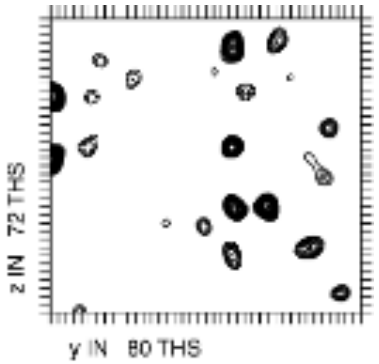
From Phases to Model

Pavel Afonine

Lawrence Berkeley National Laboratory (LBNL)

September 26th, 2024
BNL

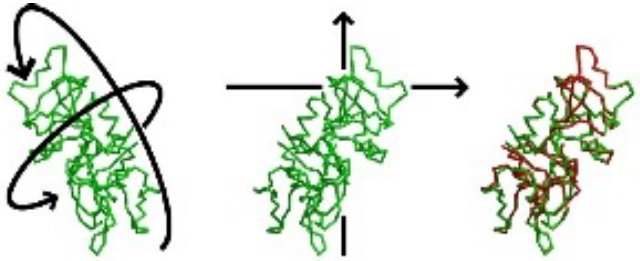
Phasing = How to recover phases



Experimentally

Exploit properties of a few special atoms:

- Anomalous scattering
- A large number of electrons

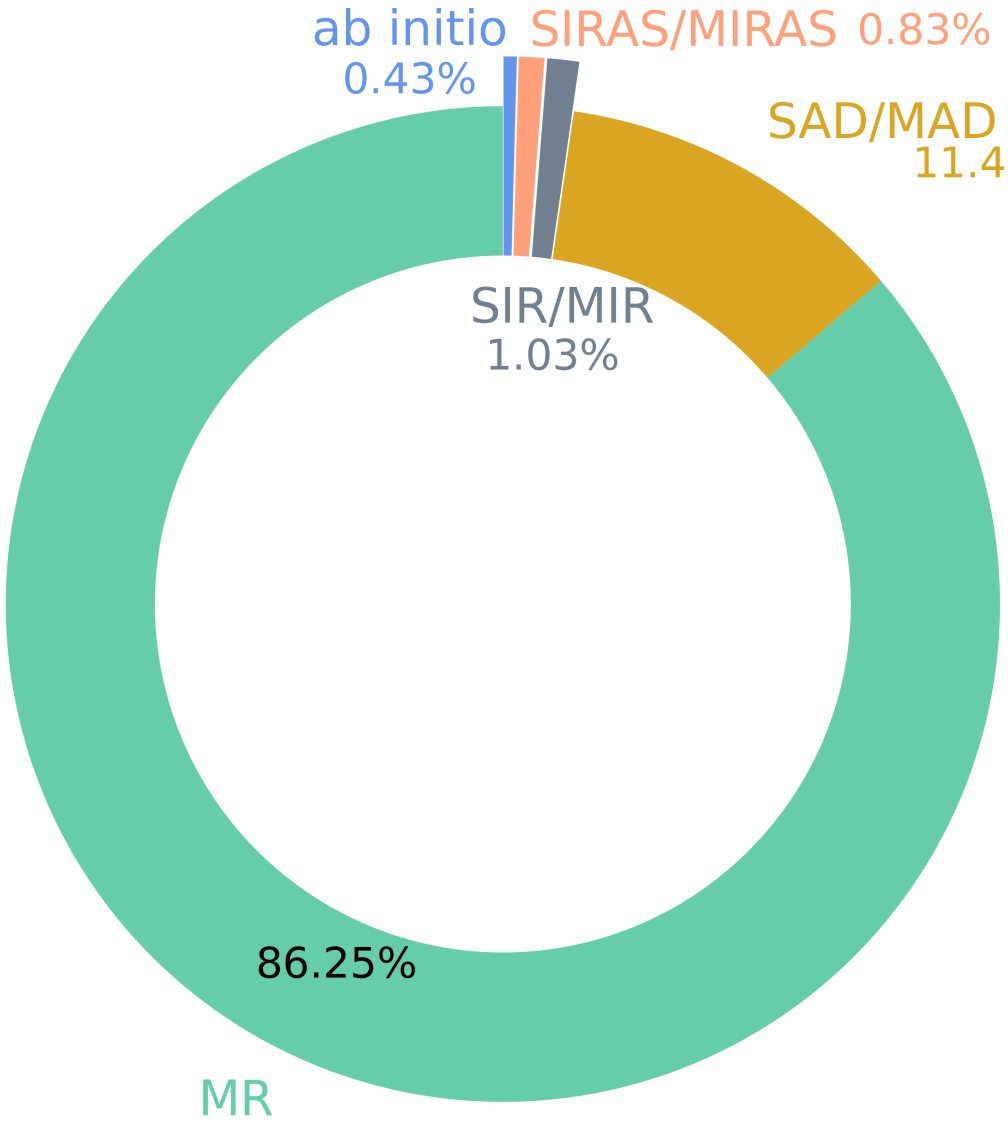


Computationally

- *Molecular Replacement (MR)*
  A previously known structure provides initial phase estimates for a new structure.

