

# Buccaneer

The buccaneer software for automated model building of protein structures across a broad range of resolutions.

Kevin Cowtan  
YSBL, University of York  
[cowtan@ysbl.york.ac.uk](mailto:cowtan@ysbl.york.ac.uk)

# Buccaneer

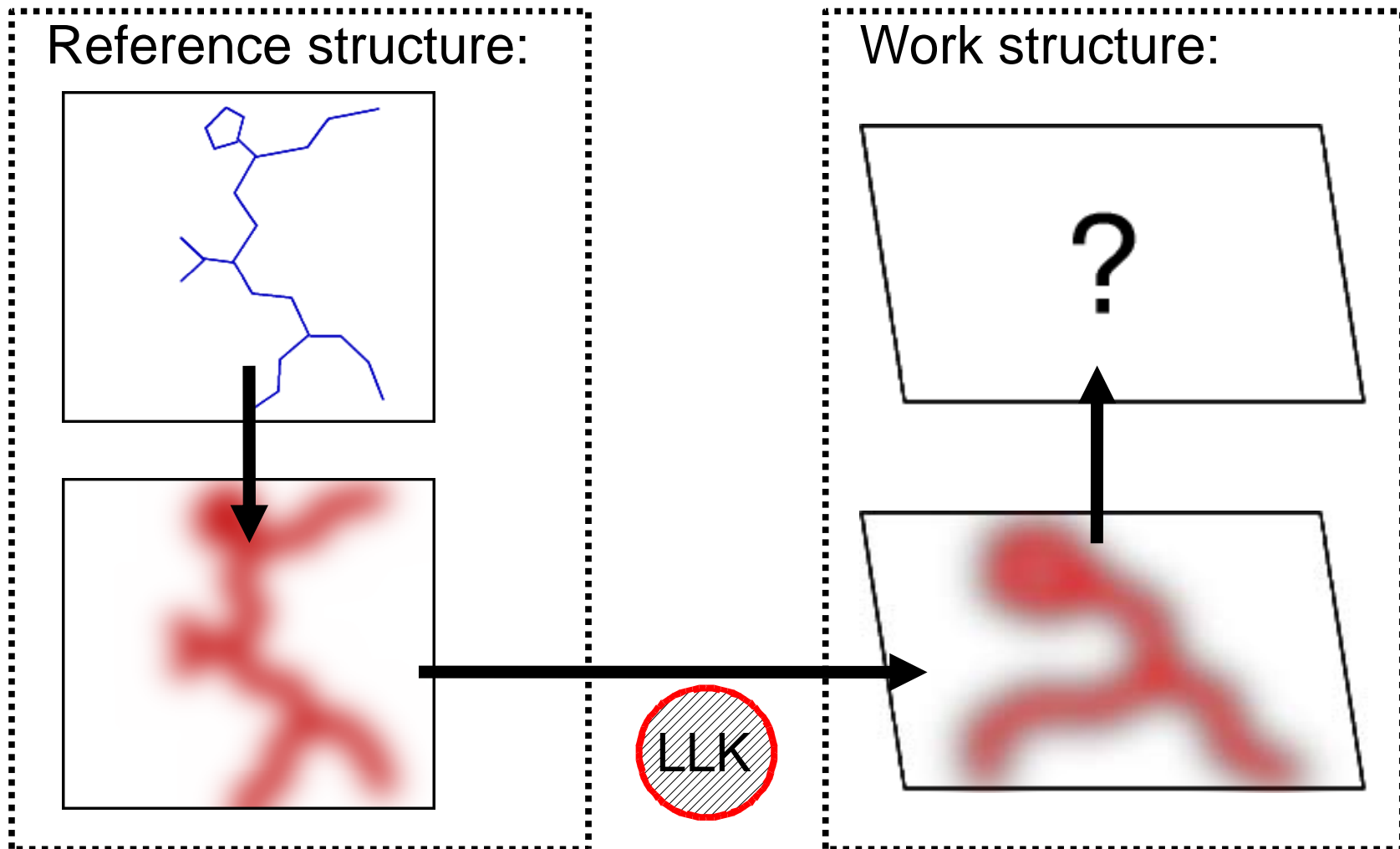
Statistical model building software based on the use of a reference structure to construct likelihood targets for protein features.

New since 2007: Version 1.0 for CCP4 v6.1

- Buccaneer-Refmac pipeline
- NCS auto-completion
- Improved sequencing

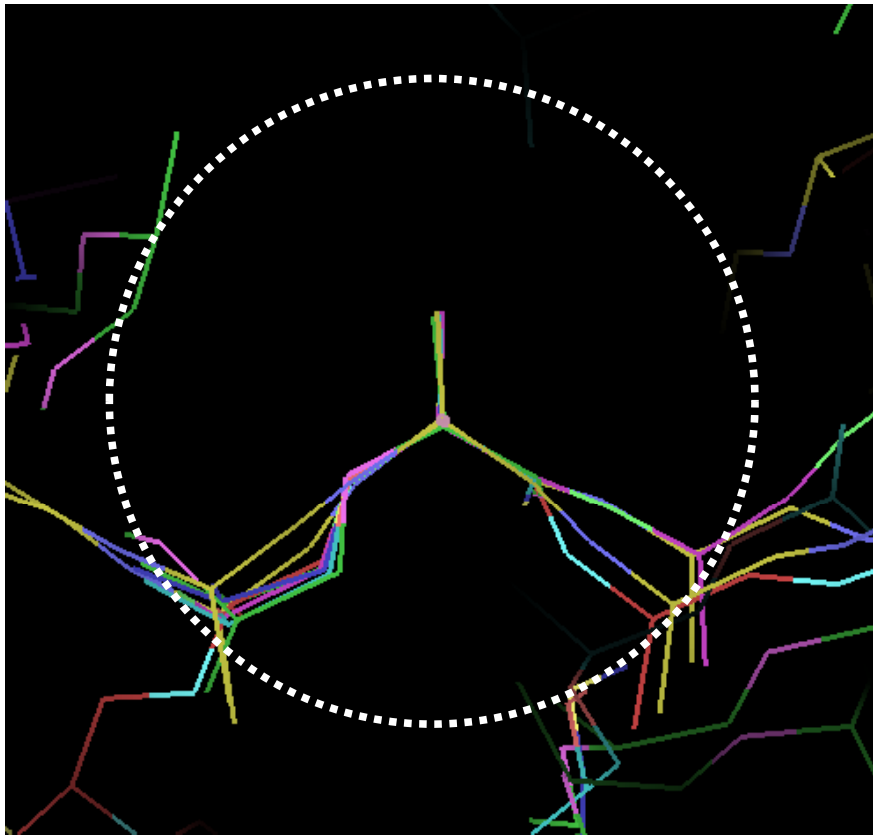
# Buccaneer: Method

- Compare simulated map and known model to obtain likelihood target, then search for this target in the unknown map.

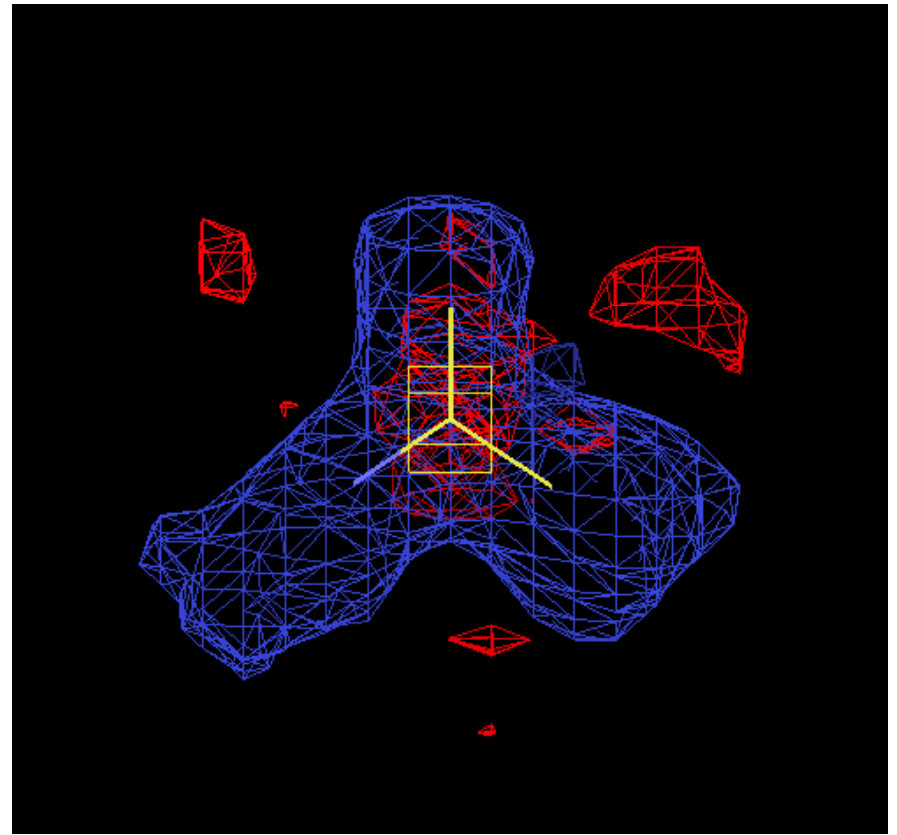


# Buccaneer: Method

- Compile statistics for reference map in 4Å sphere about  $C\alpha$  => LLK target.



- Use mean/variance.



4Å sphere about Ca also used by 'CAPRA' loeger et al. (but different target function).

