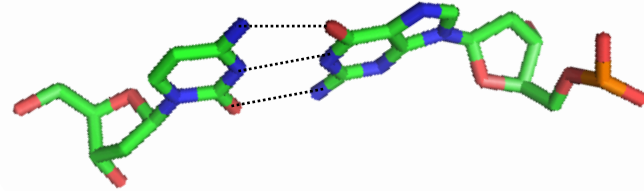
Browse View

Browse View

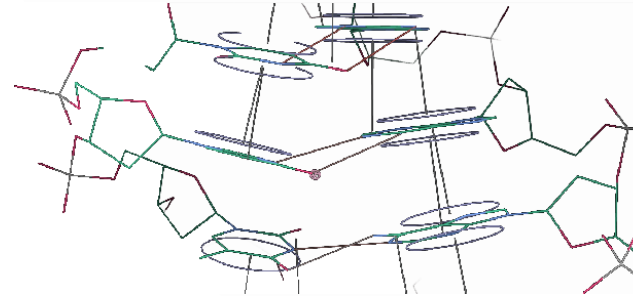
All Restraints Great and Small

LIBG: keeping nucleic acids in place

Base-pair restraints



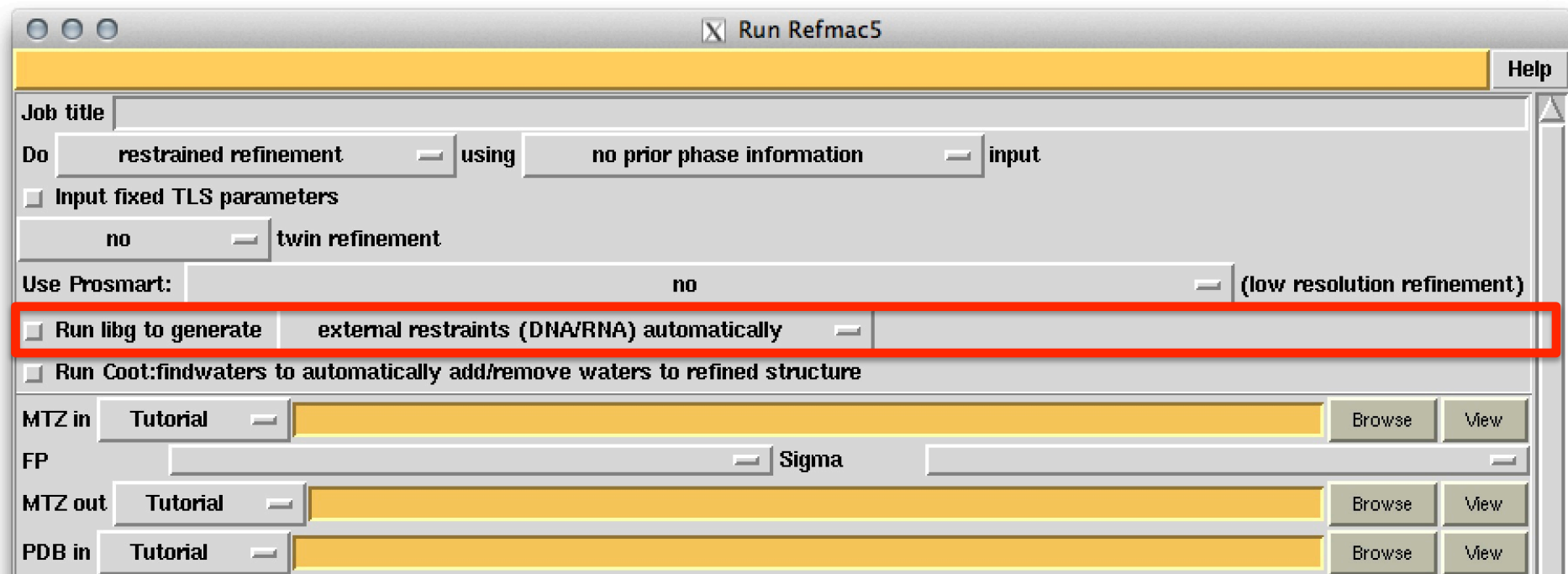
Base-stacking restraints
(parallel plane restraints)



Torsion angles restraints,
Sugar conformation restraints,
Etc.

All Restraints Great and Small

LIBG: keeping nucleic acids in place



The image shows a screenshot of the 'Run Refmac5' dialog box. The window has a title bar with standard Mac OS buttons and a close button. The main area contains several configuration options. A red rectangle highlights the checkbox 'Run libg to generate external restraints (DNA/RNA) automatically', which is currently unchecked. Other visible options include 'Do restrained refinement using no prior phase information input', 'Input fixed TLS parameters' (unchecked), 'Use Prosmart: no (low resolution refinement)', and 'Run Coot:findwaters to automatically add/remove waters to refined structure' (unchecked). At the bottom, there are input fields for 'MTZ in', 'MTZ out', and 'PDB in', each with a 'Tutorial' dropdown and a file path field, accompanied by 'Browse' and 'View' buttons. A 'Sigma' field is also present between the MTZ and PDB inputs.

Run Refmac5

Job title

Do **restrained refinement** using **no prior phase information** input

☐ Input fixed TLS parameters

no twin refinement

Use Prosmart: **no** (low resolution refinement)

☐ **Run libg to generate external restraints (DNA/RNA) automatically**

☐ Run Coot:findwaters to automatically add/remove waters to refined structure

MTZ in **Tutorial** Browse View

FP Sigma

MTZ out **Tutorial** Browse View

PDB in **Tutorial** Browse View

Automation: Magic Green Button

LORESTR: automation of refinement at low resolution

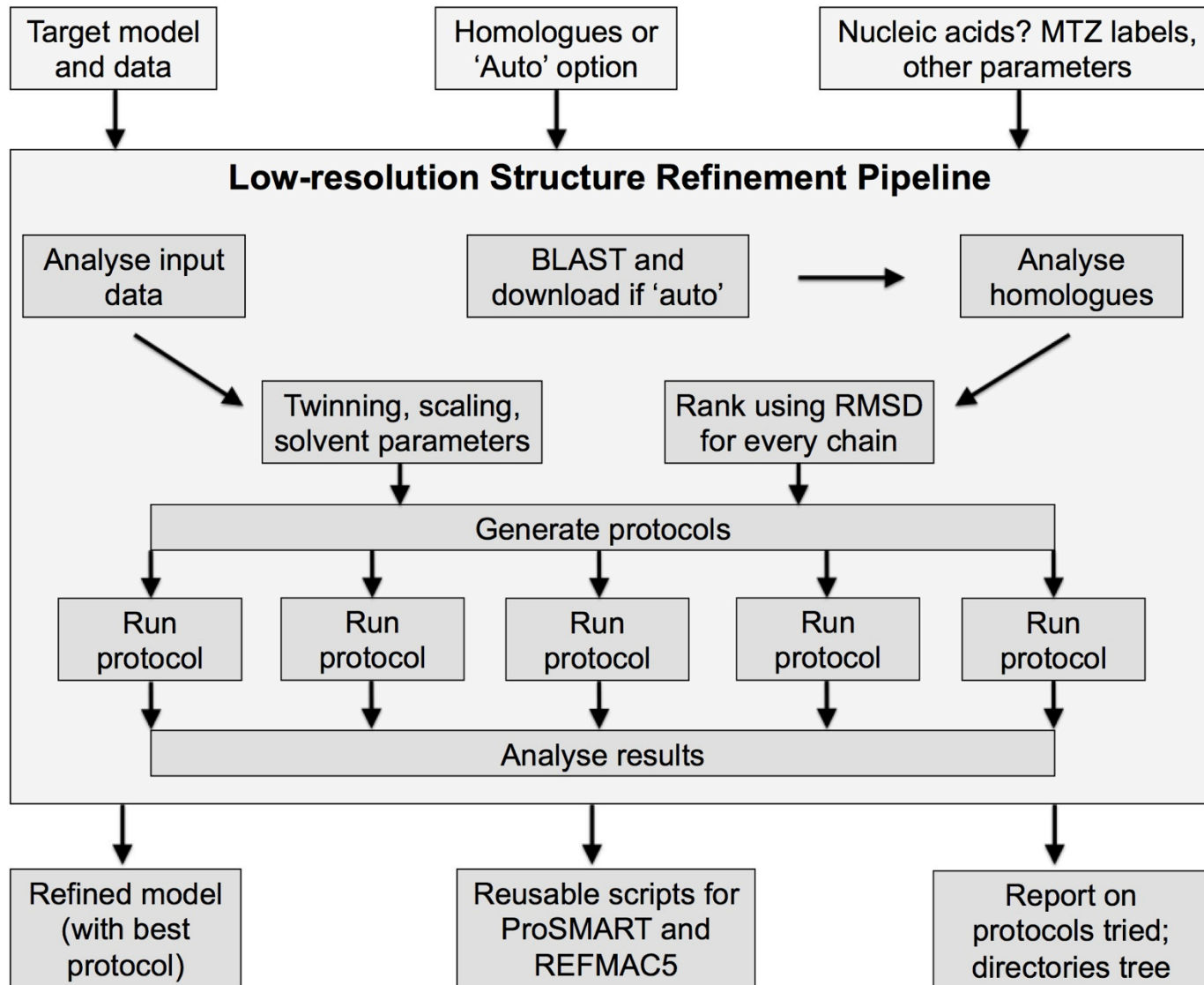
- Efficiency of ProSMART-generated restraints greatly depends on the homologues used
- If several homologues are available, substantial manual effort is required to find their optimal combination
- Other refinement parameters (scaling, solvent, etc) also affect efficiency of the process