- *Direct Methods*
  Phase relationships can be formulated by assuming the positivity and atomicity of the electron density.

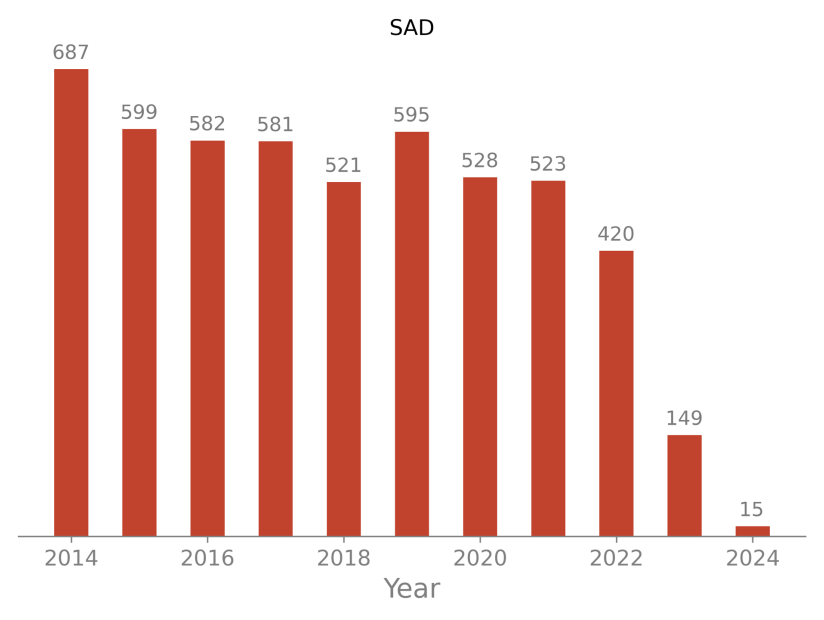
Phasing methods in the PDB



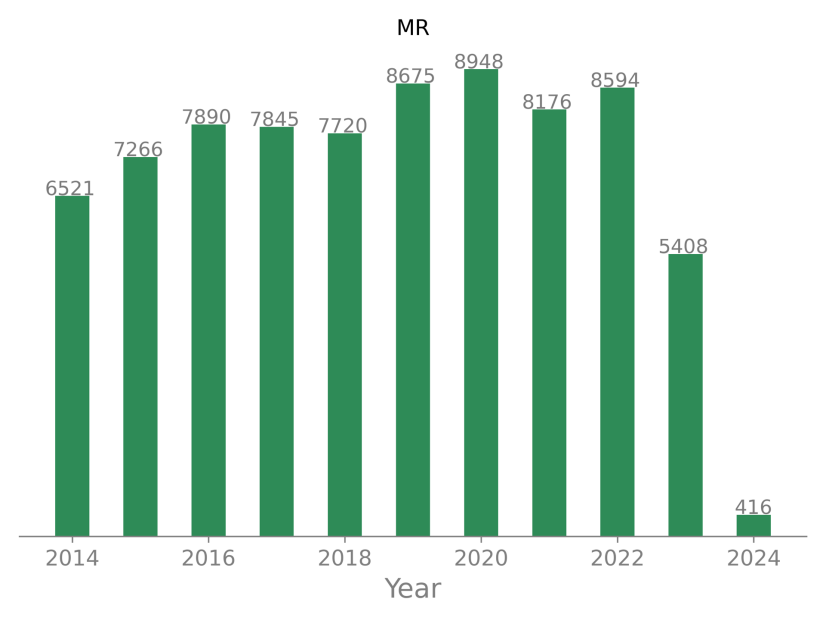
Note: Not all models in the PDB have (correct) info

Phasing methods in the PDB

SAD



Molecular replacement



- Less experimental phasing
- More and more MR structures
- Trend will continue with predicted models (AlphaFold etc.)

