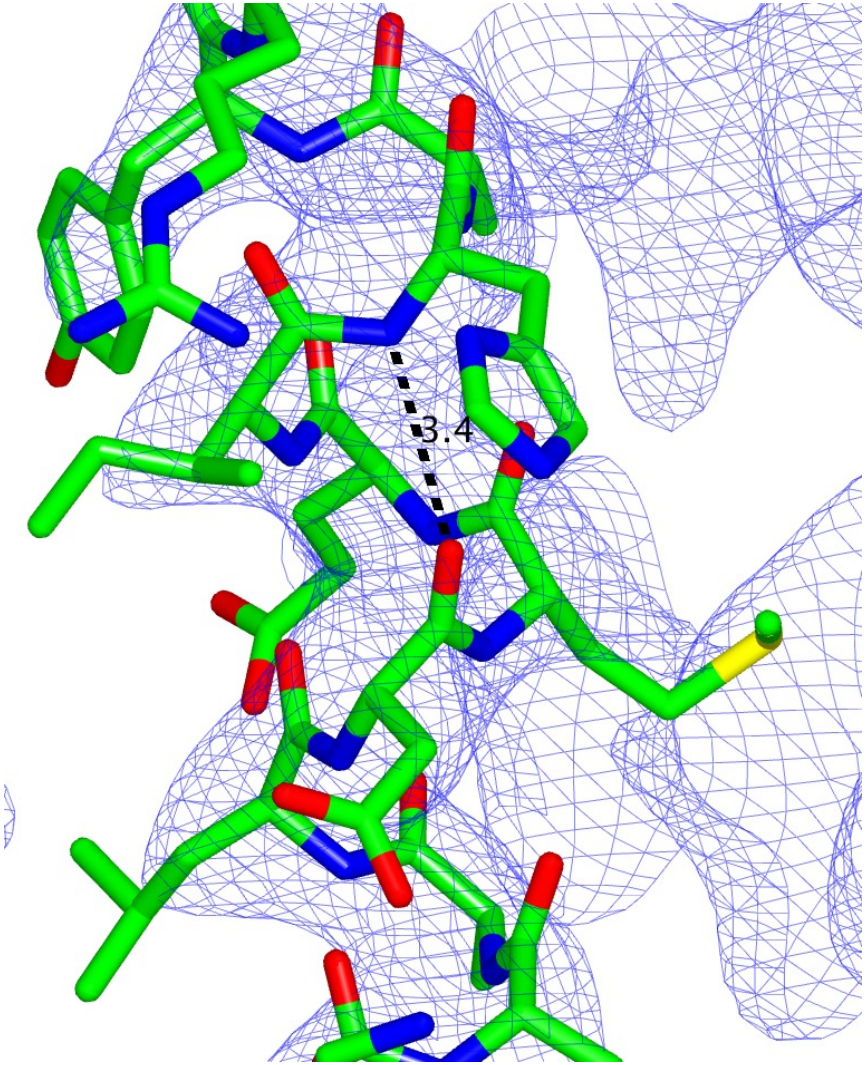
ProSMART

Injection of prior knowledge to aid new structure determination

- **External Restraints from homologous structures**
  - Protein or nucleic acid chains
- **Hydrogen bond restraints**
  - Protein backbone
- **Generic self-restraints**
  - Everything – protein, nucleic acid, ligand, water
- **Structure analysis**
  - Alignment & comparison - helps analyse differences between models

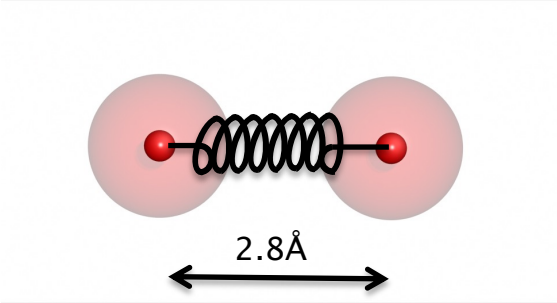
Independent of global conformation

ProSMART External Restraints



3g4w – 3.7Å

Prior information:

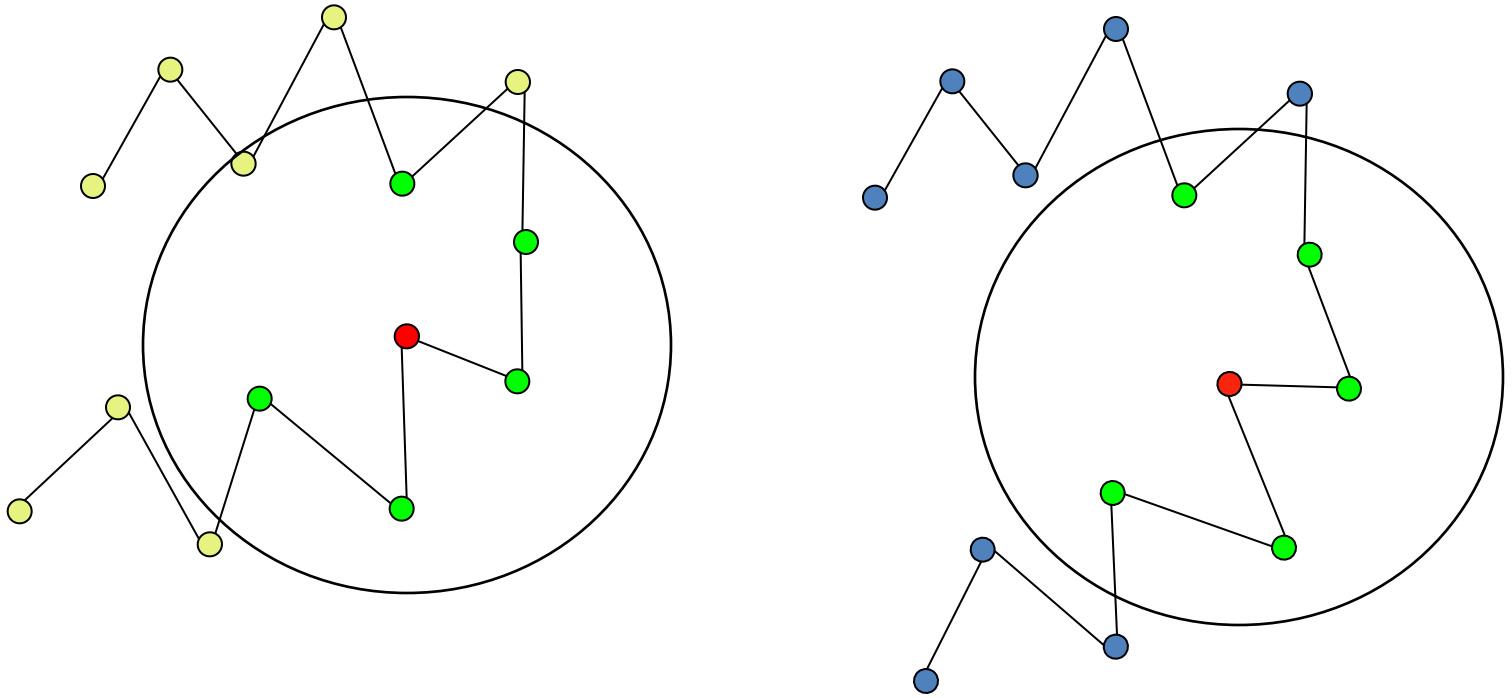


Stabilises structural features

External Restraint Generation

structure to be refined

known similar structure (prior)

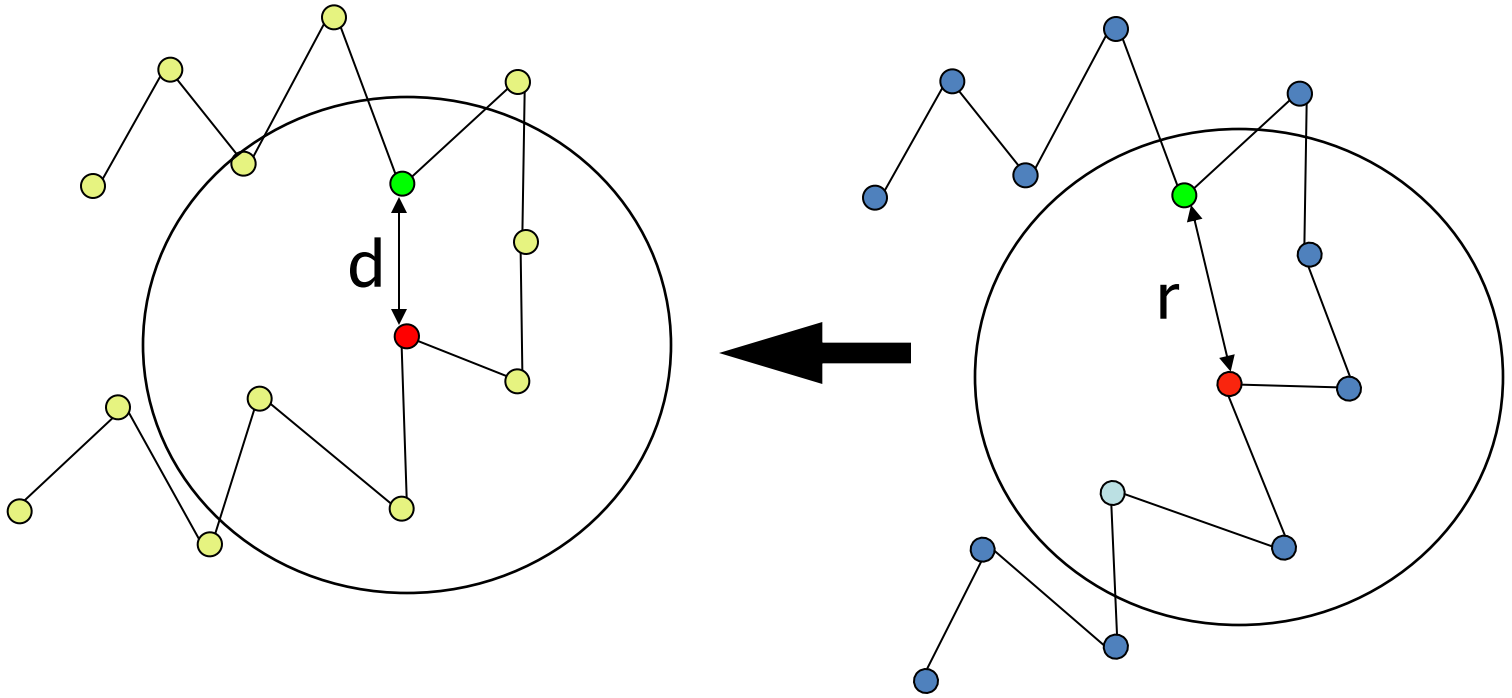


(abstract representation of an atomic model; circles = atoms)

External Restraint Generation

structure to be refined

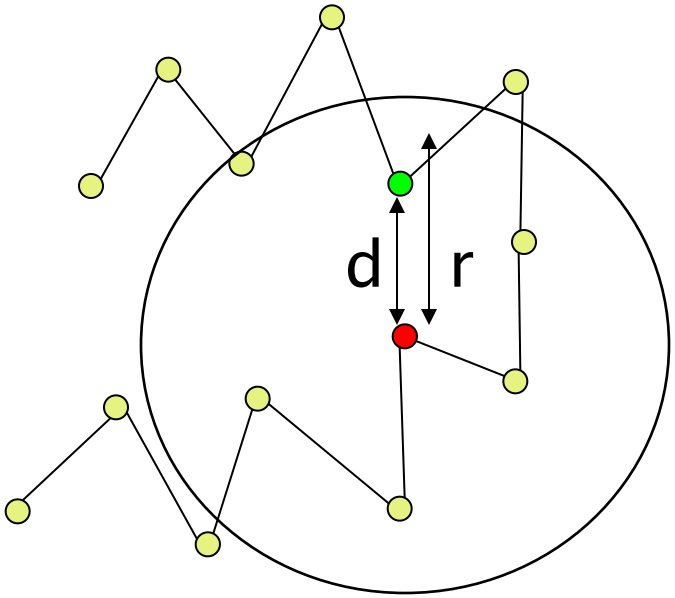
known similar structure (prior)



