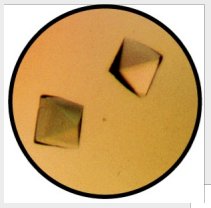
Experimental Phasing

Phenix Workshop
October 14, 2023
80th Pittsburgh Diffraction Conference

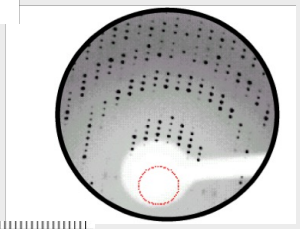
Tom Terwilliger
Los Alamos National Laboratory/New Mexico Consortium

Solving a structure with SAD phasing

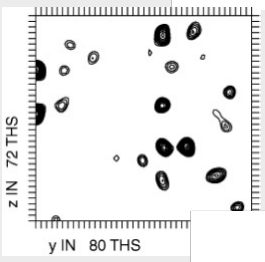
Crystals with Se/Zn/…



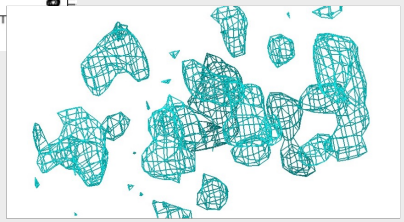
Collect anomalous SAD data



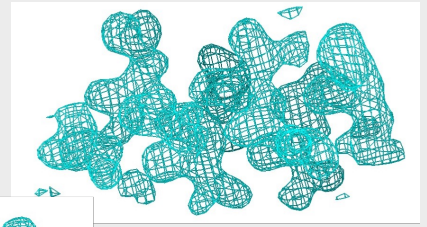
Locate substructure



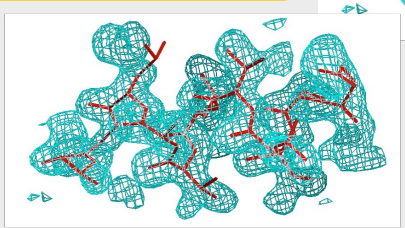
Phasing (calculate density map)



Density modification (improve map)



Model building



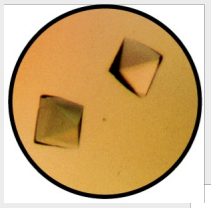
Solving a structure with SAD phasing (Se)

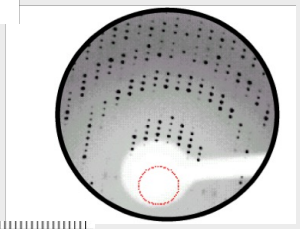
Planning the experiment

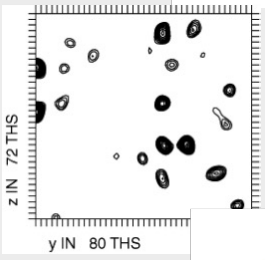
Automating the analysis

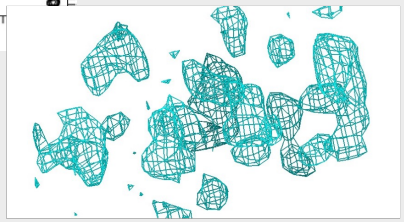
Improving the map

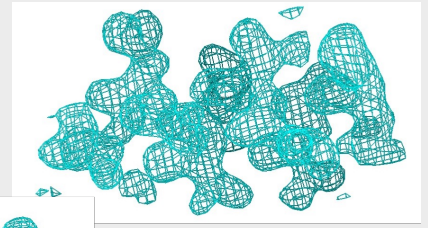
Building a model

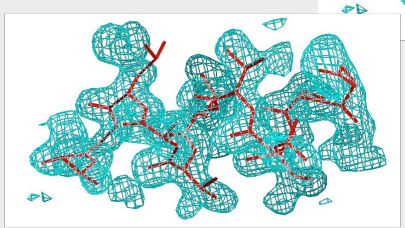












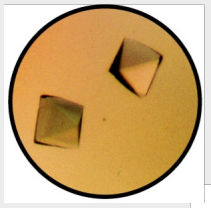
Will I solve my SAD structure?

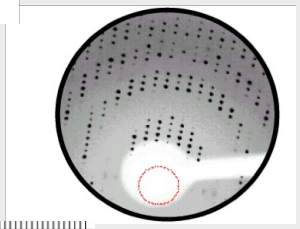
Planning the experiment

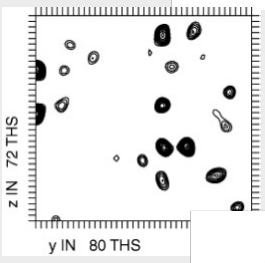
Automating the analysis

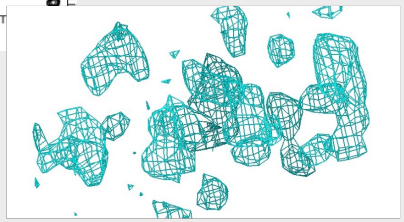
Improving the map

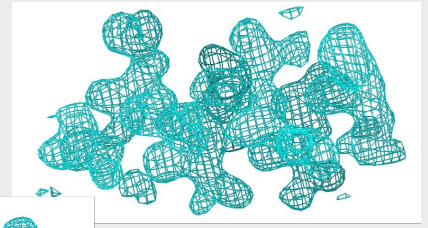
Building a model

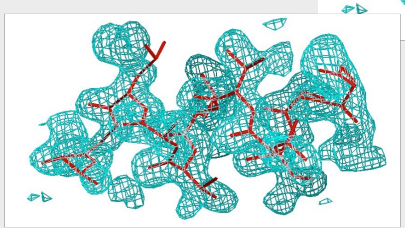












Will I find the anomalous substructure?

How many sites?

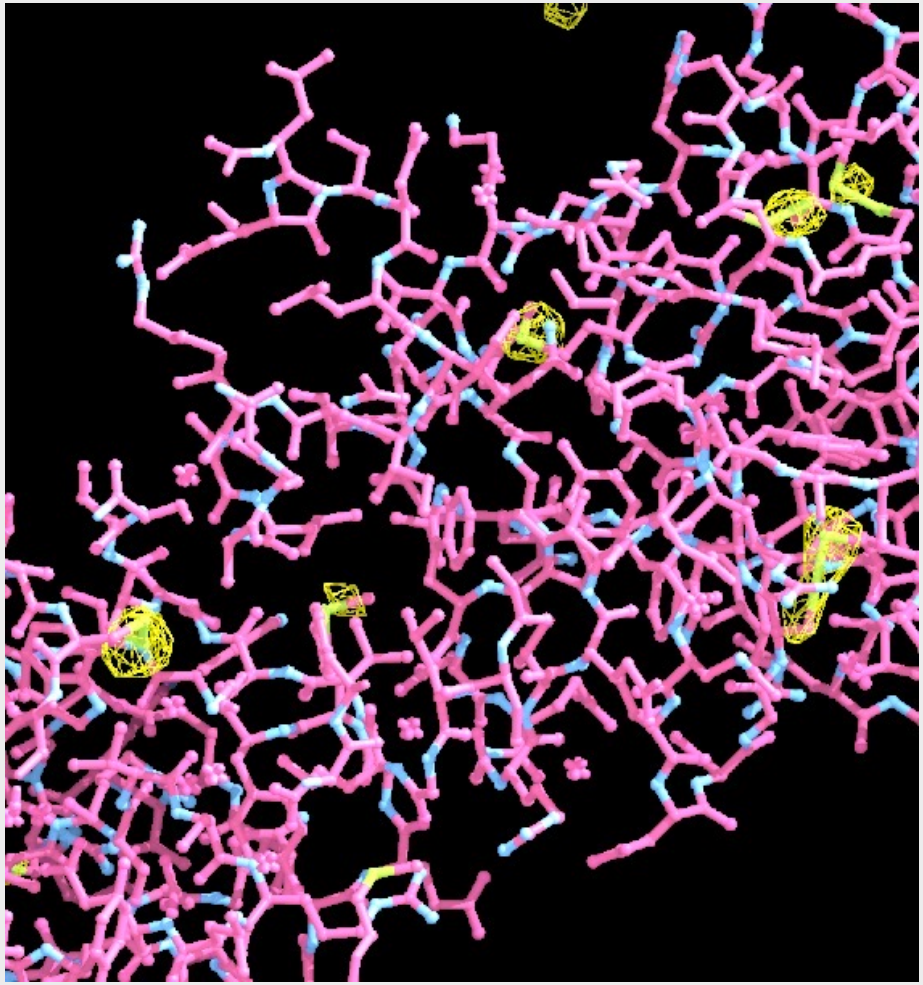
Are sites ordered?

Anomalous atom?

Wavelength?

Accurate data?

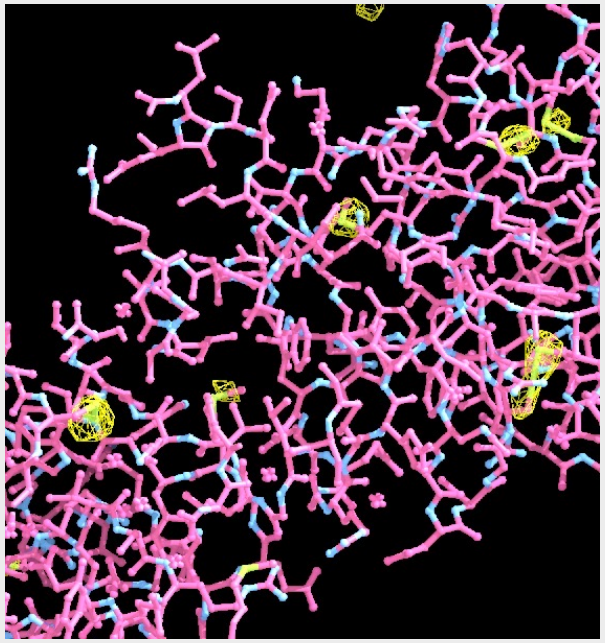
How many reflections?