Some limitations of predicted models

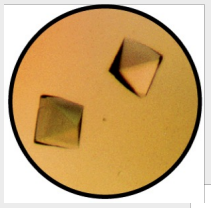
- Predicted models are great to jumpstart structure determination
- Limitations:

Only protein	→	No water, ions, covalent modifications, carbohydrates, ligands, DNA, RNA
Little information about residues that are far apart	→	Models may have distortions and incorrect domain relationships
Trained on good and poor structures	→	Parameters may systematically include poor geometry
Even high pLDDt residues may be incorrect	→	Model does not yield MR solution

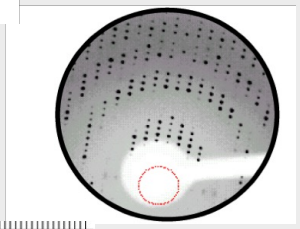
→ We still need experimental phasing

Solving a structure with SAD phasing

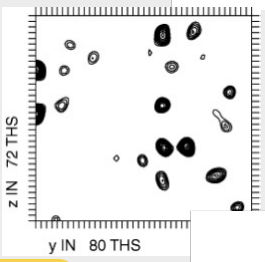
Crystals (SeMet)



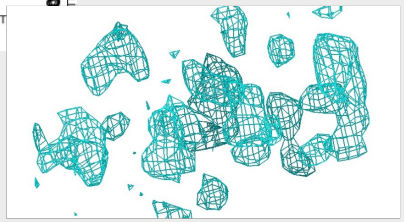
Collect anomalous SAD data



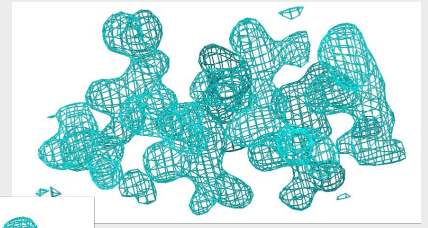
Locate Se atoms



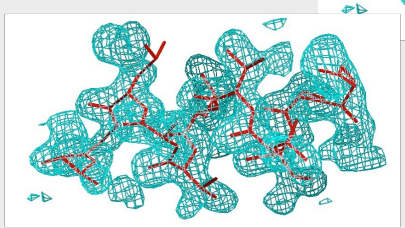
Phasing (calculate density map)



Density modification (improve map)



Model building



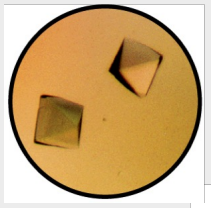
Solving a structure with SAD phasing (Se)

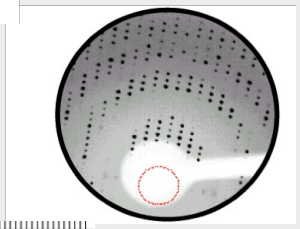
Planning the experiment

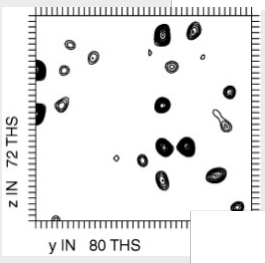
Automating the analysis

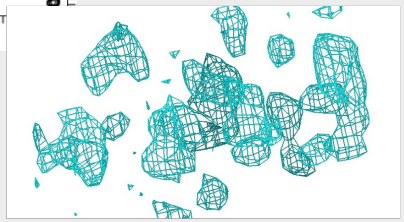
Improving the map

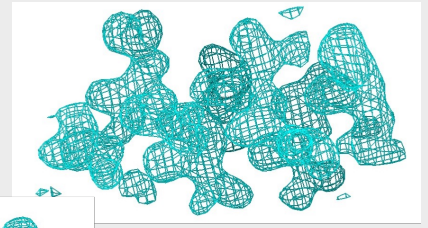
Building a model

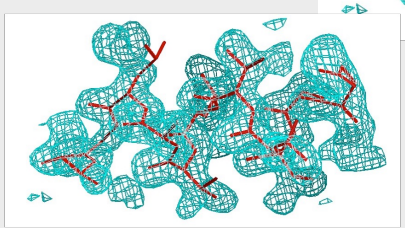












Will I find the anomalous substructure?

How many sites?