# Buccaneer

10 stages:

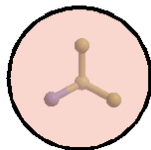
- **Find** candidate C-alpha positions
- **Grow** them into chain fragments
- **Join** and merge the fragments, resolving branches
- **Link** nearby N and C termini (if possible)
- **Sequence** the chains (i.e. dock sequence)
- **Correct** insertions/deletions
- **Filter** based on poor density
- **NCS Rebuild** to complete NCS copies of chains
- **Prune** any remaining clashing chains
- **Rebuild** side chains

# Buccaneer

Use a likelihood function based on conserved density features.

The same likelihood function is used several times. This makes the program very simple (<3000 lines), and the whole calculation works over a range of resolutions.

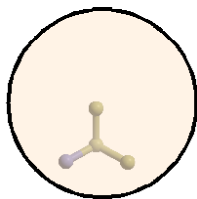
**Finding, growing:** Look for C-alpha environment



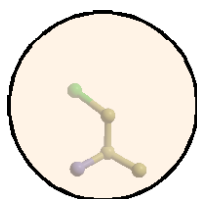
(4.0Å sphere about Cα)

**Sequencing:**

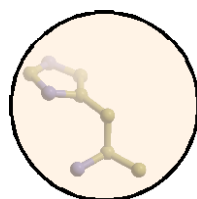
Look for C-beta environment



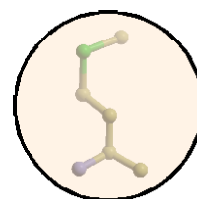
ALA



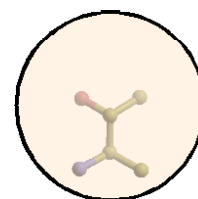
CYS



HIS



MET



THR

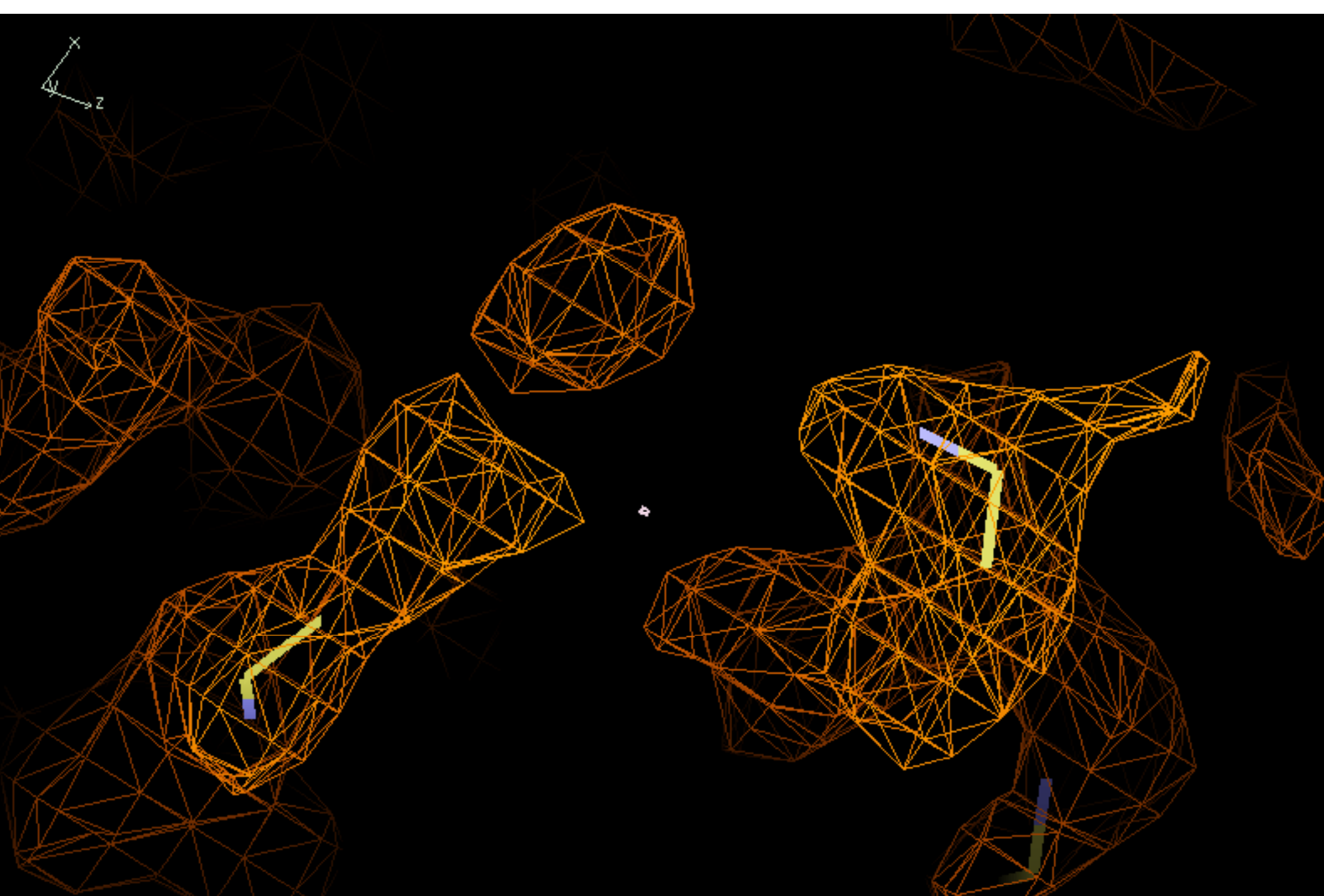
(5.5Å sphere about Cβ)