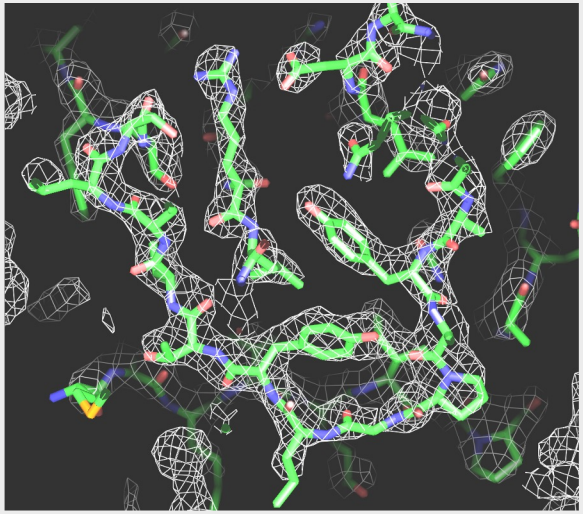


Key steps in SAD structure determination

1. Find the substructure

Anomalous signal S

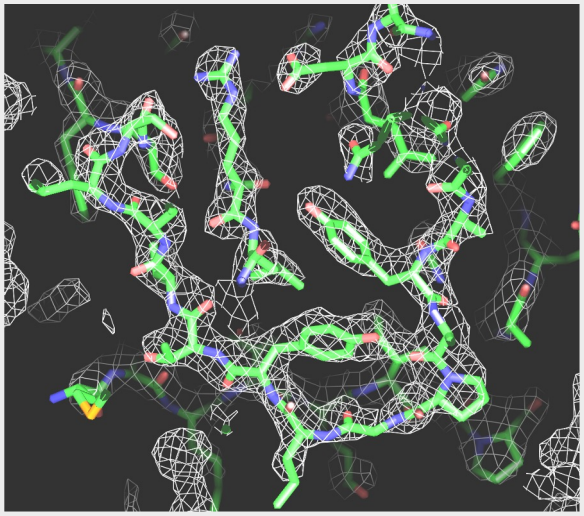




2. Calculate an interpretable map

Anomalous correlation CC^*_{ano}

Anomalous correlation

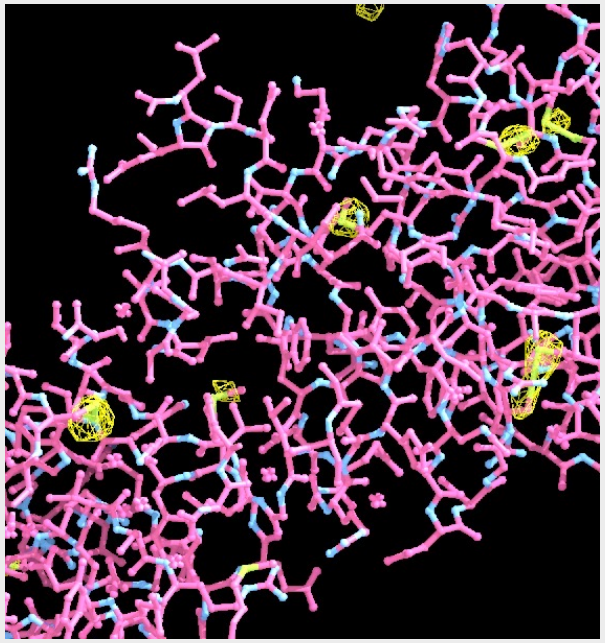


Anomalous correlation CC^*_{ano}

- Correlation of anomalous differences with ideal
- Accuracy of anomalous data
- Accuracy of phasing

Anomalous signal

Anomalous signal S



- Peak height in anomalous difference Fourier
- “Information per site” (can we find each site)
- Substructure likely to be found if S > 10

Will I find the anomalous substructure?

How many sites?

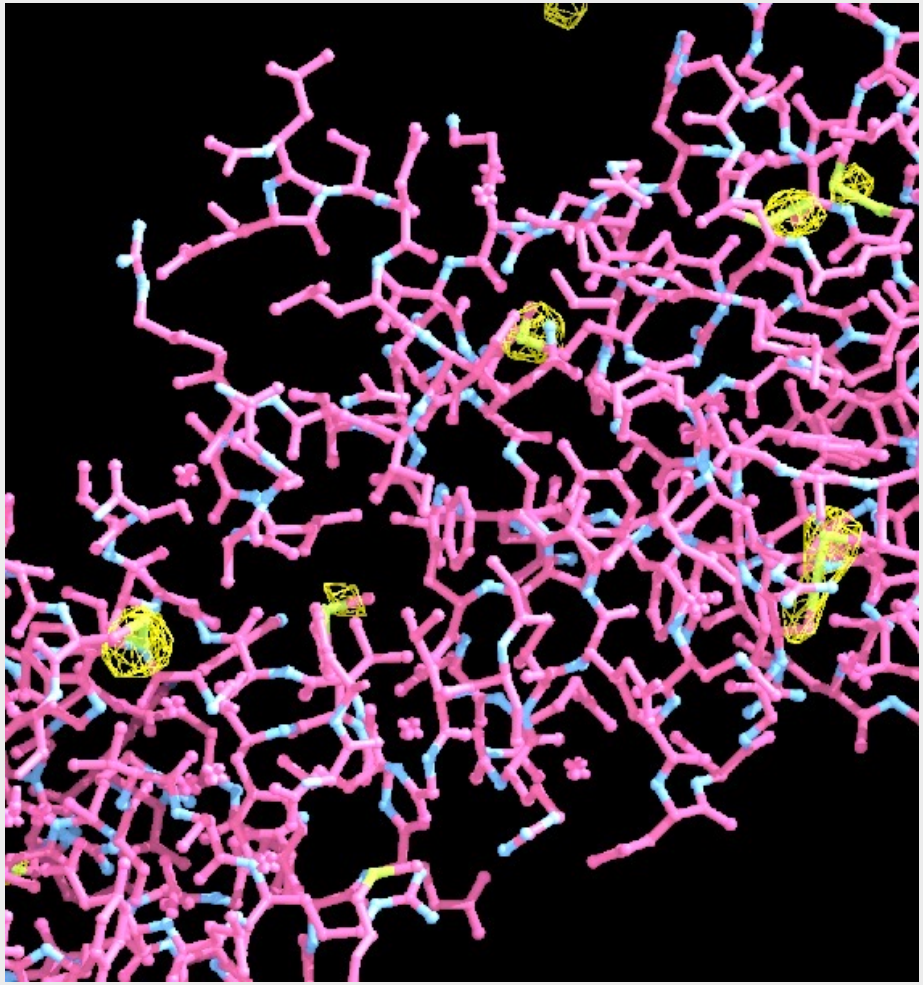
Are sites ordered?

Anomalous atom?

Wavelength?

Accurate data?

How many reflections?



Anomalous signal: information about each site
(peak height in anomalous difference Fourier)

Accuracy of the data

Anomalous correlation

Number of reflections

Anomalous signal S

$$\langle S_{ano} \rangle = CC^{*}_{ano} \cdot \frac{\sqrt{N_{refl}}}{\sqrt{n_{sites}}} \cdot \frac{1}{f^{1/2}}$$

Will I find sites?

Number of sites

B-value for anomalous sub-structure

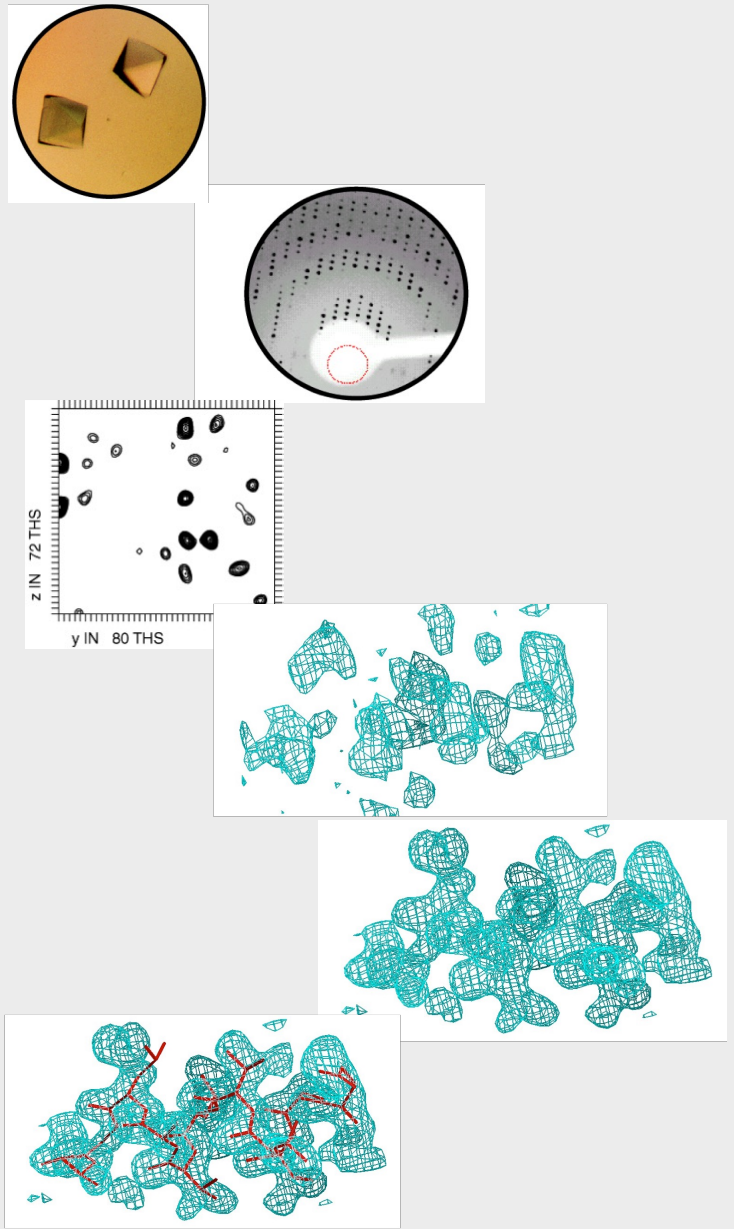
Will I solve my SAD structure?