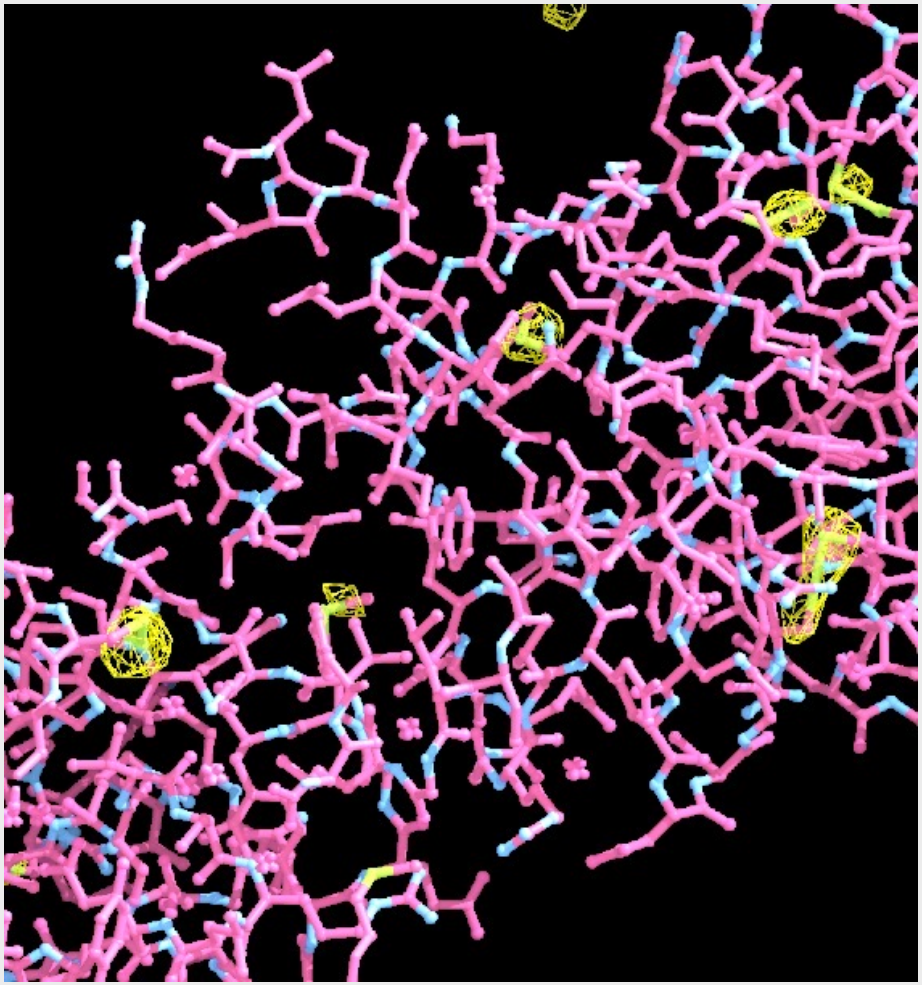
Are sites ordered?

Anomalous atom?

Wavelength?

Accurate data?

How many reflections?

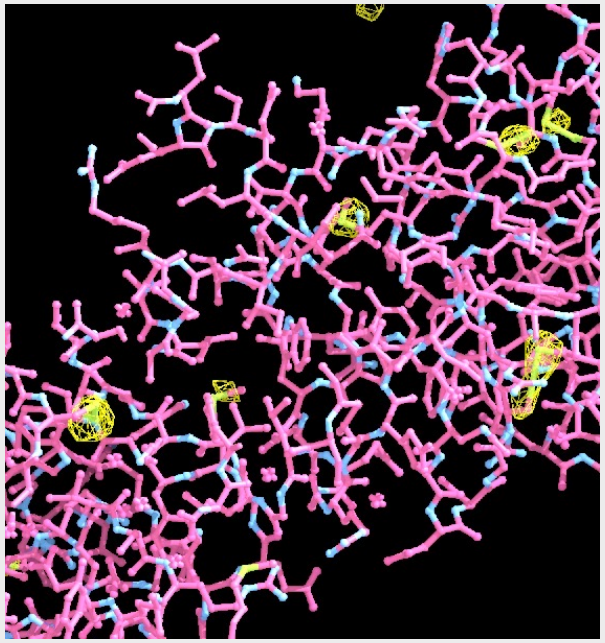


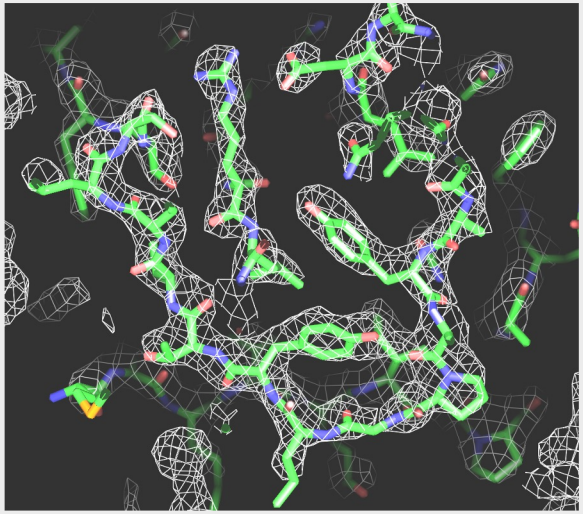
→ Plan SAD experiment: Simulate SAD experiment to find the probability that you are going to solve your structure