... x20

# Buccaneer

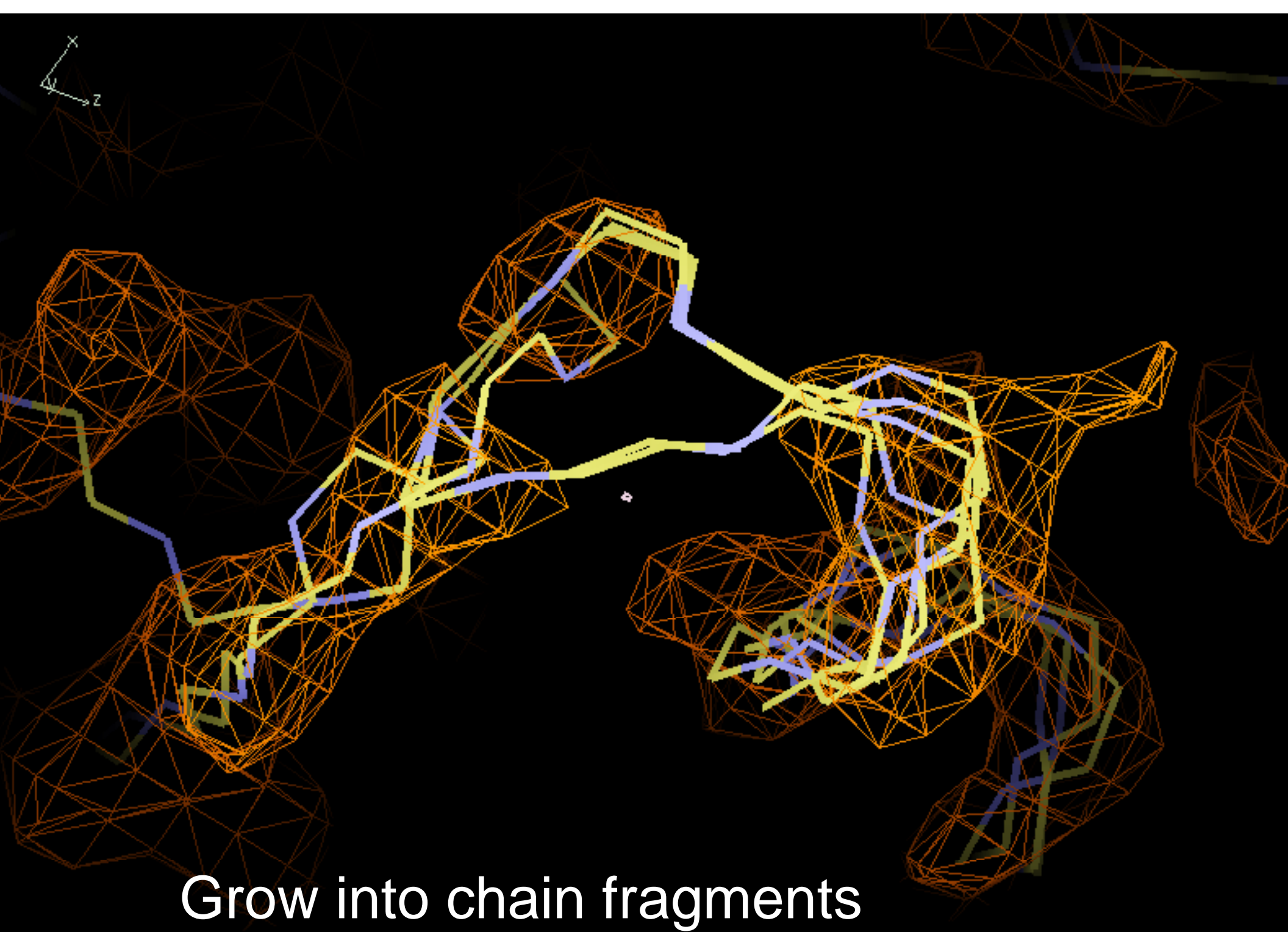
## **Case Study:**

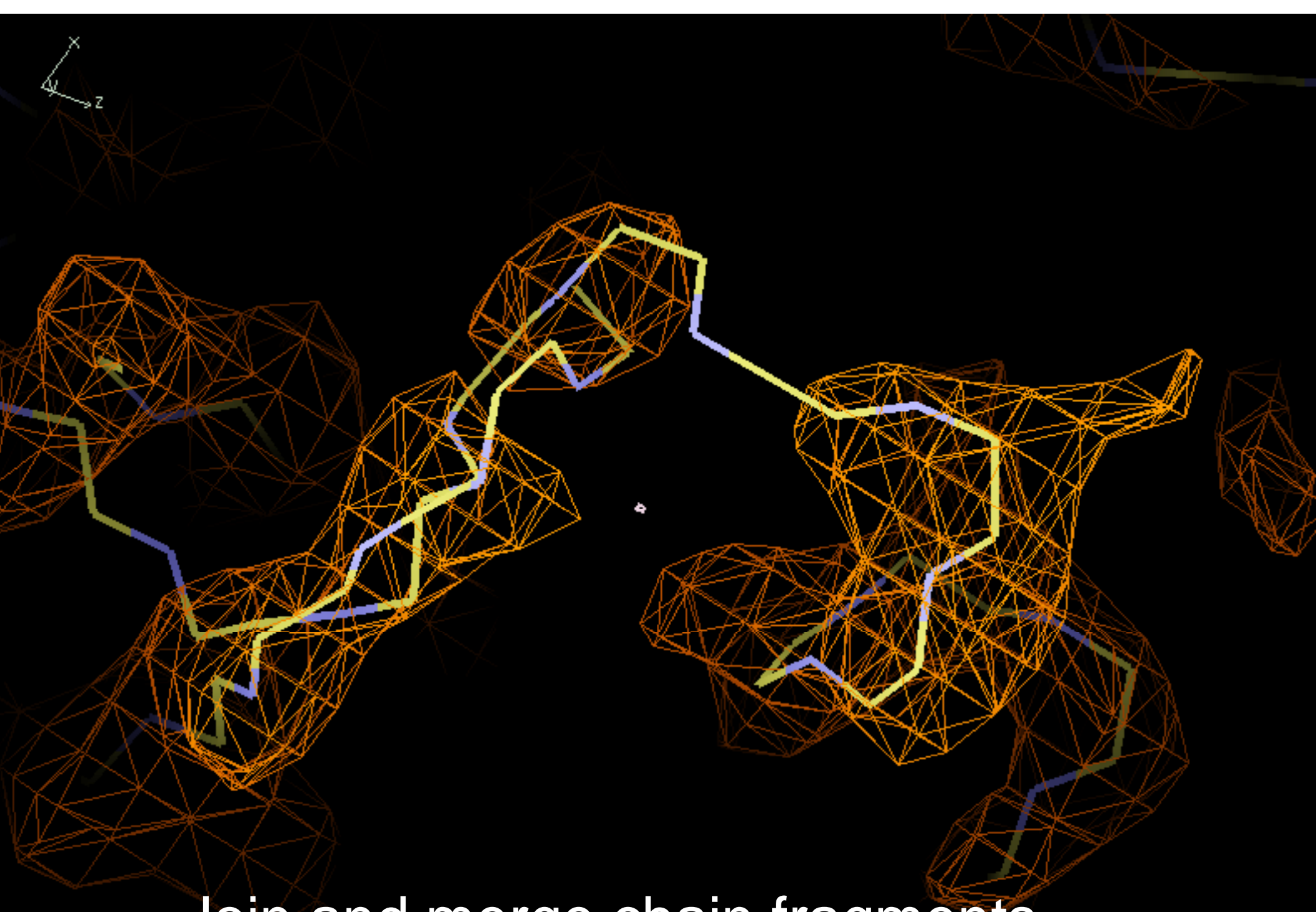
A difficult loop in a 2.9Å map, calculated using real data from the JCSG.