Planning the experiment

Automating the analysis

Improving the map

Building a model



Why automate structure determination?

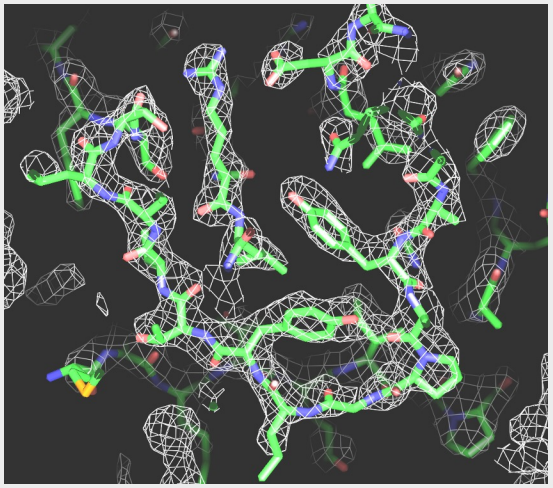
Makes straightforward cases easier

… and difficult cases feasible for experts

Speeds up the process

Reduces errors

Allows you to try more possibilities



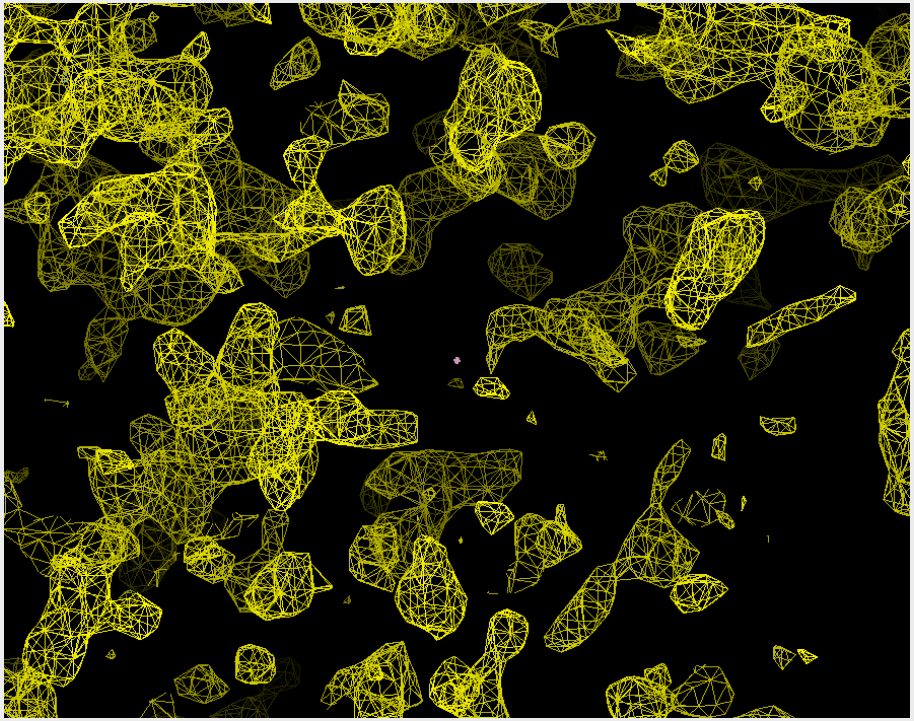
Decision-making in automation

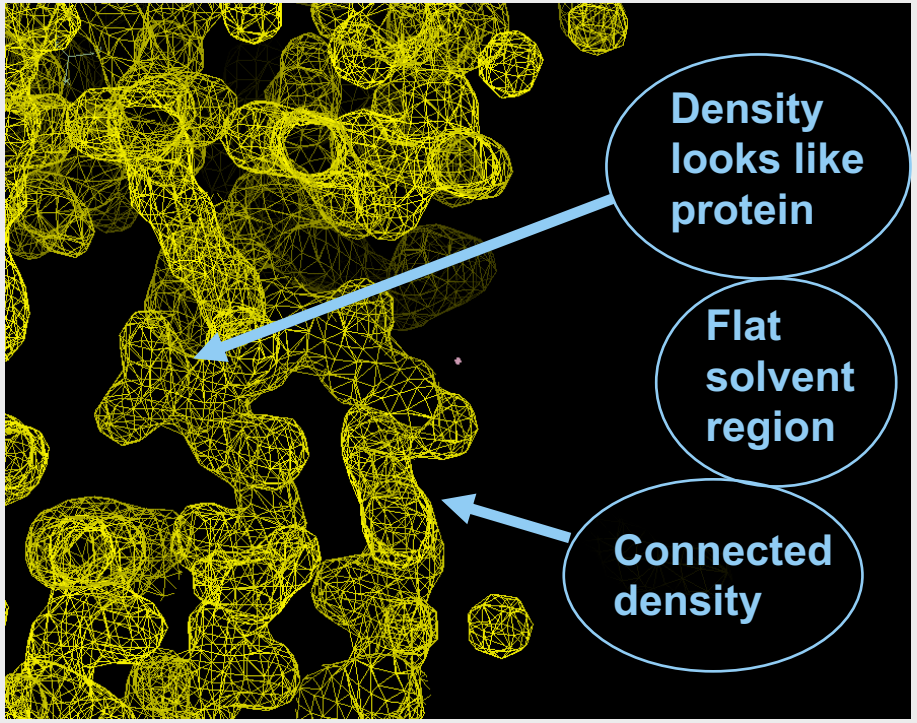
What does a good electron density map look like?

Using expected features of maps to make decisions and to improve maps

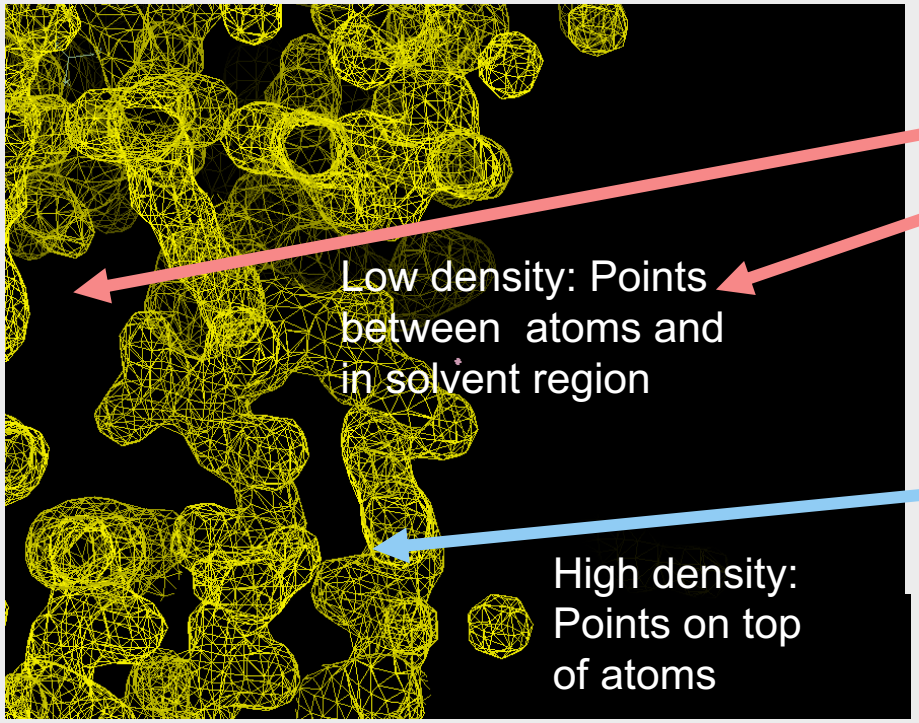
Decision-making in automation

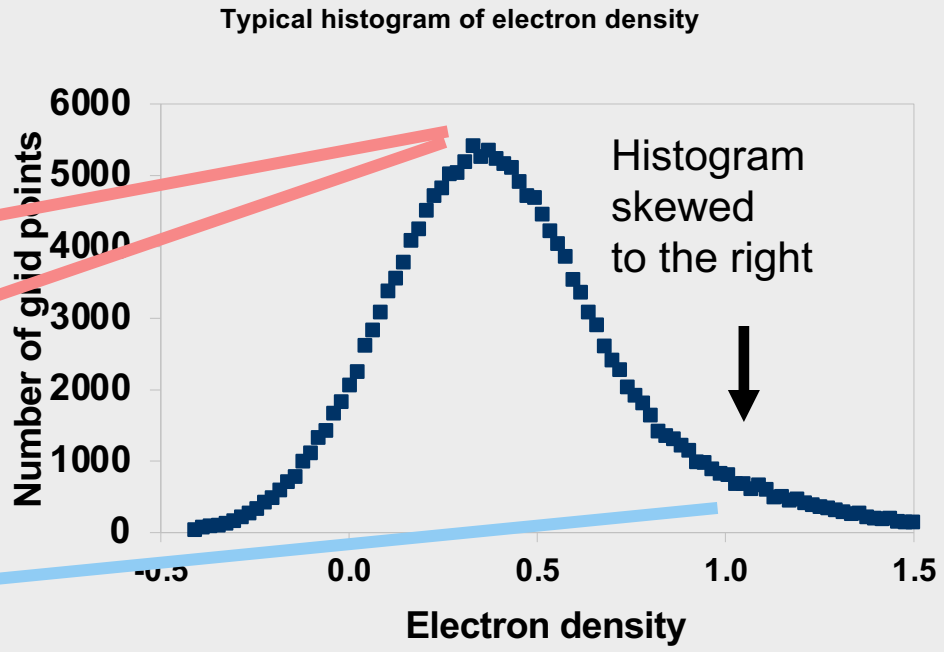
Which map is better?



