

*Phenix User Workshop, Pittsburgh Diffraction Conference,  
October 14th 2023*



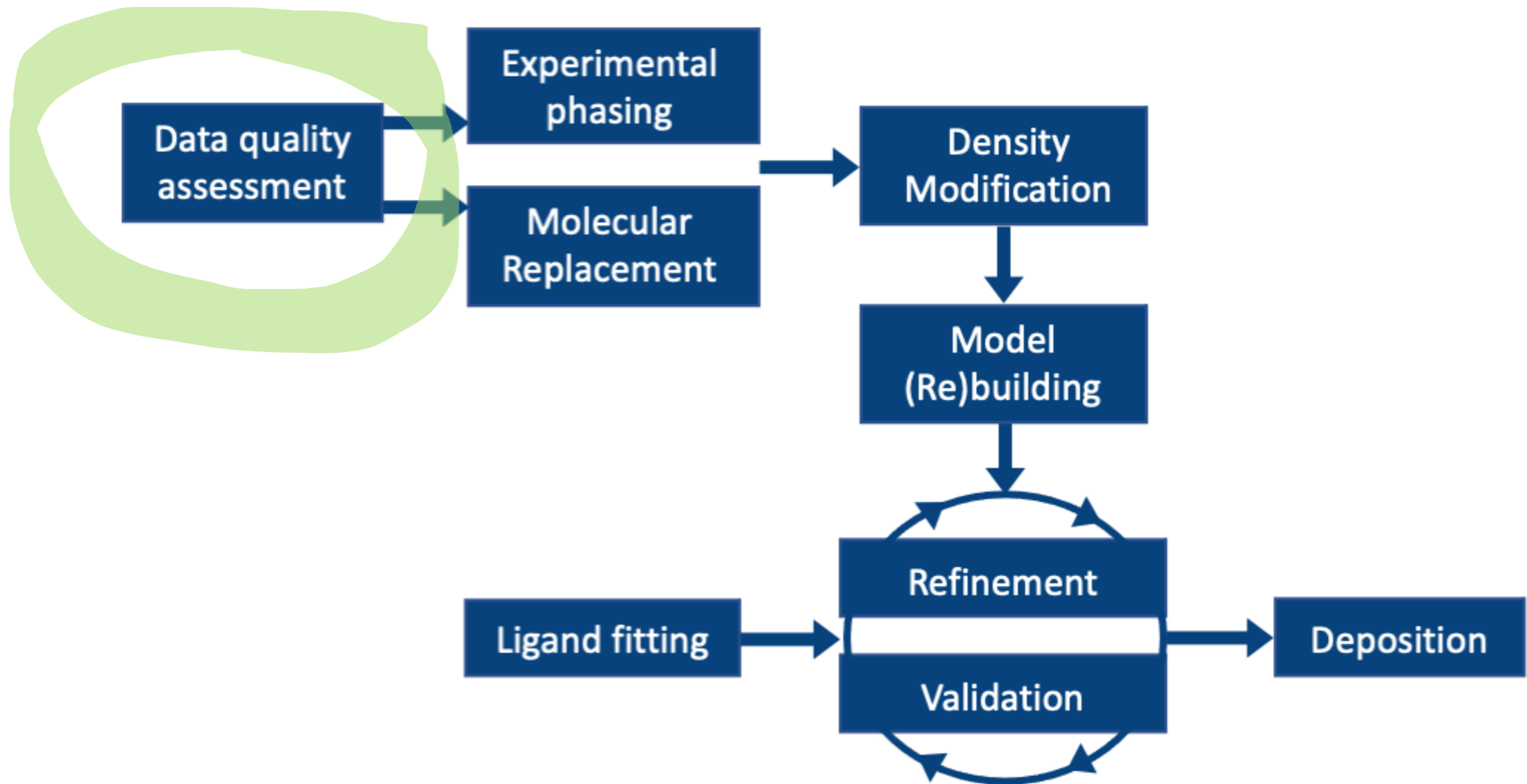
Xtrriage

Dorothee Liebschner  