Solution:

**LOW-REsolution STructure Refinement
(LORESTR)**

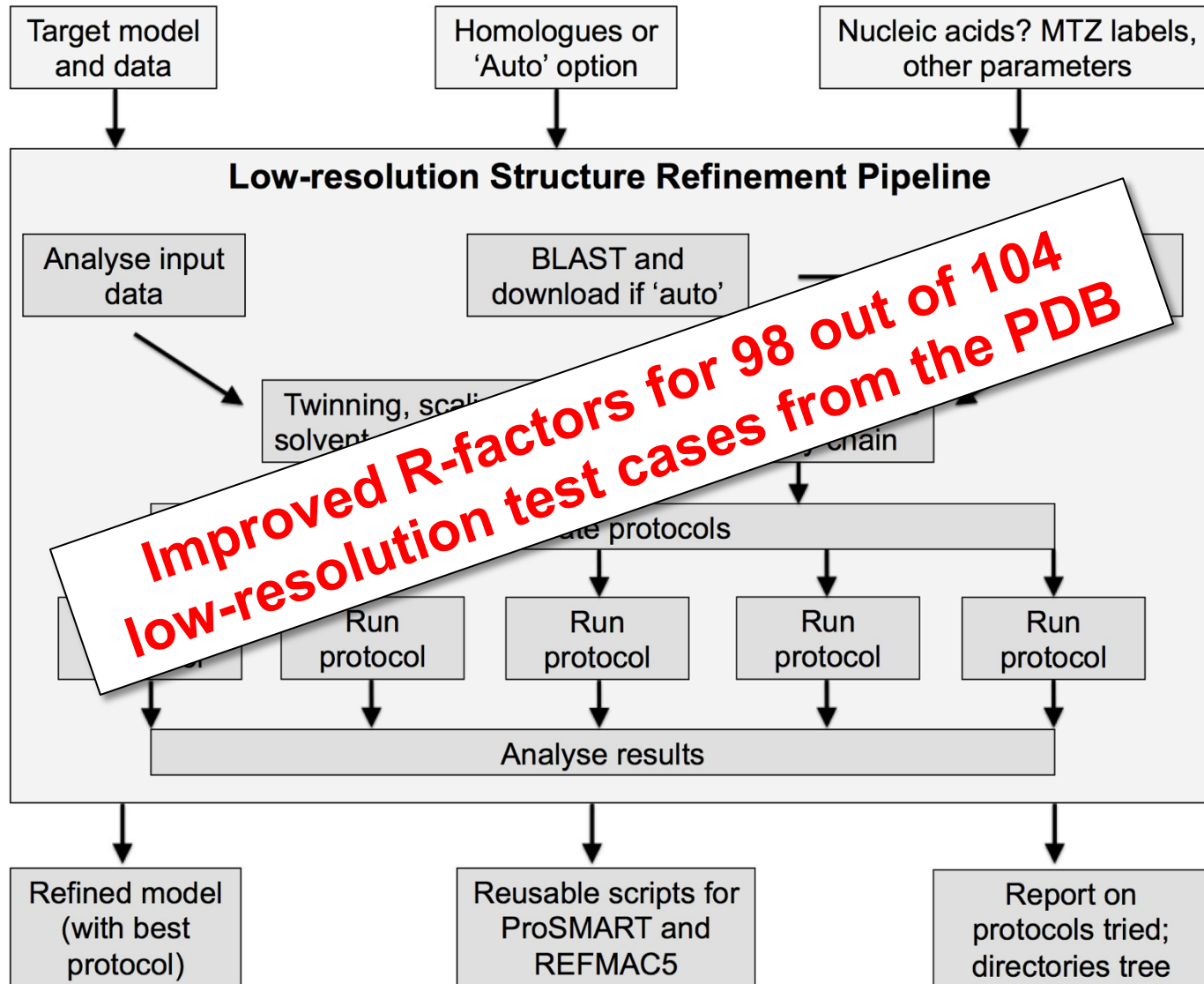
Automation: Magic Green Button

LORESTR: automation of refinement at low resolution



Automation: Magic Green Button

LORESTR: automation of refinement at low resolution



Automation: Magic Green Button

LORESTR: automation of refinement at low resolution

Run LORESTR pipeline [X] Help

Job title

Input and output files for refinement

MTZ in	Tutorial		Browse	View
PDB in	Tutorial		Browse	View
MTZ out	Tutorial		Browse	View
PDB out	Tutorial		Browse	View

☒ Run BLAST search over PDB database, fetch homologous structures and use them for external restraints

☐ Supply homologous structures for restraint generation

☐ Supply TLS file for the structure

☐ Number of CPUs to use:

Options ☒

☐ Run 100-200 cycles of jelly body refinement (if model is straight after molecular replacement)

☒ Save disk space by removing excessive ProSMART output (all restraints left)

Use chains for restraints generation (will determine number of protocols to run)

Use up to homologues for analysis

☐ Generate restraints for DNA/RNA chains

☐ CNS-style Free flag ('FREE 1' option for Refmac)

☐ Specify LABIN parameter for Refmac:

Do not download homologues with resolution lower than Angstrom

Run Save or Restore Close

Automation: Magic Green Button

LORESTR: automation of refinement at low resolution

Input Data Options Advanced Options

Job title LORESTR

Use data from job 107 Import a coordinate set – optional selection of as input below..

Main inputs

 Atomic model 27 Atomic model imported from 1.pdb by job 27  

 Reflections 37 1: unknown180412 imported by job 37  

 Free R set 37 1: unknown180412 imported by job 37  

Optional additional inputs

☐ TLS coefficients ..is not used  

 Hide list

145 of 144 Reference model imported from 1.pdb by job 144
146 of 144 Reference model imported from 2.pdb by job 144
147 of 144 Reference model imported from 2_eh.pdb by job 144

 Reference model 145 of 144 Reference model imported from 1.pdb by job :  

Input Data Options Advanced Options

Main options

☒ Auto mode: find homologous structures and use them for restraints generation.
☐ Generate external restraints for DNA/RNA chains

Number of CPUs to use: 8

After Molecular Replacement

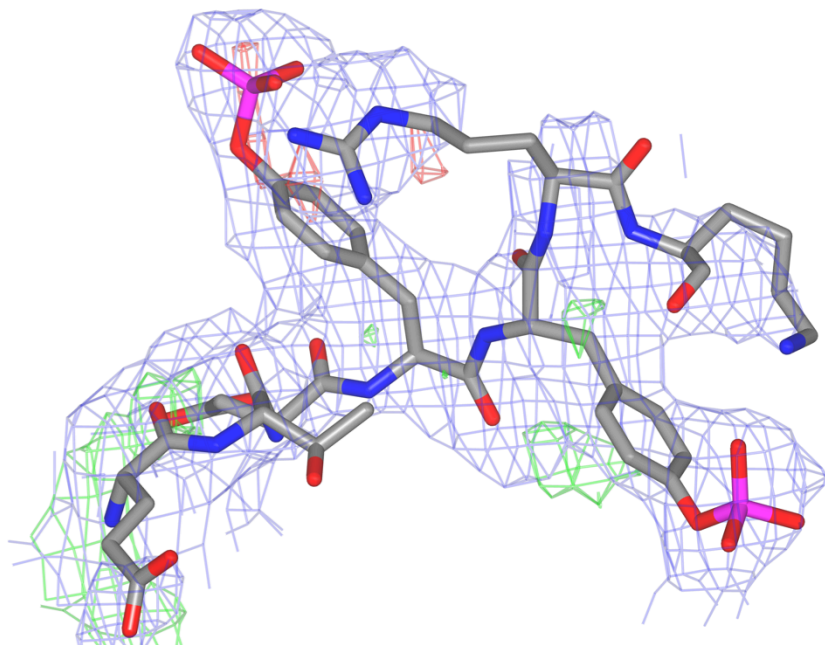
☐ Run 100–200 cycles of jelly body refinement first (if model is straight after MR)