Key steps for SAD phasing

1. Find the substructure

Anomalous signal S





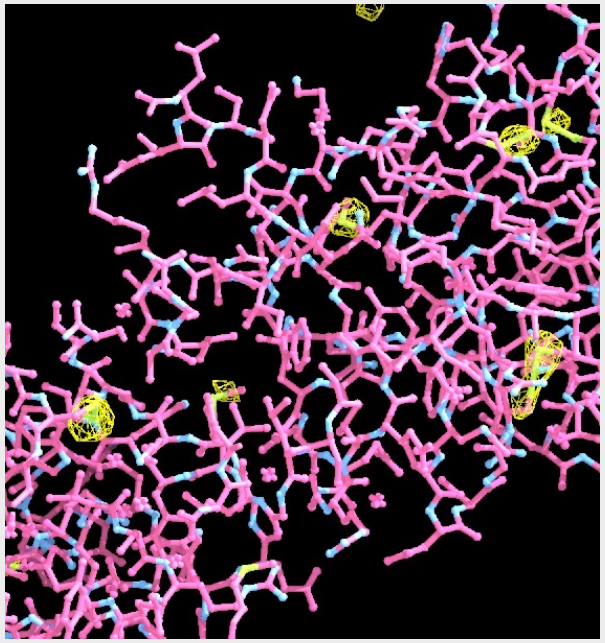
2. Calculate an interpretable map

Anomalous correlation CC_{ano}

Anomalous signal

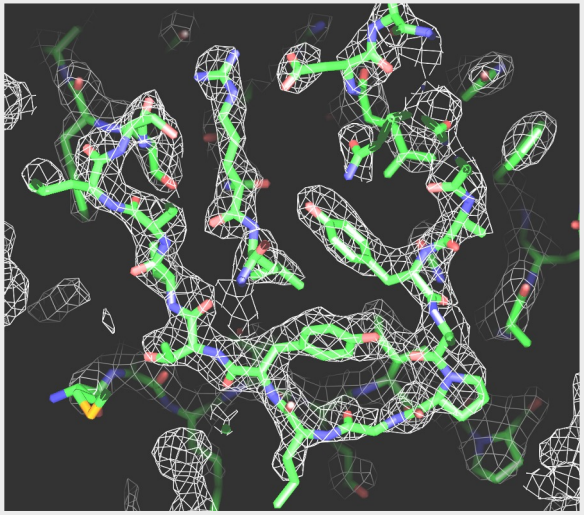
1. Find the substructure

Anomalous signal S



- Peak height in anomalous difference Fourier
- “Information per site”
- Chances are you find the substructure if S > 10

Anomalous correlation



2. Calculate an interpretable map

Anomalous correlation CC_{ano}

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

Anomalous signal: key to finding substructure

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?