Lawrence Berkeley Laboratory

# Data validation is the first step!

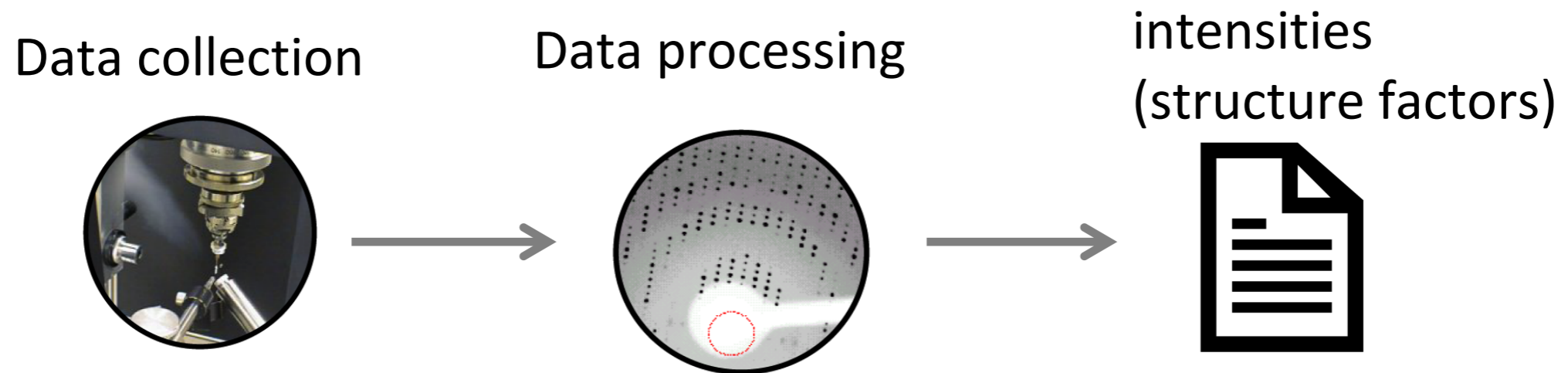
---

Before doing anything else, you should validate your data!