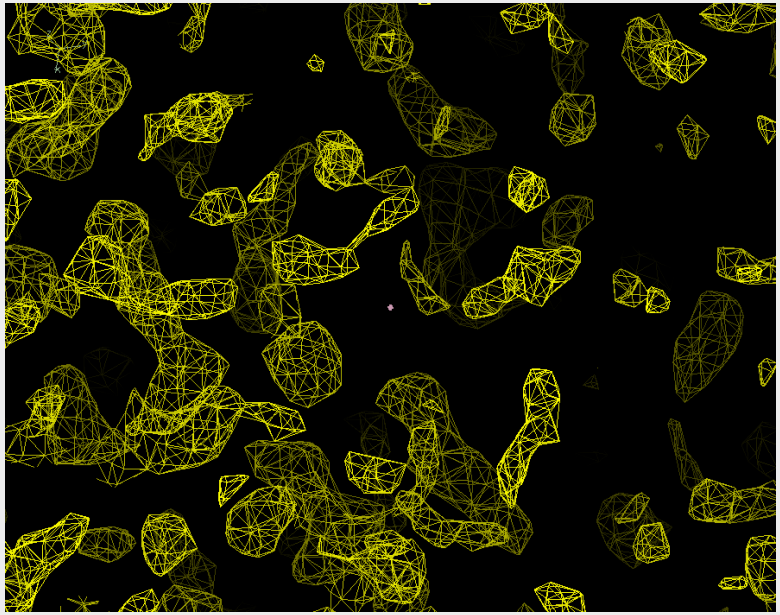


Histograms of density have positive skew

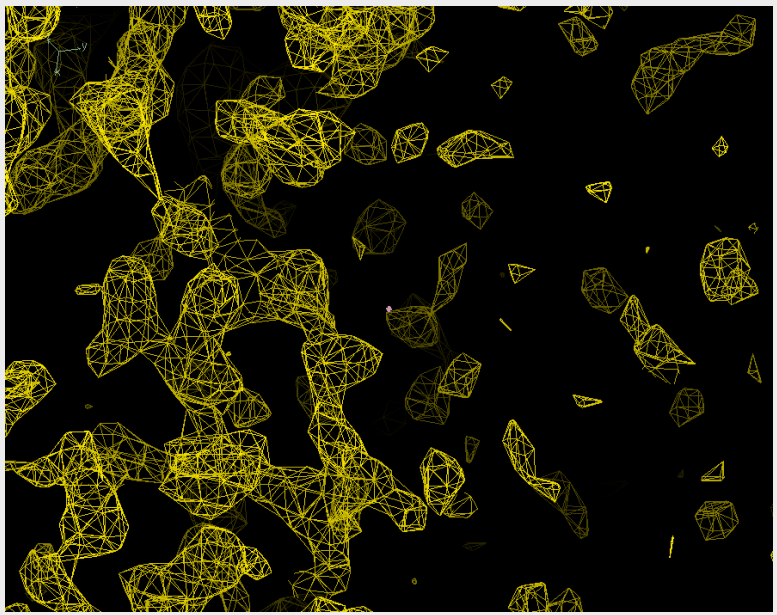




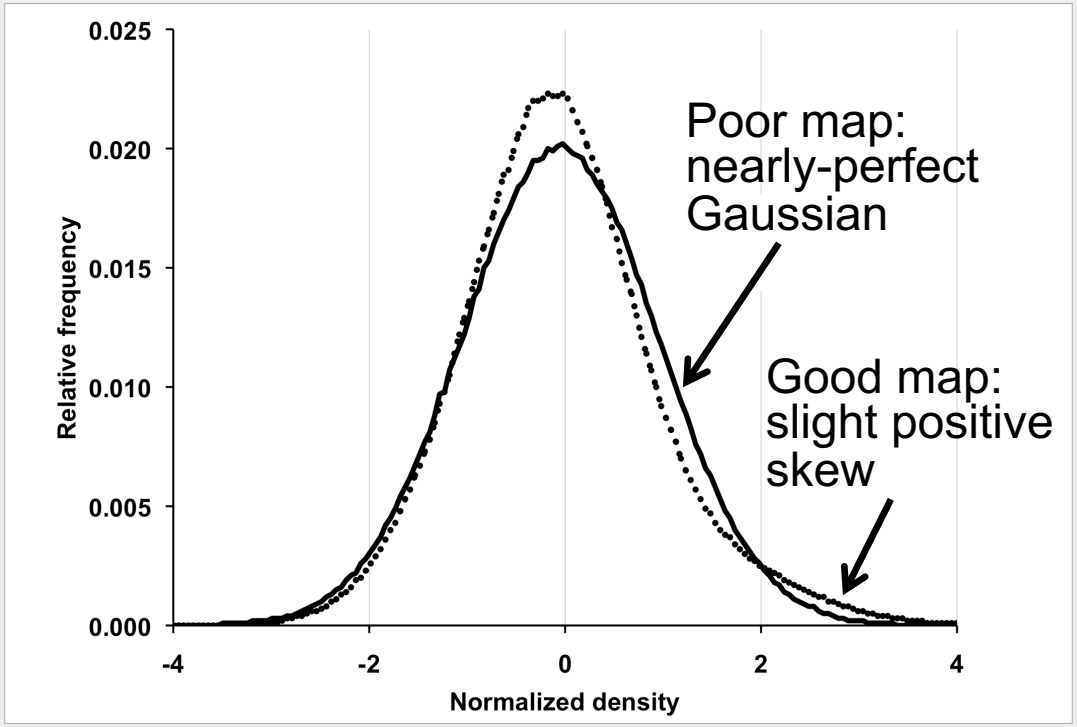
Histograms of density have positive skew



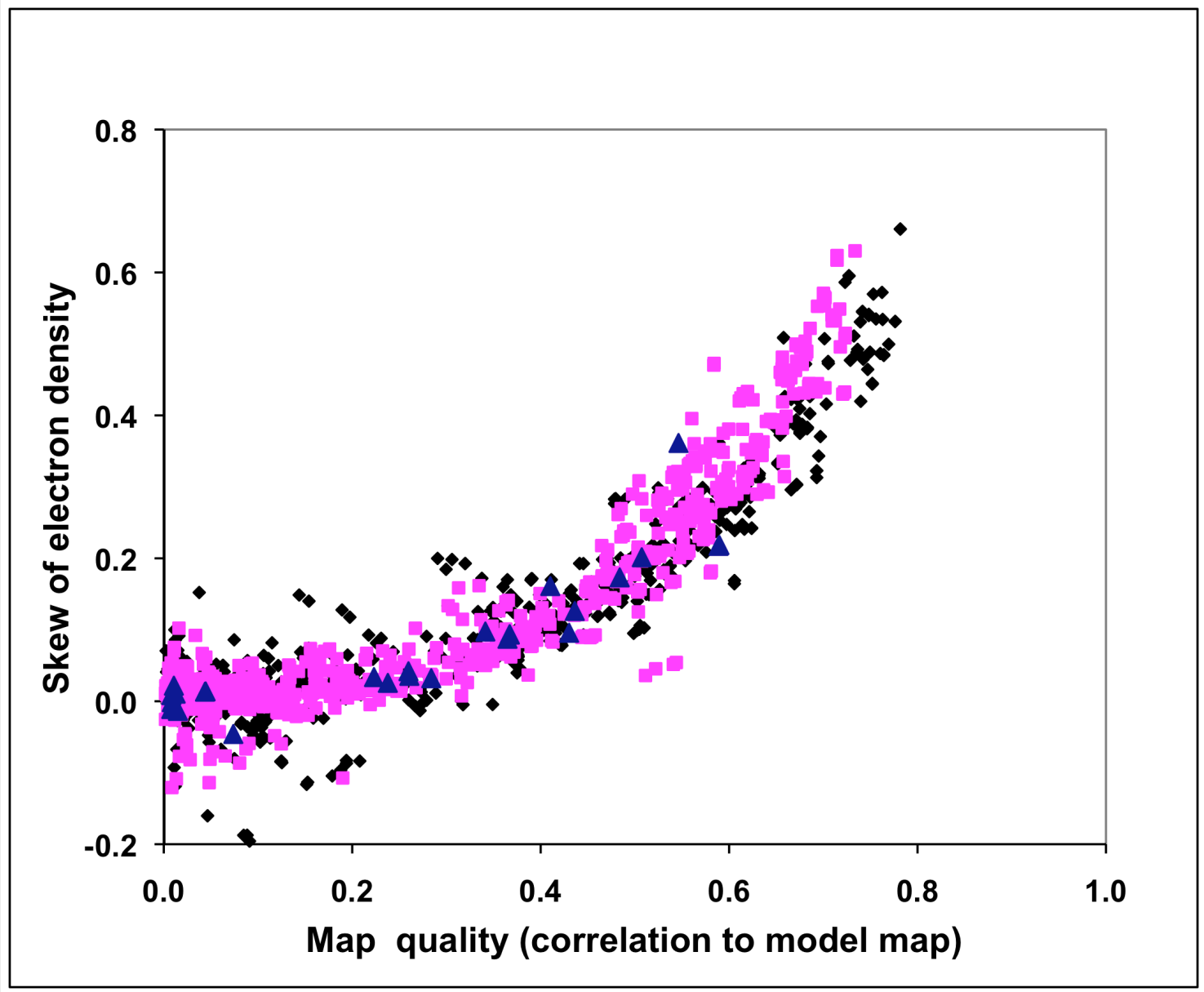
Poor map
(inverse hand)