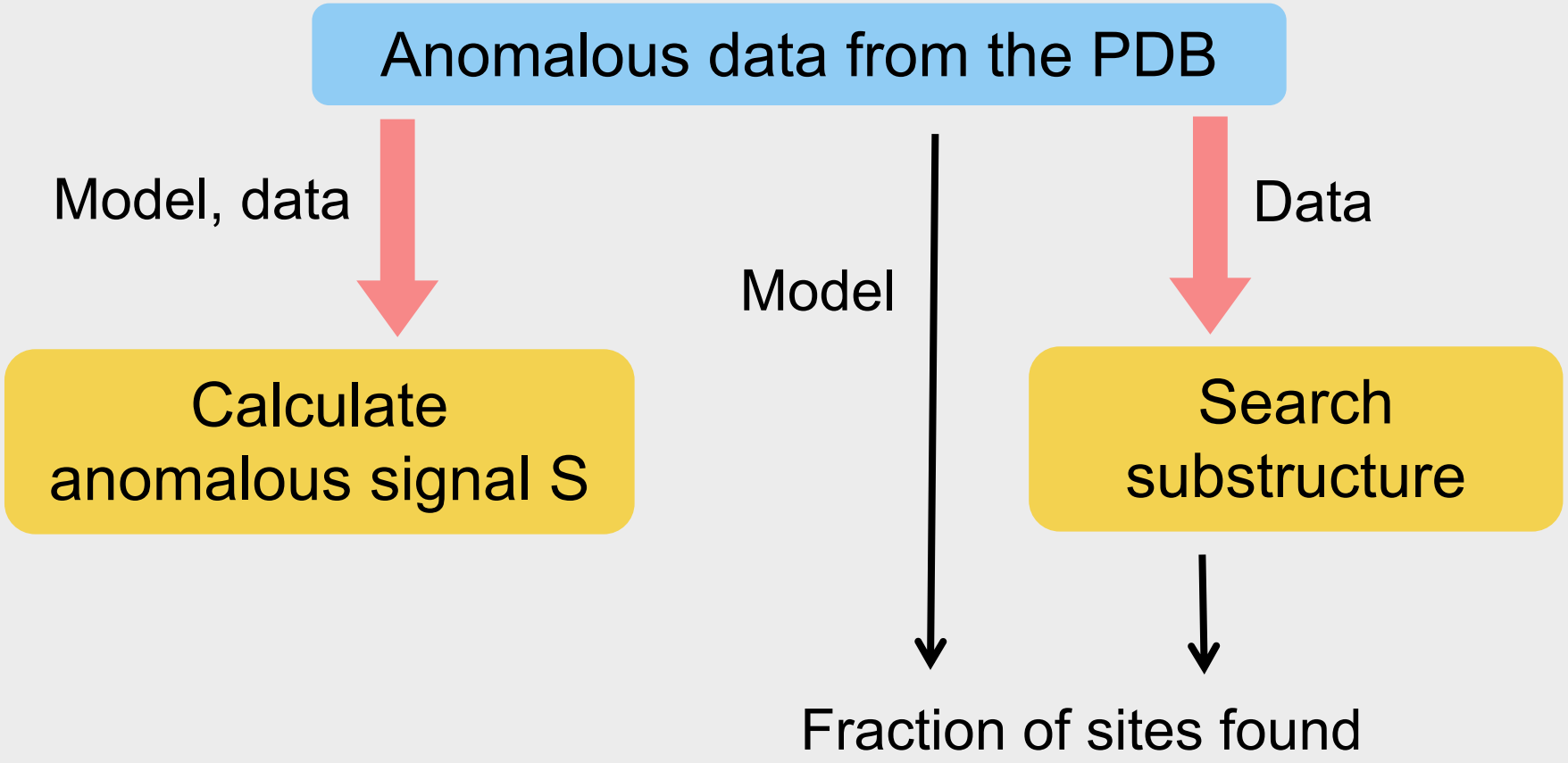
Number of sites

B-value for anomalous sub-structure

Can I solve my structure by SAD phasing? Anomalous signal in SAD phasing. Terwilliger TC, Bunkóczi G, Hung L-W, Zwart PH, Smith JL, Akey DL, Adams PD *Acta Cryst.* **D72**, 346-358 (2016).

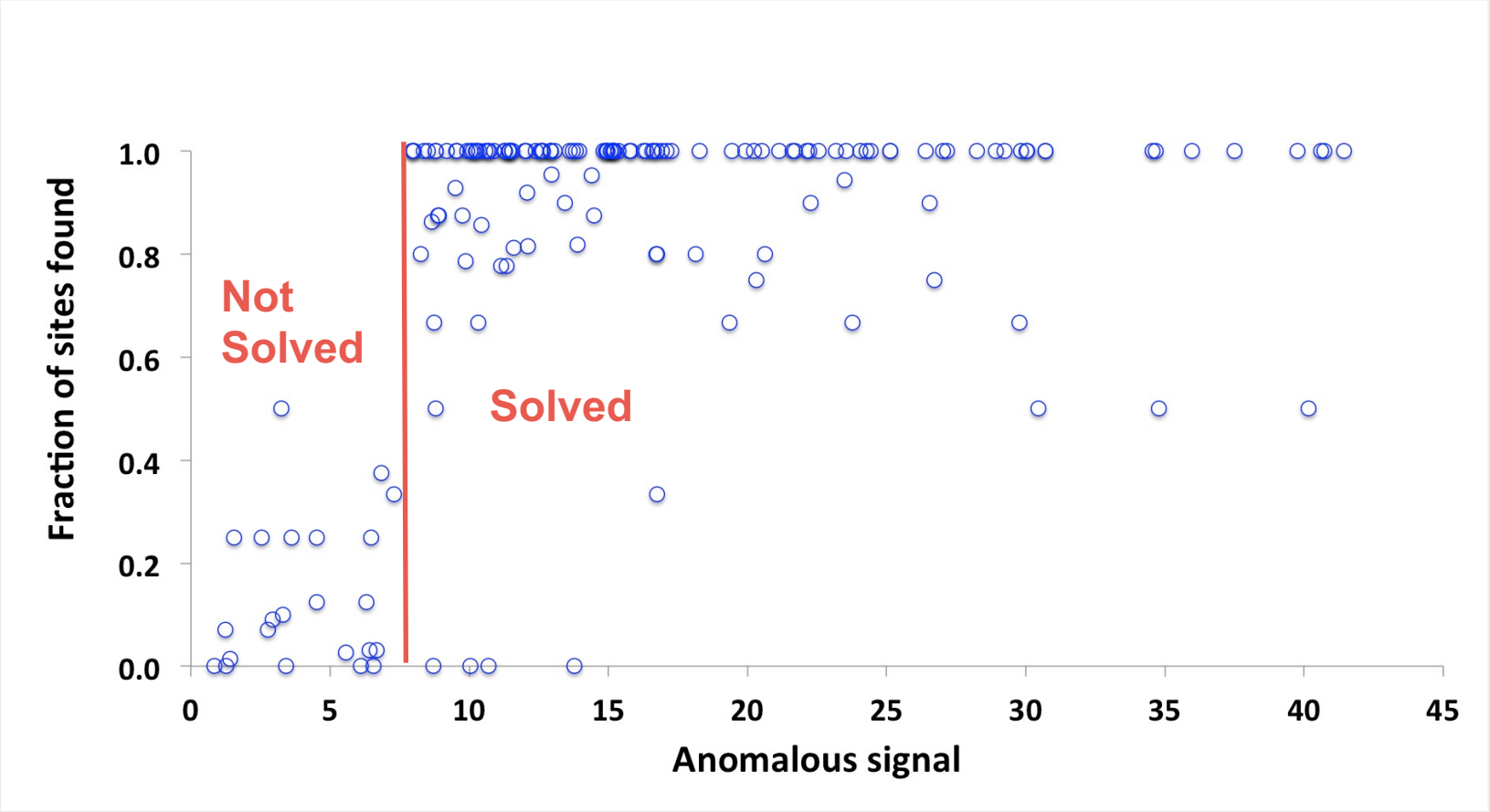
Anomalous signal: key to finding substructure

