CBMS Structural Biology Workbenches
September 2024

Phenix

Molecular Replacement

Oleg Sobolev
Lawrence Berkeley Laboratory

Los Alamos NATIONAL LABORATORY EST.1943

BERKELEY LAB

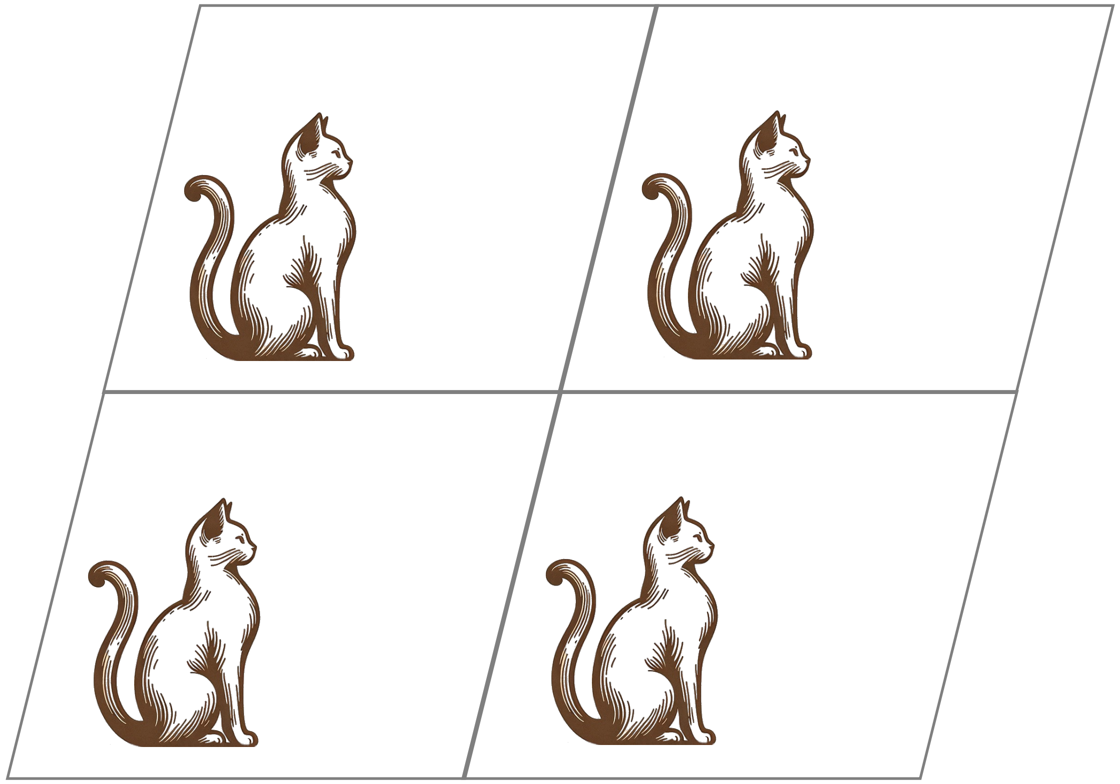
UNIVERSITY OF CAMBRIDGE

Molecular Replacement (MR)

MR = solve the unknown crystal structure of a molecule using a related known molecular model.

Known model

Crystal of unknown structure

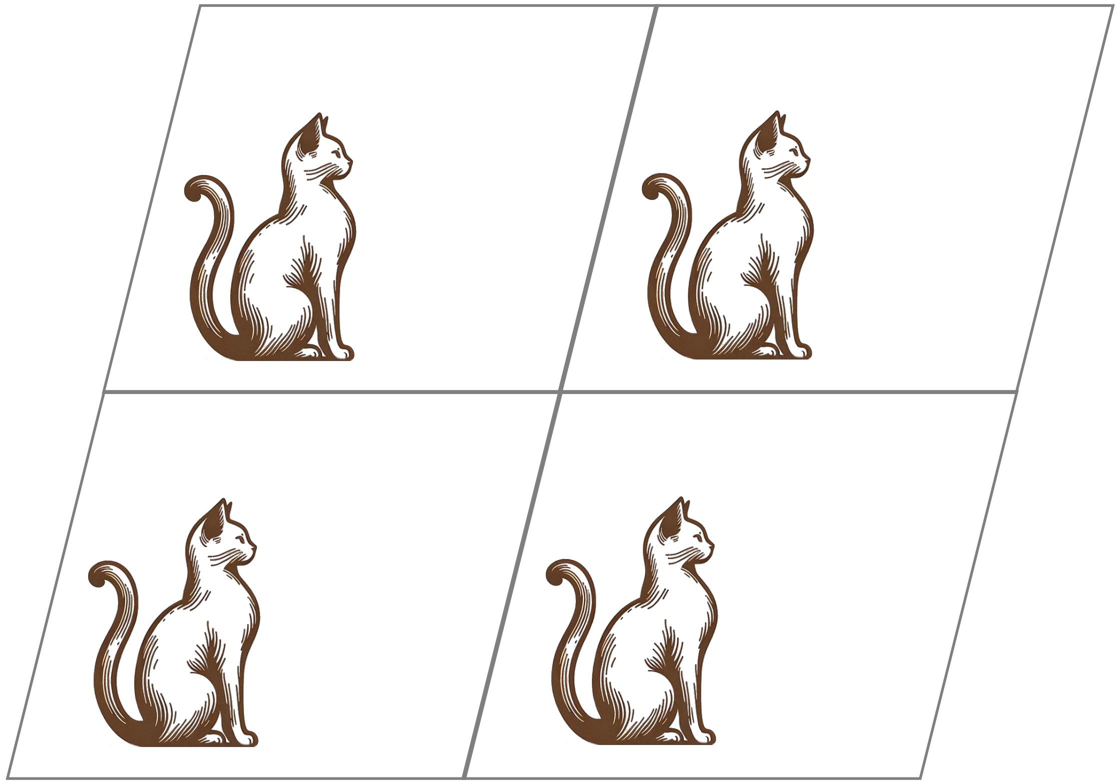


Molecular Replacement (MR)

MR = solve the unknown crystal structure of a molecule using a related known molecular model.