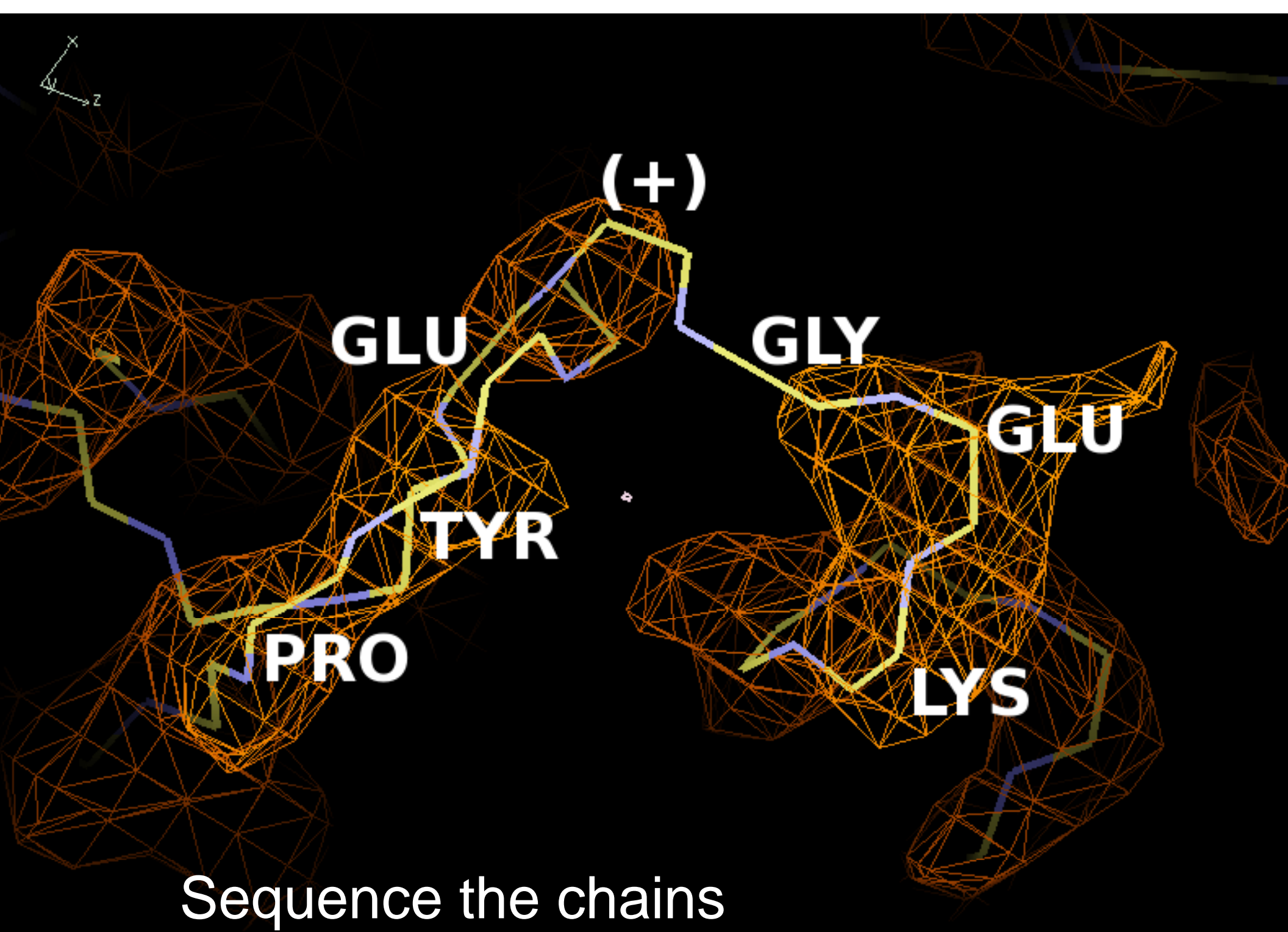
Find candidate C-alpha positions



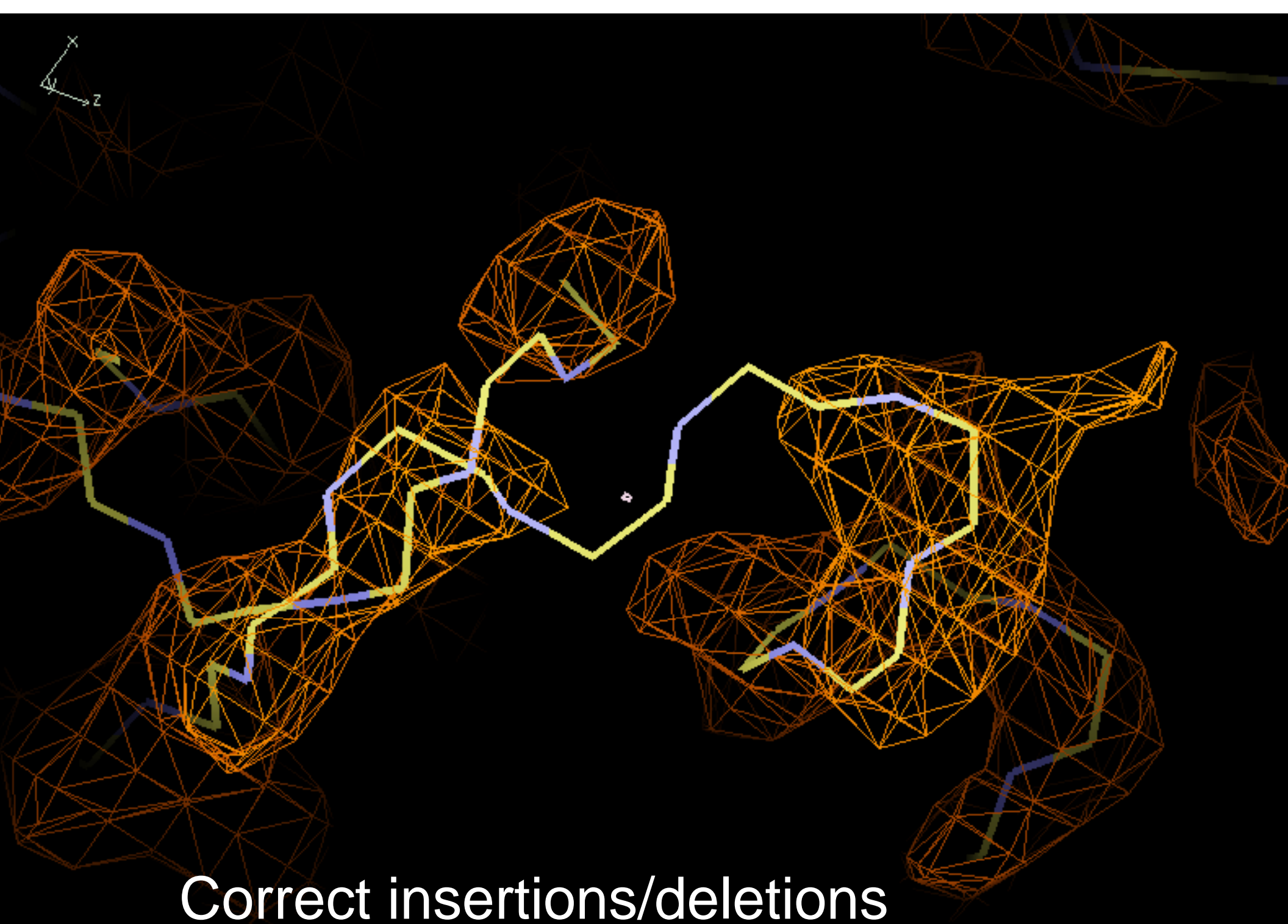


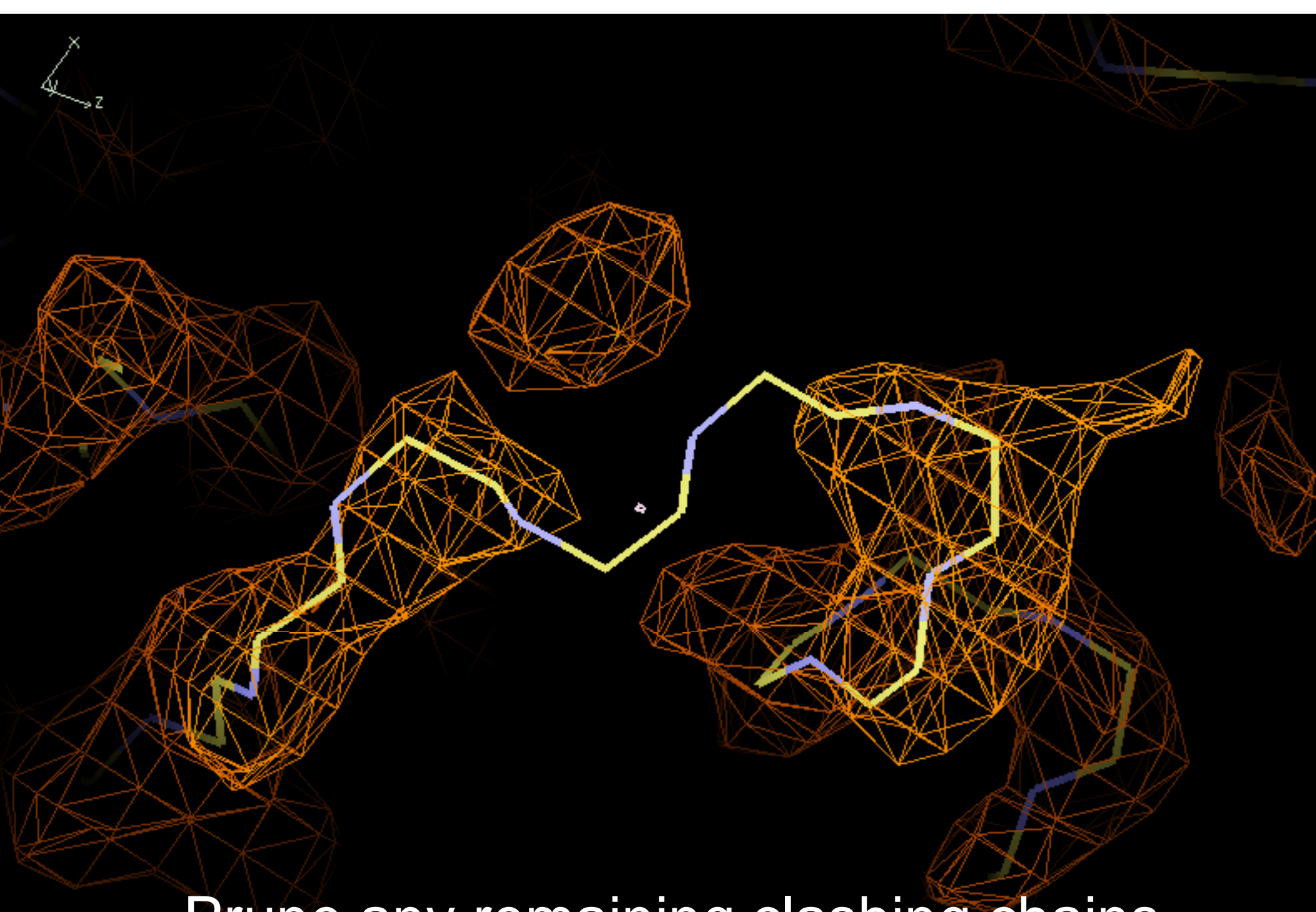


Join and merge chain fragments

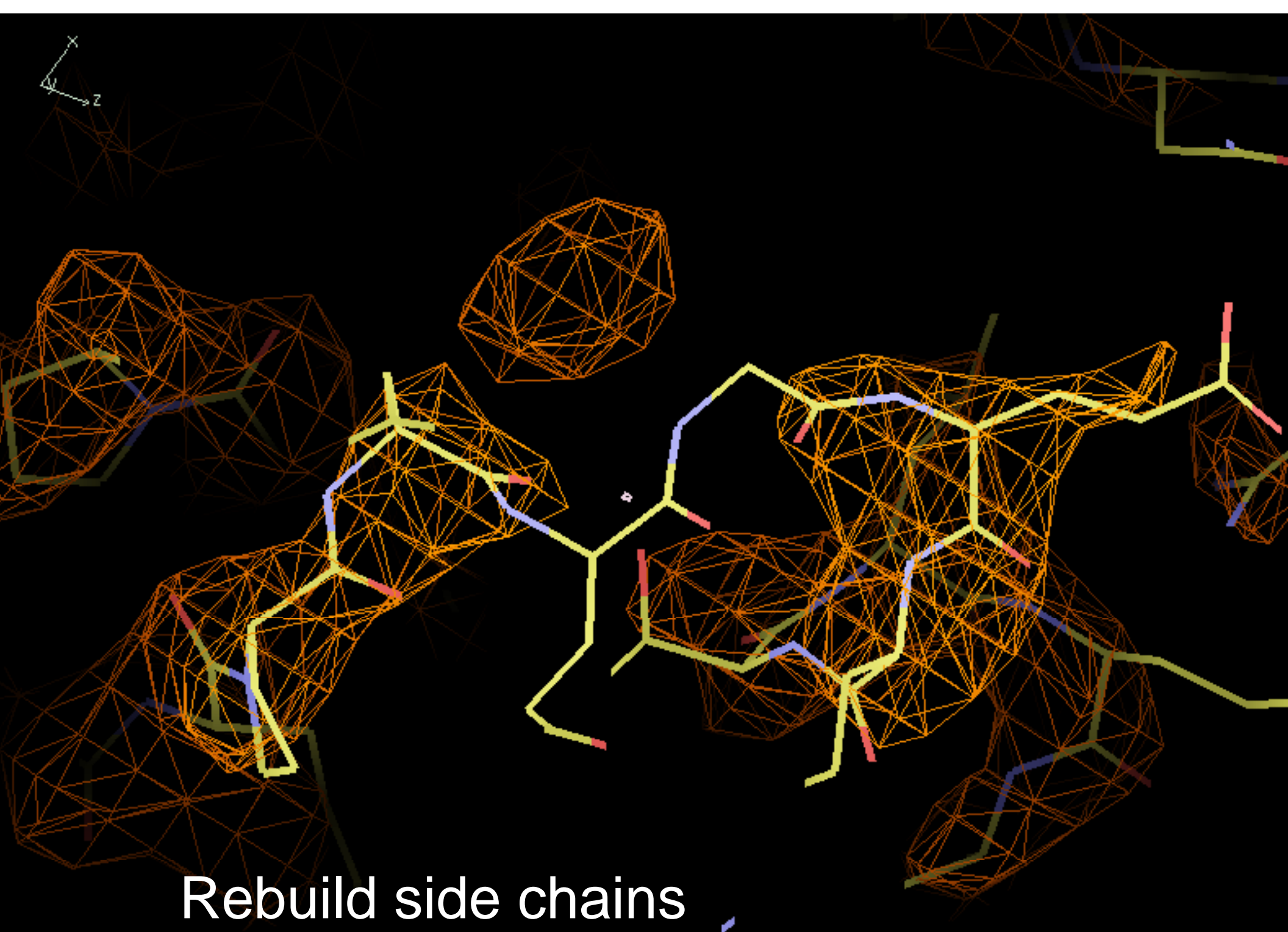