Automation: Magic Green Button

LORESTR: automation of refinement at low resolution

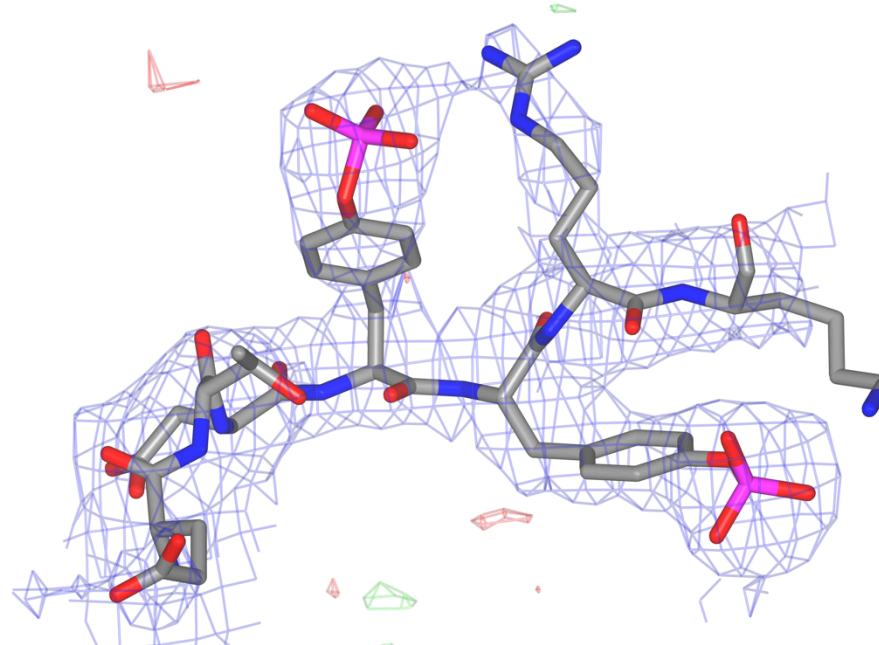
Structure	R_{work} (%)	R_{free} (%)	Difference between R_{free} and R_{work} (%)	Ramachandran favoured (%)	Ramachandran outliers (%)
2z8c	21.7	29.2	7.5	78.2	5.0
After <i>LORESTR</i> in auto mode	23.2	25.2	2.0	90.3	4.0
Resolved with 1ir3, then <i>LORESTR</i>	18.6	23.4	4.8	96.0	1.3

2z8c



4 out of 7 residues are Ramachandran outliers

re-solved using 1ir3, then LORESTR



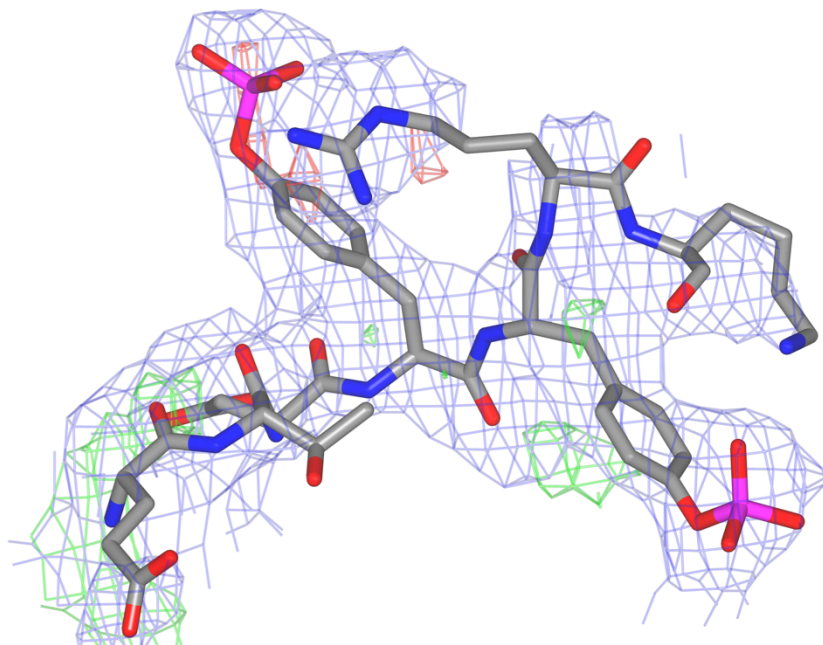
no Ramachandran outliers in this region

Automation: Magic Green Button

LORESTR: automation of refinement at low resolution

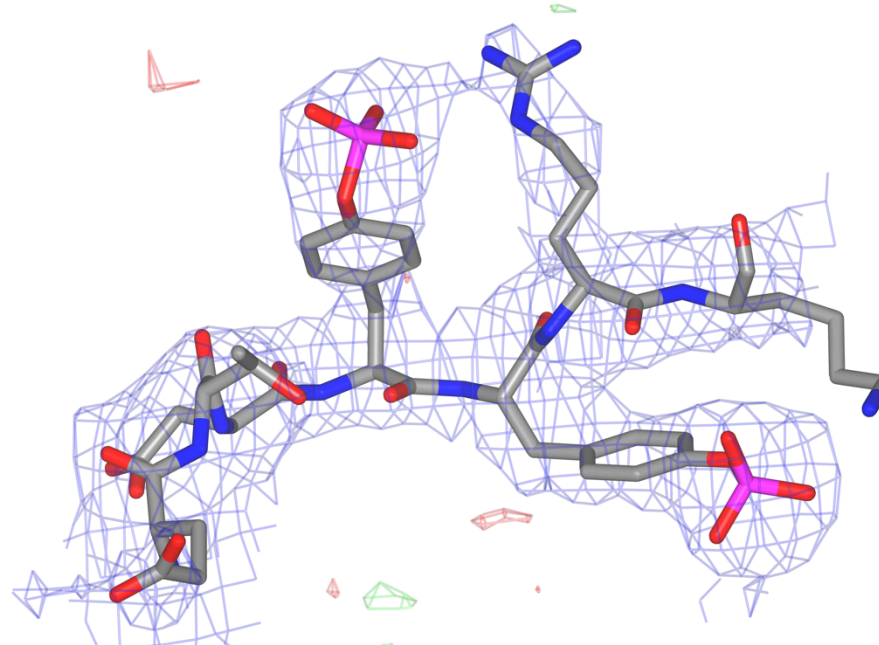
Structure	R_{work} (%)	R_{free} (%)	Difference between R_{free} and R_{work} (%)	Ramachandran favoured (%)	Ramachandran outliers (%)
2z8c	21.7	29.2	7.5	78.2	5.0
After <i>LORESTR</i> in auto mode	23.2	25.2	2.0	90.3	4.0
Resolved with 1ir3, then <i>LORESTR</i>	18.6	23.4	4.8	96.0	1.3