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

External Restraint Generation

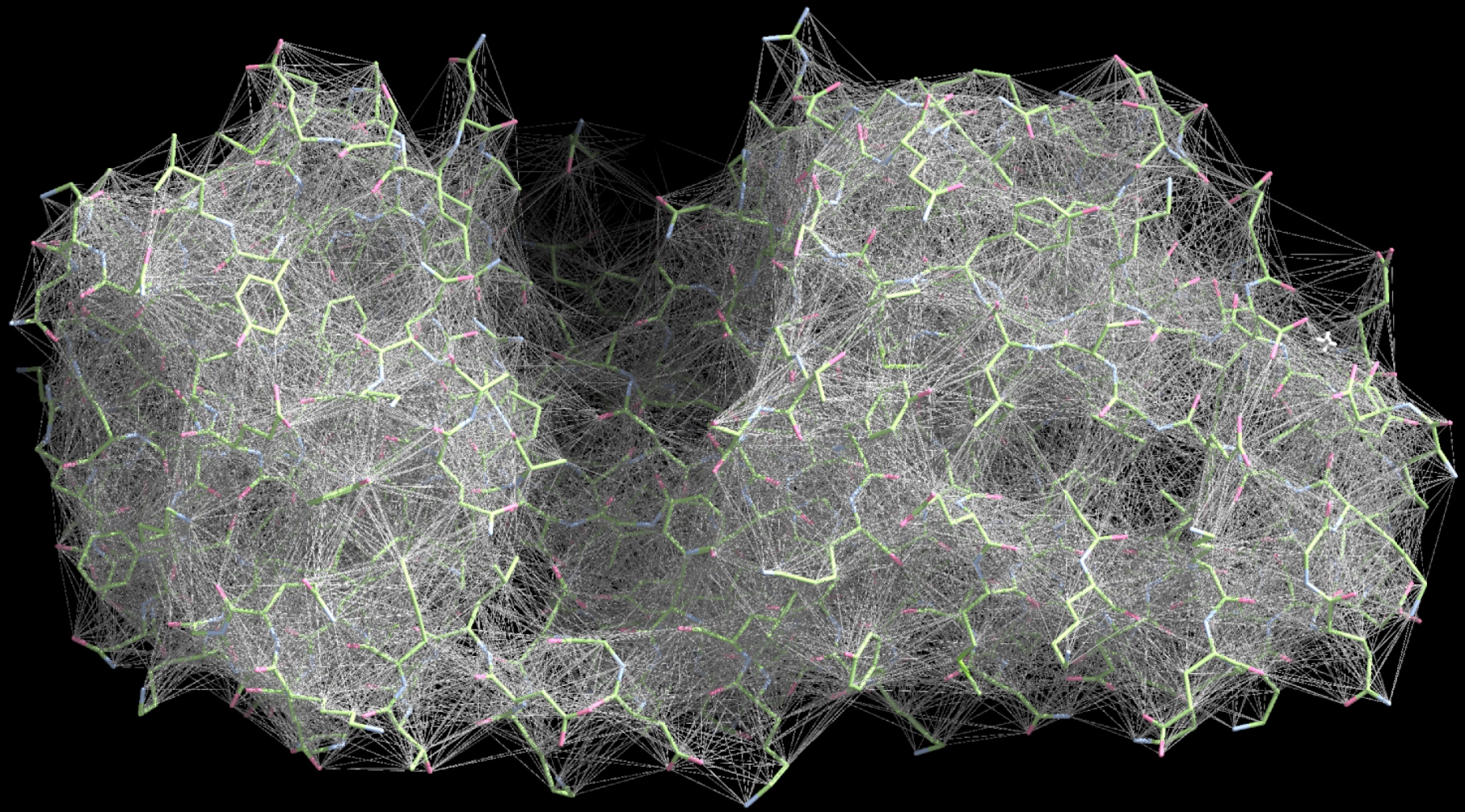
structure to be refined



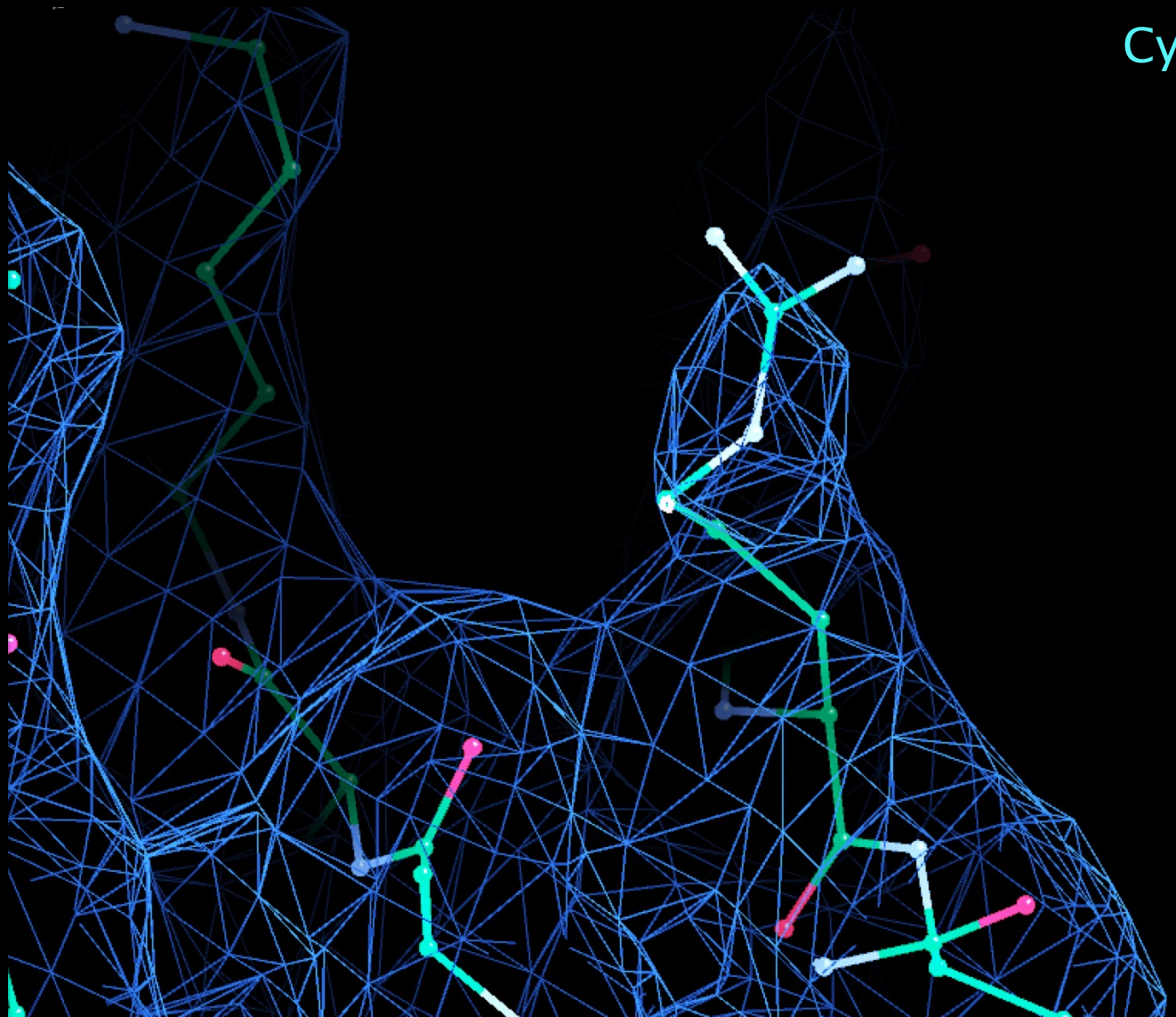
$d \sim N(r,\sigma^2)$

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

ProSMART Restraints

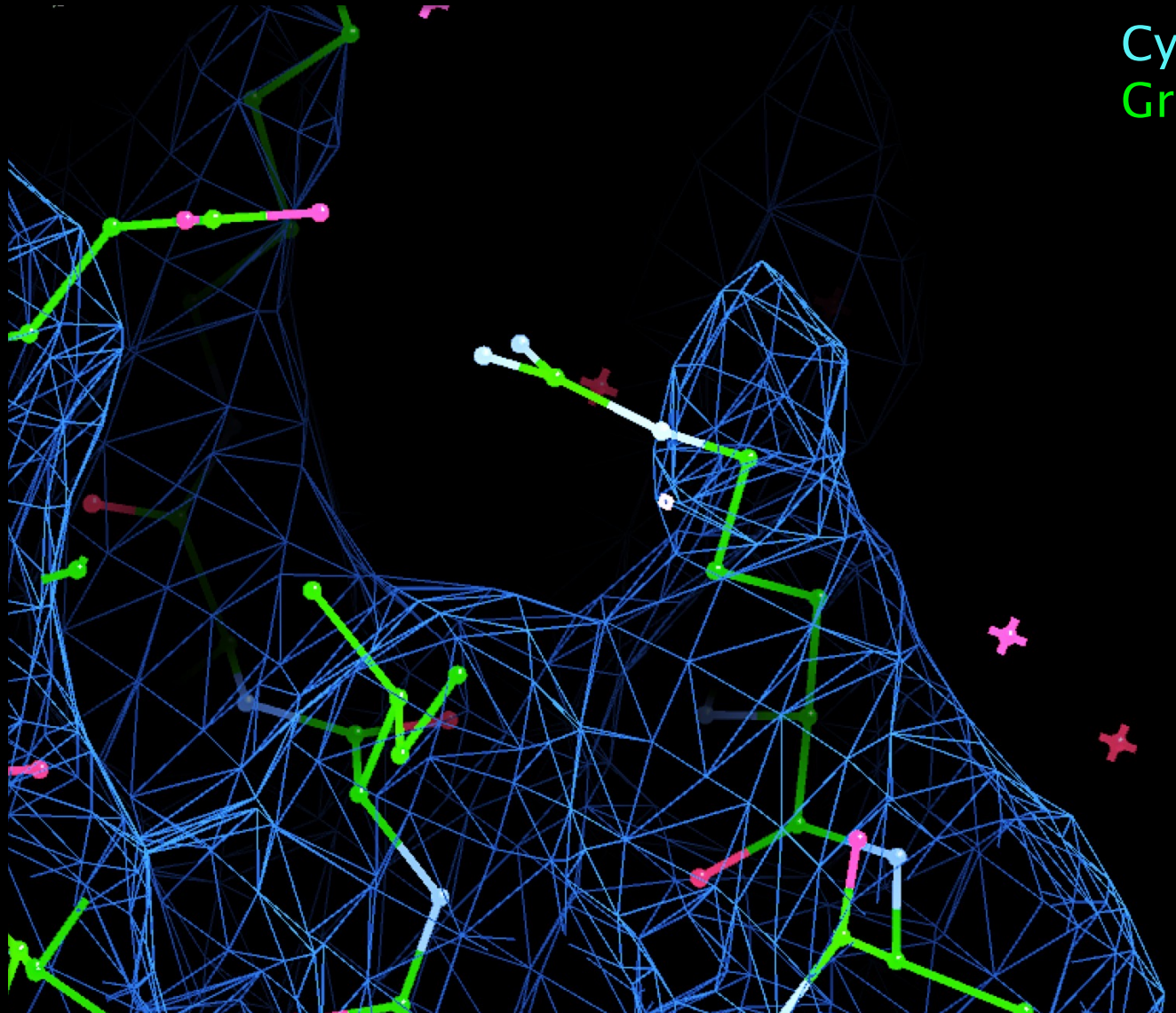


Robust Estimation



Cyan: original

Robust Estimation



Cyan: original
Green: homolog

Robust Estimation

Cyan: original
Green: homolog
Yellow: refined

Red: long
Grey: similar
Blue: short

Robust Estimation

Cyan: original
Green: homolog
Yellow: refined

Red: long
Grey: similar
Blue: short

Motivational Example

Ovotransferrin

1ryx – 3.5Å

Low-resolution refinement:

Weak signal
Noisy data

↓

Unstable refinement

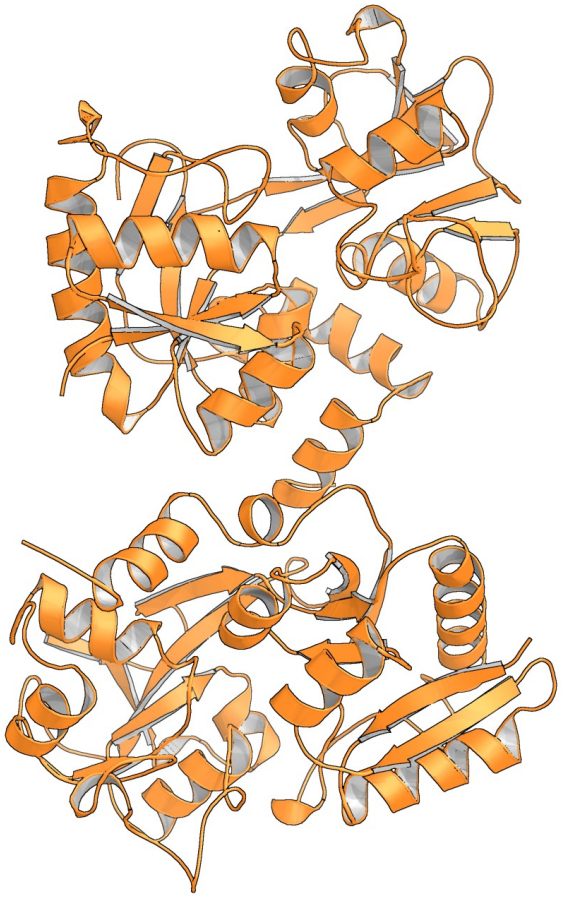
Result:
Poor quality model

Motivational Example

Ovotransferrin

1ryx – 3.5Å

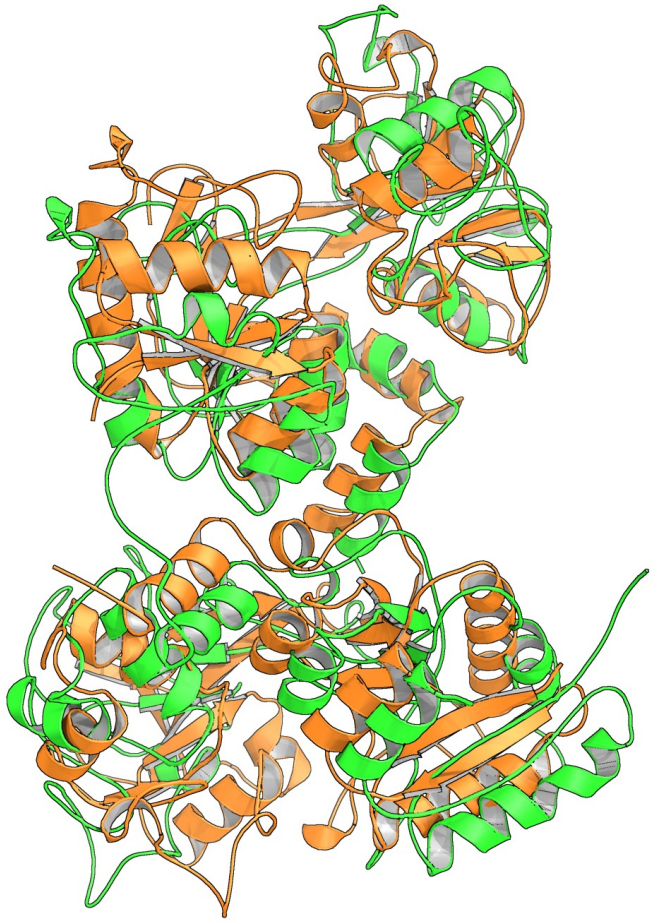
High-resolution homologue



2d3i – 2.15Å

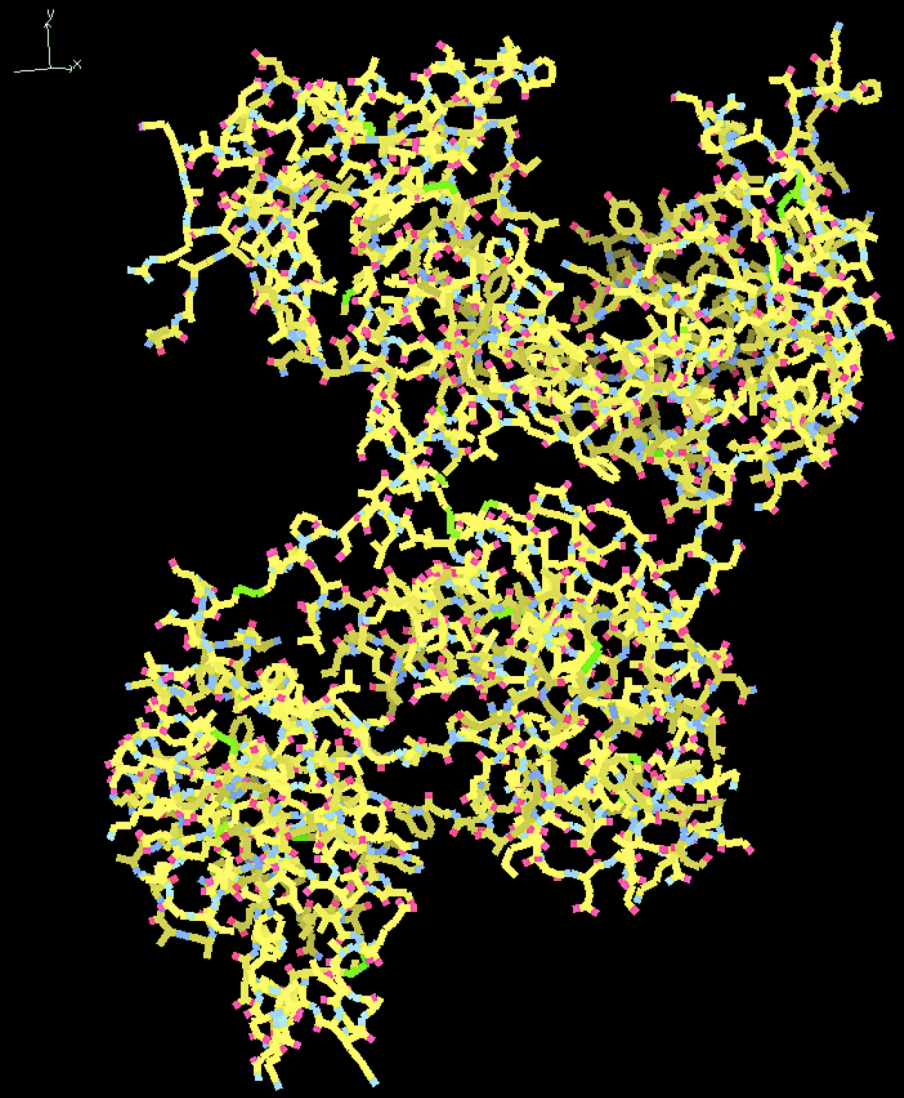
Motivational Example

Ovotransferrin



Models don’t superpose well

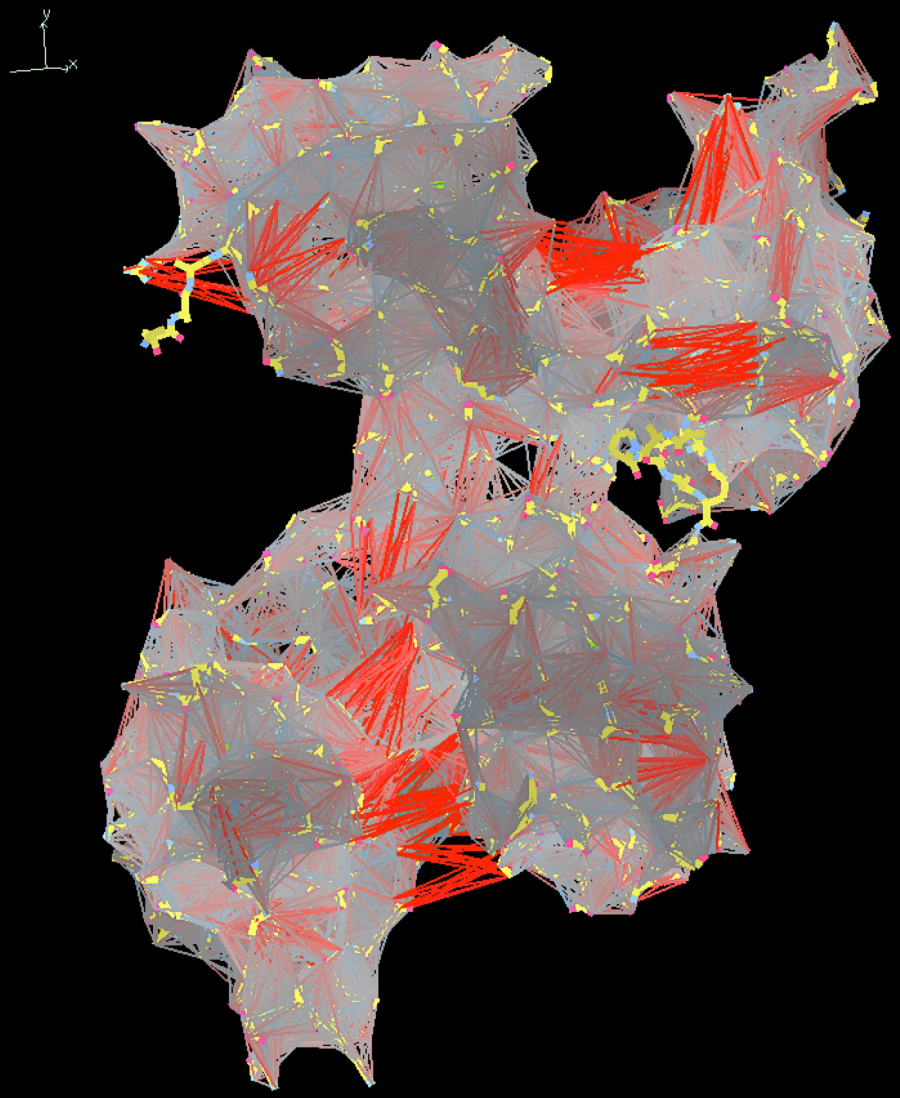
Example: Ovotransferrin



1ryx (3.5Å)