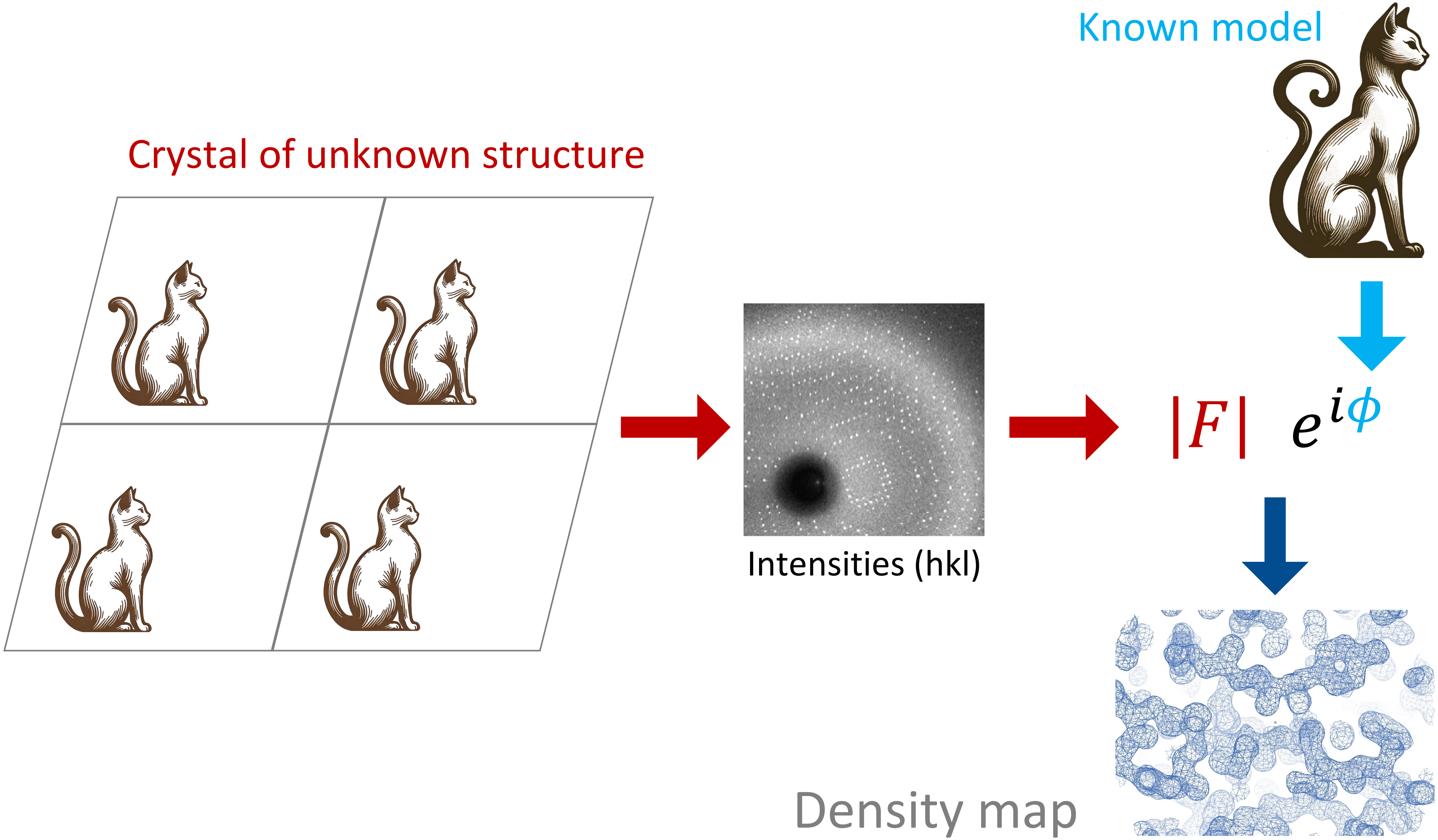
Known model

Crystal of unknown structure



Known model provides initial estimates of the phases of the unknown structure.

Molecular Replacement (MR)



Crystal of unknown structure

Intensities (hkl)

Known model

$|F| \quad e^{i\phi}$

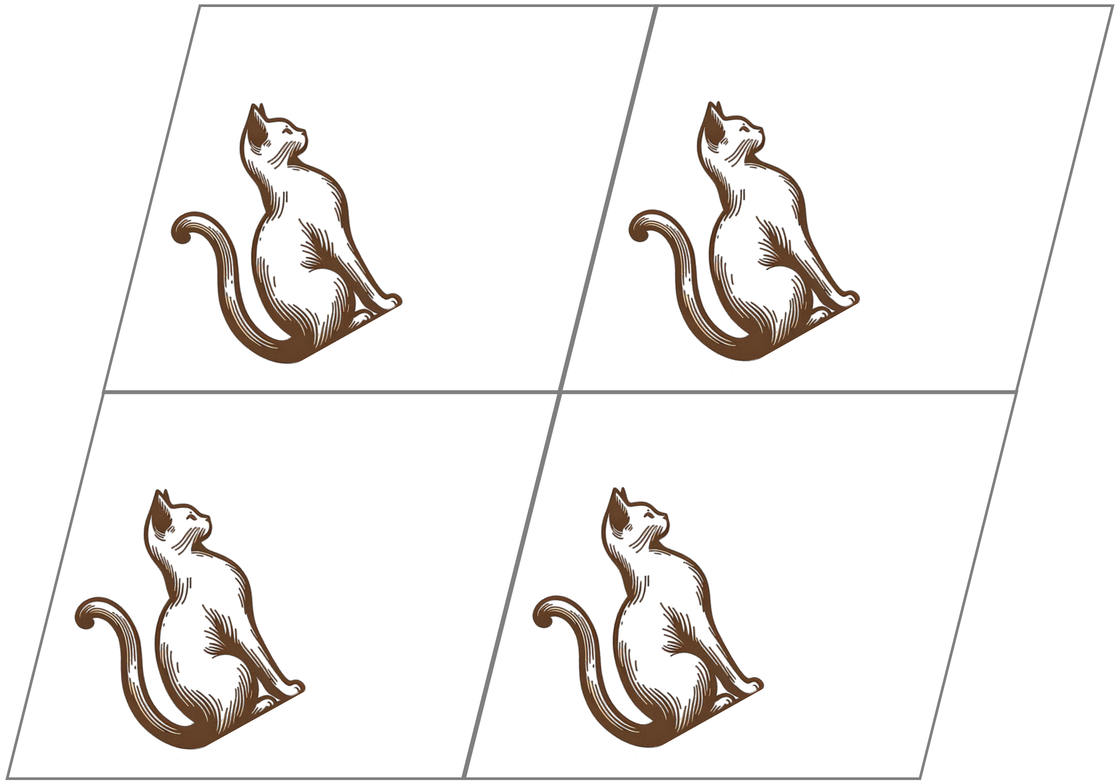
Density map

Molecular replacement: Approach

Try to match the known model with the unknown structure.

Search model

Crystal of unknown structure



Find the rotation and translation of the search model so that it matches the unknown structure.

The search model

- Finding a suitable search model is critical step in MR.
- Should provide a high proportion of the scattering from the target structure with high accuracy (low r.m.s.d.).

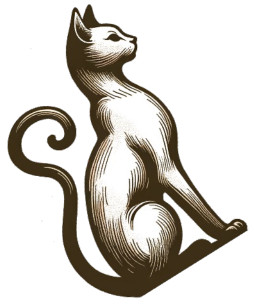
Not similar to the target

Crystal of unknown structure

Molecular replacement: Scoring

Compare observed and calculated diffraction.

Poor score



Good score

Different approaches:

- Patterson function
- Maximum-likelihood Methods (Phaser)

MR Scoring: Maximum Likelihood Method

"For any postulated orientation and position of the model, what is the probability of obtaining the structure amplitudes that we observe?"

Explicitly models errors
- Experimental uncertainties
- r.m.s. coordinate error of the search model

→ Likelihood methods are more robust and generally give clearer solutions in difficult cases