2z8c



4 out of 7 residues are Ramachandran outliers

re-solved using 1ir3, then LORESTR



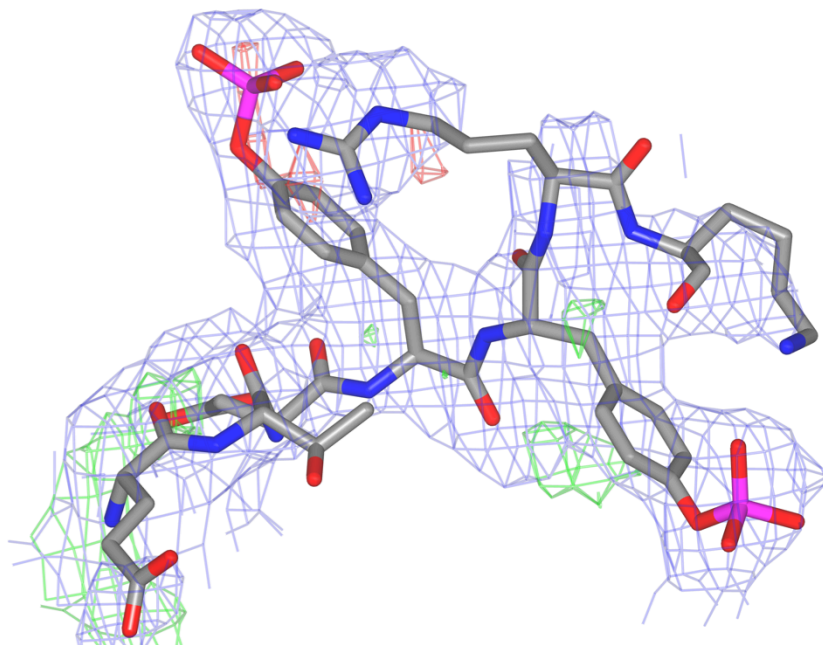
no Ramachandran outliers in this region

Automation: Magic Green Button

LORESTR: automation of refinement at low resolution

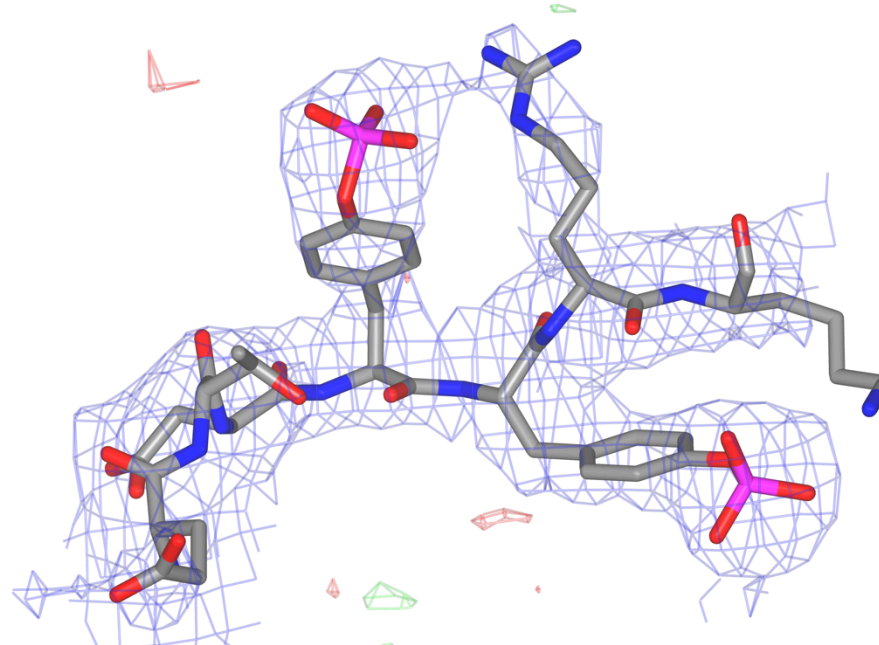
Structure	R_{work} (%)	R_{free} (%)	Difference between R_{free} and R_{work} (%)	Ramachandran favoured (%)	Ramachandran outliers (%)
2z8c	21.7	29.2	7.5	78.2	5.0
After <i>LORESTR</i> in auto mode	23.2	25.2	2.0	90.3	4.0
Resolved with 1ir3, then <i>LORESTR</i>	18.6	23.4	4.8	96.0	1.3

2z8c



4 out of 7 residues are Ramachandran outliers

re-solved using 1ir3, then LORESTR



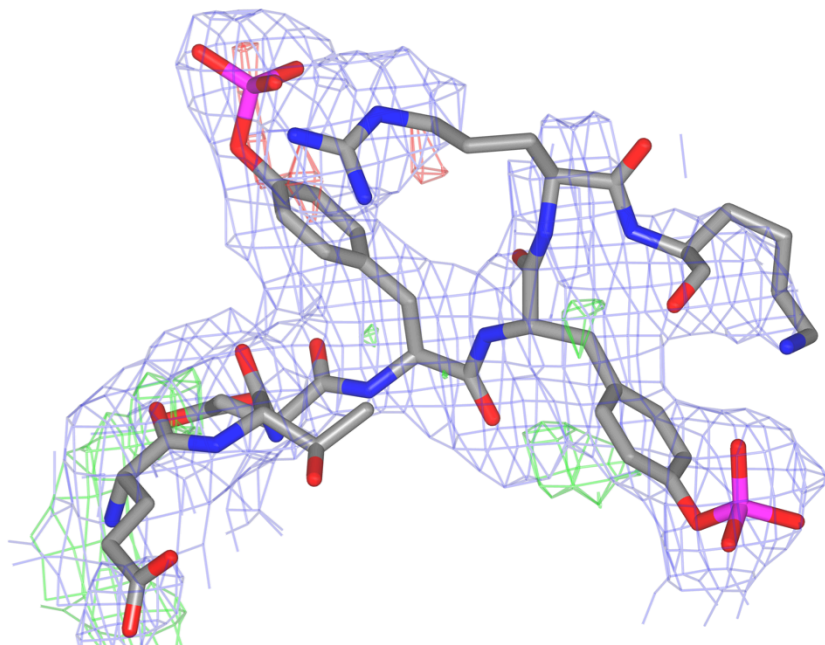
no Ramachandran outliers in this region

Automation: Magic Green Button

LORESTR: automation of refinement at low resolution

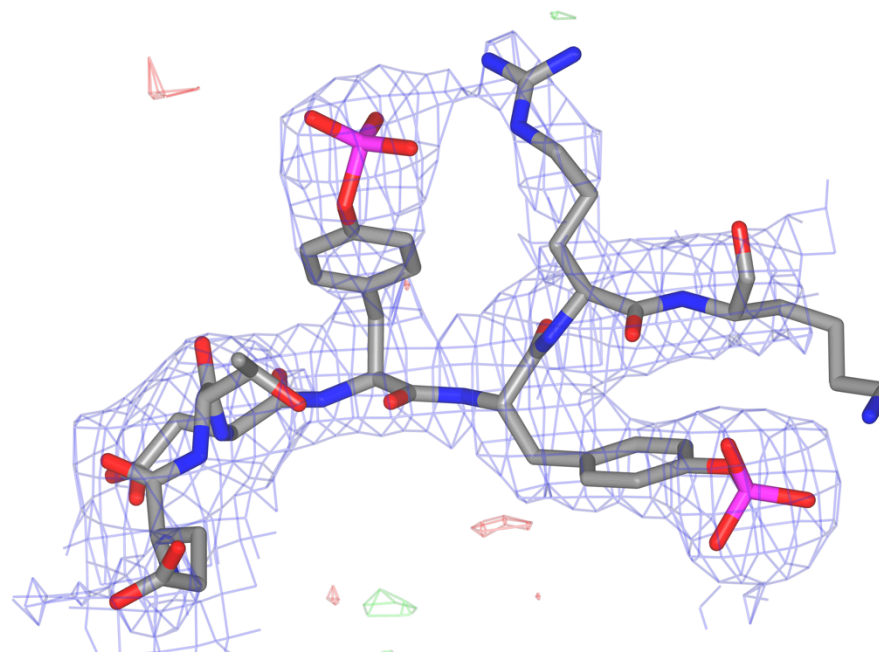
Structure	R_{work} (%)	R_{free} (%)	Difference between R_{free} and R_{work} (%)	Ramachandran favoured (%)	Ramachandran outliers (%)
2z8c	21.7	29.2	7.5	78.2	5.0
After <i>LORESTR</i> in auto mode	23.2	25.2	2.0	90.3	4.0
Resolved with 1ir3, then <i>LORESTR</i>	18.6	23.4	4.8	96.0	1.3