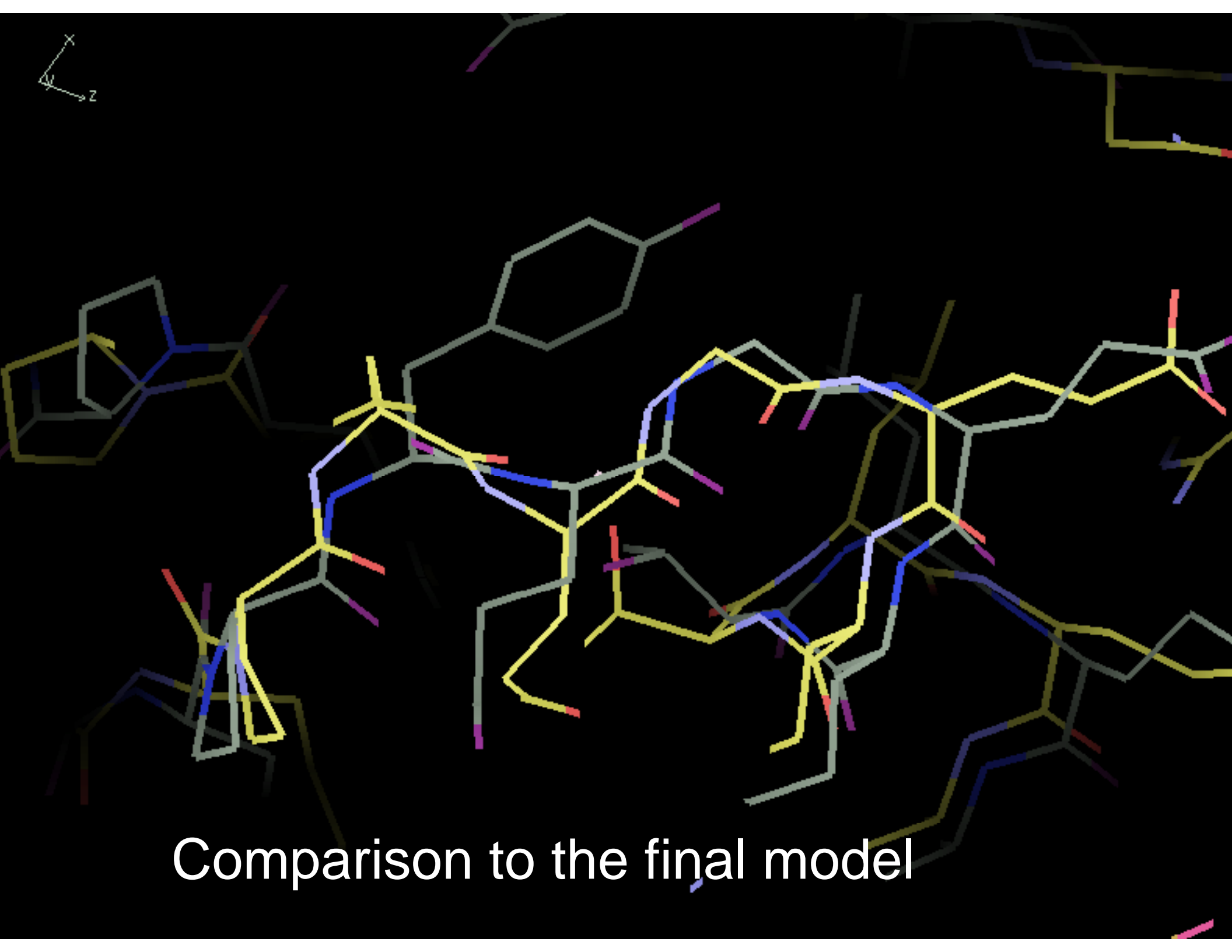






Prune any remaining clashing chains



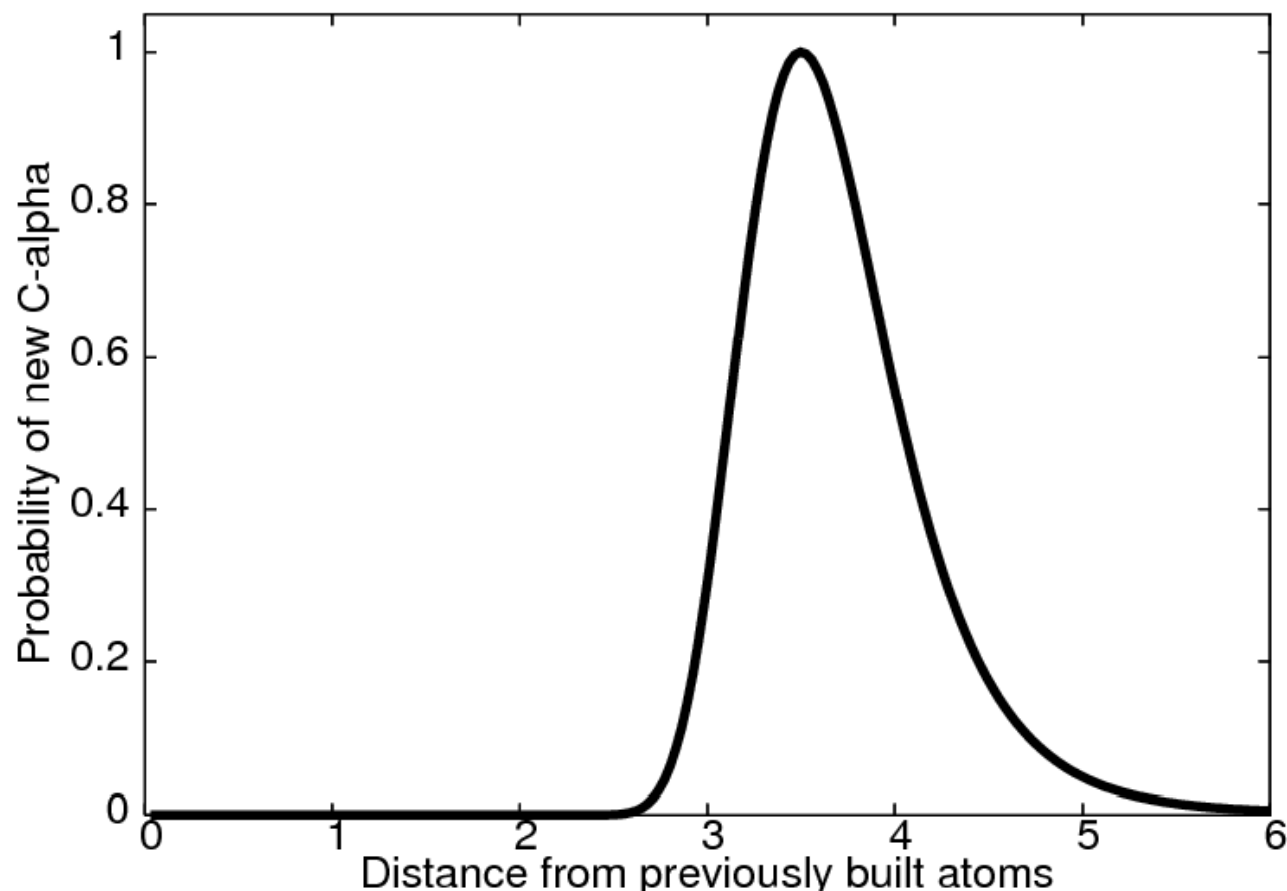


Comparison to the final model

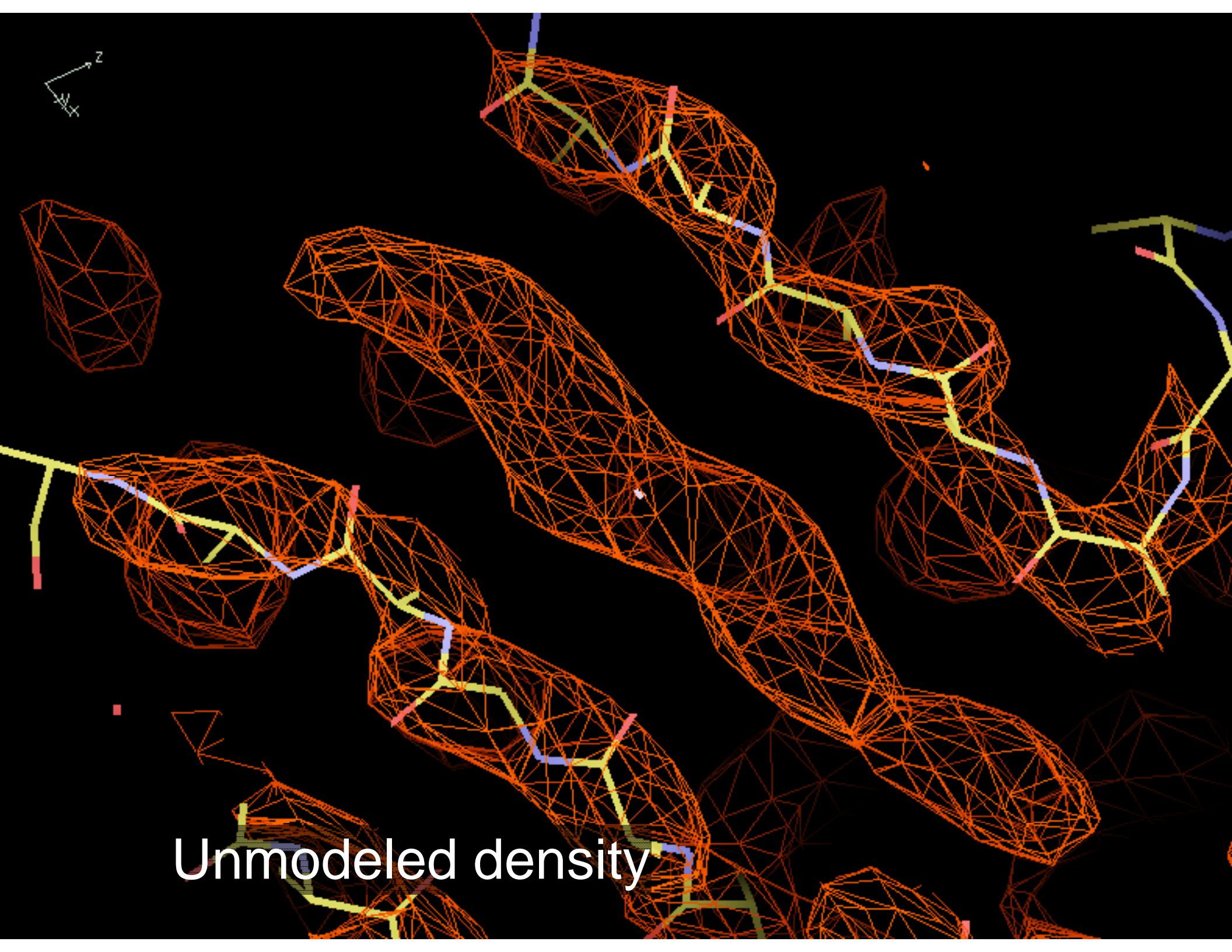
# Buccaneer

Model completion uses “**Lateral growing**”:

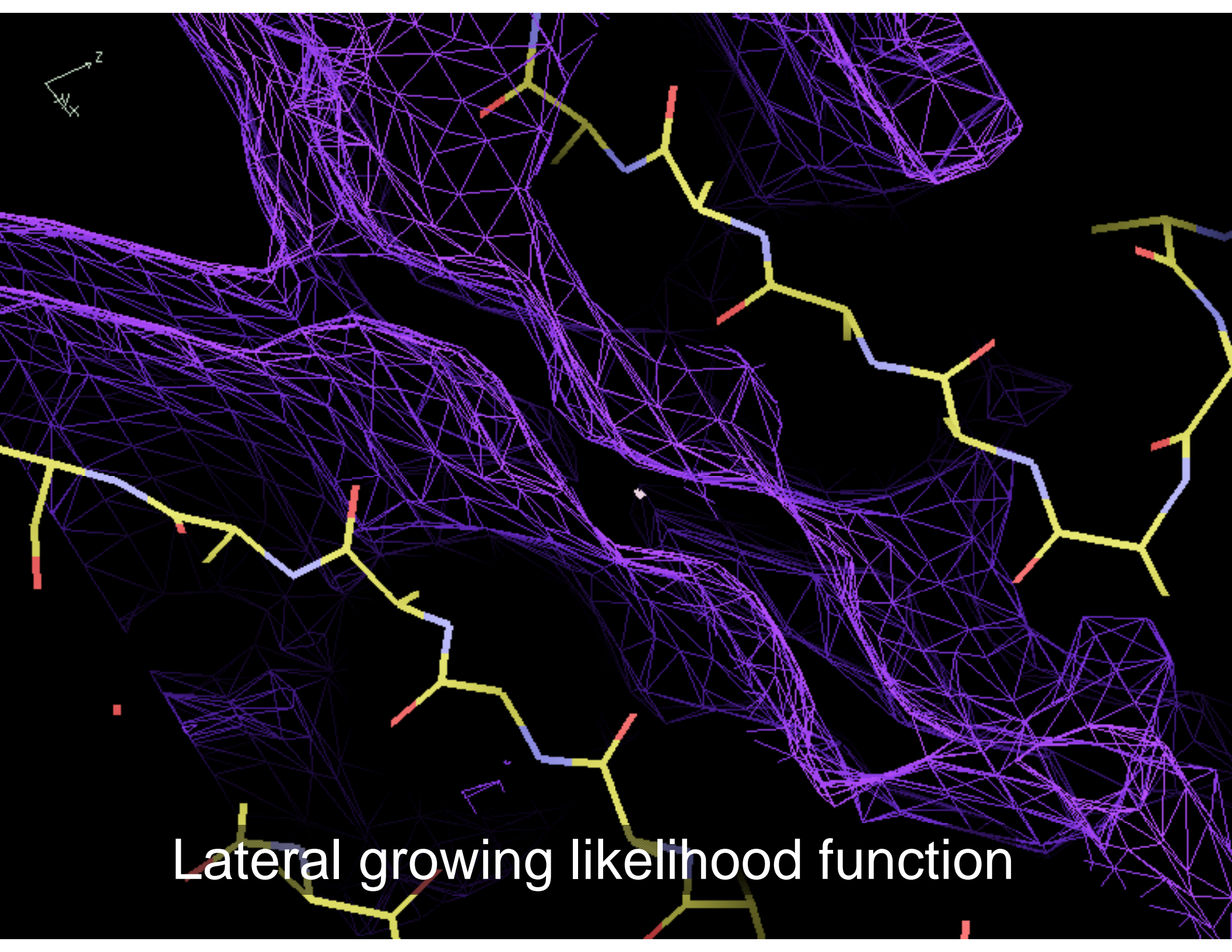
Grow sideways from existing chain fragments by looking for new C-alphas at an appropriate distance “sideways” from the existing chain:





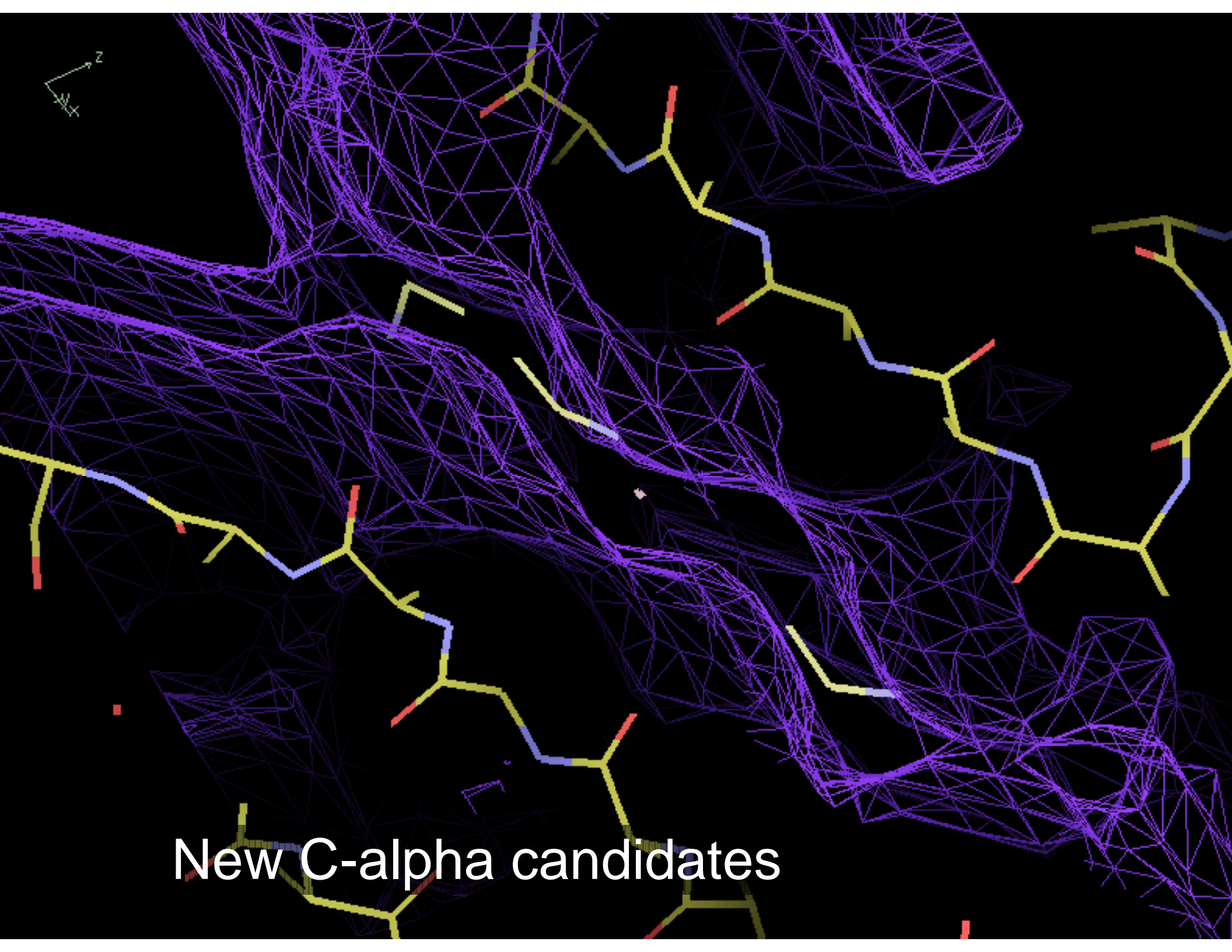


Unmodeled density

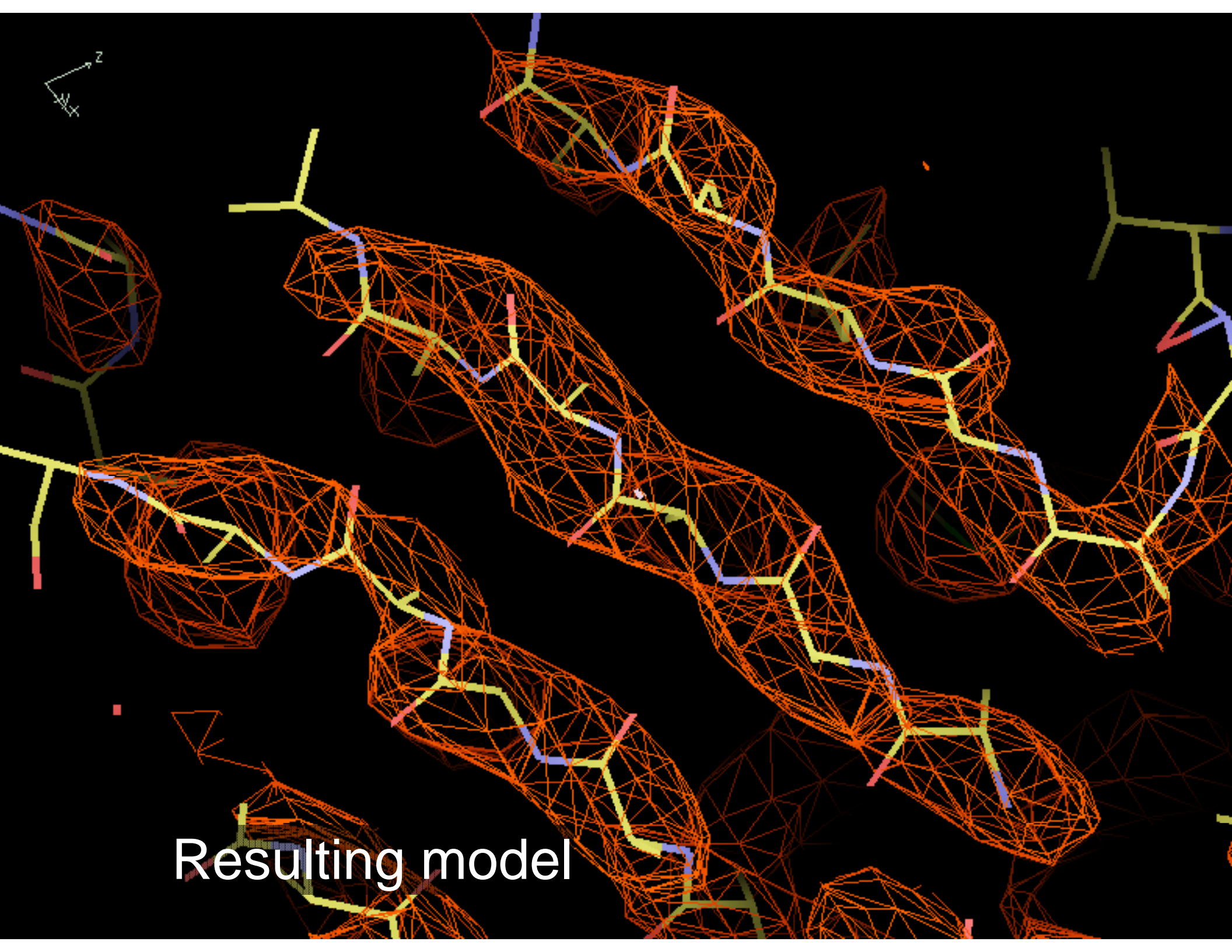


Lateral growing likelihood function





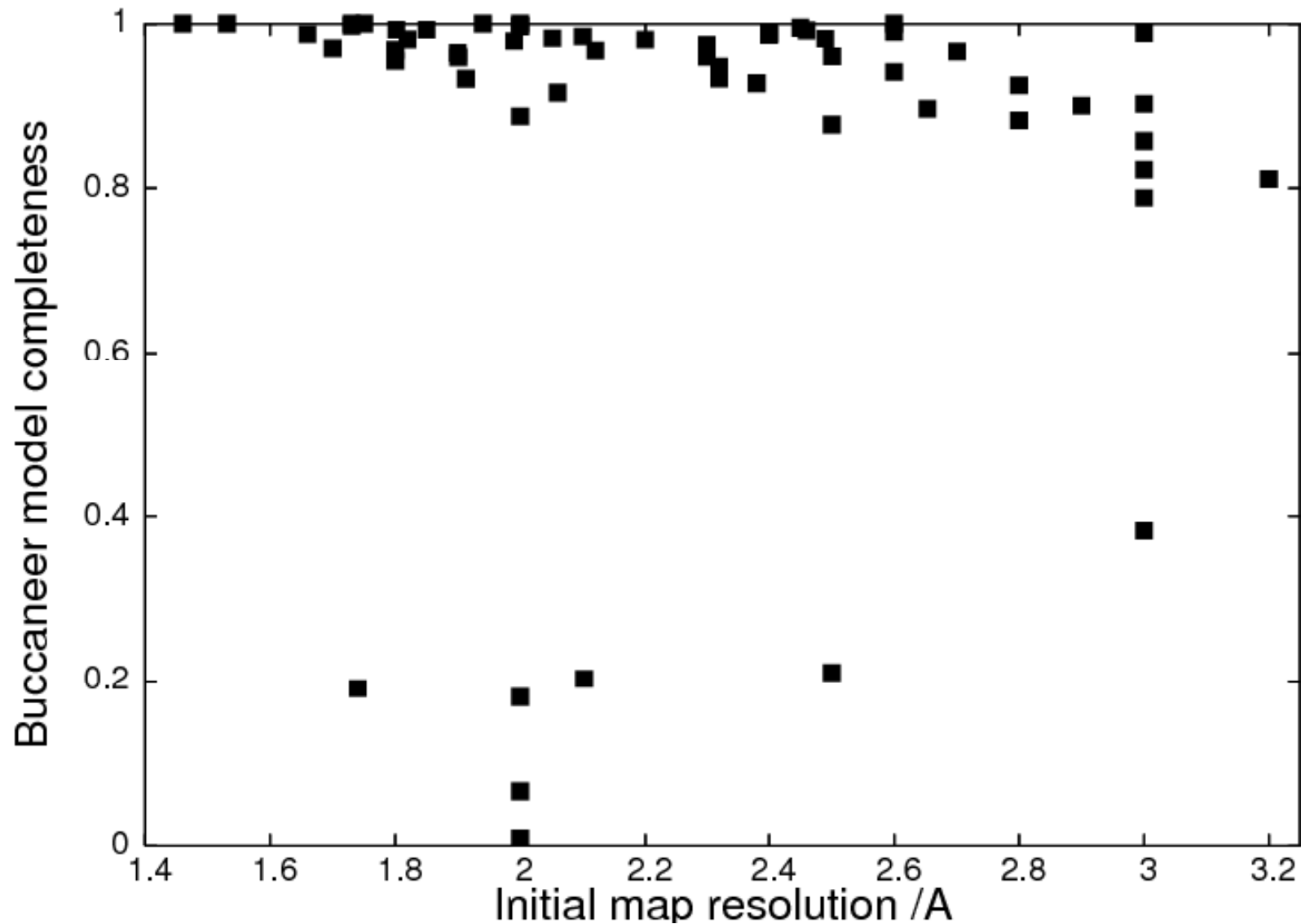
New C-alpha candidates



Resulting model

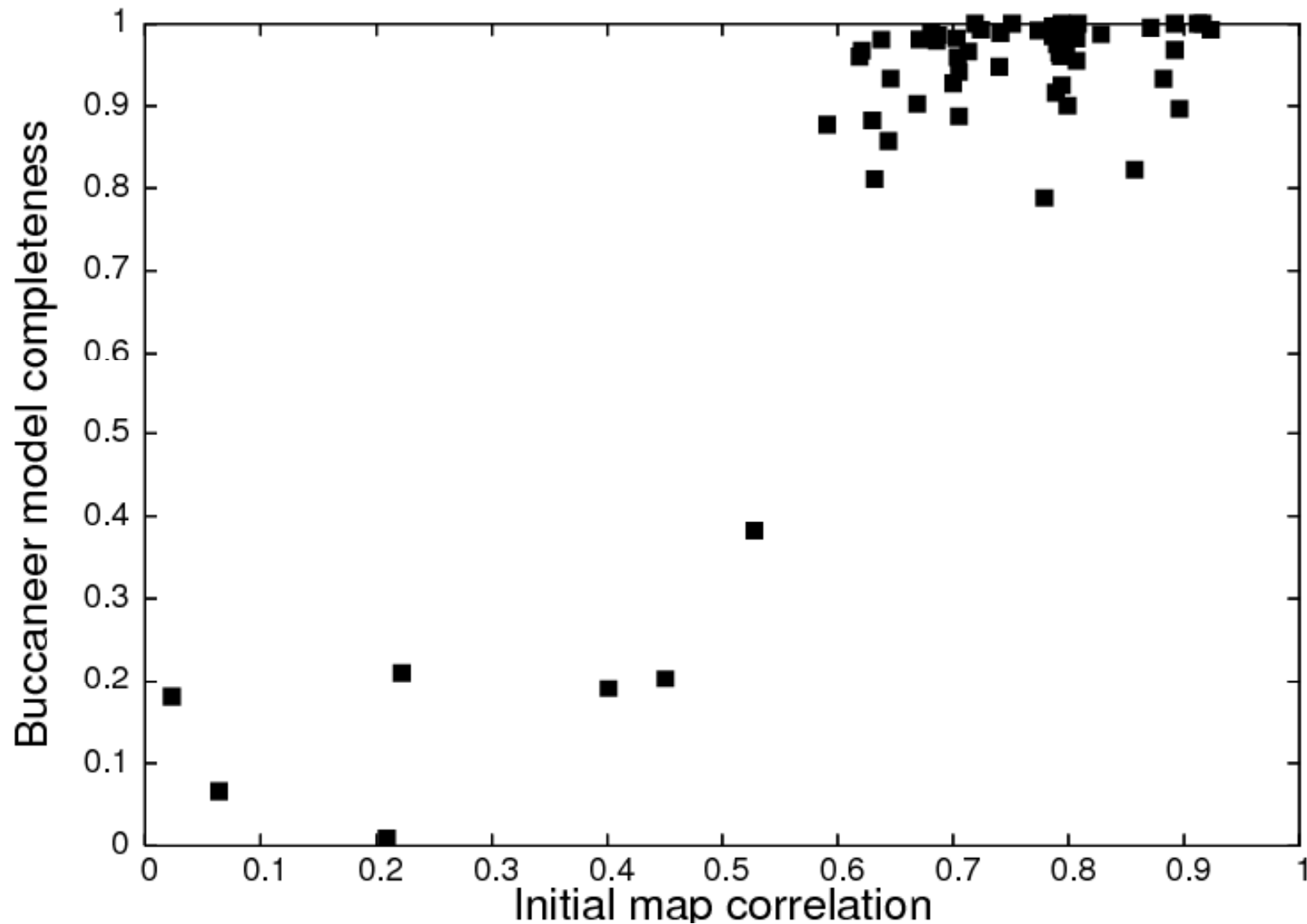
# Buccaneer: Results

Model completeness not very dependent on resolution:



# Buccaneer: Results

Model completeness dependent on initial phases:



# Buccaneer

CCP4 Program Suite 6.0.99c CCP4Interface 2.0 running

**Model Building**

- Buccaneer - autobuild/refine
- Buccaneer - fast build only
- Rapper - conformer modelling
- Sequins - sequence validation
- FFFear - Fragment Searching
- FFJoin - Merge fragments
- XtalView/xfit
- Buccaneer1.2 - autobuild/refine
- Buccaneer1.2 - fast build only

Cycle BUCCANEER and REFMAC for most complete model

185 27 Aug 08 FINISHED

Single run of BUCCANEER only (more options) quick assessment/advanced use

180 27 Aug 08 FINISHED  
179 27 Aug 08 FINISHED  
178 27 Aug 08 FINISHED  
175 27 Aug 08 FINISHED

Latest versions from York as above  
Faster and better results

# Buccaneer

Chain tracing/refinement using Buccaneer/Refmac

Help

Job title

*Data for (unsolved) work structure:* (Note: perform phase improvement/density modification first)

☐ Specify an initial model to be extended.