Maximum Likelihood Scoring in Phaser

LLG = Log Likelihood gain

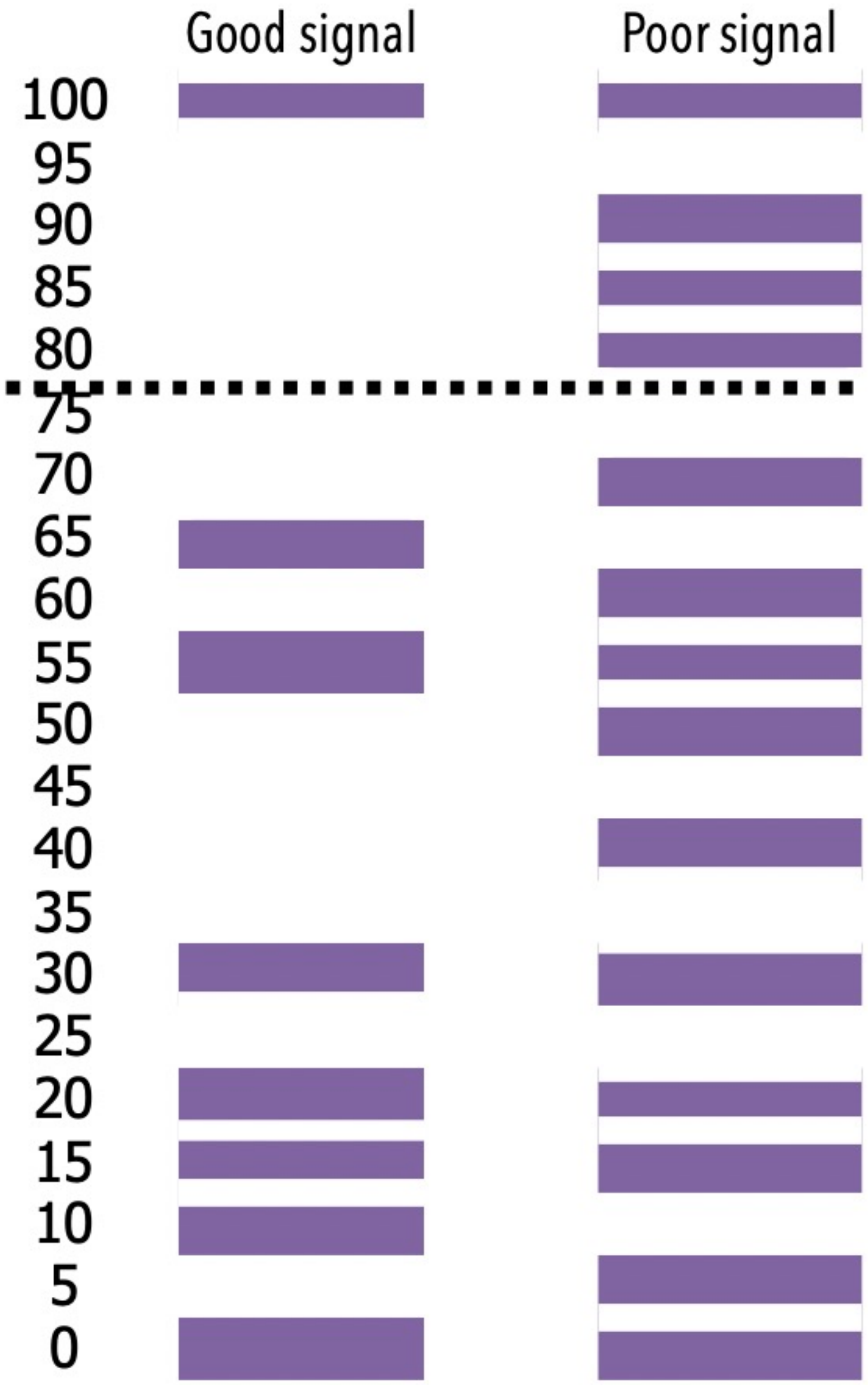
→ It measures how much better the data can be predicted with the search model than with a random distribution of the same atoms.

TF-Z = how many standard deviations your solution is above the mean (the higher the better).

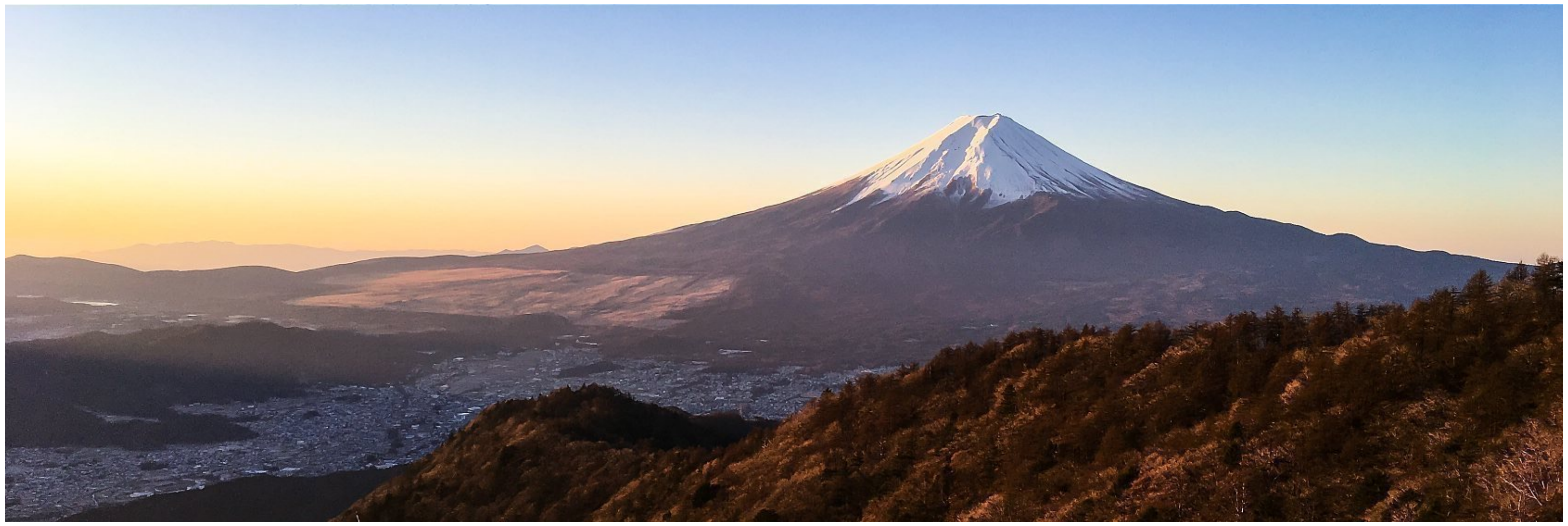
Maximum Likelihood Scoring in Phaser

Select solutions that are over 75% of the difference between the top peak and the mean.