Better map

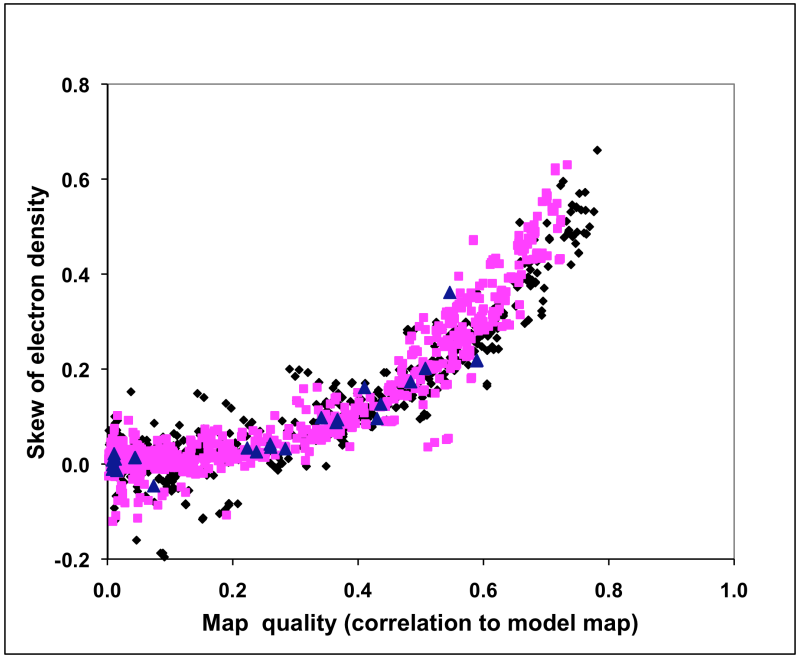


Positive skew in good maps

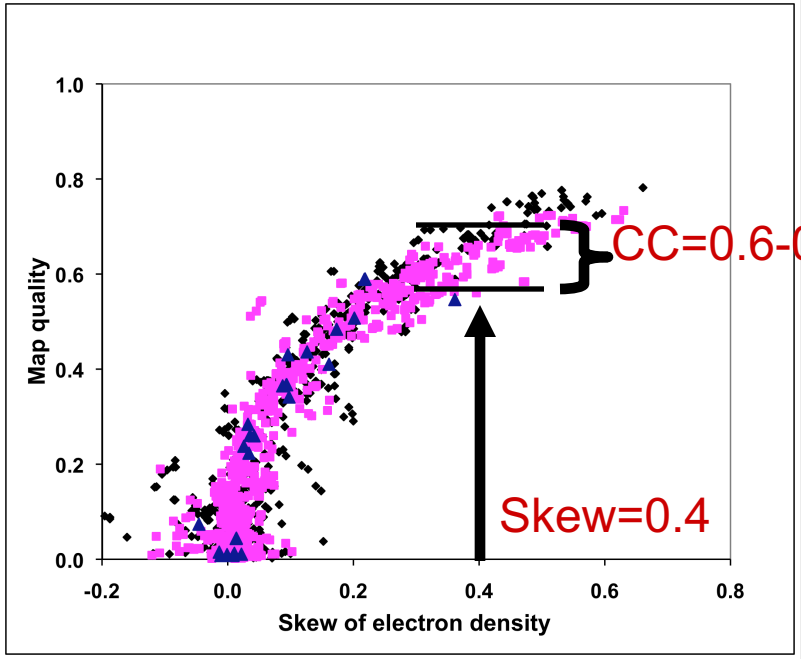


Estimate map quality from skew

Skew depends on map quality



Estimate map quality from skew



Density modification

What does a good density map look like?

Use expected features of maps to improve map quality

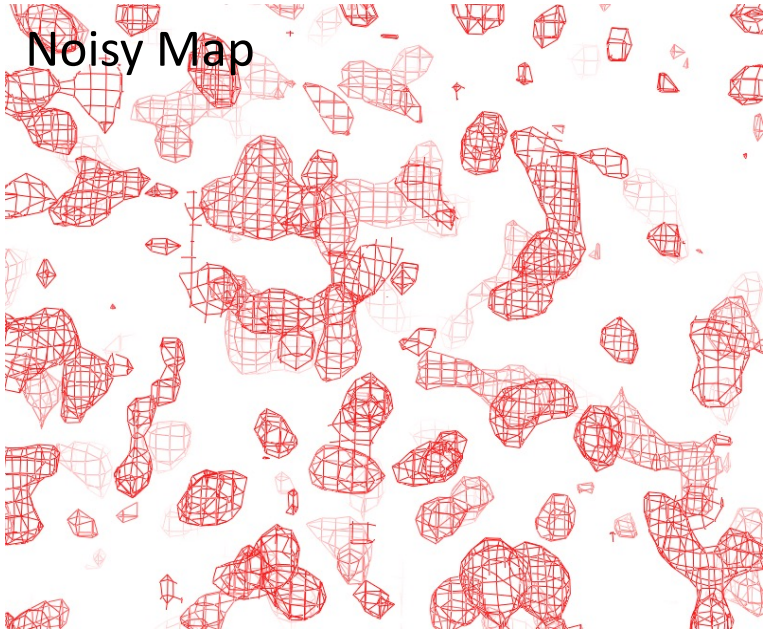
Key feature of this process: improving density anywhere can improve it everywhere

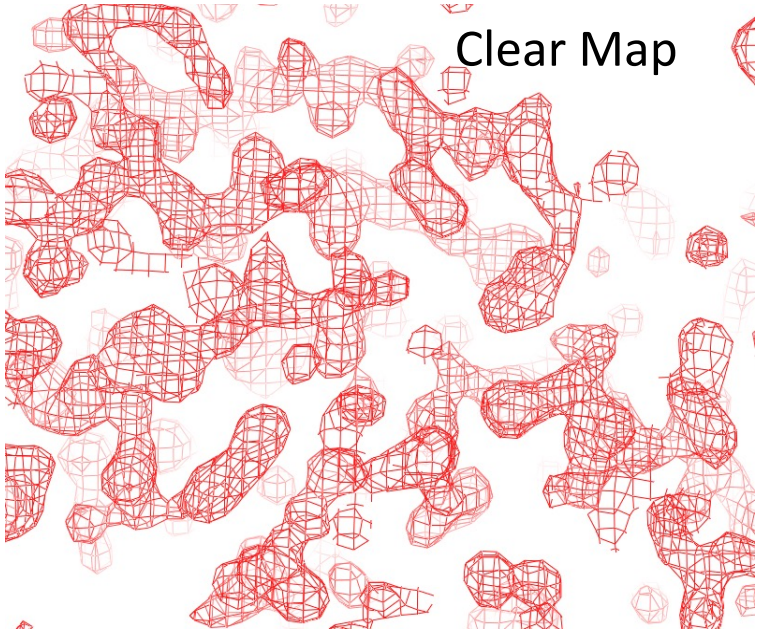
X-ray density modification: “phase improvement”