Example: Ovotransferrin

Restraints:
Backbone
Side chains



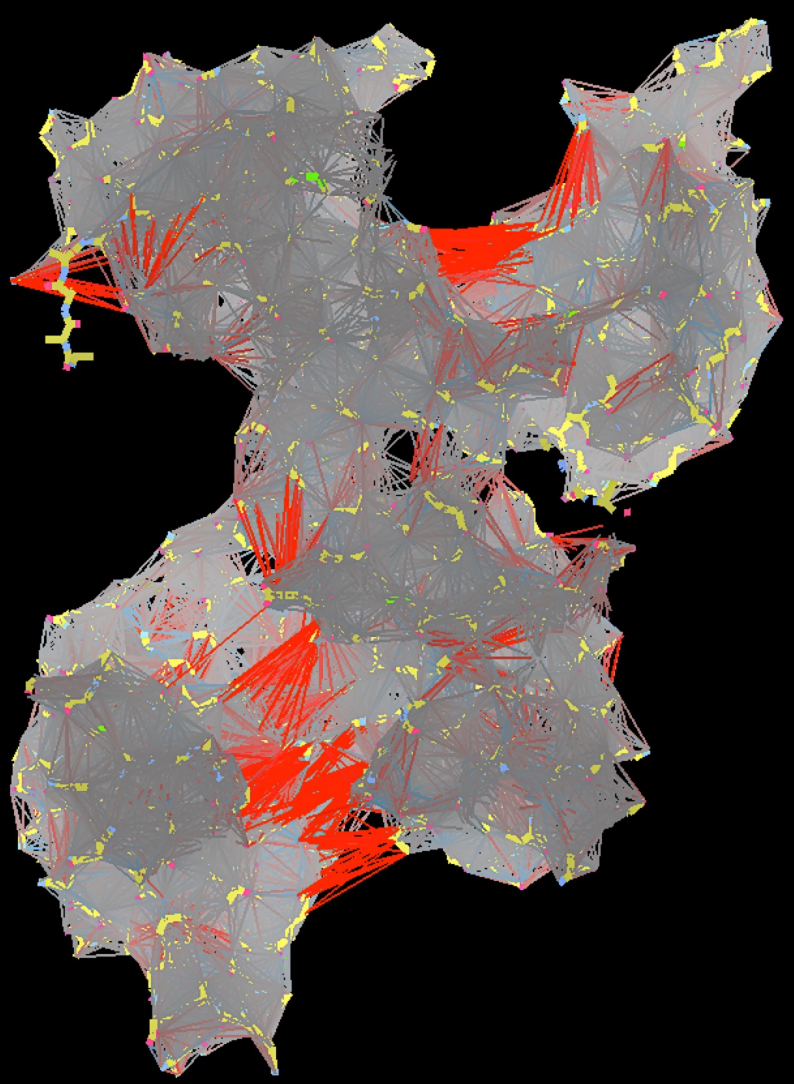
1ryx (3.5Å)
restrained to
2d3i (2.15Å)

Red: long
Grey: similar
Blue: short

Example: Ovotransferrin

Restraints:
- Backbone
- Side chains

After re-refinement

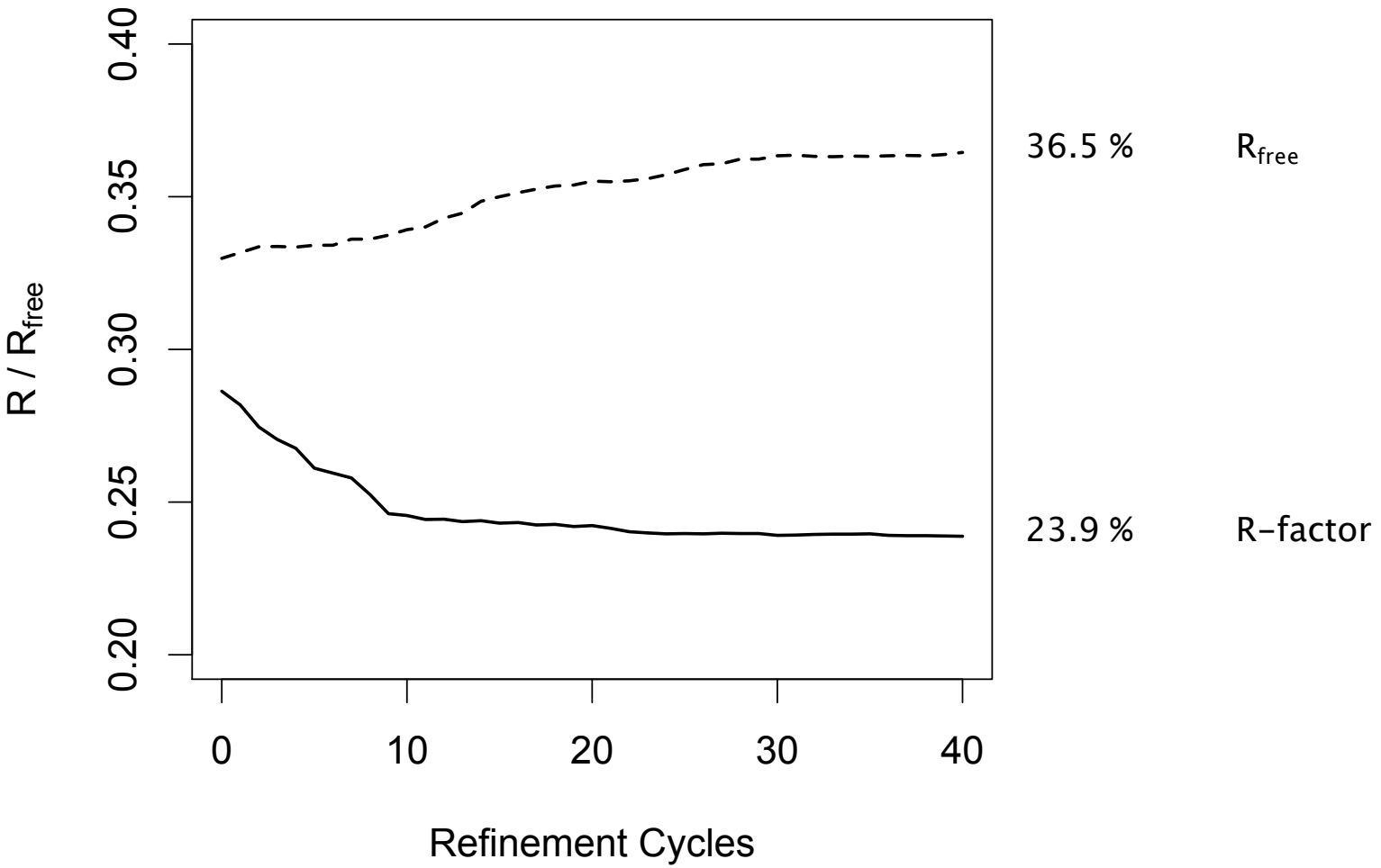


1ryx (3.5Å)
restrained to
2d3i (2.15Å)