2z8c



4 out of 7 residues are Ramachandran outliers

re-solved using 1ir3, then LORESTR

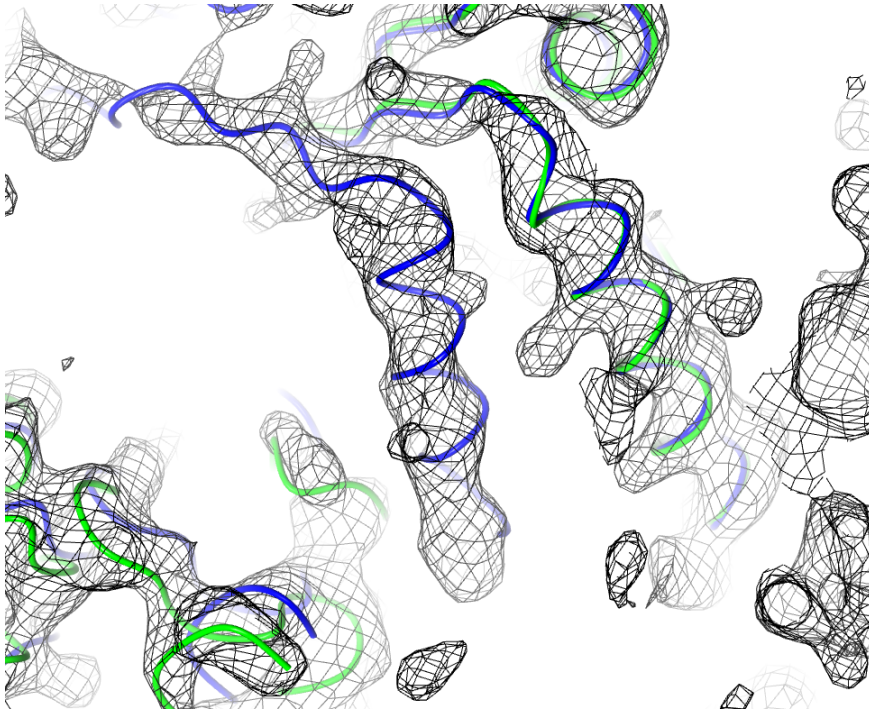


no Ramachandran outliers in this region

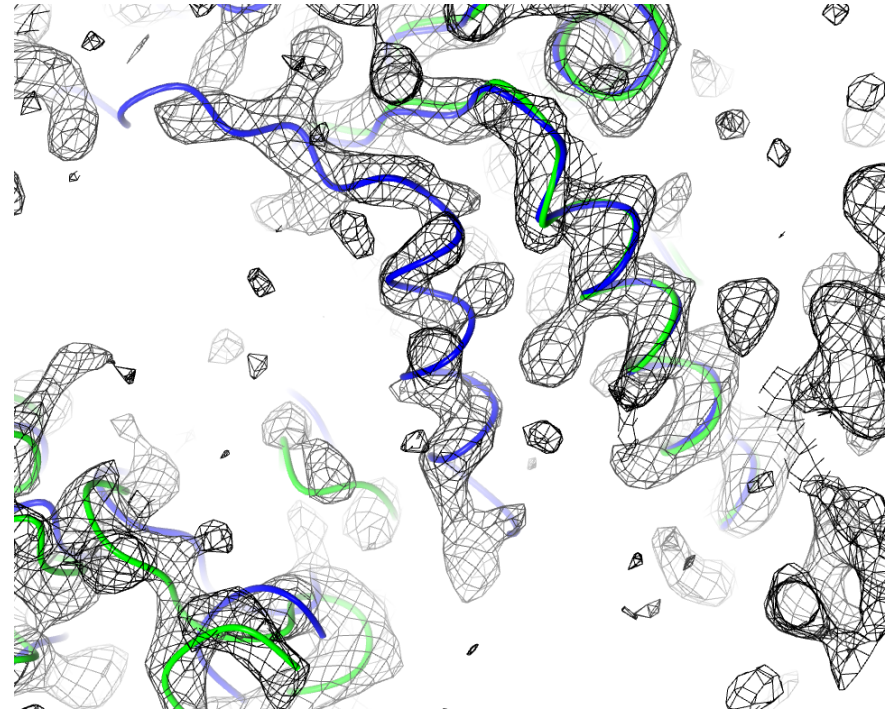
REFMAC Anisotropic Map Sharpening

Idea – remove an overall B value

Original Map



Sharpened map from REFMAC



Green: original structure

Blue: homologous structure

2r6c (4.0Å) – helix unmodelled

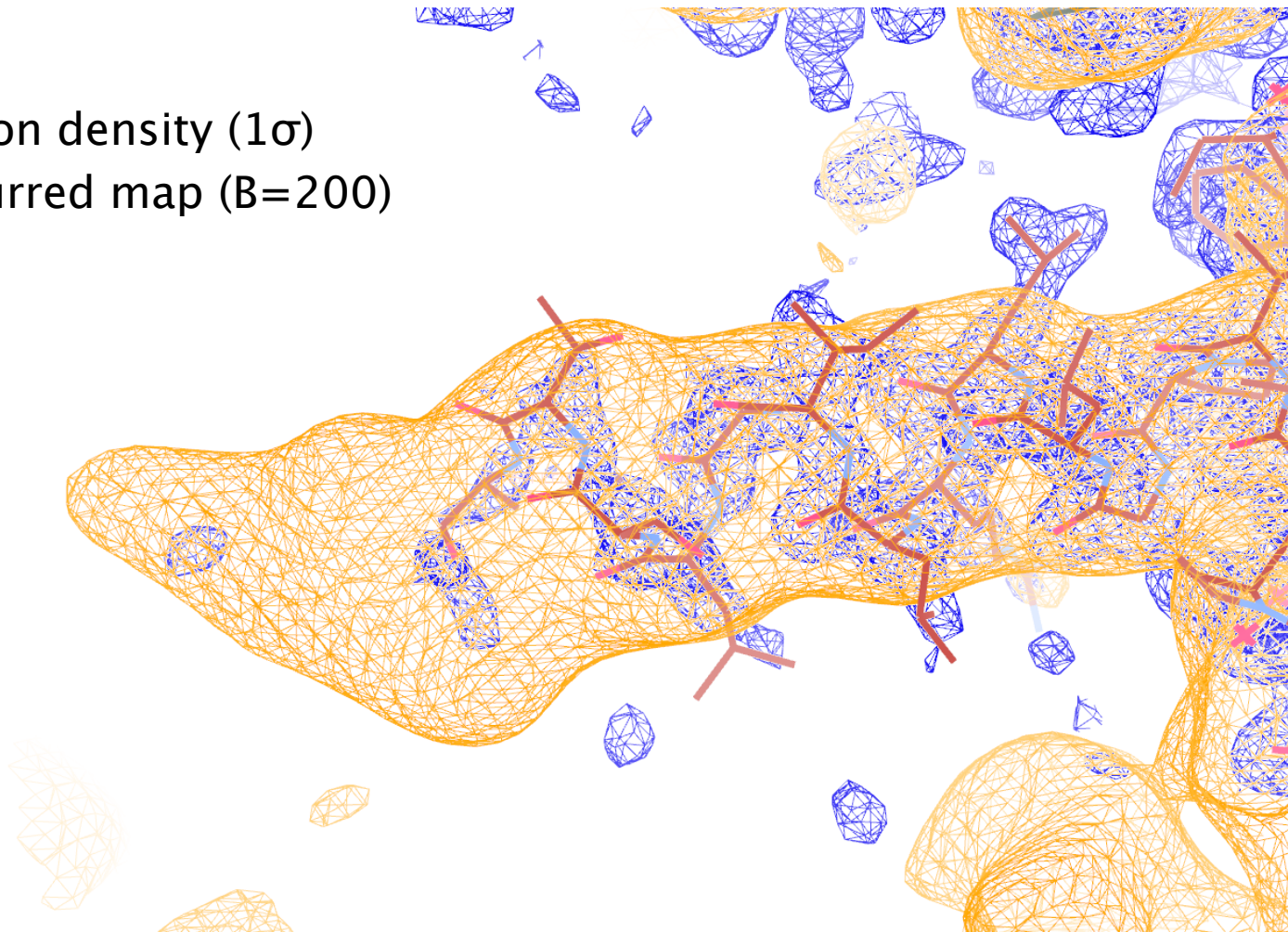
2r6a (2.9Å)

Map Blurring

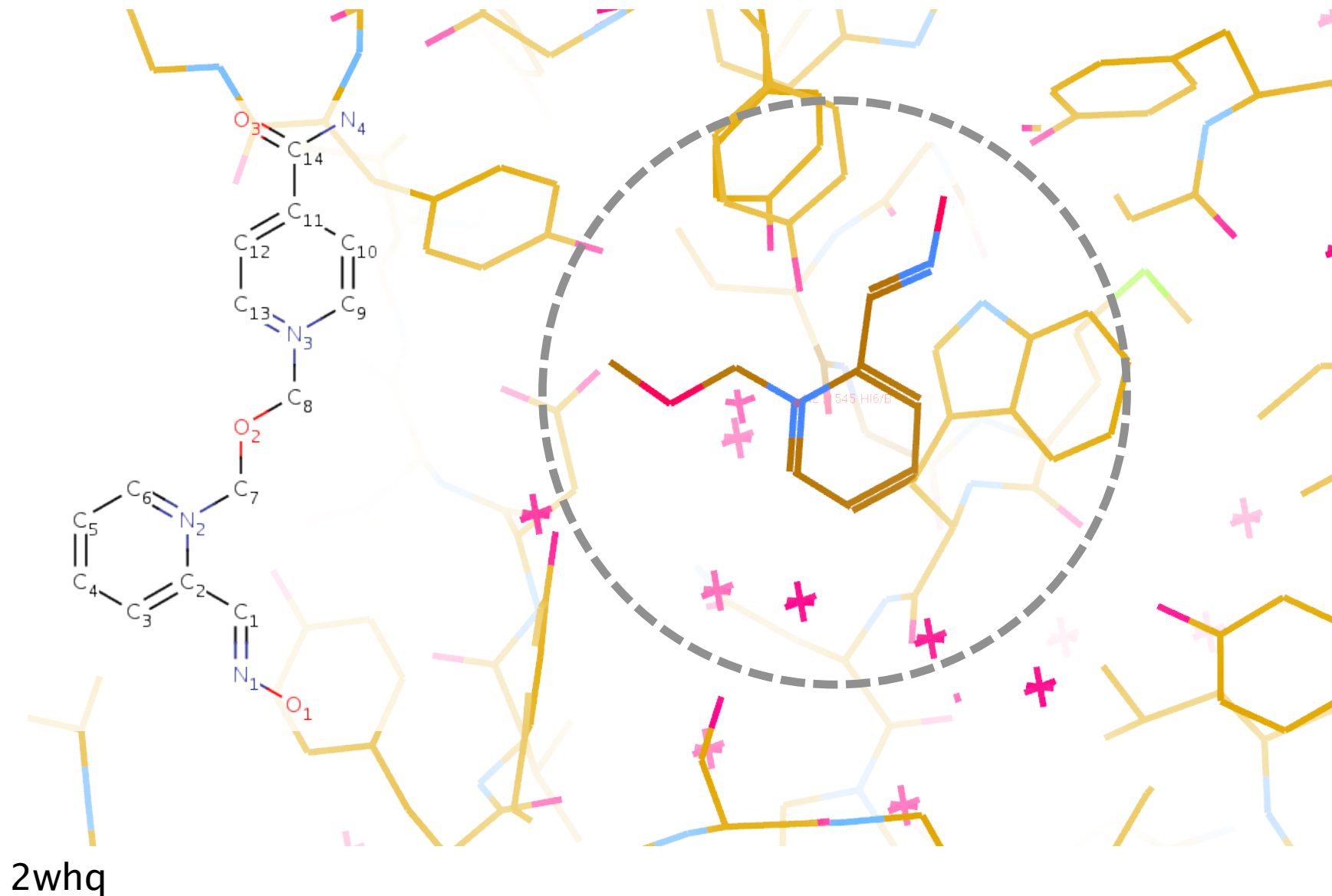
Idea – apply an overall B value

Blue – electron density (1σ)