- Good signal, few potential solutions
- Poor signal, many potential solutions

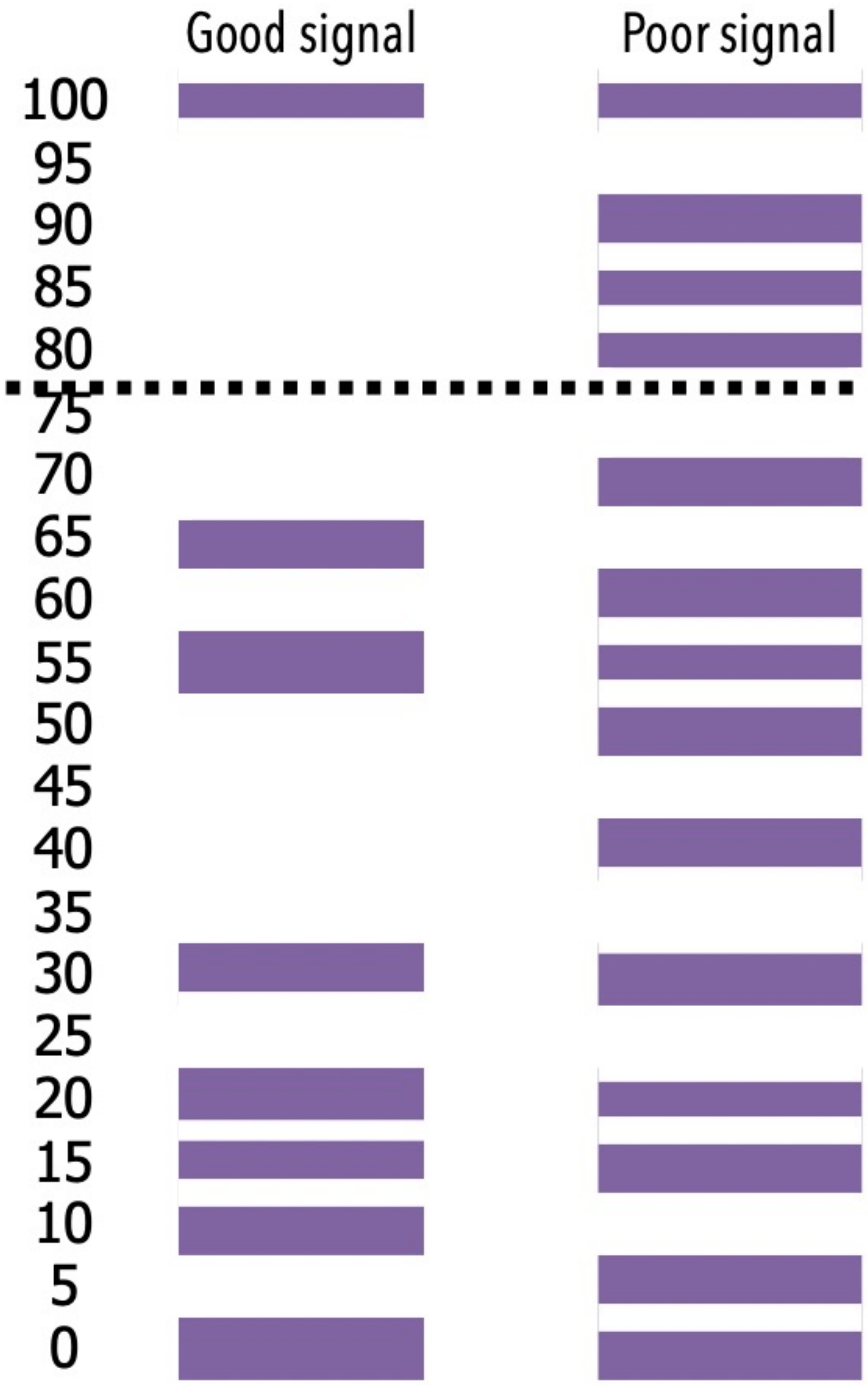


Maximum Likelihood Scoring in Phaser



Good signal

Poor signal



3) The scoring function: ML Method

TF Z-score	LLG score	Solved?
< 5	< 25	no
5 - 6	25 - 36	unlikely
6 - 7	36 - 49	possibly
7 - 8	49 - 64	probably
> 8	> 64	definitely

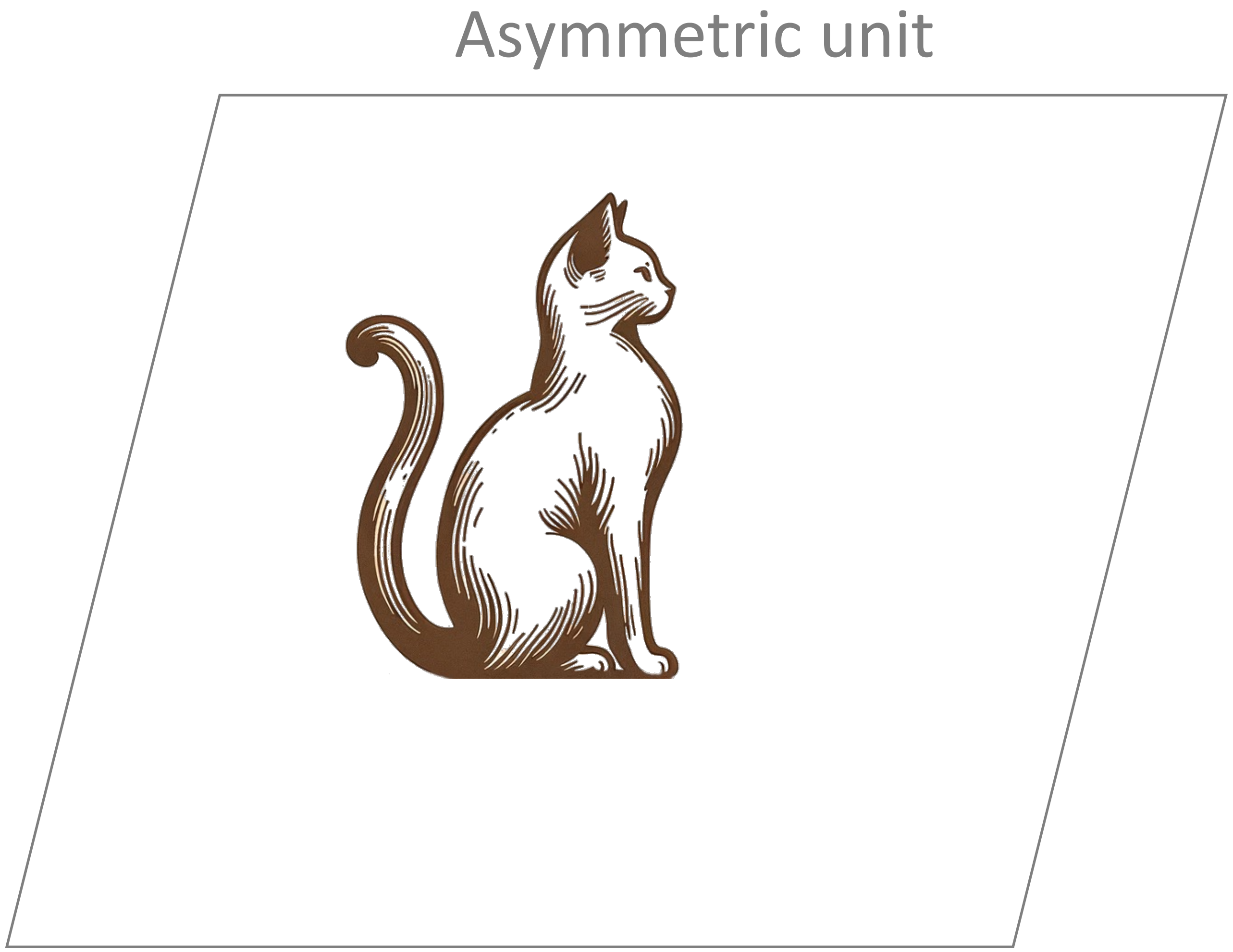
What is needed to run Phaser MR

- Reflection data
- Search model
- Error estimation of the search model
  - Homologue: sequence identity
  - Predicted model: r.m.s.d. (1.0 Å)
- ASU content
- Twinning
- (tNCS)

ASU content

- Sequence of your construct
- How many copies of the the molecule(s)?