Red: long
Grey: similar
Blue: short

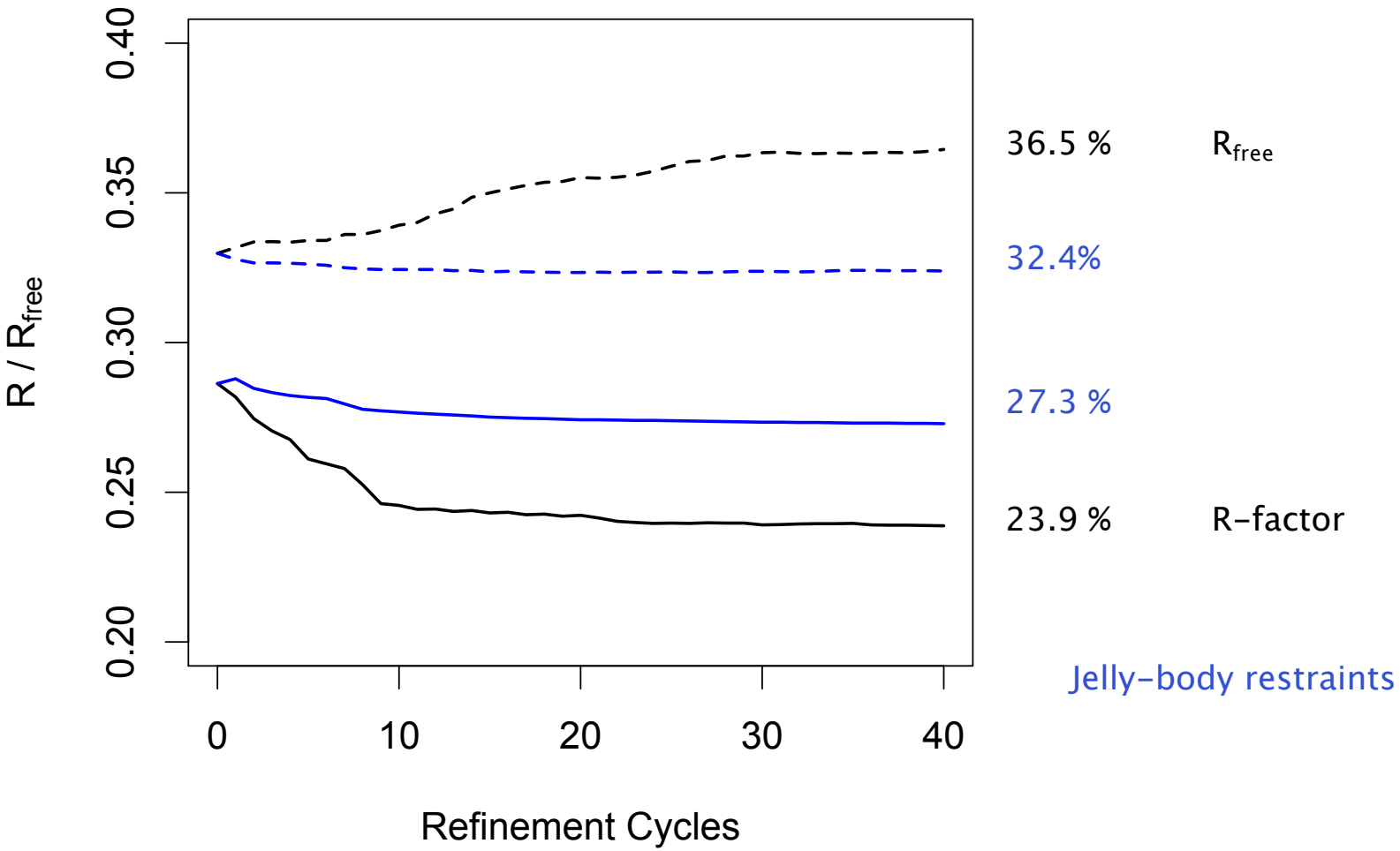
External Restraints

Ovotransferrin



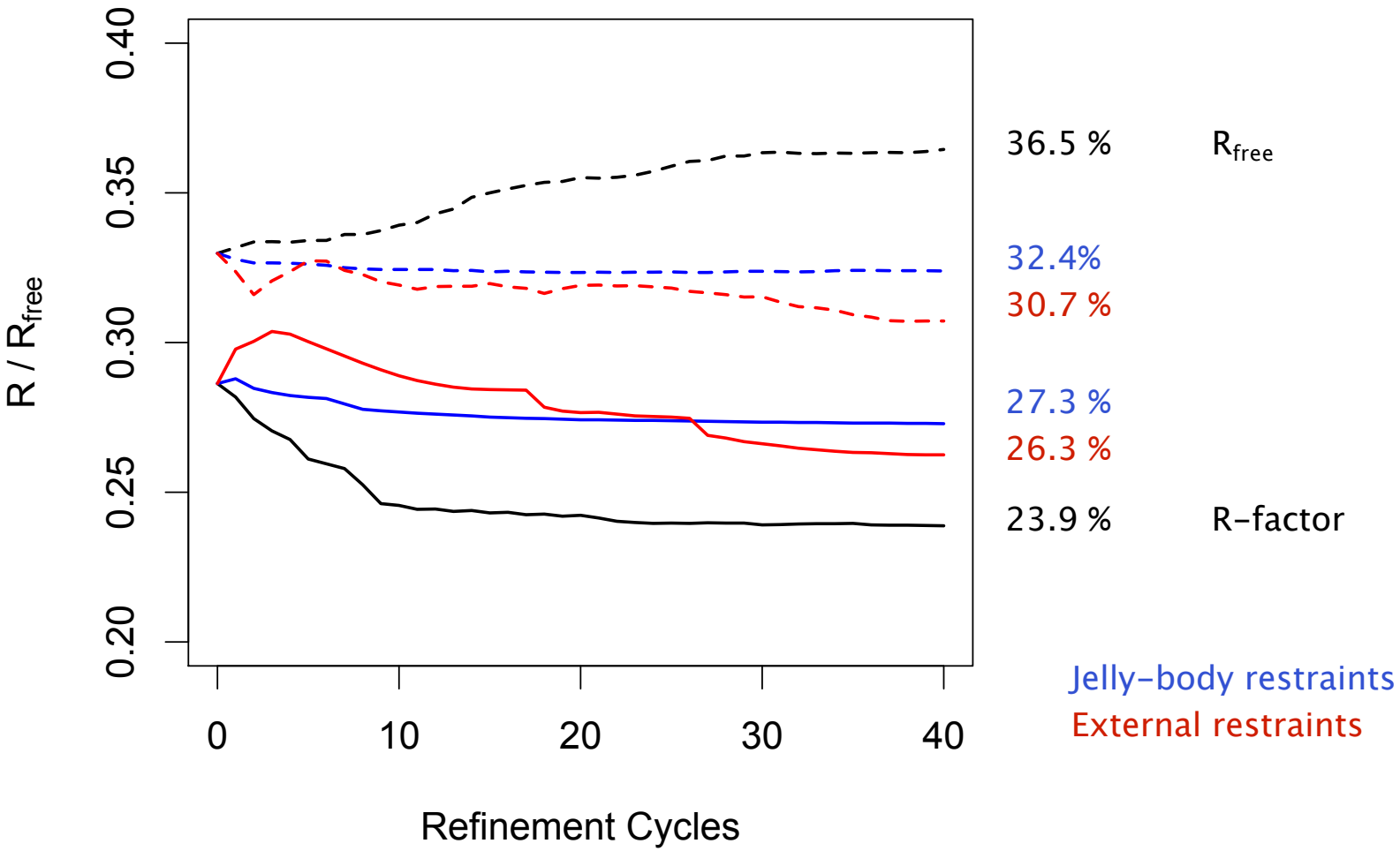
External Restraints

Ovotransferrin



External Restraints

Ovotransferrin

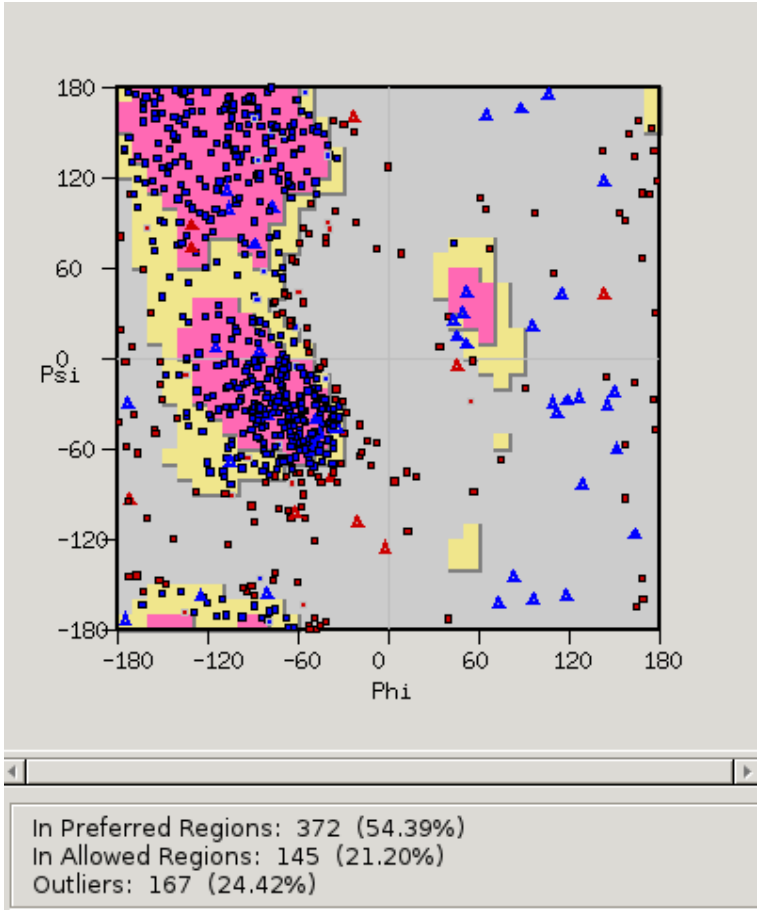


External Restraints

Ovotransferrin

Original Structure

R/R_{free} : 0.286/0.330

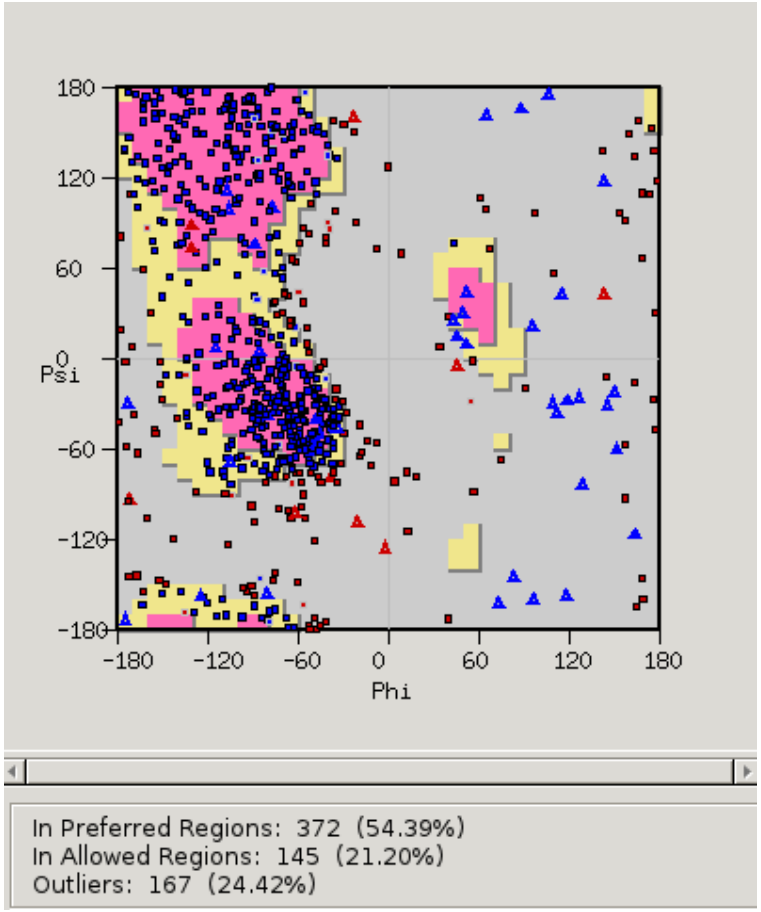


External Restraints

Ovotransferrin

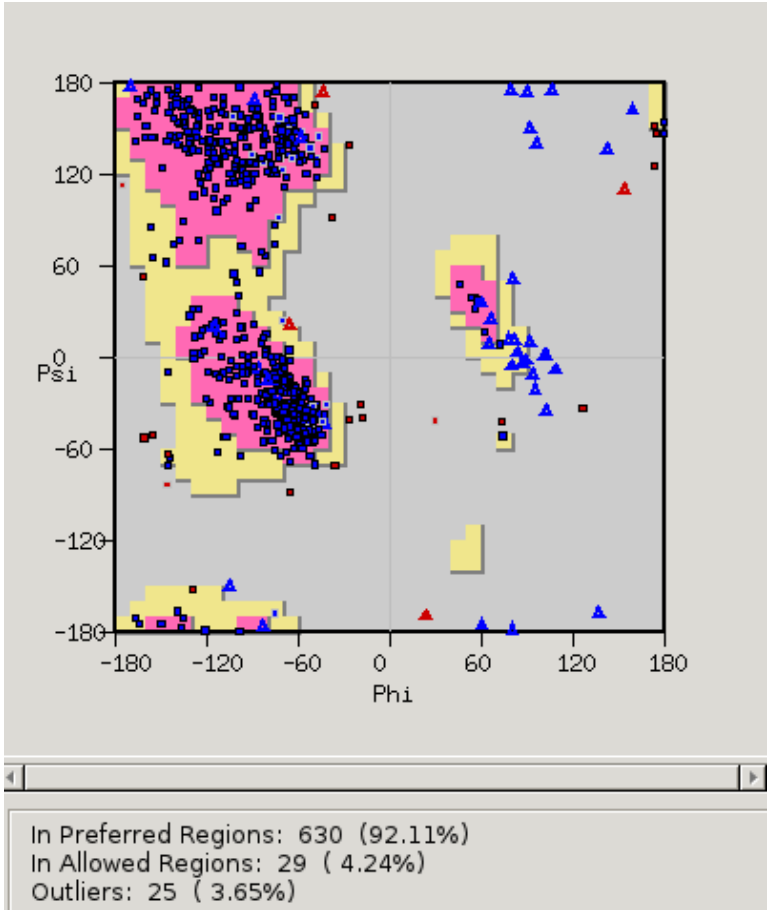
Original Structure

R/R_{free} : 0.286/0.330

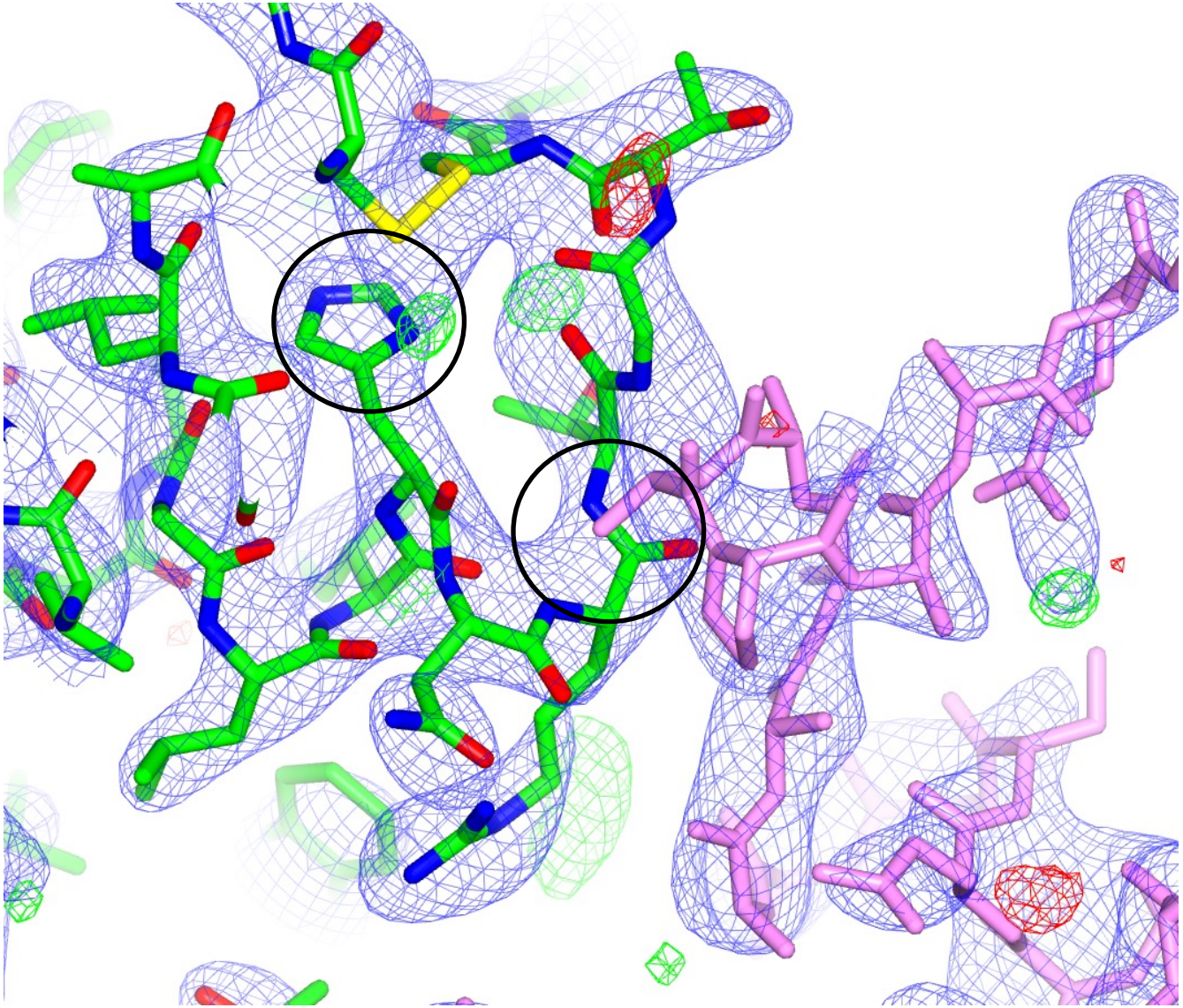


Re-refined with External Restraints

R/R_{free} : 0.263/0.307



External Restraints



1.3σ

Original Structure

R/R_{free} : 0.286/0.330

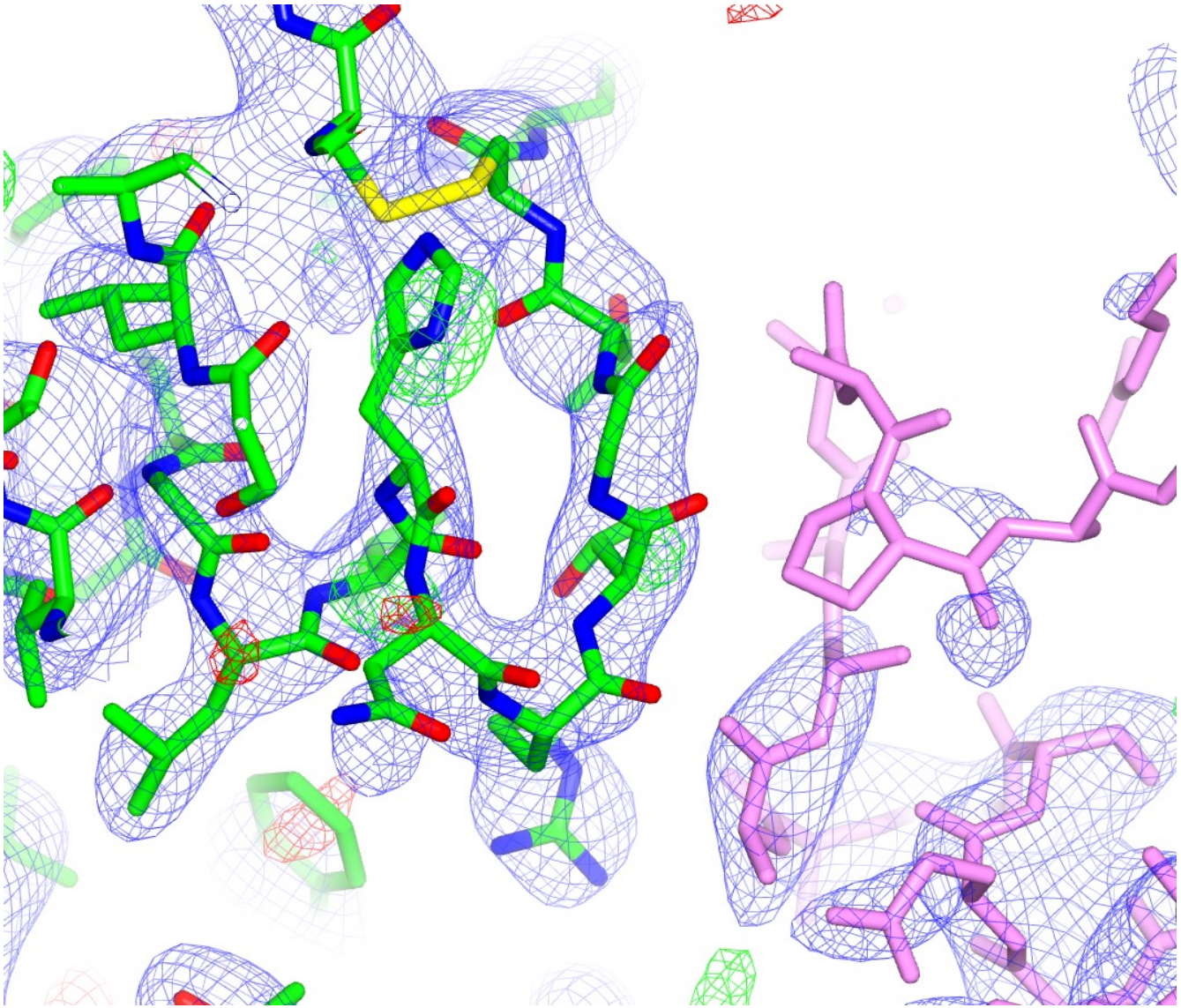
↓

External restraints

(40 cycles)

R/R_{free} : 0.263/0.307

External Restraints



1.3σ

Original Structure

R/R_{free} : 0.286/0.330

↓

External restraints

(40 cycles)