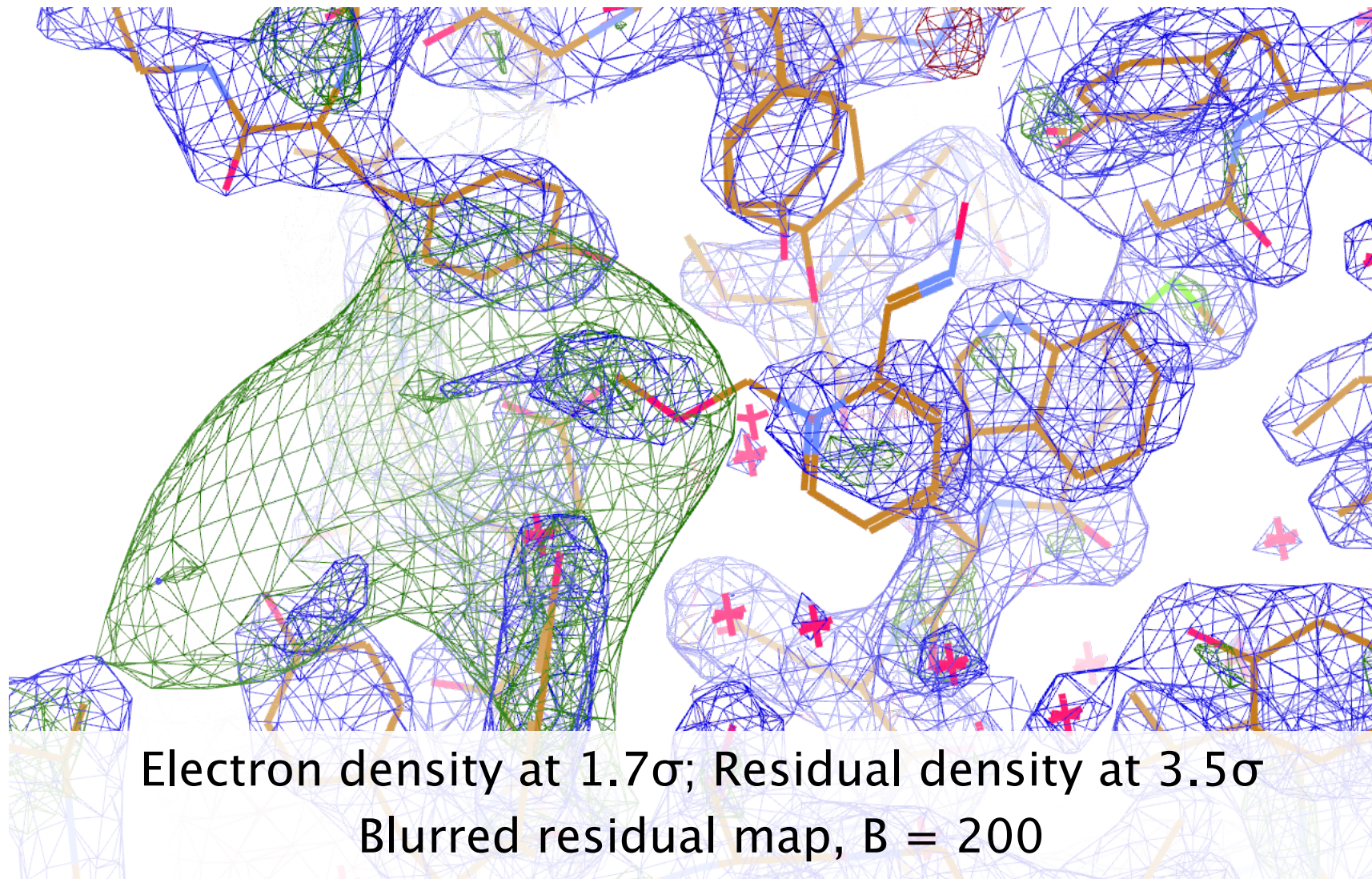
Orange – blurred map (B=200)



Map Blurring



Map Blurring

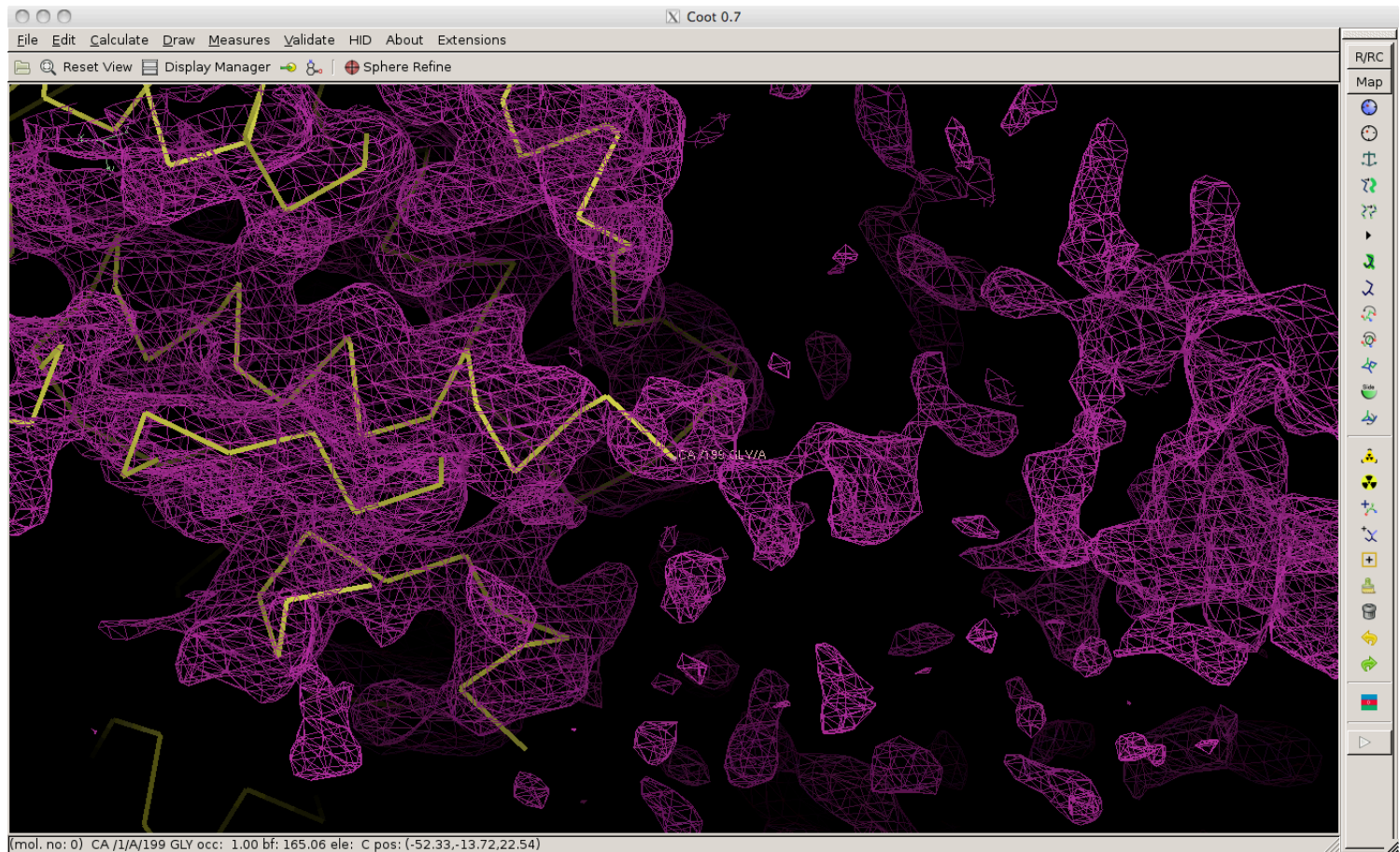


Map Sharpening/Blurring

Output Array of Sharpened Maps

Standard:

4iwo (2.6Å)