N = 1

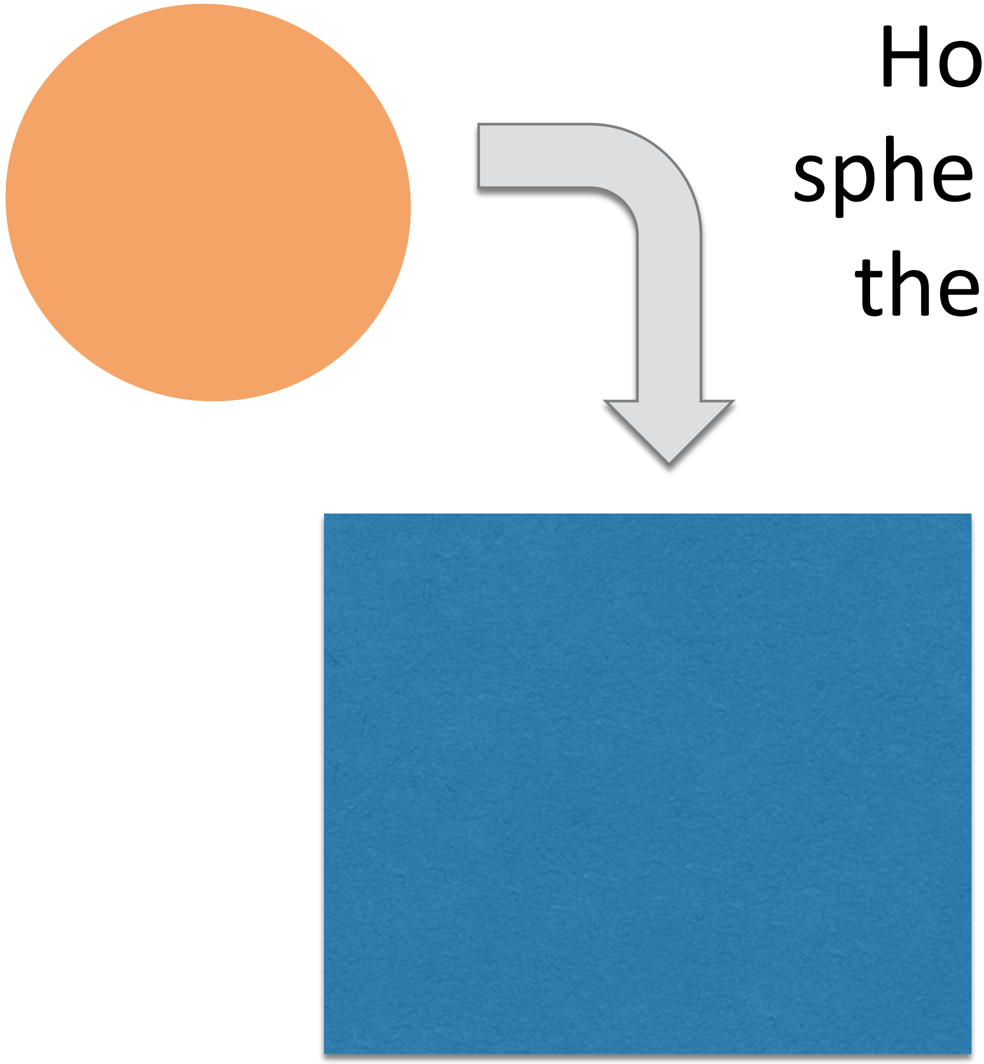


Asymmetric unit

One copy in the asymmetric unit.

Estimating the number of molecules in the ASU

Matthews coefficient: $$V_M = \frac{\text{volume of asymmetric unit}}{\text{molecular weight}}$$

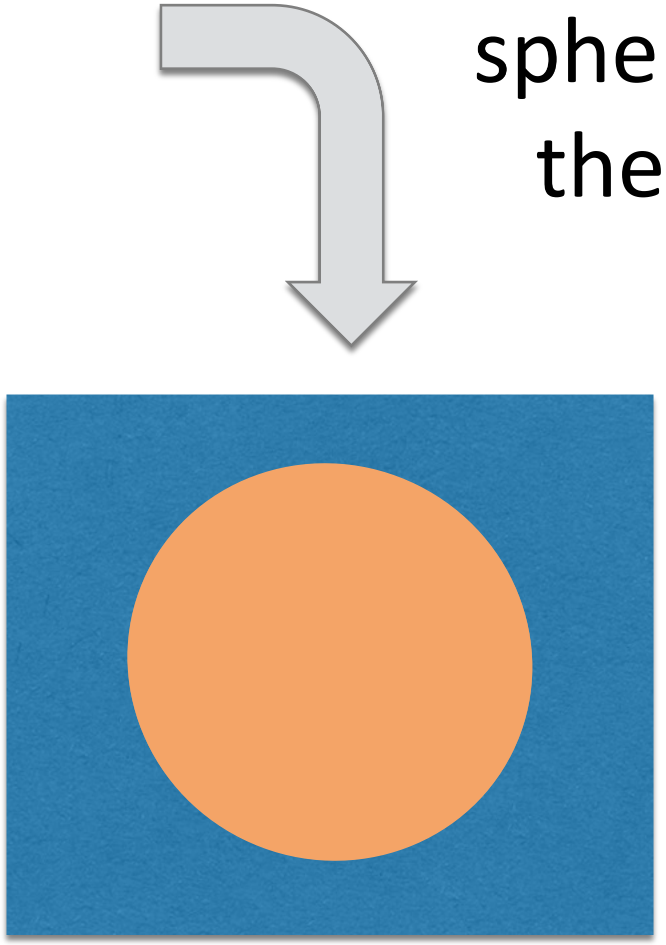


How many spheres fit into the square?

Estimating the number of molecules in the ASU

Matthews coefficient: $$V_M = \frac{\text{volume of asymmetric unit}}{\text{molecular weight}}$$

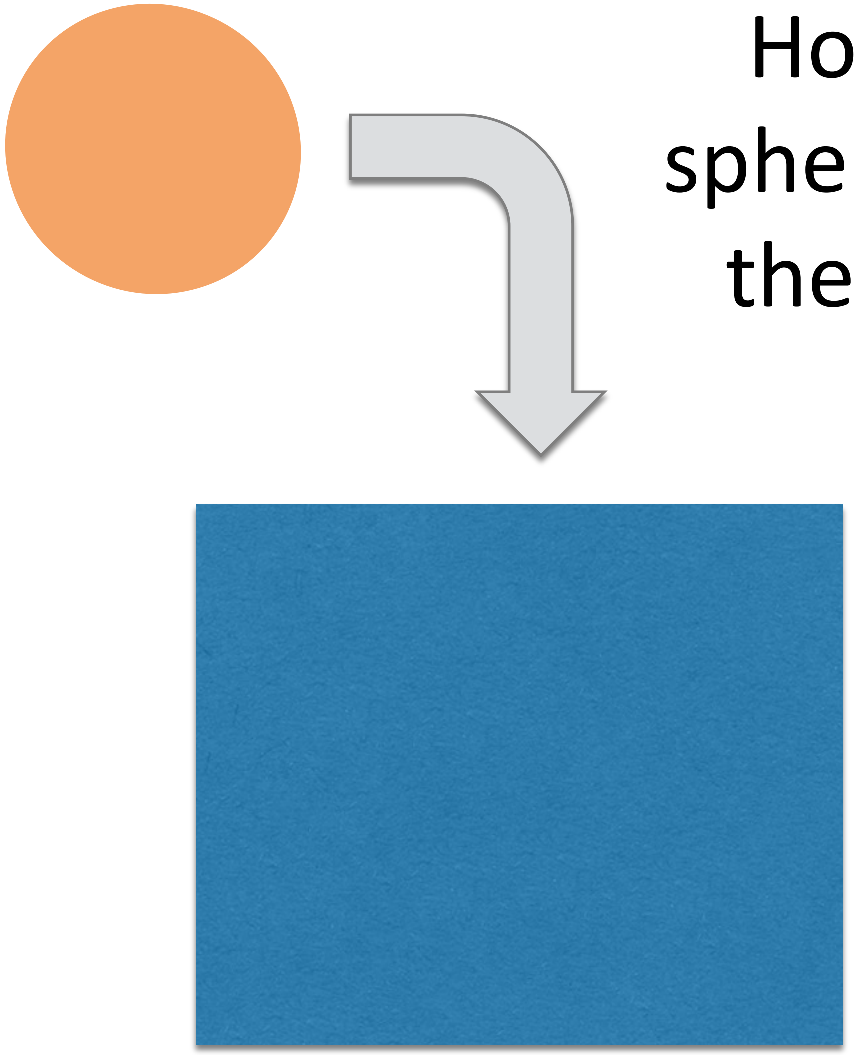
How many spheres fit into the square?