R/R_{free} : 0.263/0.307

↓

Modify

Real Space Refine

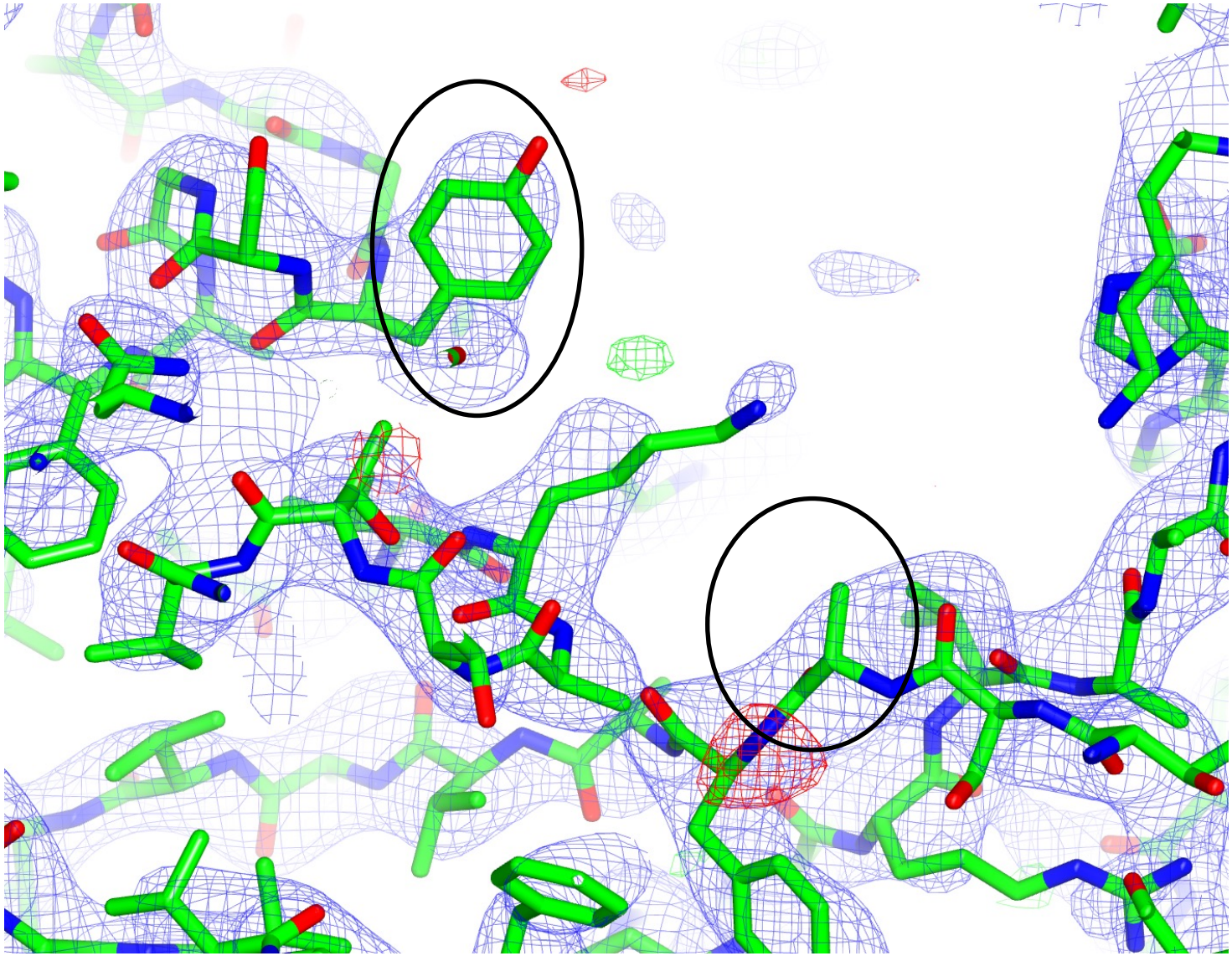
↓

Jelly body

(40 cycles)

R/R_{free} : 0.253/0.304

External Restraints



1.3σ

Original Structure

R/R_{free} : 0.286/0.330

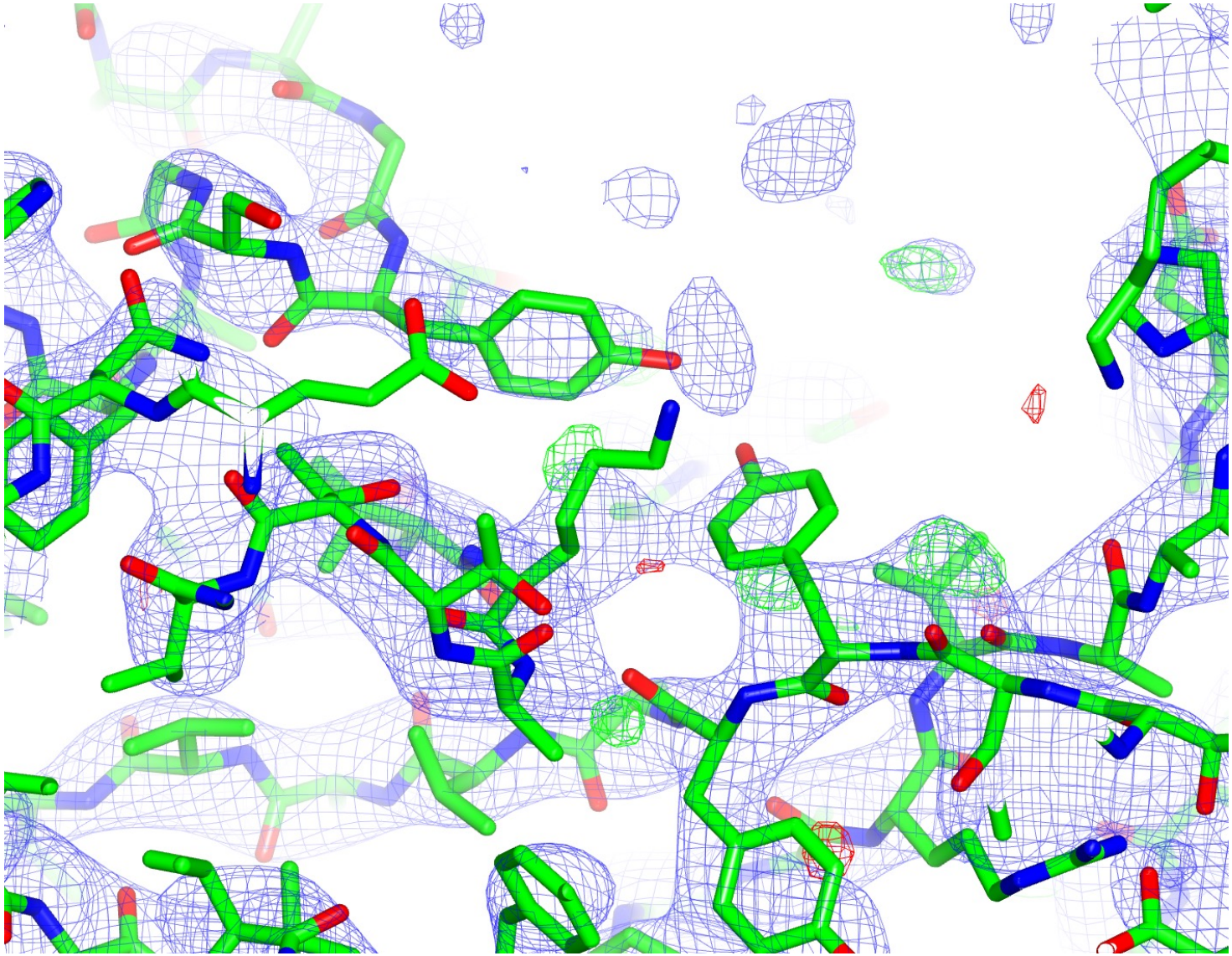
↓

External restraints

(40 cycles)

R/R_{free} : 0.263/0.307

External Restraints



1.3σ

Original Structure

R/R_{free} : 0.286/0.330

↓

External restraints

(40 cycles)

R/R_{free} : 0.263/0.307

↓

Build TYR92

Modify LYS209

↓

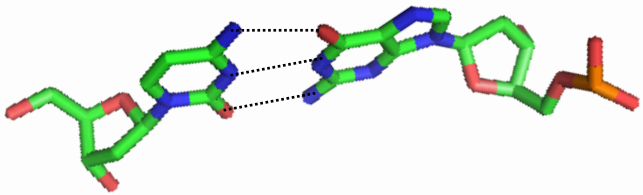
Jelly body

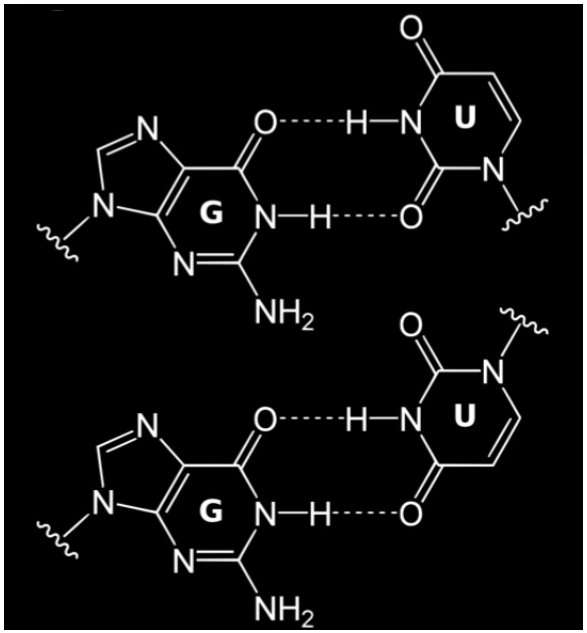
(40 cycles)

R/R_{free} : 0.252/0.307

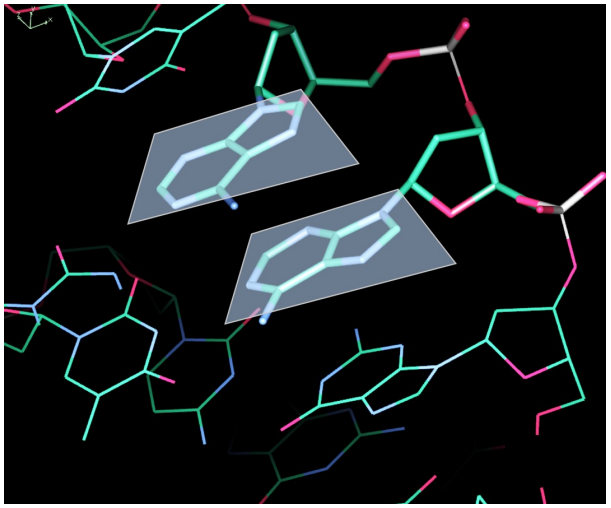
LibG Nucleic Acid Restraints

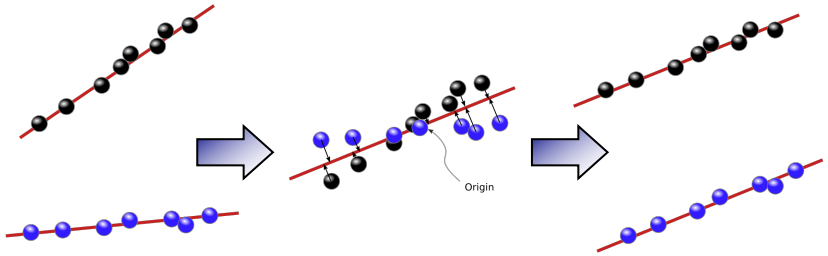
Base-pair restraints



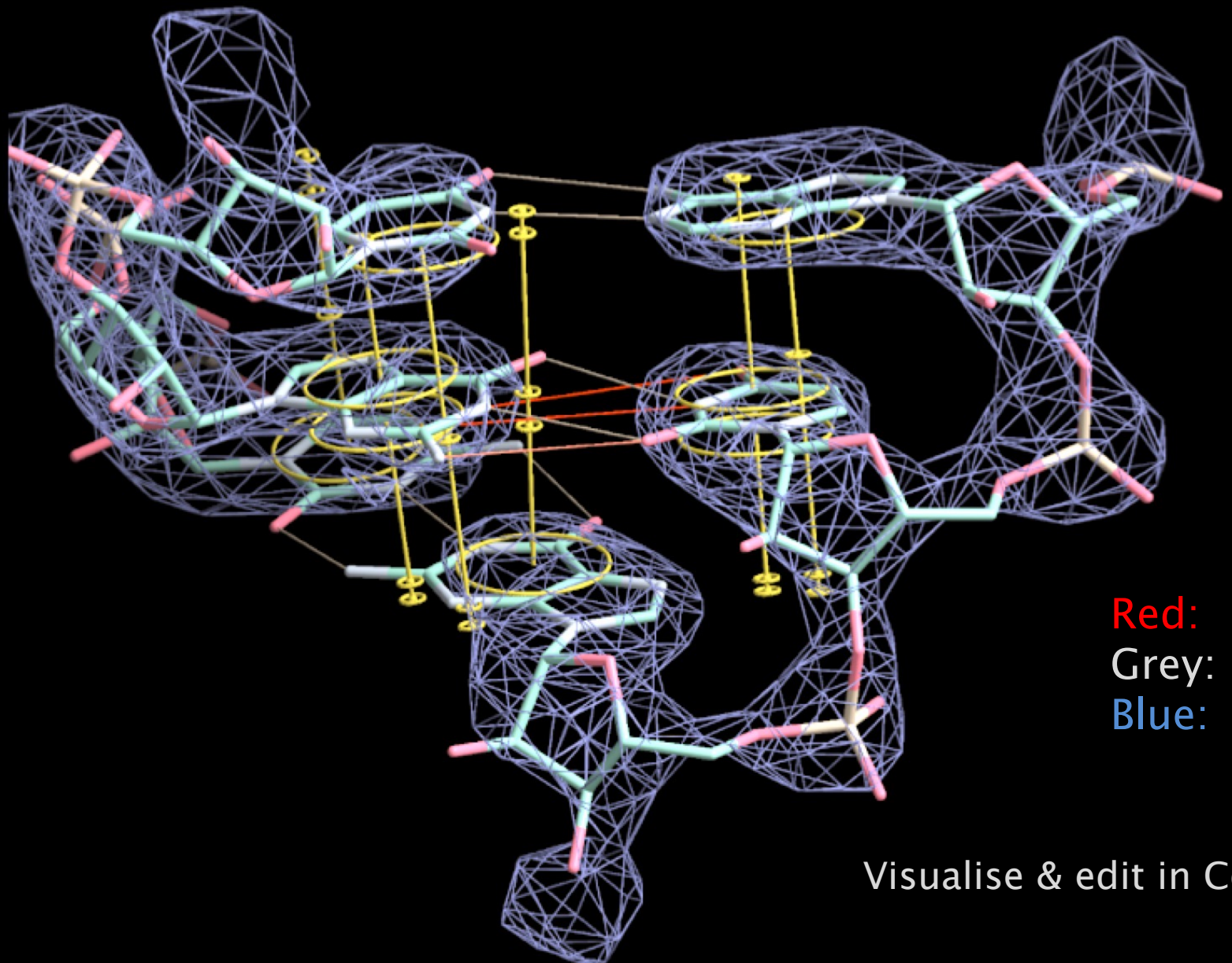


Parallel plane restraints





LibG Nucleic Acid Restraints



Red: long
Grey: similar
Blue: short

Visualise & edit in Coot

External Restraints

When refining at low resolution, check:

- Refinement statistics – ***Not always conclusive***
- Geometry – ***Not always conclusive***
- Electron density – ***Not always reliable***