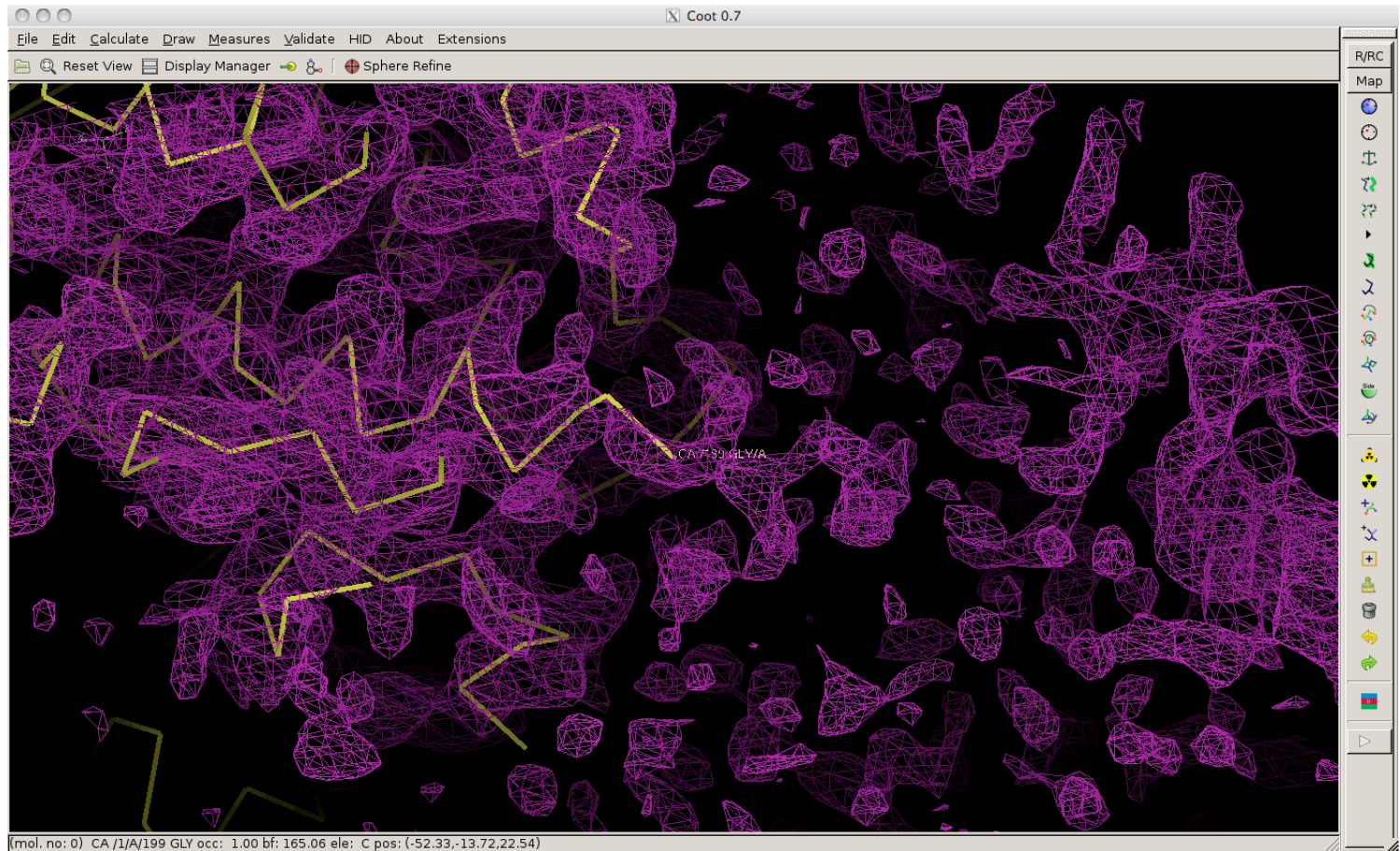


Map Sharpening/Blurring

Output Array of Sharpened Maps

Sharpen 50:

4iwo (2.6Å)

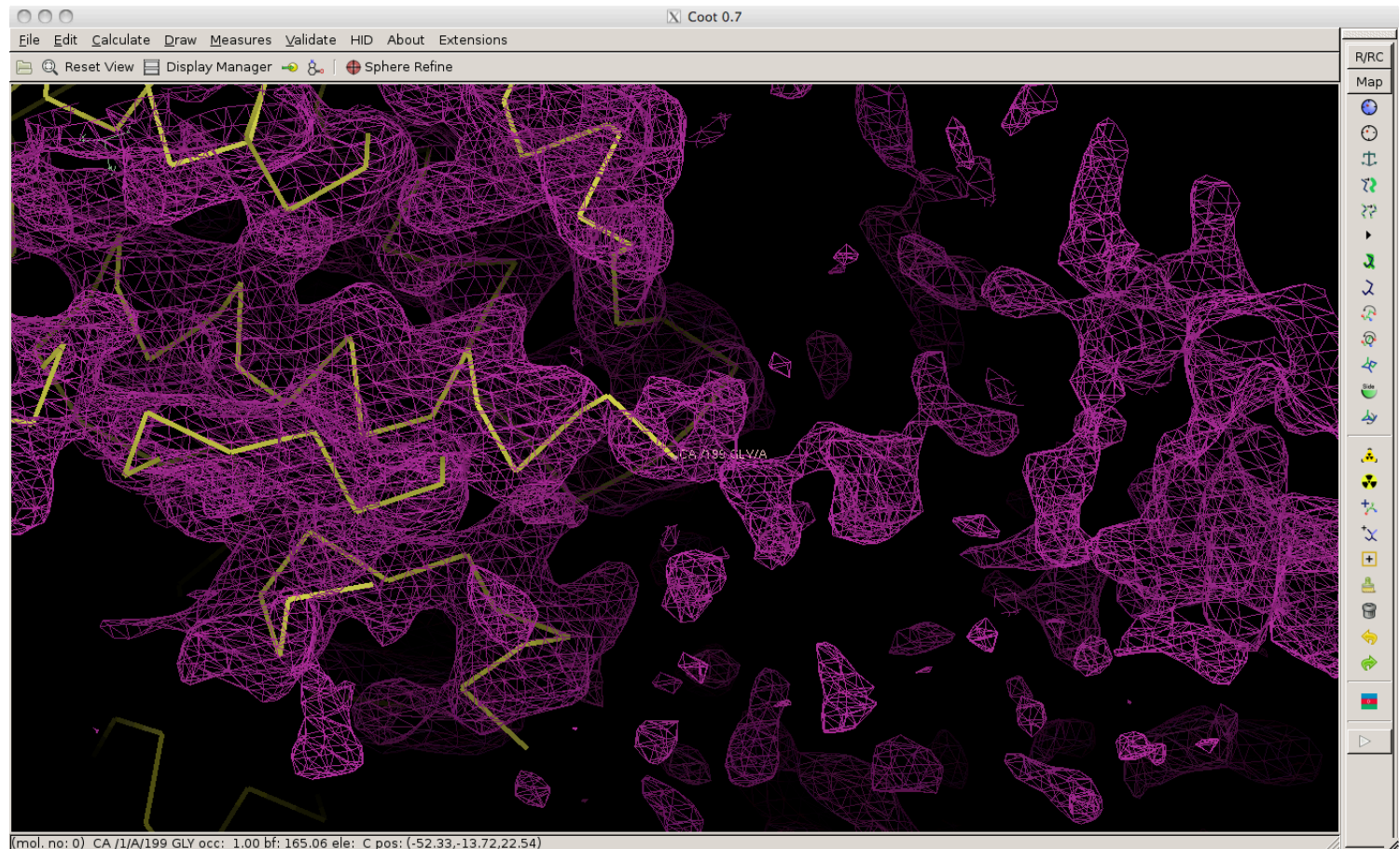


Map Sharpening/Blurring

Output Array of Sharpened Maps

Standard:

4iwo (2.6Å)