Clearly only one sphere

Estimating the number of molecules in the ASU

Matthews coefficient: $V_M = \frac{\text{volume of asymmetric unit}}{\text{molecular weight}}$

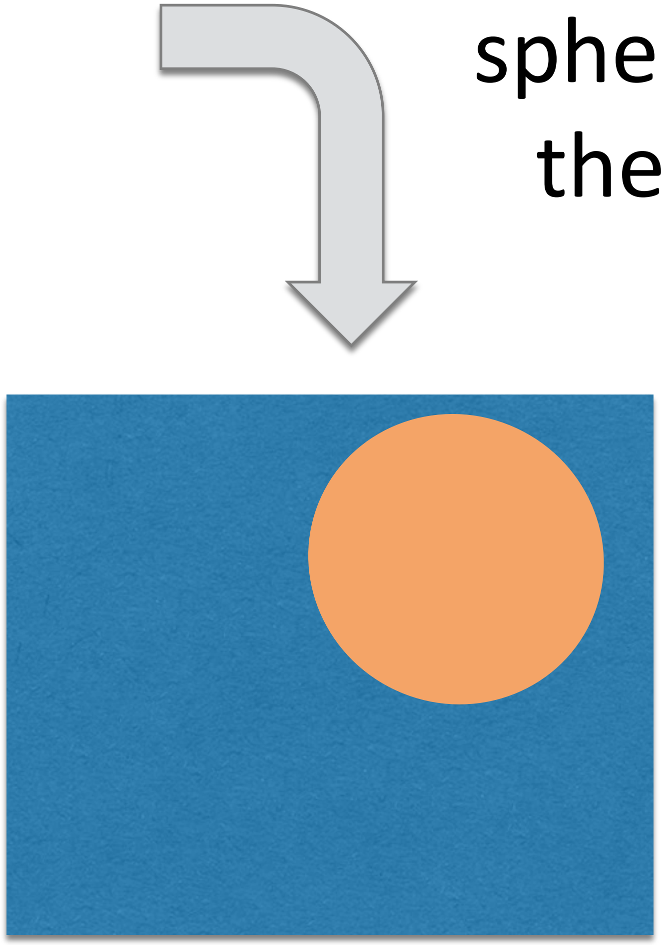


How many spheres fit into the square?

Estimating the number of molecules in the ASU

Matthews coefficient: $V_M = \dfrac{\text{volume of asymmetric unit}}{\text{molecular weight}}$

How many spheres fit into the square?

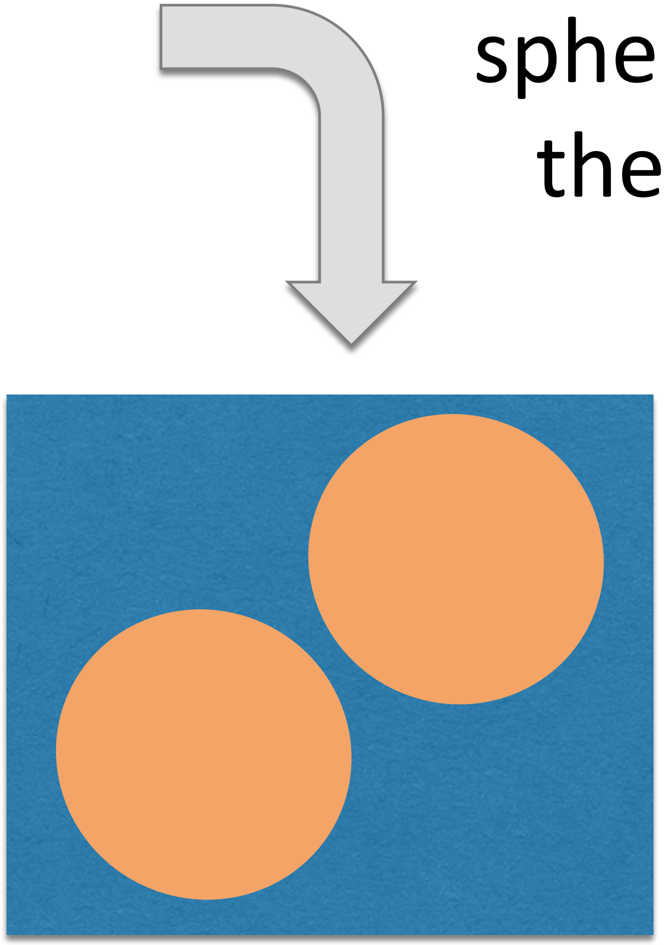


There could be only one…

Estimating the number of molecules in the ASU

Matthews coefficient: $$V_M = \frac{\text{volume of asymmetric unit}}{\text{molecular weight}}$$

How many spheres fit into the square?

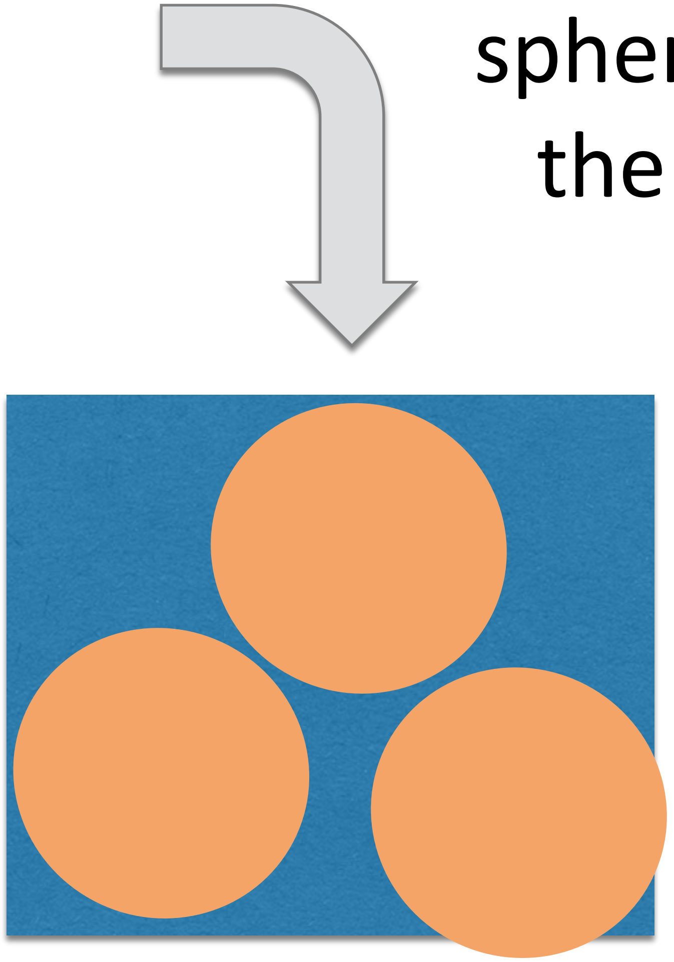


… or maybe two

Estimating the number of molecules in the ASU

Matthews coefficient: $V_M = \dfrac{\text{volume of asymmetric unit}}{\text{molecular weight}}$