Relationship between the anomalous signal and the solution of the anomalous substructure



Anomalous signal: key to finding substructure

Relationship between the anomalous signal and the solution of the anomalous substructure



Simulating the anomalous signal

I/sigma (accuracy of data)

Anomalous atom (wavelength, f")

Number of sites

Resolution

Sequence

Anomalous signal S

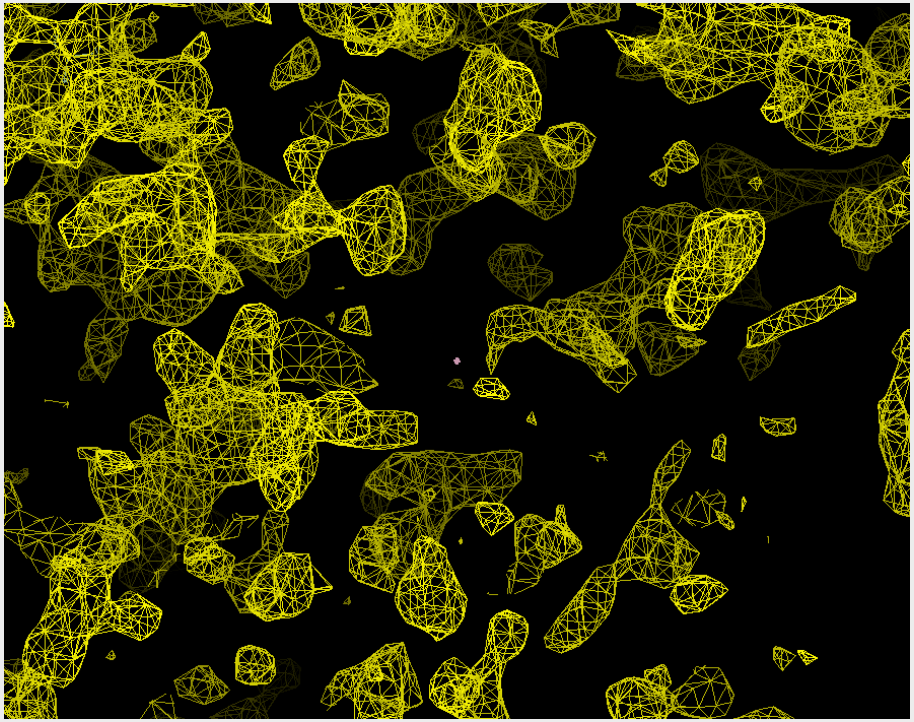
Can I solve my structure by SAD phasing? Planning an experiment, scaling data and evaluating the useful anomalous correlation and anomalous signal. Terwilliger TC, Bunkóczi G, Hung L-W, Zwart PH, Smith JL, Akey D, Adams PD Acta Cryst. D72, 359-374 (2016).

Summary

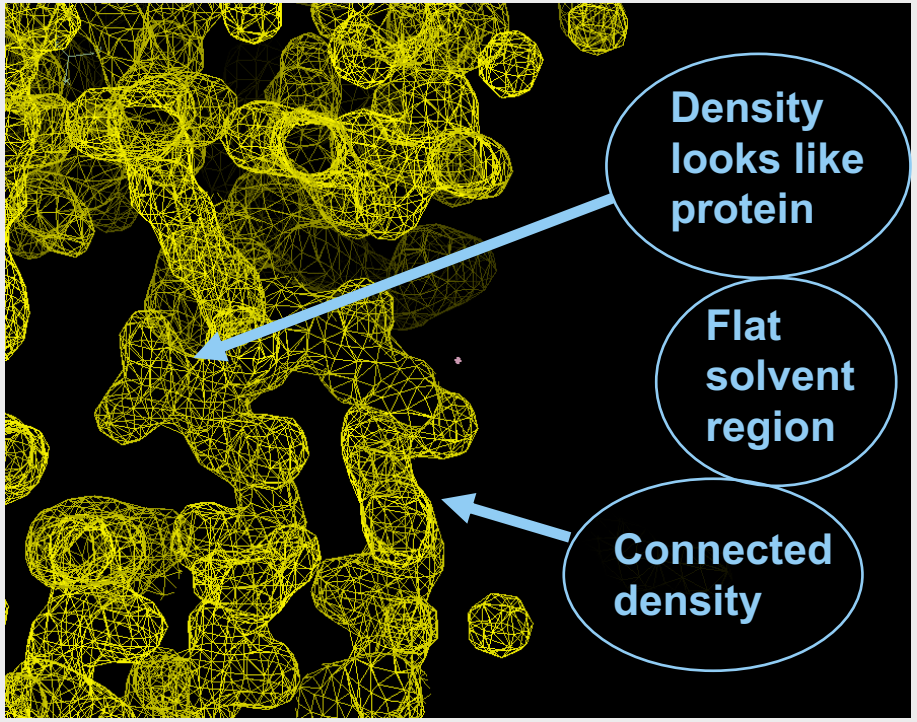
- We can estimate the anomalous signal S from the data
- If S > 10 → substructure is likely to be found
- We can simulate the anomalous signal (before doing the experiment)