Work SEQ in PROJECT Browse View

Work MTZ in PROJECT Browse View

FP SIGFP

HLA HLB

HLC HLD

Free R flag

Use Free-R flag: ☒ Use map coefficients: ☐ Use PHI/FOM instead of HL coefficients: ☐

Work PDB out PROJECT buccaneer.pdb Browse View

*Options* ☒

Number of cycles of building/refinement to run: 3

*Buccaneer parameters* ☐

*Refmac parameters* ☐

Run Save or Restore Close



# Buccaneer

What it does:

- Trace protein chains (trans-peptides only)
- Link across small gaps
- Sequence
- Apply NCS
- Build side chains (roughly)
- Refine (if recycled)
- WORK AT LOW RESOLUTIONS
  - 3.6Å (4Å?) with good phases

# Buccaneer

What it does not do (yet):

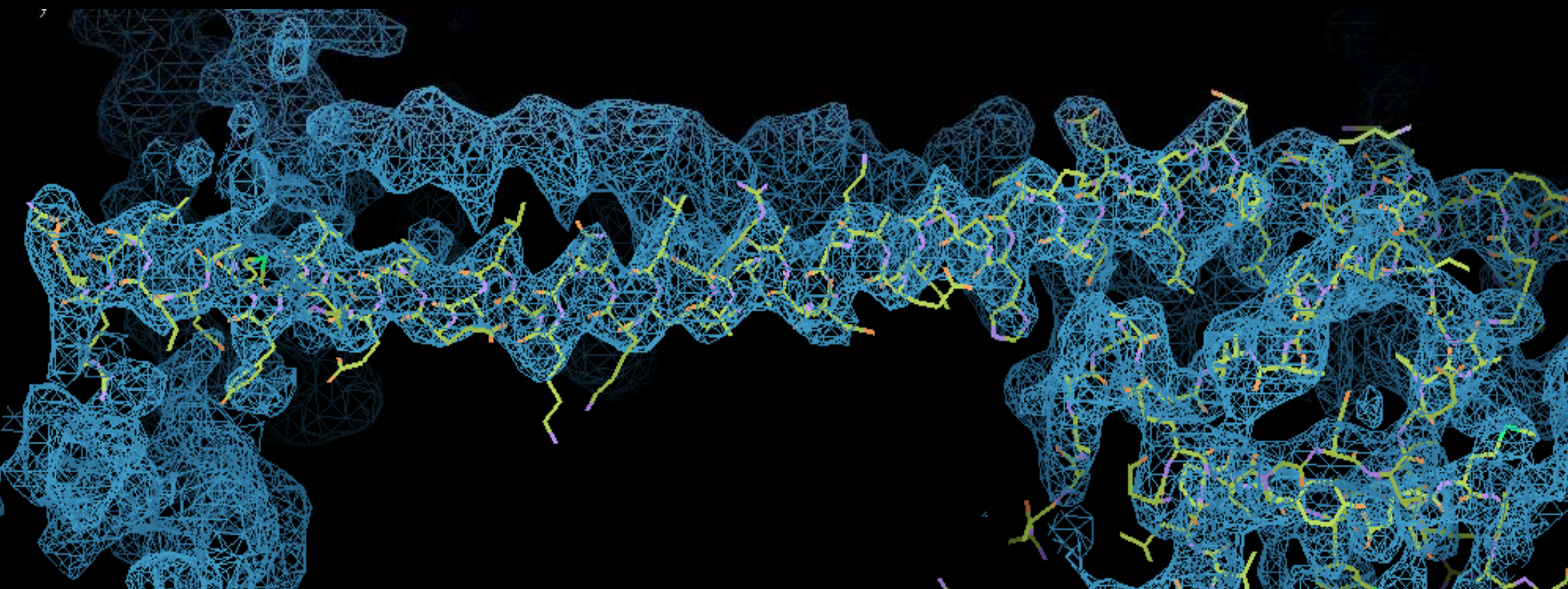
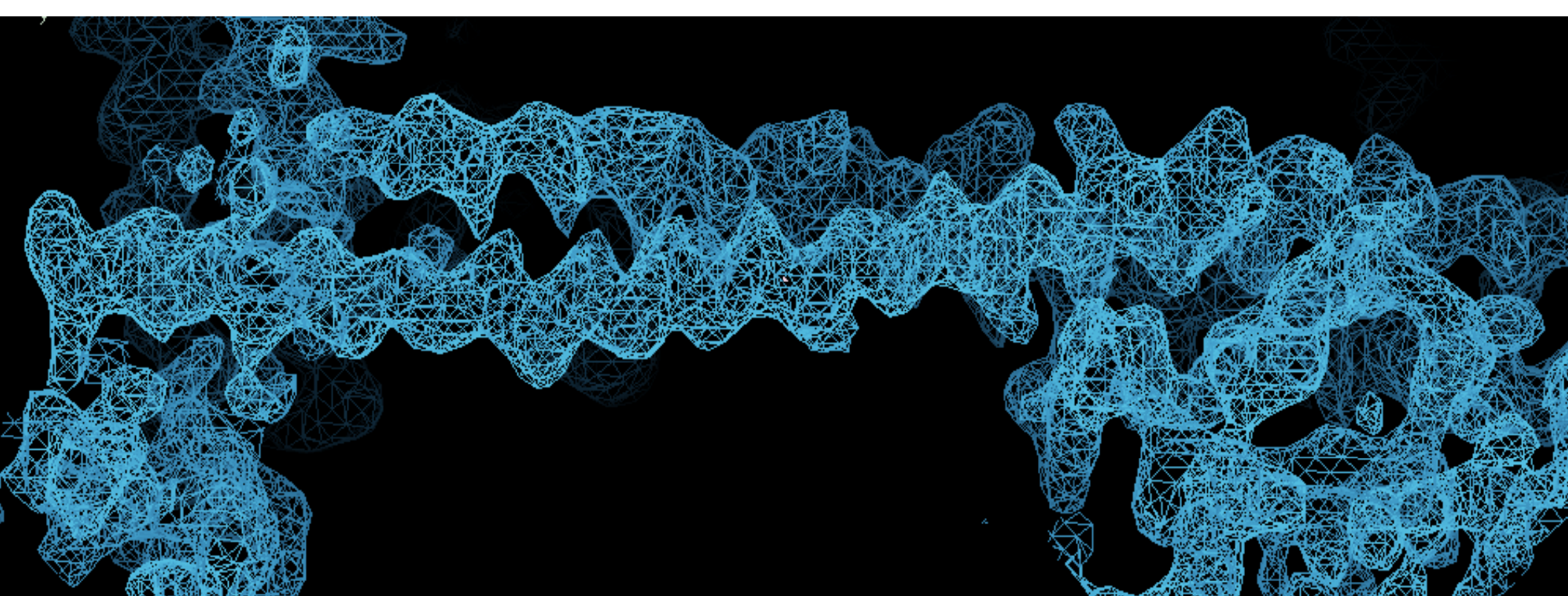
- Cis-peptides
- Waters
- Ligands
- Loop fitting
- Move C-alpha to fit rotamer

# Buccaneer

What you need to do afterwards:

- Tidy up with Coot.
  - Or ARP/wARP when resolution is good.
  - Buccaneer/ARP/wARP better+ faster than ARP/wARP.
- Typical Coot steps:
  - Connect up any broken chains.
  - Use density fit and rotamer analysis to check rotamers.
  - Check Ramachandran, molprobity, etc.
  - Add waters, ligands, check un-modeled blobs..
  - Re-refine, examine difference maps.





# Buccaneer: Latest

## Buccaneer 1.2

- Use of Se atoms, MR model in sequencing.
- Improved numbering of output sequences (ins/del)
- Favour more probable sidechain rotamers
- Prune clashing side chains
- Optionally fix the model in the ASU
- Performance improvements (1.5 x)
  - Including 'Fast mode' (2-3 x for good maps)
- *Multi-threading*



# Buccaneer: Summary

A simple, fast, easy to use (i.e. MTZ and sequence) method of model building which is robust against resolution.

User reports for structures down to 3.7Å when phasing is good.

Results can be further improved by iterating with refinement in refmac (and in future, density modification).

Proven on real world problems.

# Achnowledgements

## Help:

- JCSG data archive: [www.jcsg.org](http://www.jcsg.org)
- Eleanor Dodson, Paul Emsley,  
Randy Read, Clemens Vonrhein,  
Raj Pannu

## Funding:

- The Royal Society