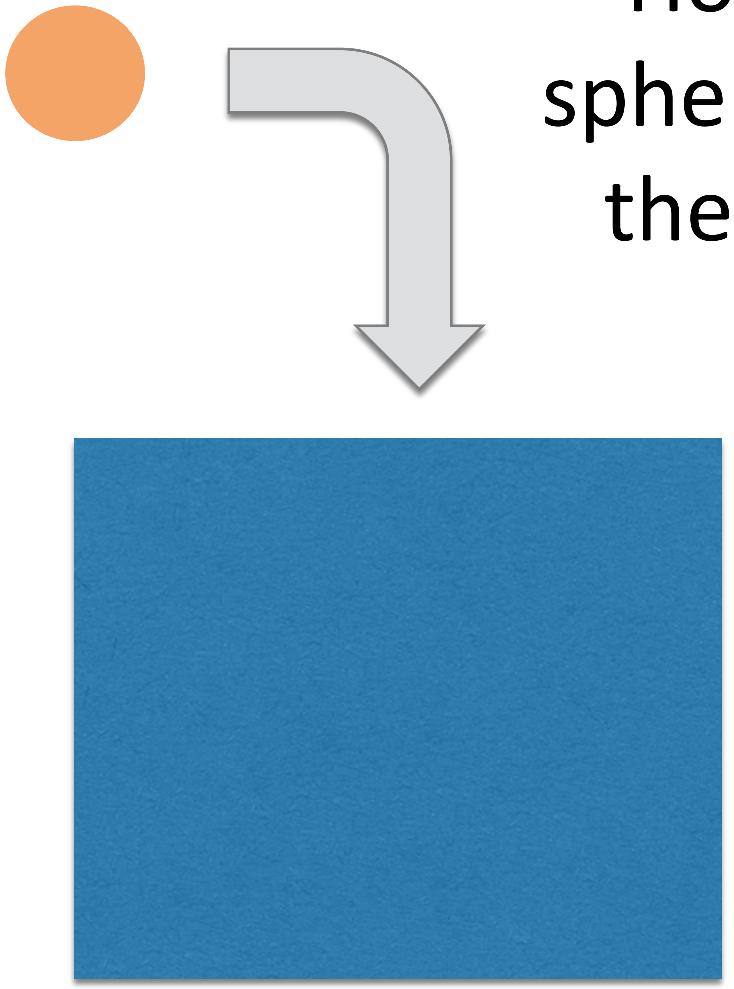
How many spheres fit into the square?



...but not three

Estimating the number of molecules in the ASU

Matthews coefficient: $$V_M = \frac{\text{volume of asymmetric unit}}{\text{molecular weight}}$$

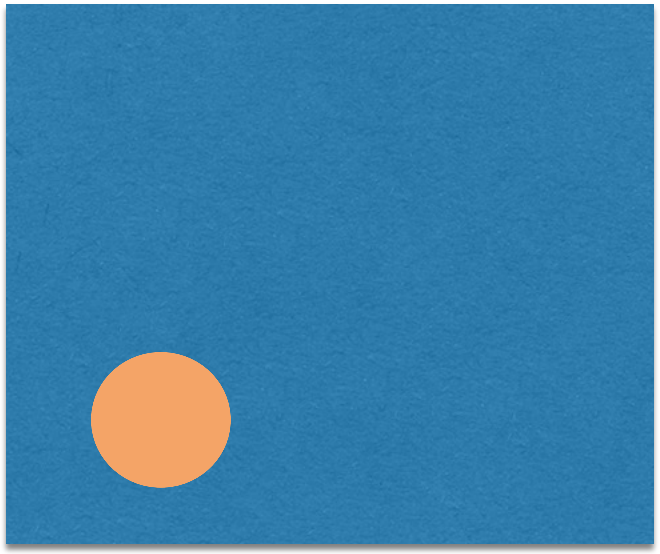


How many spheres fit into the square?

Estimating the number of molecules in the ASU

Matthews coefficient: $$V_M = \frac{\text{volume of asymmetric unit}}{\text{molecular weight}}$$

How many spheres fit into the square?

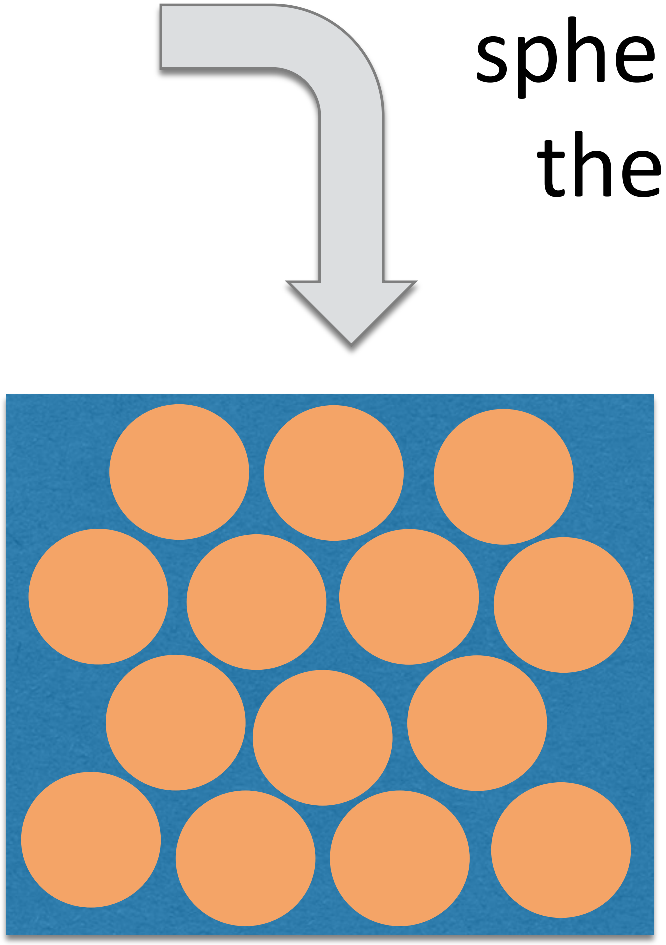


Only one

Estimating the number of molecules in the ASU

Matthews coefficient: $$V_M = \frac{\text{volume of asymmetric unit}}{\text{molecular weight}}$$

How many spheres fit into the square?



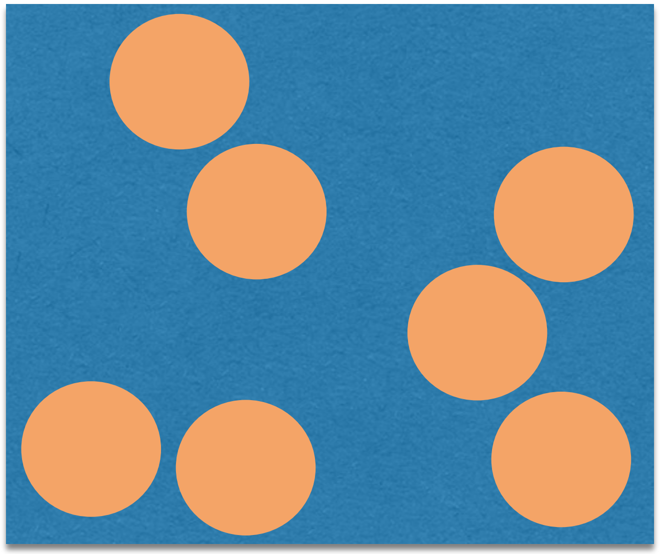
Many (14)

Estimating the number of molecules in the ASU

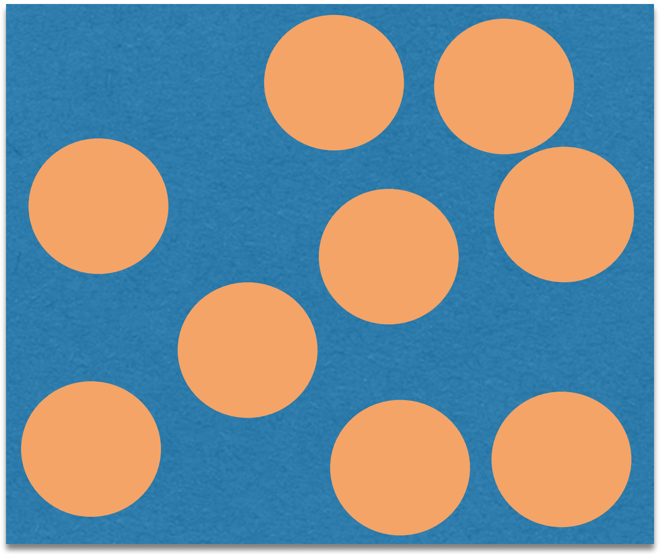
Matthews coefficient:

$$V_M = \frac{\text{volume of asymmetric unit}}{\text{molecular weight}}$$

How many spheres fit into the square?