Map improvement by density modification

Original map



DM map

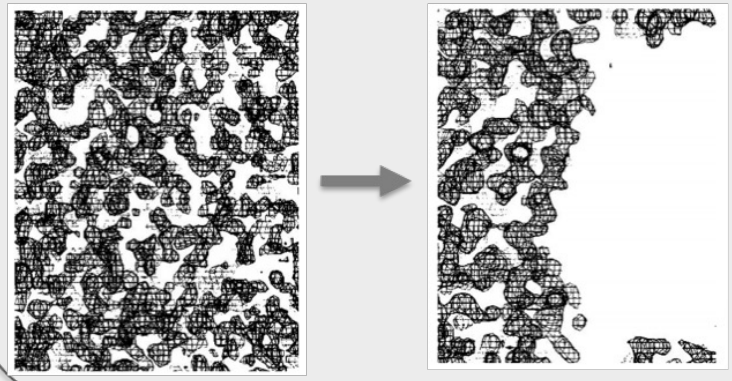


Density modification: change map so it is most consistent with what we know about macromolecules

Density modification

Crystallography

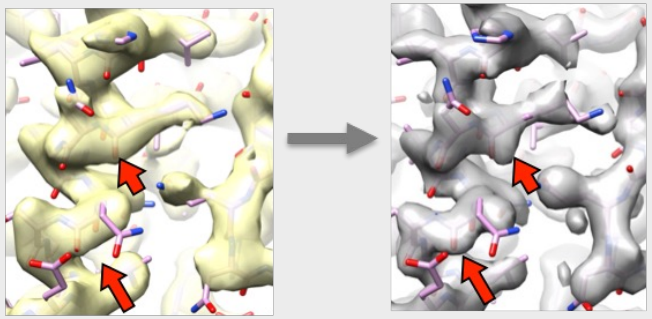
From uninterpretable to interpretable map



Cryo-EM

Effect is less dramatic as in crystallography

- Can increase map resolution (0.05-0.3 Å)
- Can improve map clarity for interpretation



nature methods

ARTICLES

https://doi.org/10.1038/s41592-020-0914-9

Check for updates

Improvement of cryo-EM maps by density modification

Thomas C. Terwilliger[1,2], Steven J. Ludtke[3], Randy J. Read[4], Paul D. Adams[5,6] and Pavel V. Afonine[5]

Automated model building

Automated model building, facts:

- No automated model building produces 100% complete and accurate model
- Produces initial model for further manual building
- The lower the resolution, the less complete and accurate the auto built model

References

phenix-online.org

Phenix documentation

Tutorials with sample data

Video tutorials

www.youtube.com/c/phenixtutorials

Terwilliger, T. C. (2000). *Acta Cryst. D*. **56**, 965–972.

Terwilliger, T. C. (2002). *Acta Cryst. D*. **58**, 1937–1940

Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L.-W., Read, R. J. & Adams, P. D. (2008). *Acta Cryst. D*. **64**, 61–69.

Terwilliger, T. C., Adams, P. D., Read, R. J., McCoy, A. J., Moriarty, N. W., Grosse-Kunstleve, R. W., Afonine, P. V., Zwart, P. H. & Hung, L.-W. (2009). *Acta Cryst. D***65**, 582–601.

Terwilliger, T. C., Bunkóczi, G., Hung, L.-W., Zwart, P. H., Smith, J. L., Akey, D. L. & Adams, P. D. (2016). *Acta Cryst. D***72**, 346–358.

Terwilliger, T. C., Bunkóczi, G., Hung, L.-W., Zwart, P. H., Smith, J. L., Akey, D. L. & Adams, P. D. (2016). *Acta Cryst. D***72**, 359–374.