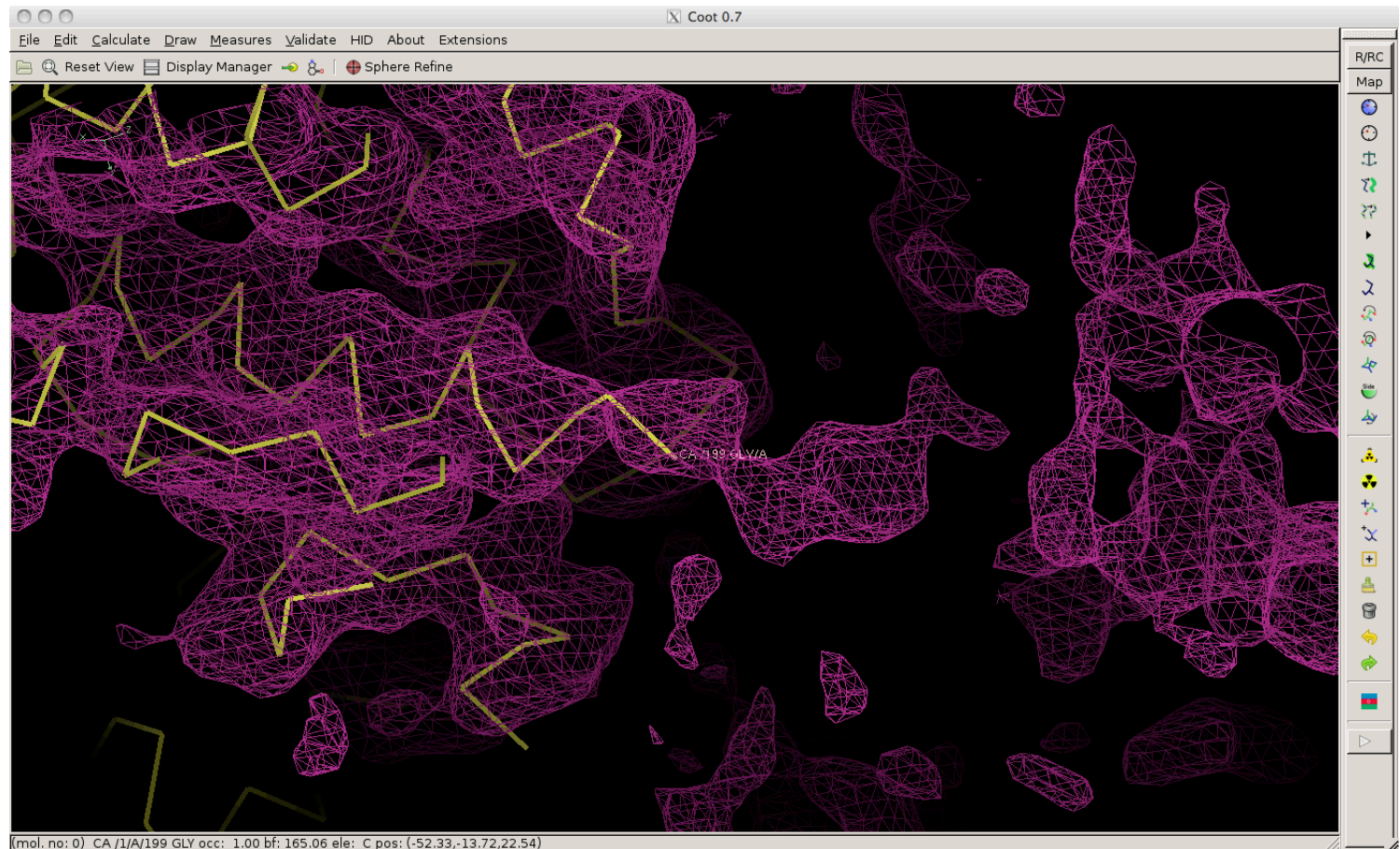


Map Sharpening/Blurring

Output Array of Sharpened Maps

Blur 50:

4iwo (2.6Å)

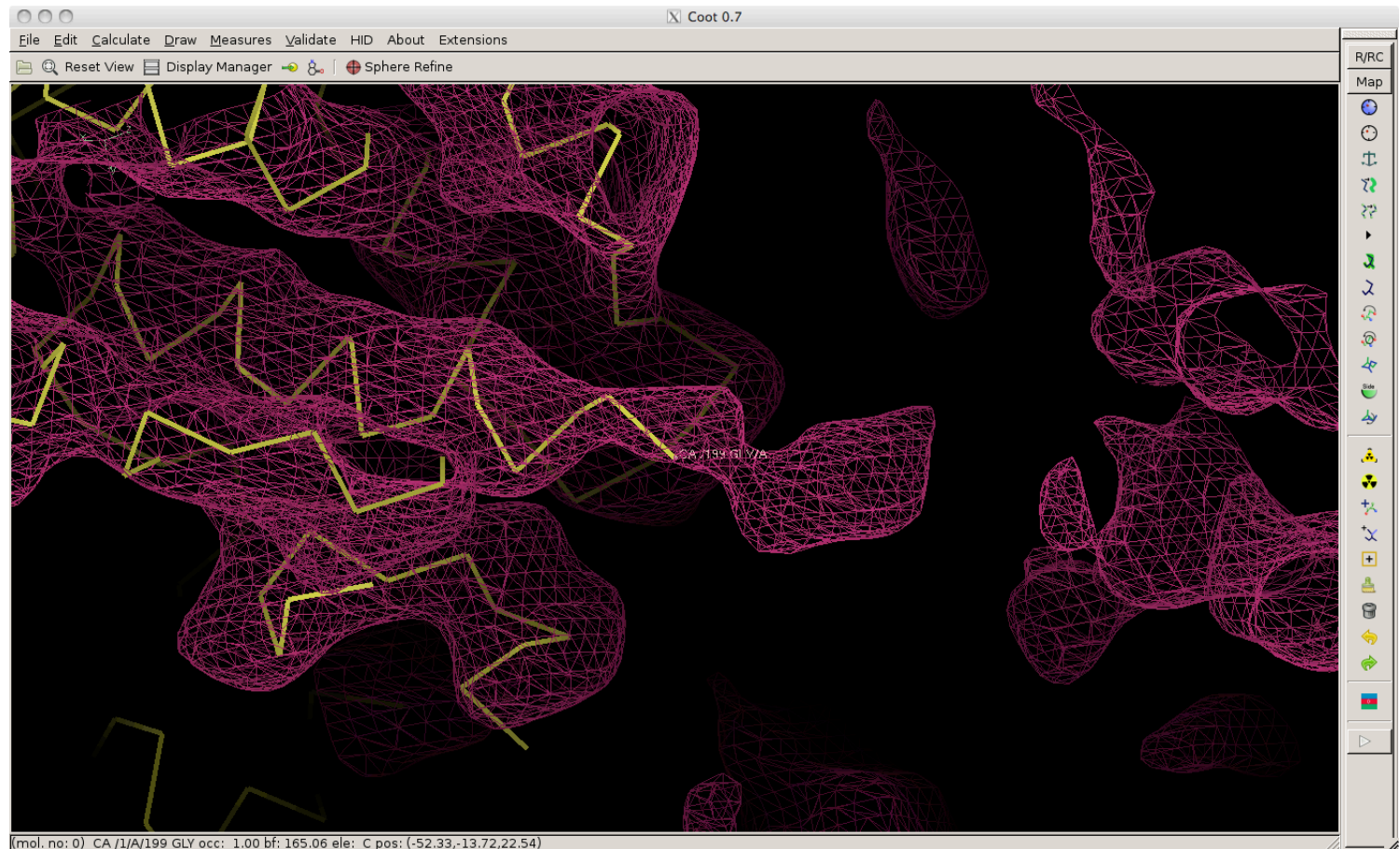


Map Sharpening/Blurring

Output Array of Sharpened Maps

Blur 150:

4iwo (2.6Å)



What and When

Early stages (e.g. straight after MR)

- Jelly body – sometimes up to 200 cycles

Medium stages – during model building

- Auto local NCS – wherever possible
- External restraints (40 cycles) – homologue available
- Otherwise, jelly body... but not together
- H-bond restraints – no homologue available
- Secondary structure conformation restraints – model building tool
- Add hydrogens (?)

Medium-final stages

- TLS – at medium resolutions
- Twin refinement – only if you are sure

Final stages of refinement

- Jelly body – around 20 cycles

Summary

CCP4 tools for refinement:

REFMAC: Refinement, map calculation

ProSMART: External restraints, comparative analysis

LIBG: Nucleic acid restraints

LORESTR: Automated low-resolution refinement pipeline (*released soon!*)

ACEDRG: Generating ligand description dictionaries and conformers

COOT: Automated & manual real-space refinement
Visualisation & manipulation of restraints, map blurring
...also morphing, jiggle-fit, backrub rotamers...

Many tools are applicable to cryo-EM as well as MX

Acknowledgements

Computational Crystallography Group:

Garib Murshudov

Paul Emsley

Oleg Kovalevskiy

Fei Long

Rob Nicholls

Michal Tykac

Many Thanks:

CCP4 core team

Colleagues from MRC-LMB and CCP4

Azzurra Carlon and colleagues from CERM

Users for feedback

Computational Crystallography Group website:

www2.mrc-lmb.cam.ac.uk/groups/murshudov/

REFMAC5:

Murshudov *et al.* (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Cryst.* D67, 355-67.

Low-resolution refinement with REFMAC5 and ProSMART:

Nicholls *et al.* (2013) Recent Advances in Low Resolution Refinement Tools in REFMAC5. *Adv. Methods for Bio. Xtallography*, 231–58.

Nicholls *et al.* (2012) Low Resolution Refinement Tools in REFMAC5. *Acta Cryst.* D68, 404-17.

ProSMART:

Nicholls *et al.* (2014) Conformation-Independent Structural Comparison of Macromolecules with ProSMART. *Acta Cryst.* D70, 2487-99.

Nicholls *et al.* (2012) An Overview of ProSMART. *CCP4 Newsletter on Protein Crystallography*.

LORESTR:

Kovalevskiy *et al.* (2016) Automated refinement ... at low resolution using prior information. *Acta Cryst.* D72, 1149–1161.