Conclusion: **At low resolution, everything has to add up!**

Take care; reflect

Quality of prior information is important – **consider manual re-refinement**

– **PDB-REDO is useful**

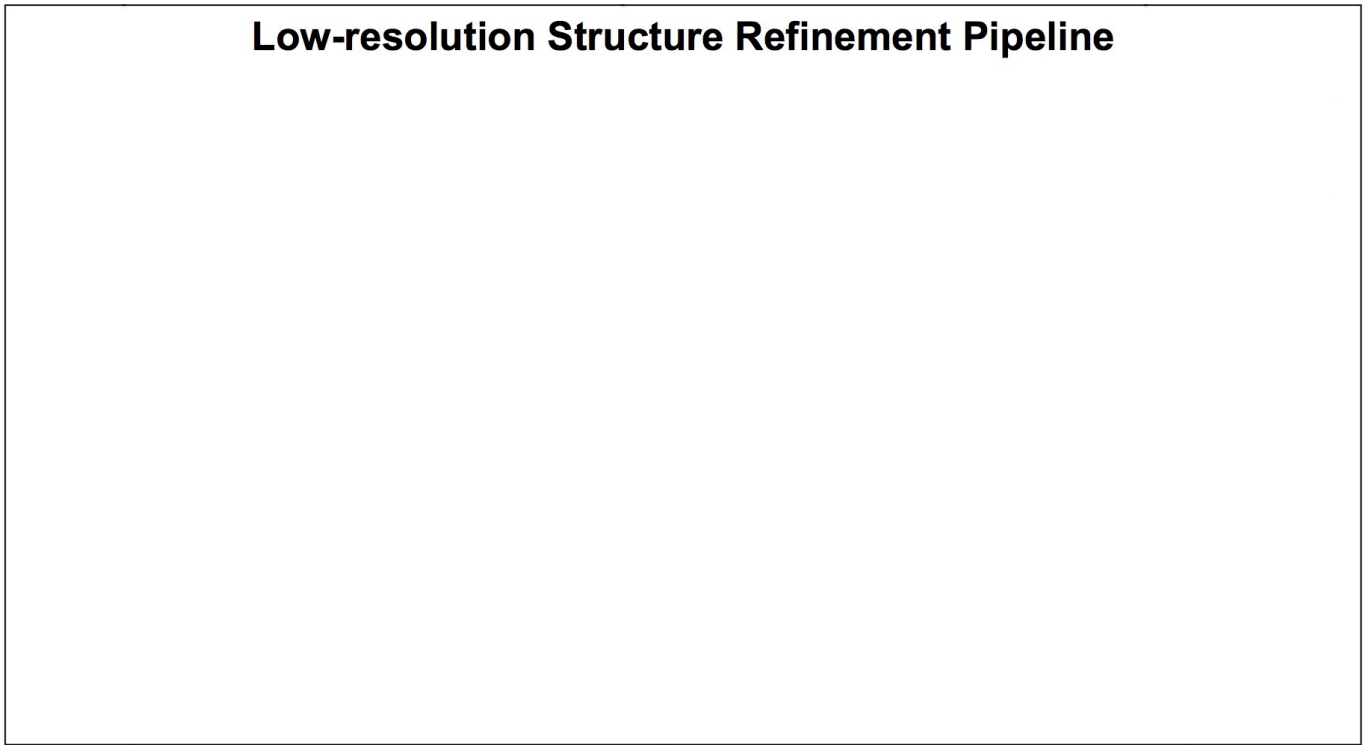
Automated pipeline - LORESTR

- 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

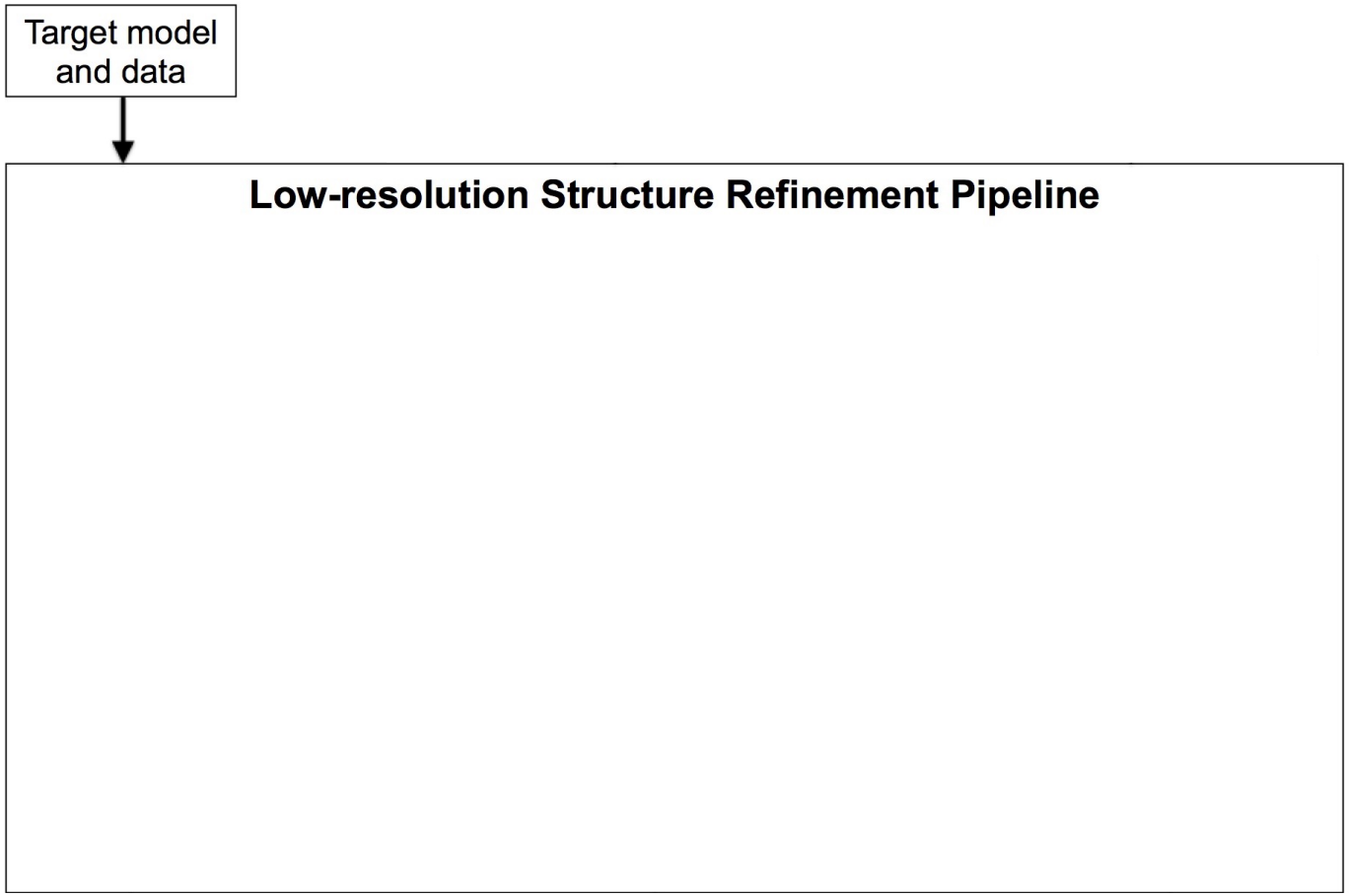
Solution:

LOw-**RE**solution **ST**ructure **R**efinement

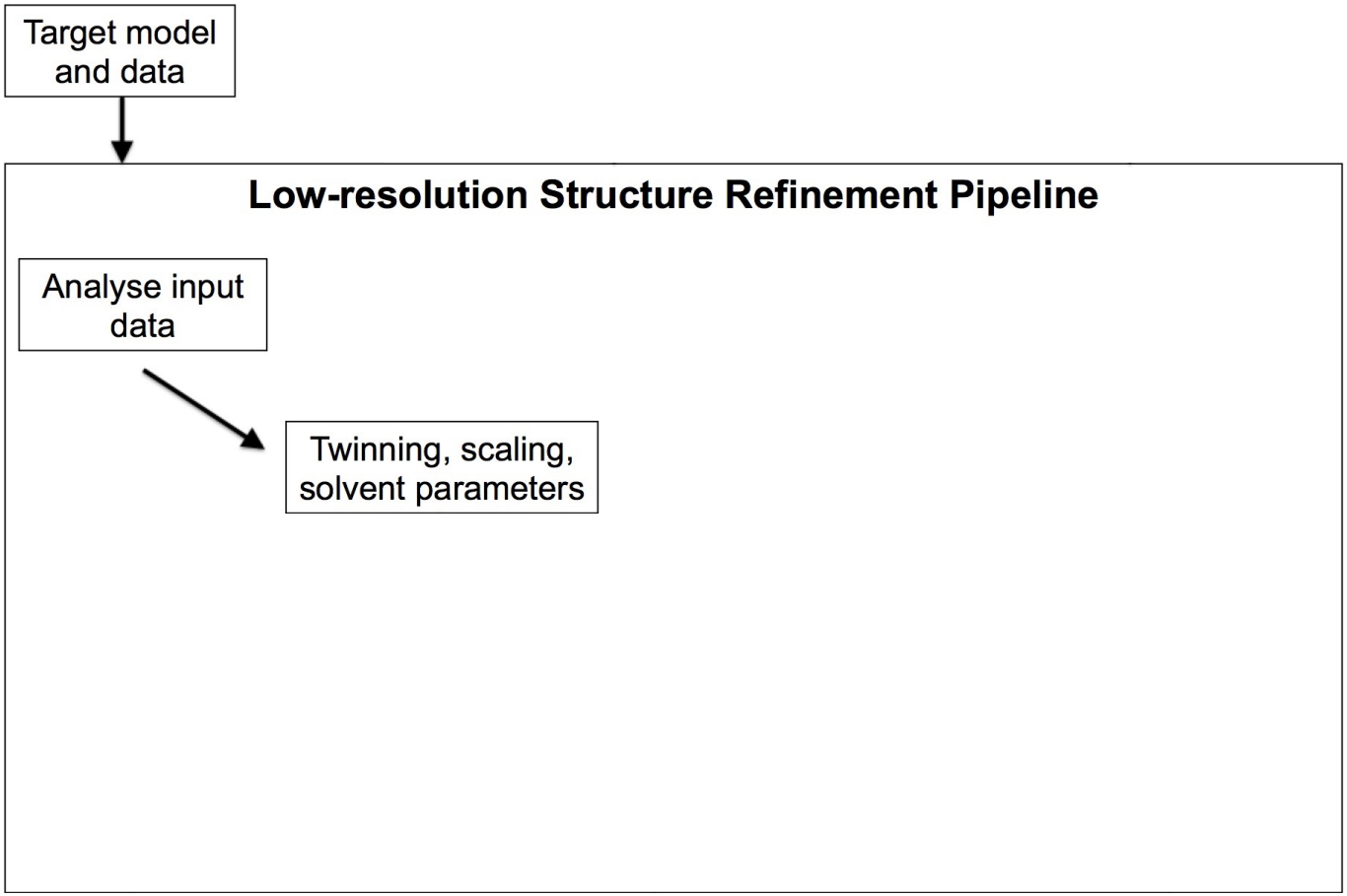
Automated pipeline - LORESTR



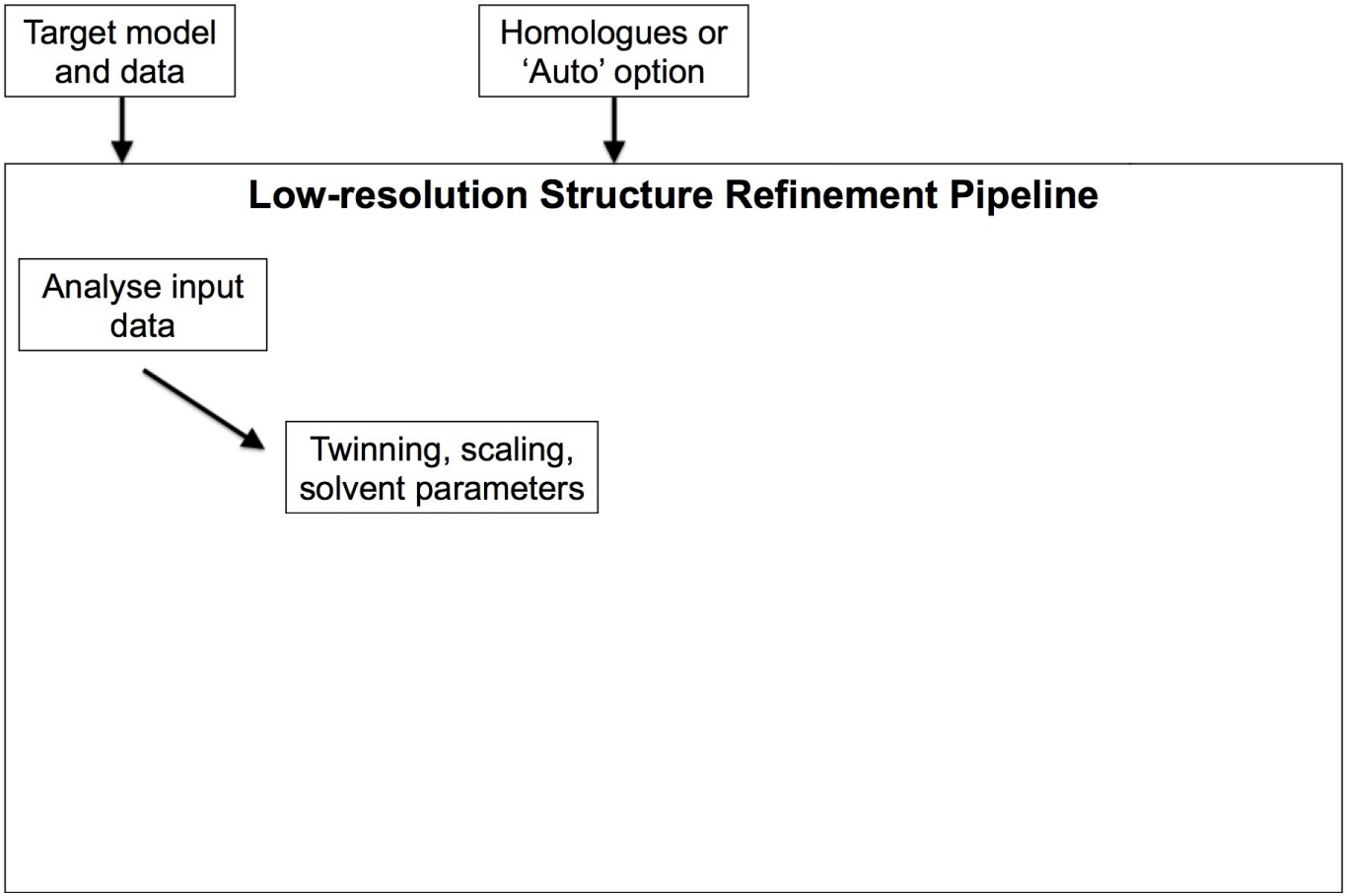
Automated pipeline - LORESTR



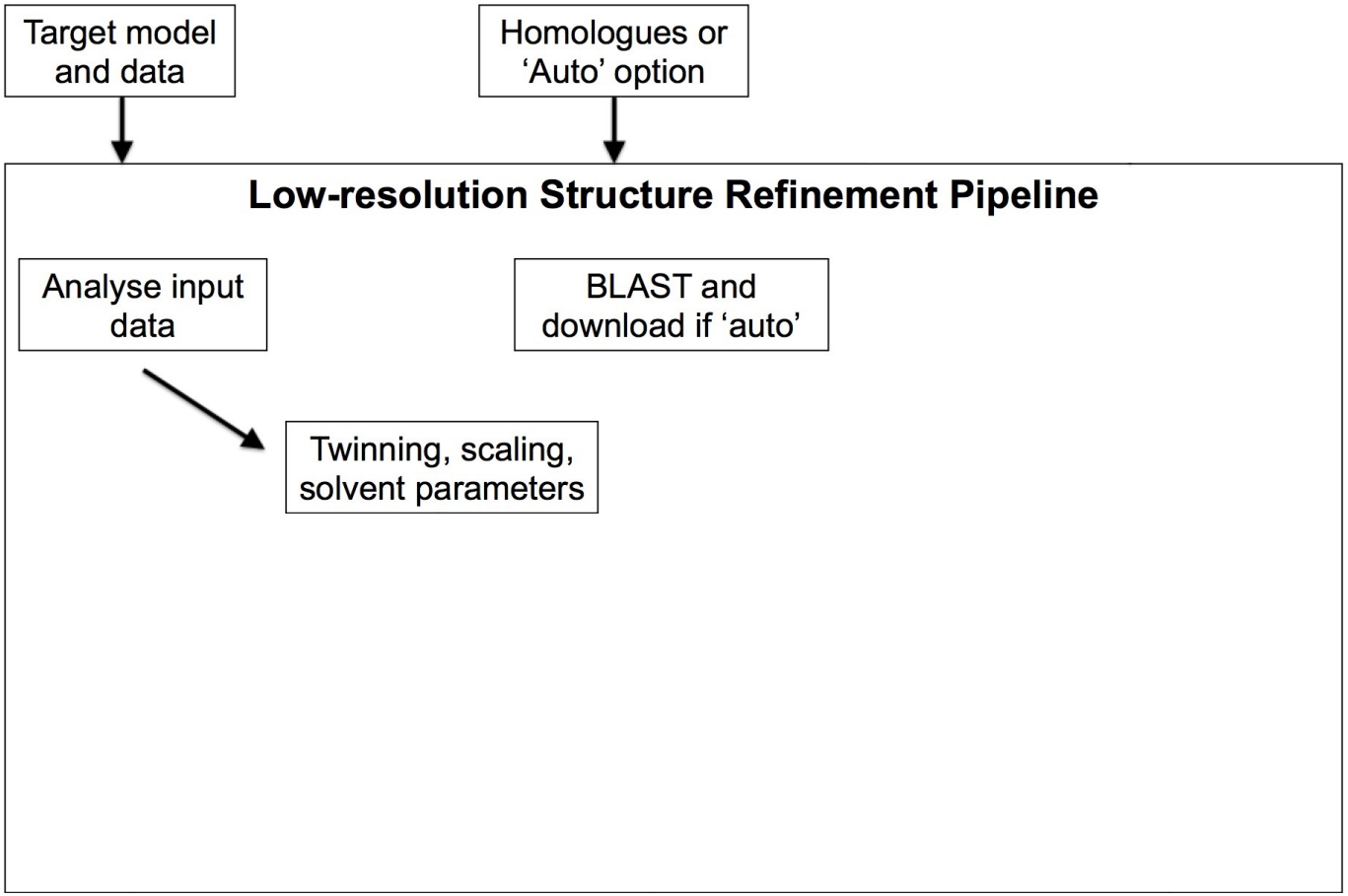
Automated pipeline - LORESTR



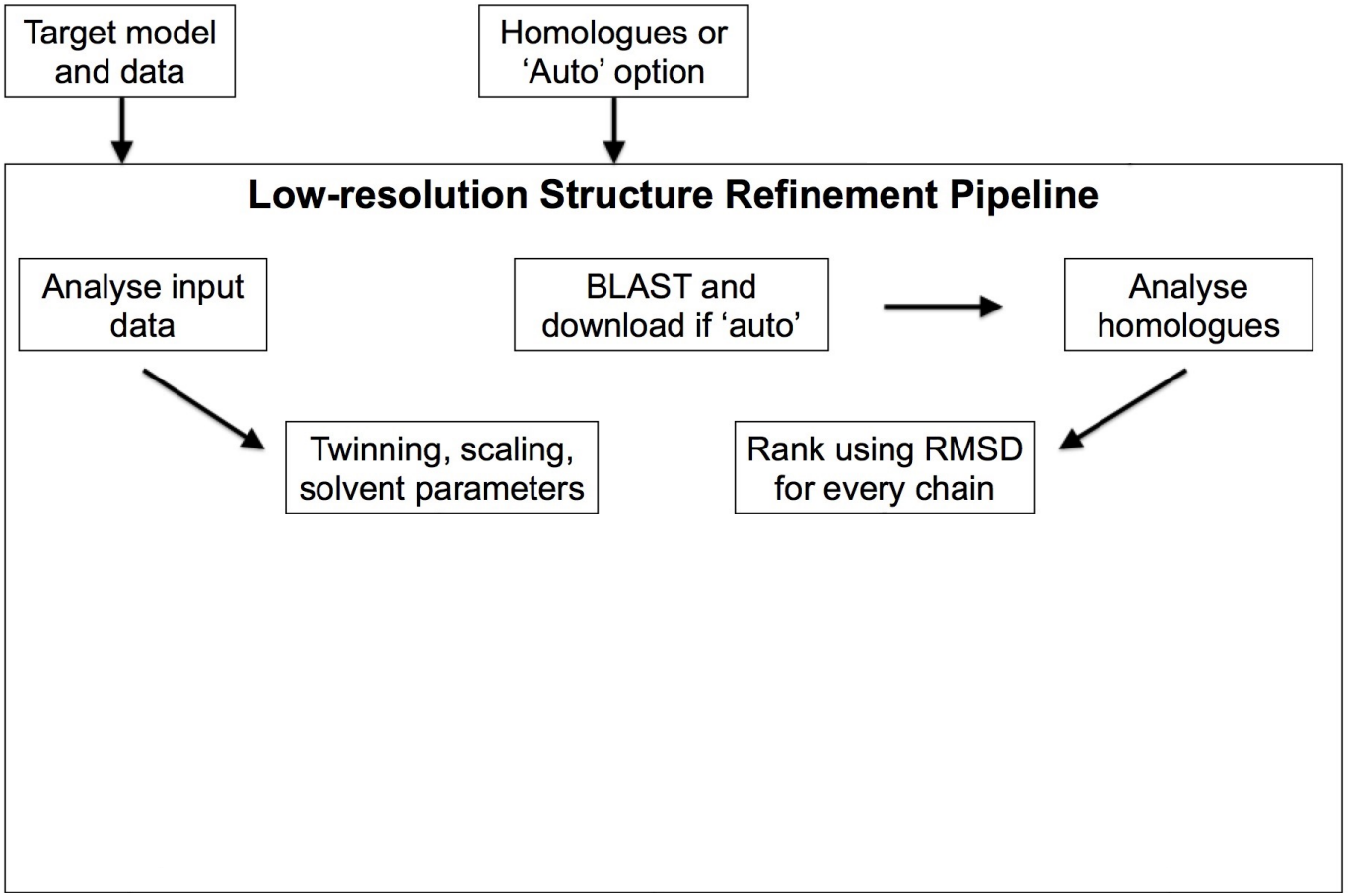
Automated pipeline - LORESTR



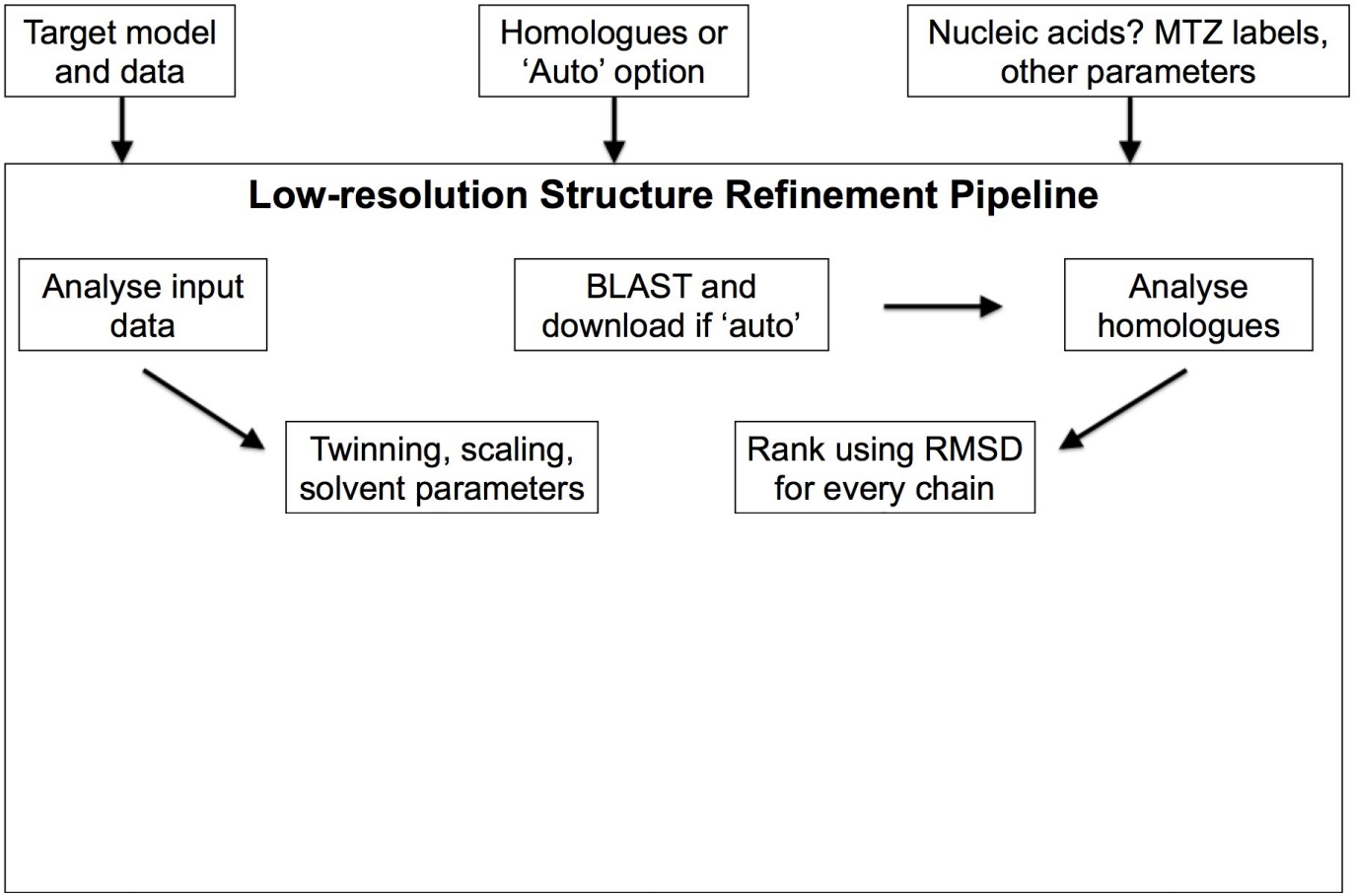
Automated pipeline - LORESTR



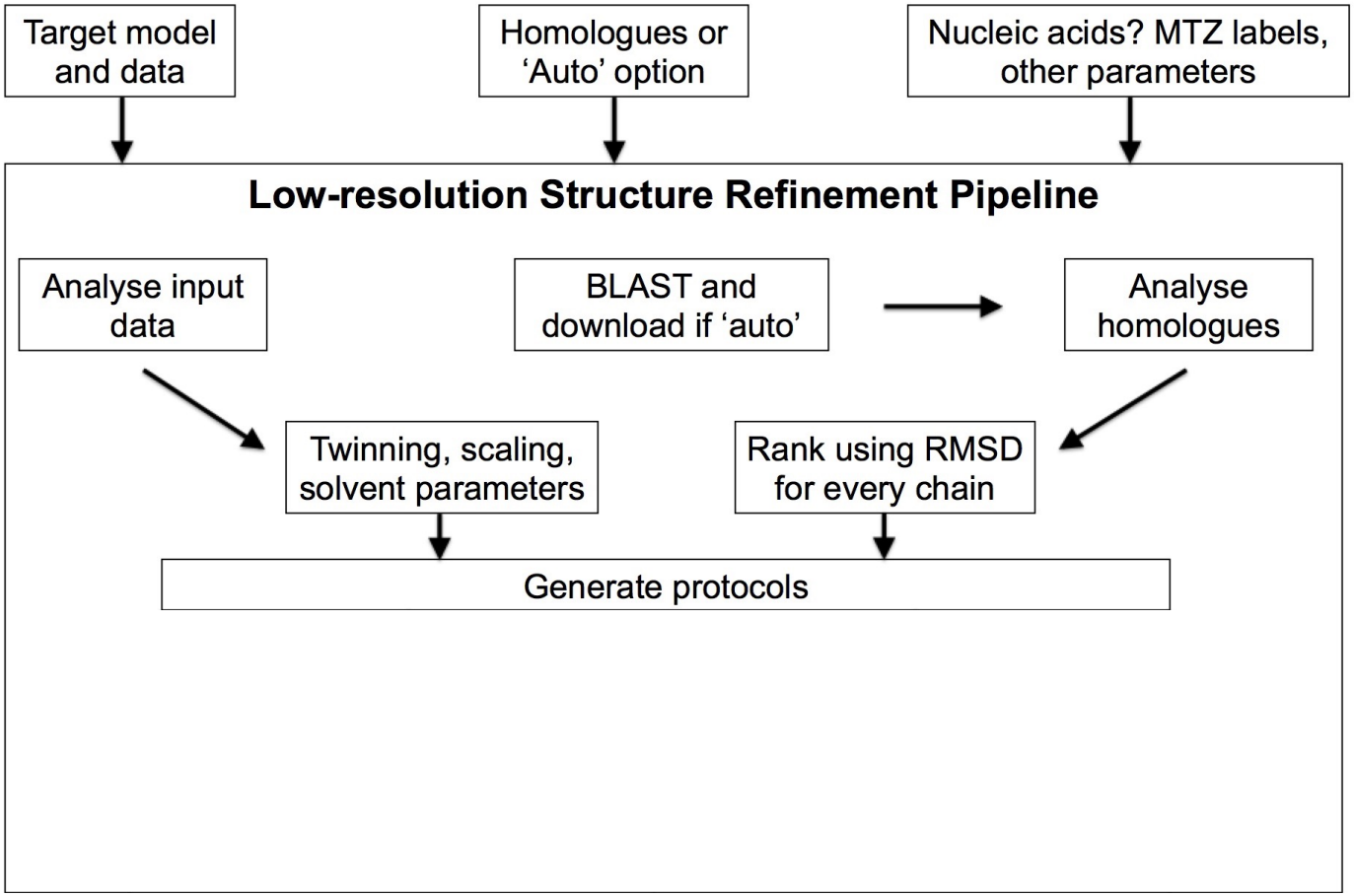
Automated pipeline - LORESTR



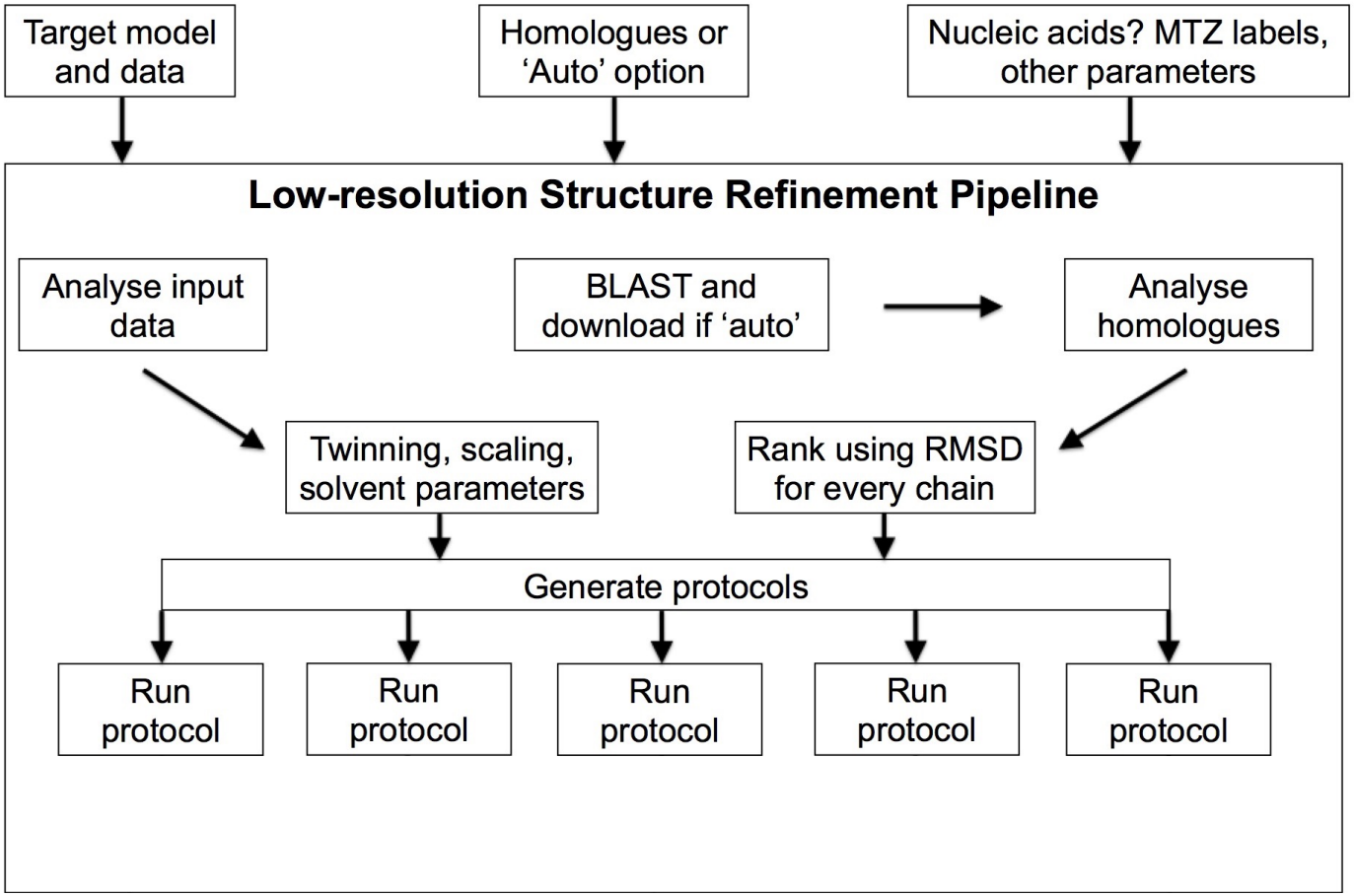
Automated pipeline - LORESTR



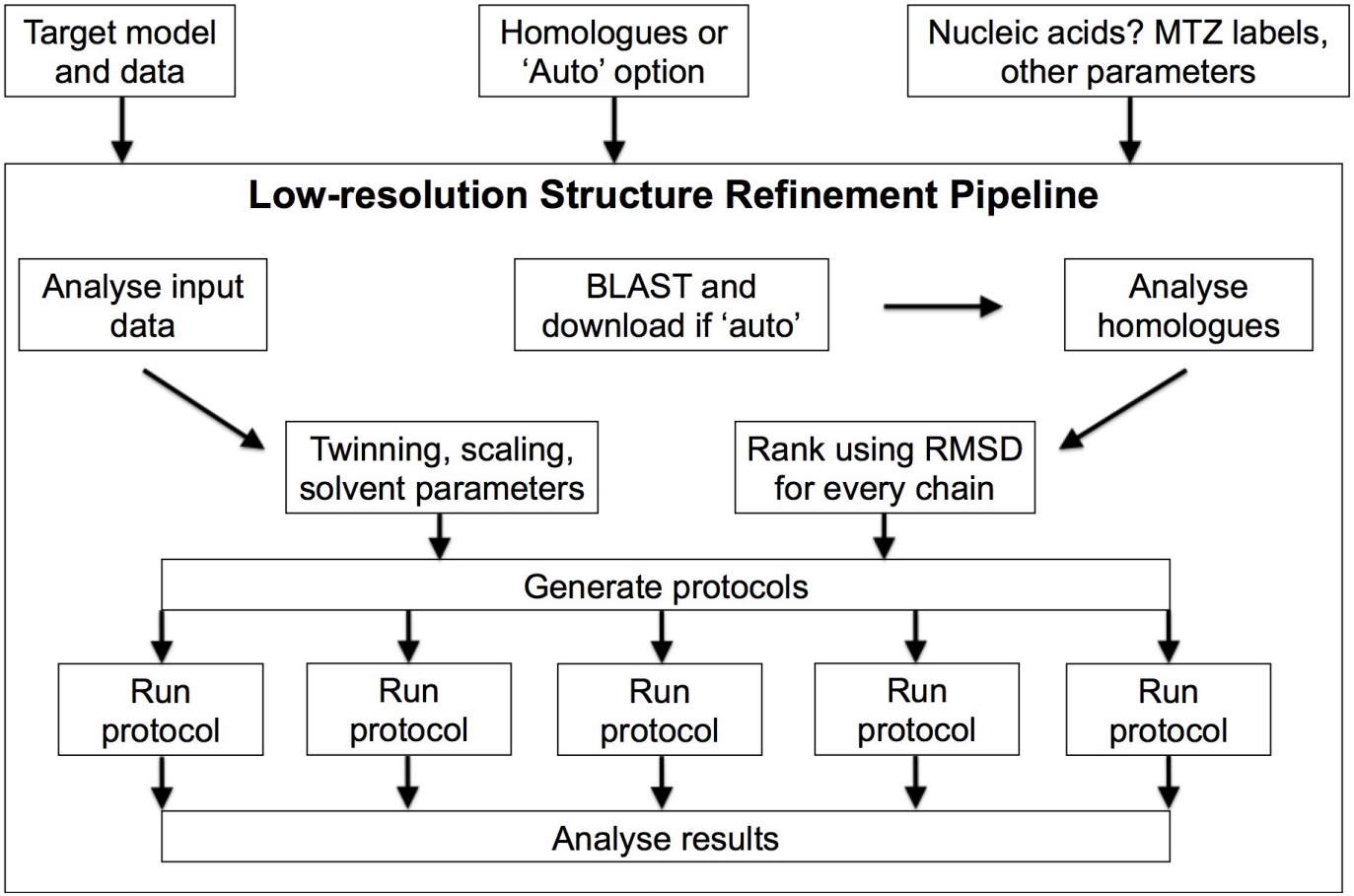
Automated pipeline - LORESTR



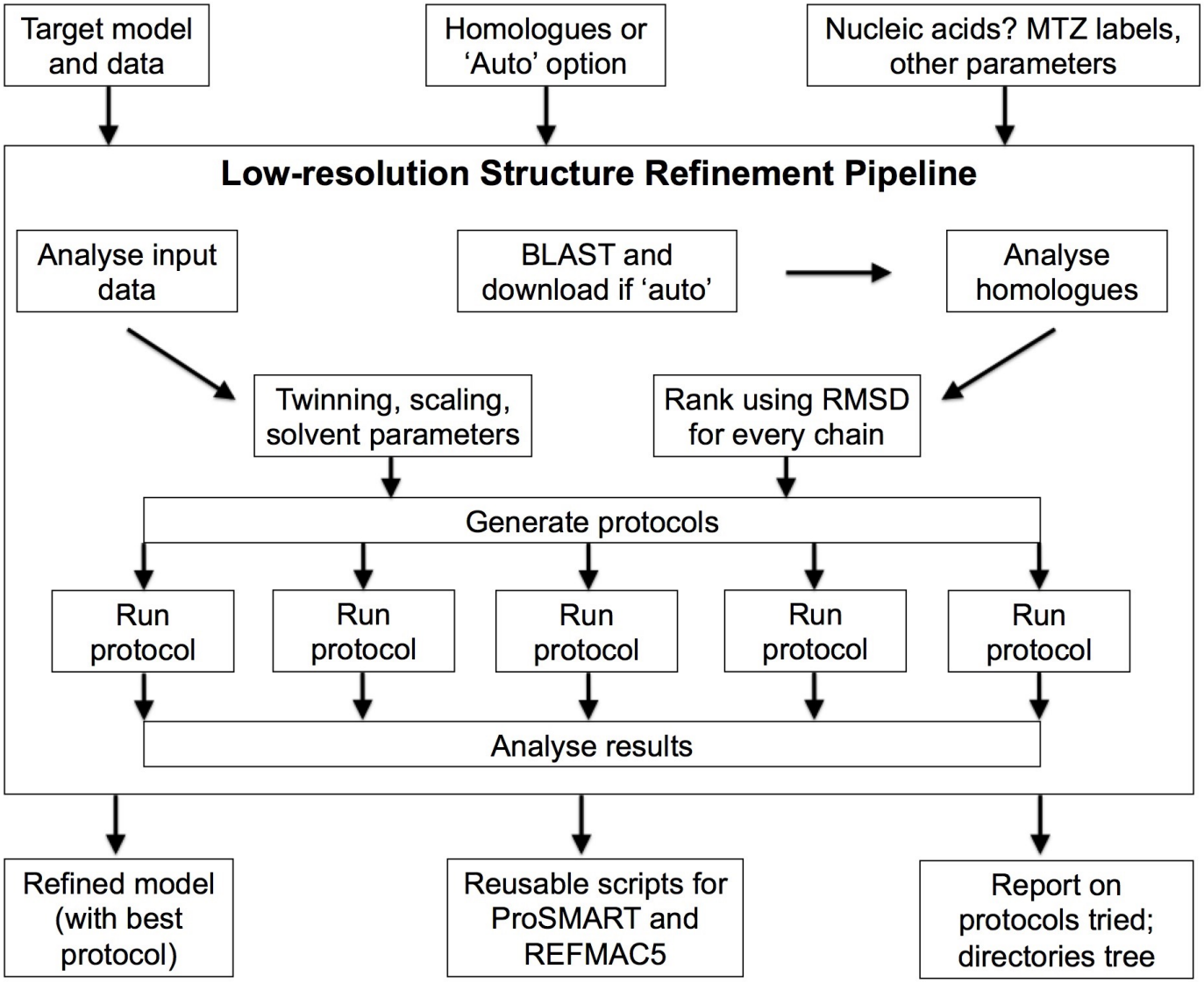
Automated pipeline - LORESTR



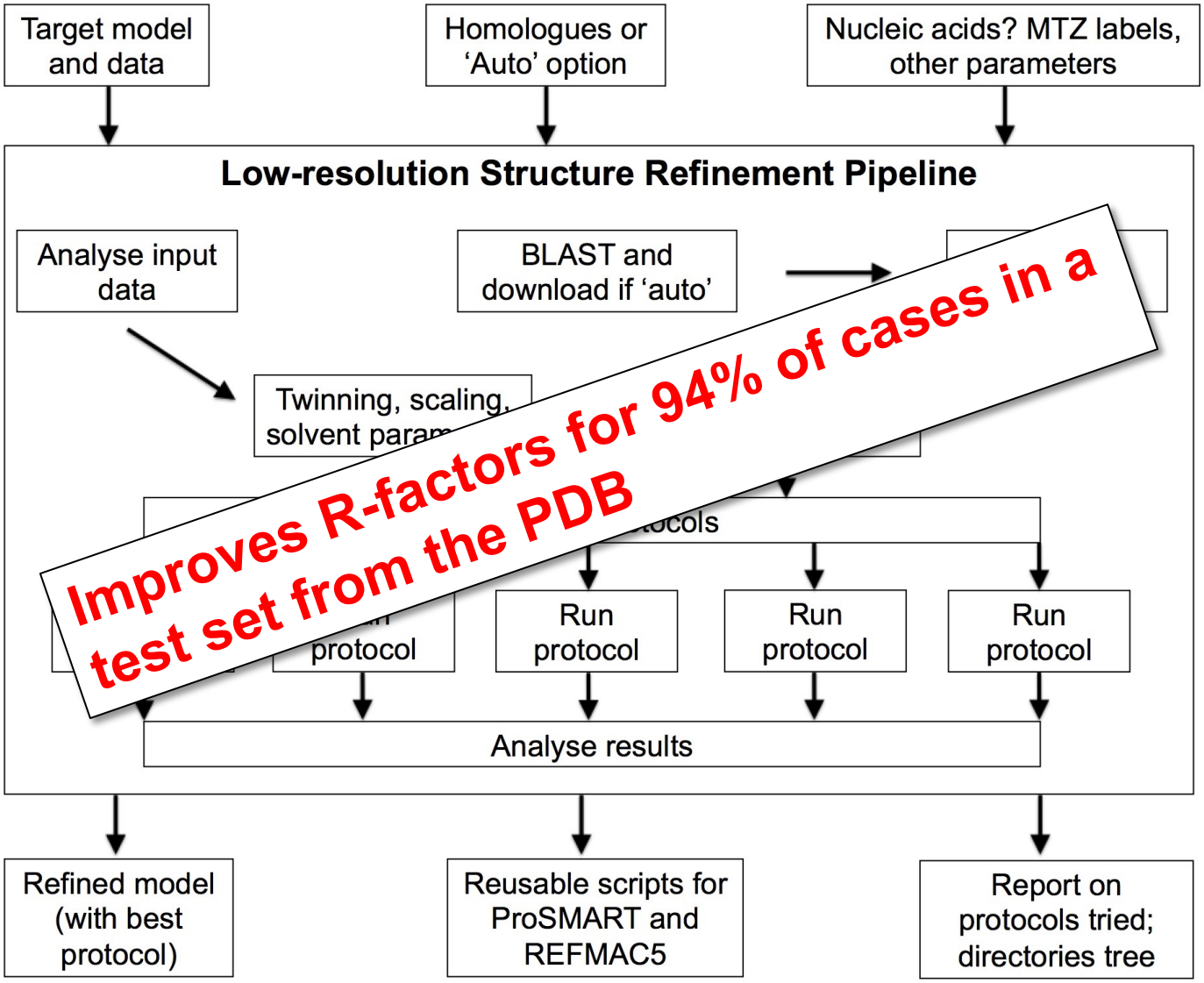
Automated pipeline - LORESTR



Automated pipeline - LORESTR



Automated pipeline - LORESTR



Summary

Tools to help with model building and refinement:

REFMAC5: Refinement, jelly body restraints, map sharpening/blurring

ProSMART: External restraints, comparative analysis

LibG: Nucleic acid restraints

LORESTR: Automated low-resolution pipeline

AceDRG: Ligand dictionary and conformer generation

Coot: Visualisation & manipulation of restraints, map blurring
…also morphing, jiggle-fit, backrub rotamers…

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

What and When

Early stages (e.g. straight after MR)

- Rigid body refinement
- 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 and DNA/RNA restraints – no homologue available
- Secondary structure conformation restraints – model building tool
- Add hydrogens (?)

Medium-final stages

- TLS – at medium resolutions
- Anisotropic B-factors – only at high resolution
- Twin refinement – only if you are sure

Final stages of refinement

Jelly body – around 20 cycles

Relevant Publications

Low-resolution refinement with REFMAC5, ProSMART, LibG & LORESTR:

Nicholls *et al.* (2017)) Low Resolution Refinement of Atomic Models Against Crystallographic Data. *Protein Crystallography*, 565-93.

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.