Experimental Data

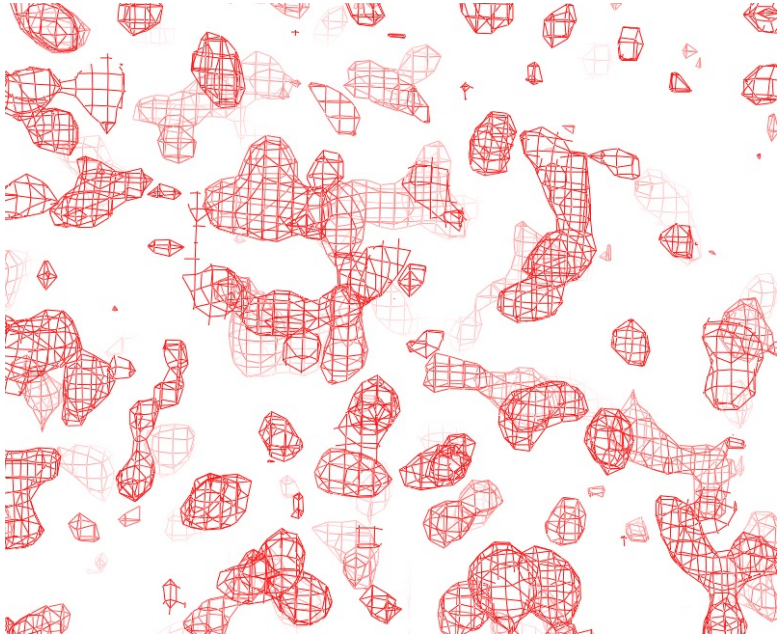
Initial phases

Improved phases

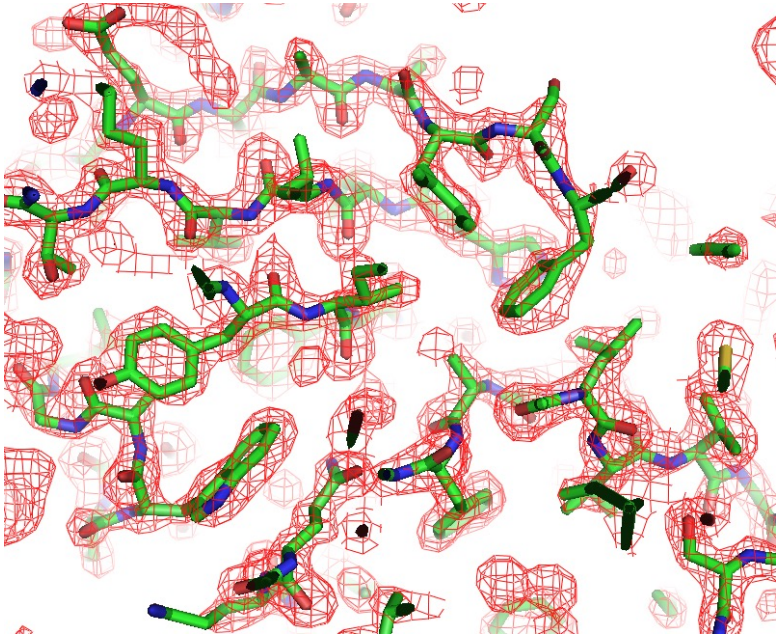




Basis of density modification



Noisy map



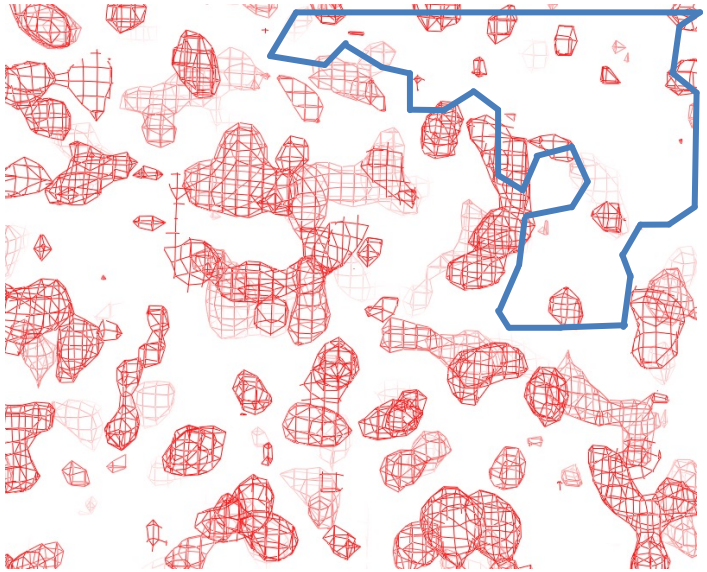
Clear map

1. We know a good map when we see it

2. Improvement anywhere ***improves the phases*** so there is improvement everywhere

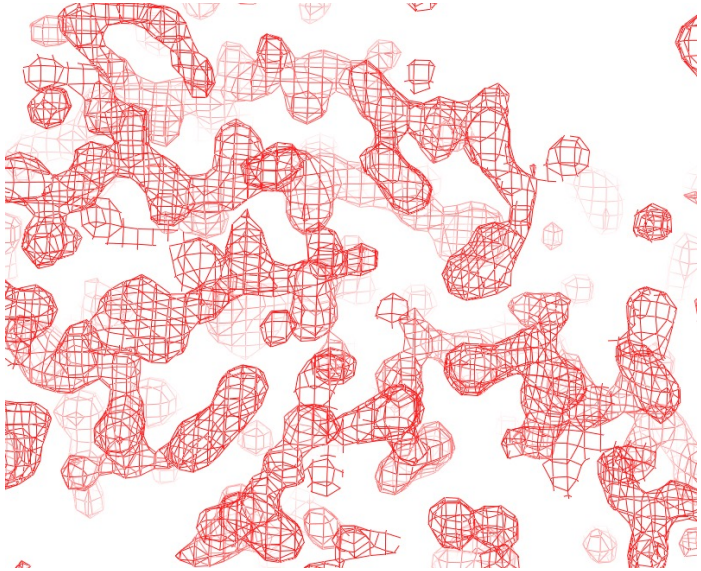
Density modification

Identify local expected density



Noisy map

Find phases consistent with **experiment** and **expected density**



Clear map

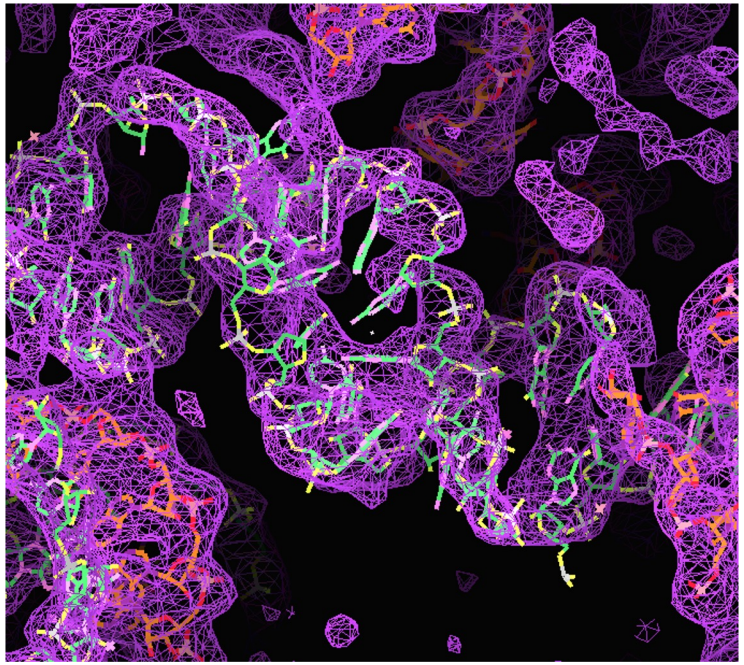
Density everywhere is improved

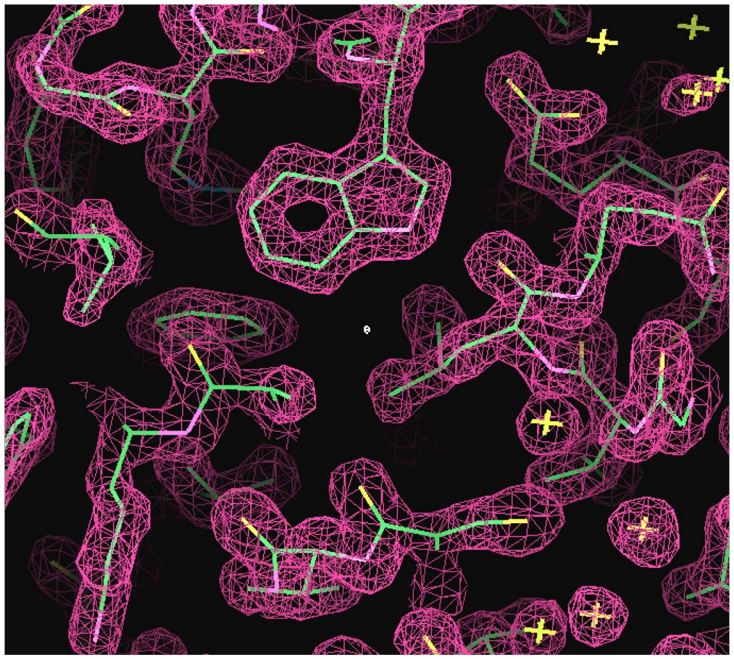
Expected density can include…

- Flat solvent
- NCS
- Density matches model
- Histogram matching
- Connectivity
- Expected shapes