7



9

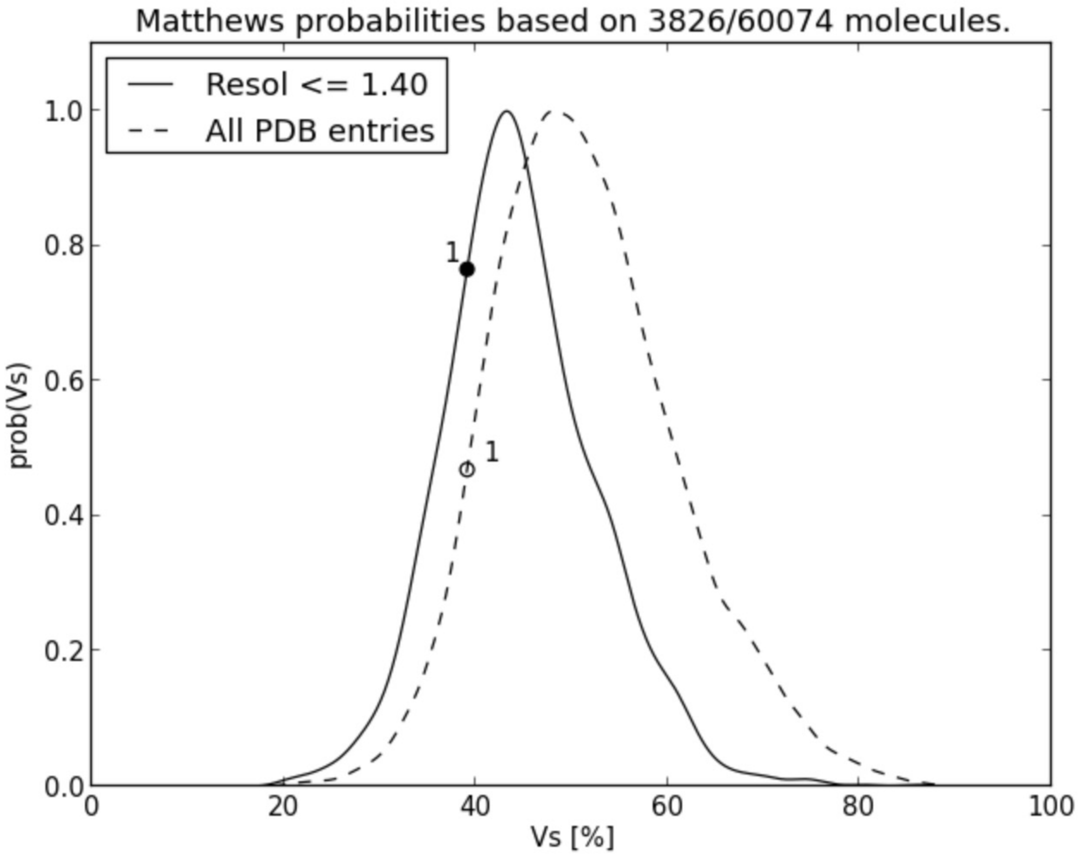
Most likely: something in between

Estimating the number of molecules in the ASU

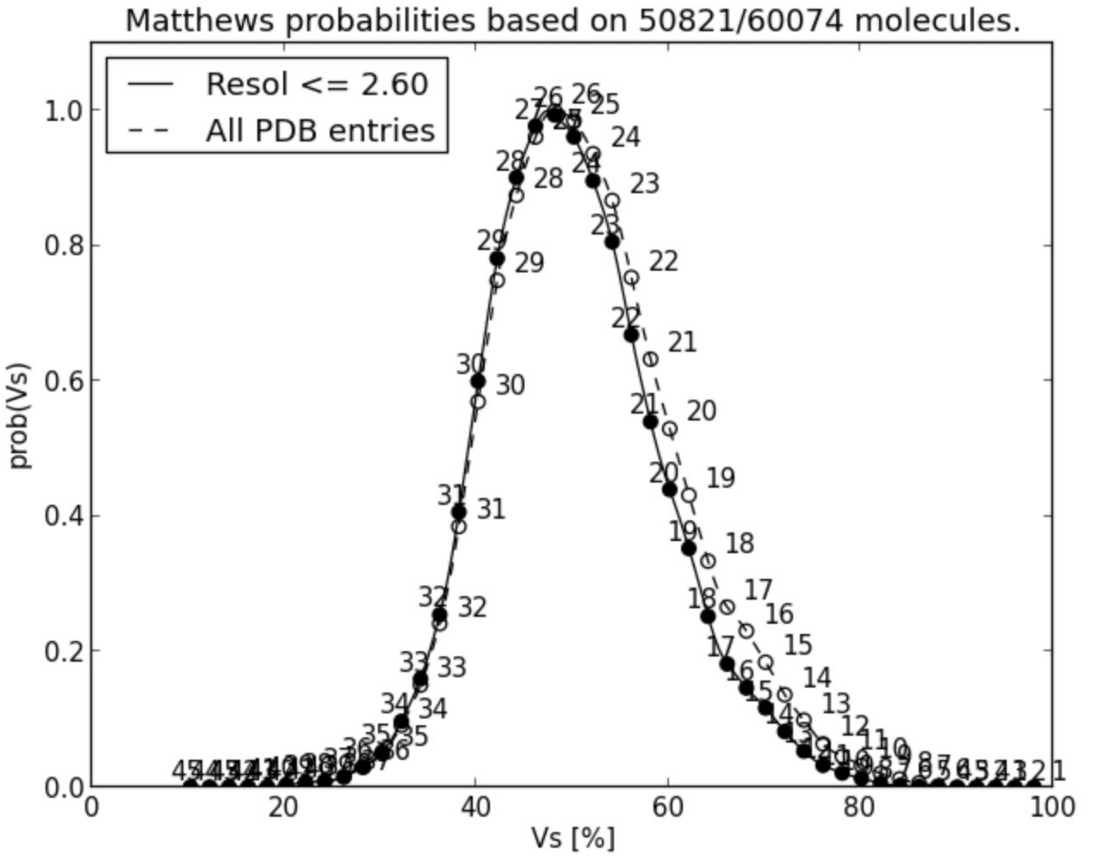
Matthews coefficient: $$V_M = \frac{\text{volume of asymmetric unit}}{\text{molecular weight}}$$

Compare your value of V_M with histograms for known structures (from PDB) → choose most probable.

Few possibilities



Many possibilities



Key points

1) MR: Use a previously known structure to solve a new structure

2) Known structure can be a homologue or a predicted model

3) Known structures may need to be modified

4) Estimate the ASU content

5) Be aware of data pathologies (twinning, tNCS)

The Phenix Project

Lawrence Berkeley Laboratory

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

Los Alamos National Laboratory
New Mexico Consortium

Tom Terwilliger, Li-Wei Hung

UTHealth

Matt Baker

University of Cambridge

Randy Read, Airlie McCoy, Alisia Fadini

Duke University

Jane & David Richardson, Christopher Williams, Vincent Chen

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