Tools for cryo-EM model fitting & refinement:

Nicholls *et al.* (2018) Current approaches for the fitting and refinement of atomic models into cryo-EM maps using CCP-EM. *Acta Cryst.* D74, 492-505.

Murshudov (2016) Refinement of atomic structures against cryo-EM maps. *Methods in Enzymology*, 277-305.

Brown *et al.* (2015) Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta Cryst.* D71, 136-53.

Tools for ligand fitting & validation:

Nicholls (2017) Ligand fitting with CCP4. *Acta Cryst.* D73, 158-170.

Emsley (2017) Tools for ligand validation in Coot. *Acta Cryst.* D73, 203-10.

Debreczeni & Emsley (2012) Handling ligands with Coot. *Acta Cryst.* D68, 425-30.

Cooperative utilisation of information from Xtal and NMR:

Kovalevskiy et al. (2018) Overview of refinement procedures within REFMAC5: Utilising Data from Different Sources. *Acta Cryst*. D74, 215-27.

Carlon et al. (2016) How to tackle protein structural data from solution and solid state: An integrated approach. *Progress in nuclear magnetic resonance spectroscopy*. 92, 54-70.

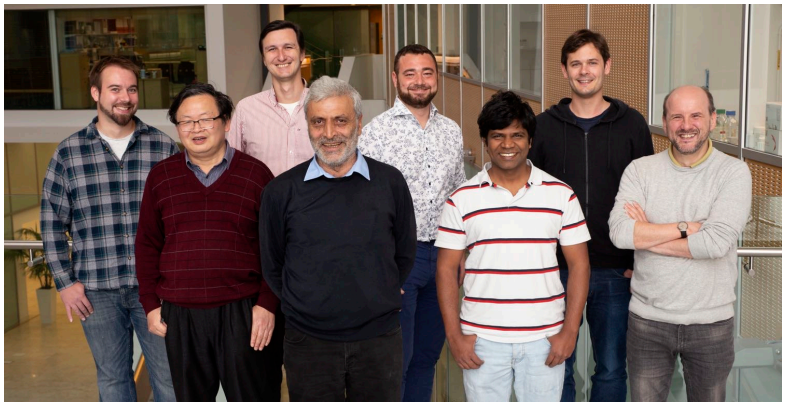
Effect of Twinning on R-factors:

Murshudov GN (2011) Some properties of Crystallographic Reliability Index – Rfactor: Effect of Twinning. *Appl. & Comp. Math*., 10, 250-61.

Acknowledgements

Contact: **lucrezia@mrc-lmb.cam.ac.uk**
www2.mrc-lmb.cam.ac.uk/groups/murshudov/

MRC-LMB Computational Structural Biology Group



Rob Nicholls
Fei Long
Oleg Kovalevskiy
Garib Murshudov
Rangana Warshamanage
Paul Emsley
Keitaro Yamahita

CCP4 Core
Eugene Krissinel
Andrey Lebedev
Charles Ballard
Ronan Keegan
Ville Uski

Global Phasing
Marcin Wojdyr

CCP4i2
Martin Noble
Stuart McNicholas
Jon Agirre
Liz Potterton

CCP-EM
Martyn Wynn
Tom Burnley
Colin Palmer

Collaborators
Marcus Fischer
Robbie Joosten
Andrea Thorn
Roberto Steiner
Alan Brown
Jude Short
Ana Casañal
Rafiga Masmaliyeva
Azzurra Carlon

Computing
Jake Grimmett
Toby Darling

All colleagues from
MRC-LMB, CCP4, CCP-EM
Users for feedback!

UKRI MRC Laboratory of Molecular Biology
CCP4
BBSRC bioscience for the future
Science & Technology Facilities Council