Automated model-building

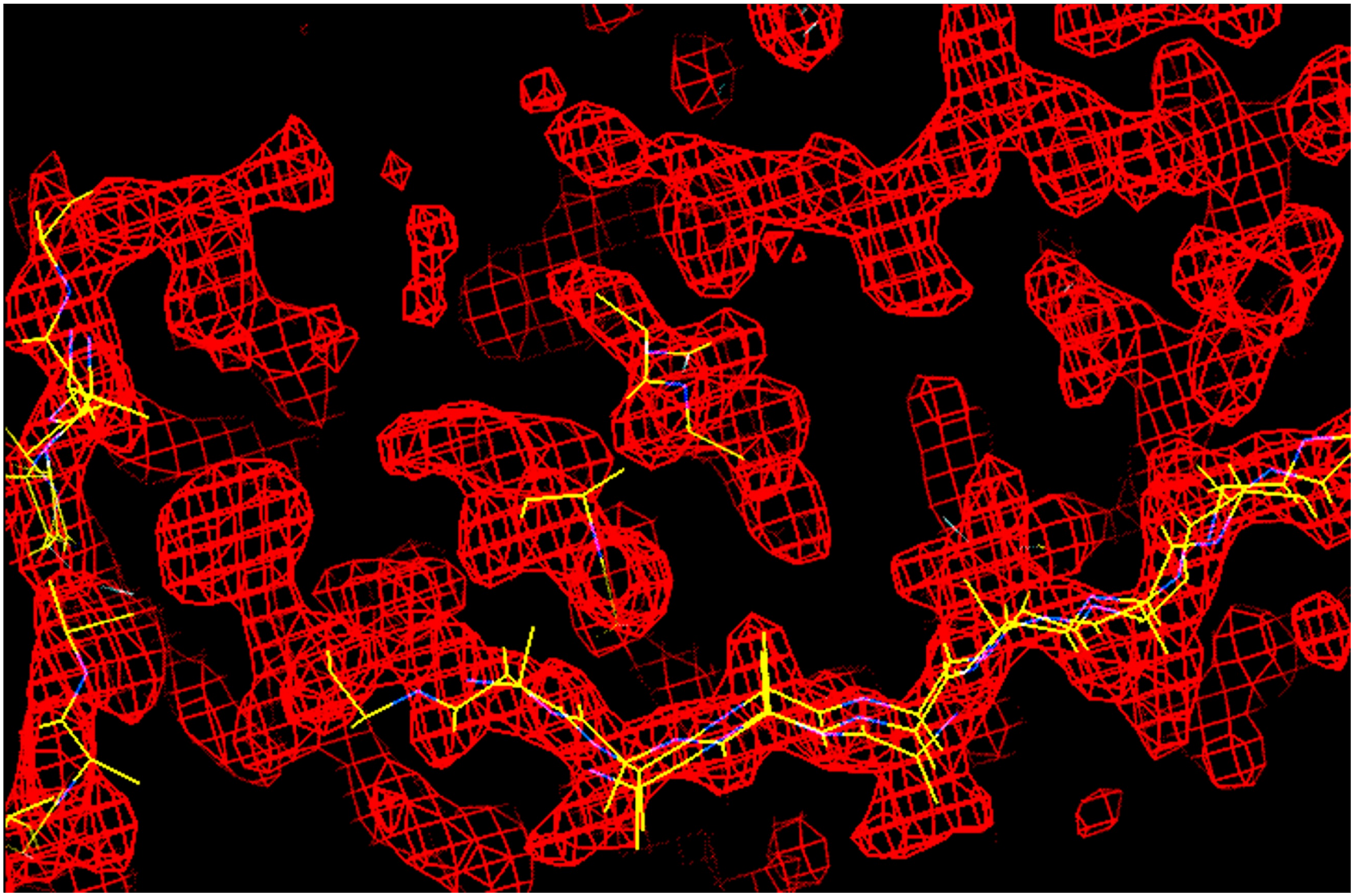
Examples



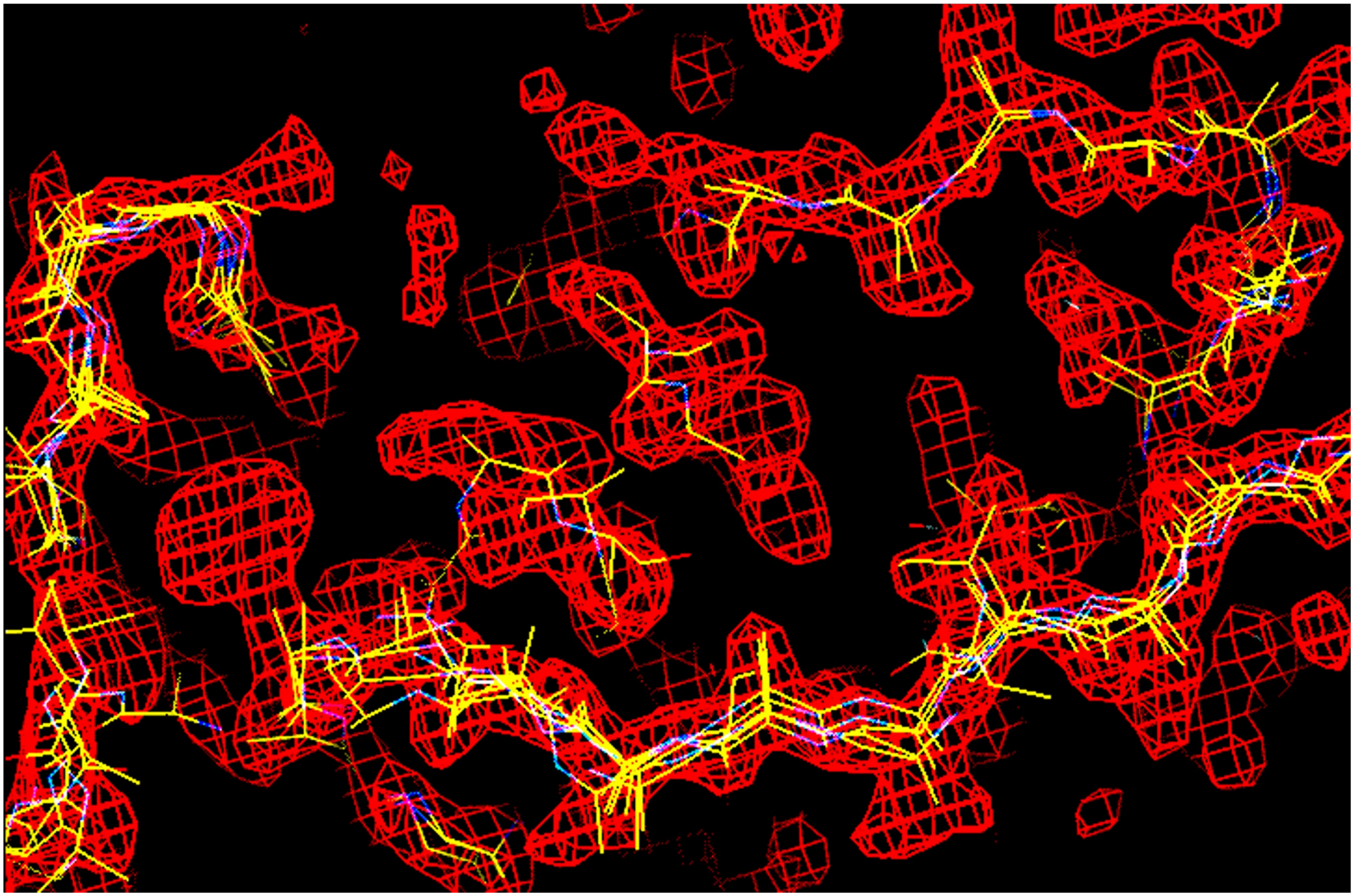


- Shape-based identification of regular secondary structure
- Extension with short fragments from high-resolution structures
- Probabilistic sequence alignment

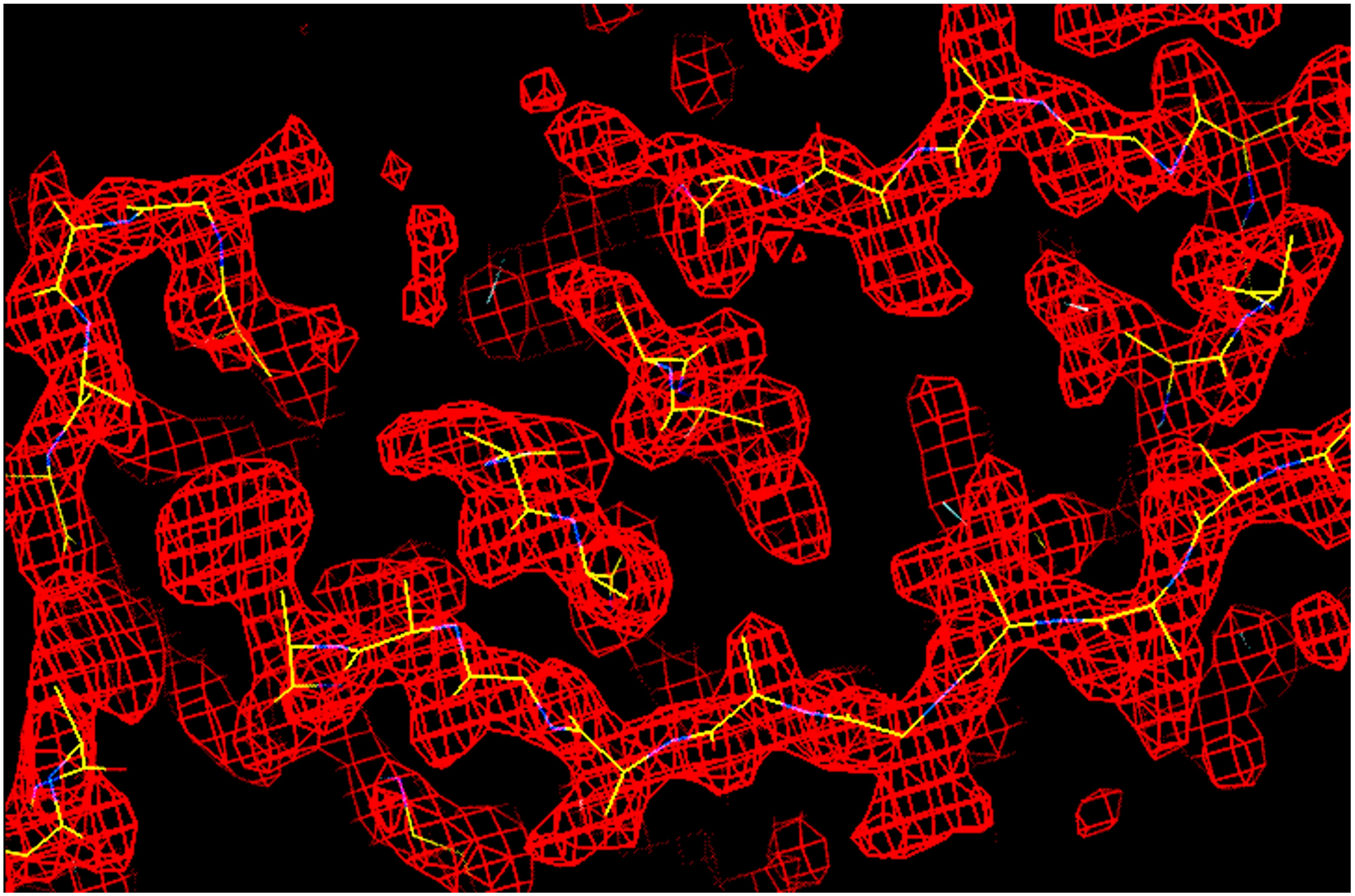
Finding regular protein structure



Extending with short fragments from PDB



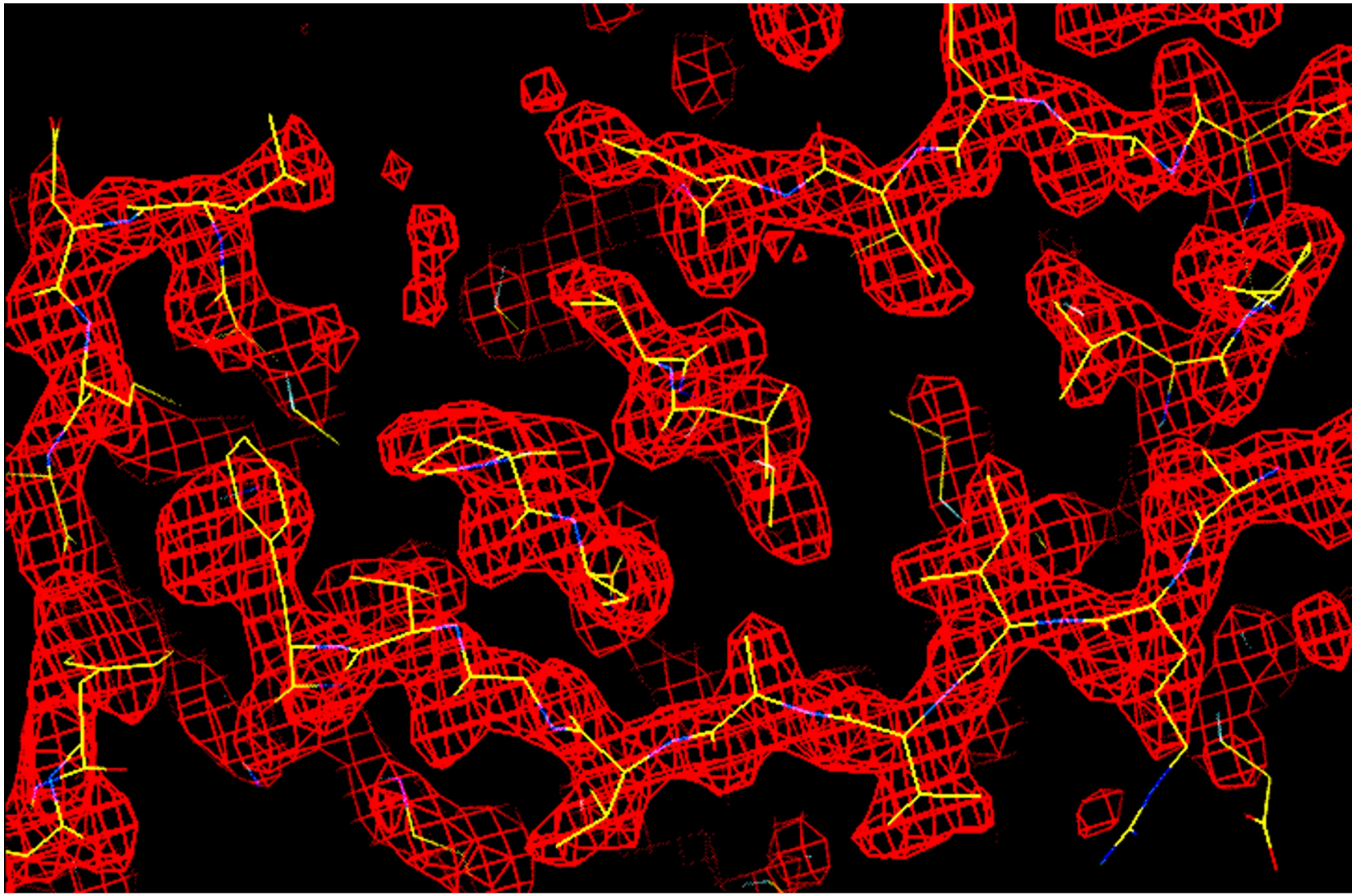
Assembling best model



Identifying residue type at each position

#	G	A	S	V	I	L	M	C	F	Y	K	R	W	H	E	D	Q	N	P	T
1	6	5	4	18	18	6	1	1	1	2	6	2	2	1	9	6	1	0	1	4
2	4	11	14	37	5	2	0	2	0	0	2	3	0	0	1	2	0	0	0	6
3	11	23	5	12	5	3	2	0	1	3	7	3	1	0	5	3	2	0	2	2
4	7	9	6	16	8	5	2	0	1	3	8	4	1	0	7	6	2	0	3	4
5	31	7	3	7	4	2	1	0	1	3	5	4	1	0	6	2	2	0	11	1
6	1	3	3	41	14	8	0	0	0	0	2	1	0	0	2	4	0	0	1	9
7	0	0	0	0	0	0	0	0	15	63	1	0	17	1	0	0	0	0	0	0
8	2	3	6	23	10	6	2	1	0	1	4	3	0	0	5	16	1	0	1	6
9	96	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Inserting side chains based on sequence



Automated structure solution

`phenix.autosol`

Experimental data, sequence, anomalously-scattering atom, wavelength(s)

Find heavy-atom sites with direct methods or likelihood (HYSS)

Calculate phases (Phaser/Solve)

Improve phases, find NCS, build model (phase_and_build)

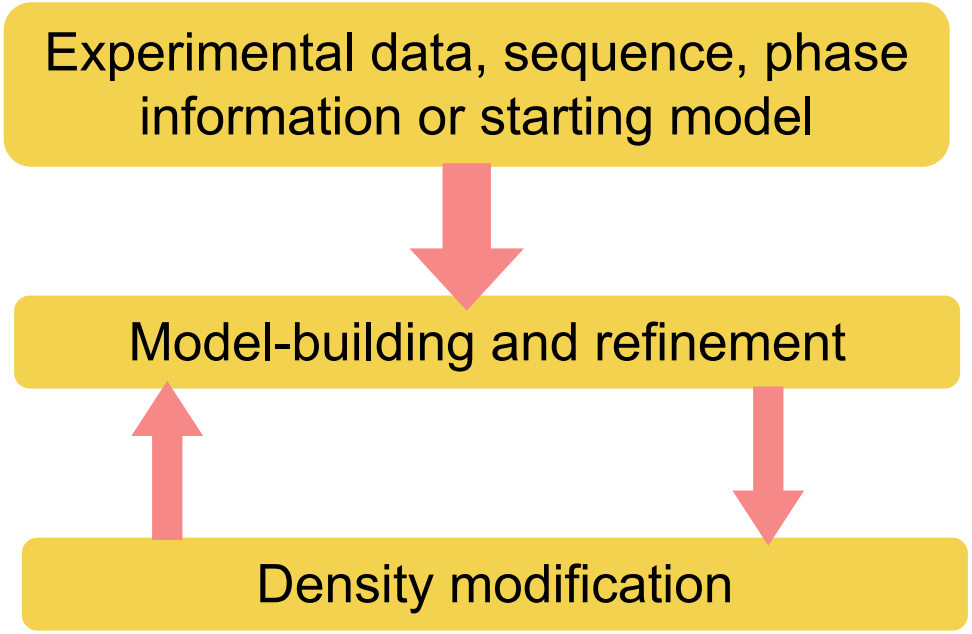
Decision to be made:

Multiple solutions, different derivatives or wavelengths

Alternative hands of space-group and substructure

Iterative map and model improvement

`phenix.autobuild`



- Resolve building
- Secondary-structure only
- Connect chains
- Fit loops
- Build outside model

The Phenix Project

Lawrence Berkeley Laboratory

Paul Adams, Pavel Afonine, Dorothee Liebschner, Nigel Moriarty, Billy Poon, Oleg Sobolev, Christopher Schlicksup

University of Cambridge

Randy Read, Airlie McCoy, Rob Oeffner

Los Alamos National Laboratory
New Mexico Consortium

Tom Terwilliger, Li-Wei Hung

UTHealth

Matt Baker

Duke University

Jane Richardson, Vincent Chen, Michael Prisant, Christopher Williams,

An NIH/NIGMS funded Program Project

Liebschner D, *et al*., Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in *Phenix*.
Acta Cryst. 2019 **D75**:861–877