# Data quality assessment

---



## **Macromolecular crystals are prone to pathologies:**

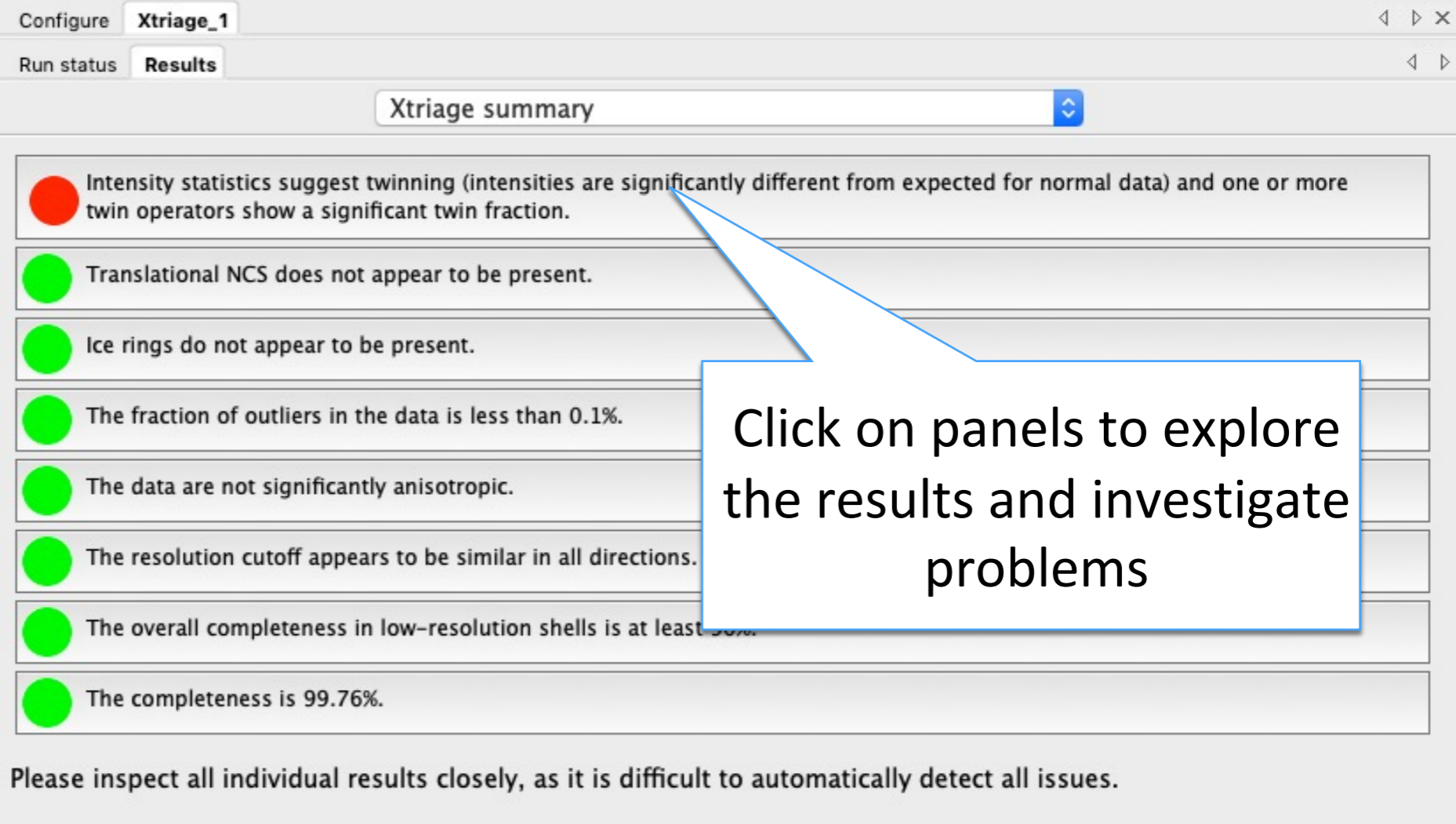
- Twinning: two or more crystals are intergrown (orientations are related by twin operation)
- tNCS: more than one copy of a molecule is in a similar orientation in the asymmetric unit

# Data quality assessment

Data anomalies can prevent structure solution!

→ It is important to check your data before phasing, model building and refinement.

**Xtrriage** does diagnostics for major pathologies and data properties (Wilson plot, completeness, symmetry).



The screenshot shows the Xtrriage software interface. At the top, there are tabs for 'Configure' and 'Xtrriage\_1', and 'Run status' and 'Results'. Below this is a search bar containing 'Xtrriage summary'. The main content area displays a list of diagnostic results, each with a colored circular icon and a text description:

- Intensity statistics suggest twinning (intensities are significantly different from expected for normal data) and one or more twin operators show a significant twin fraction.** (Red icon)
- Translational NCS does not appear to be present. (Green icon)
- Ice rings do not appear to be present. (Green icon)
- The fraction of outliers in the data is less than 0.1%. (Green icon)
- The data are not significantly anisotropic. (Green icon)
- The resolution cutoff appears to be similar in all directions. (Green icon)
- The overall completeness in low-resolution shells is at least 50%. (Green icon)
- The completeness is 99.76%. (Green icon)

A callout box with a blue border and a white background points to the first result, containing the text: 'Click on panels to explore the results and investigate problems'. At the bottom of the interface, a note reads: 'Please inspect all individual results closely, as it is difficult to automatically detect all issues.'

# Checking data quality with Xtrriage

---

## Are the data going to be useful?

- How accurate are the data?
- Is there an anomalous signal?

## Do the data have any unusual characteristics?

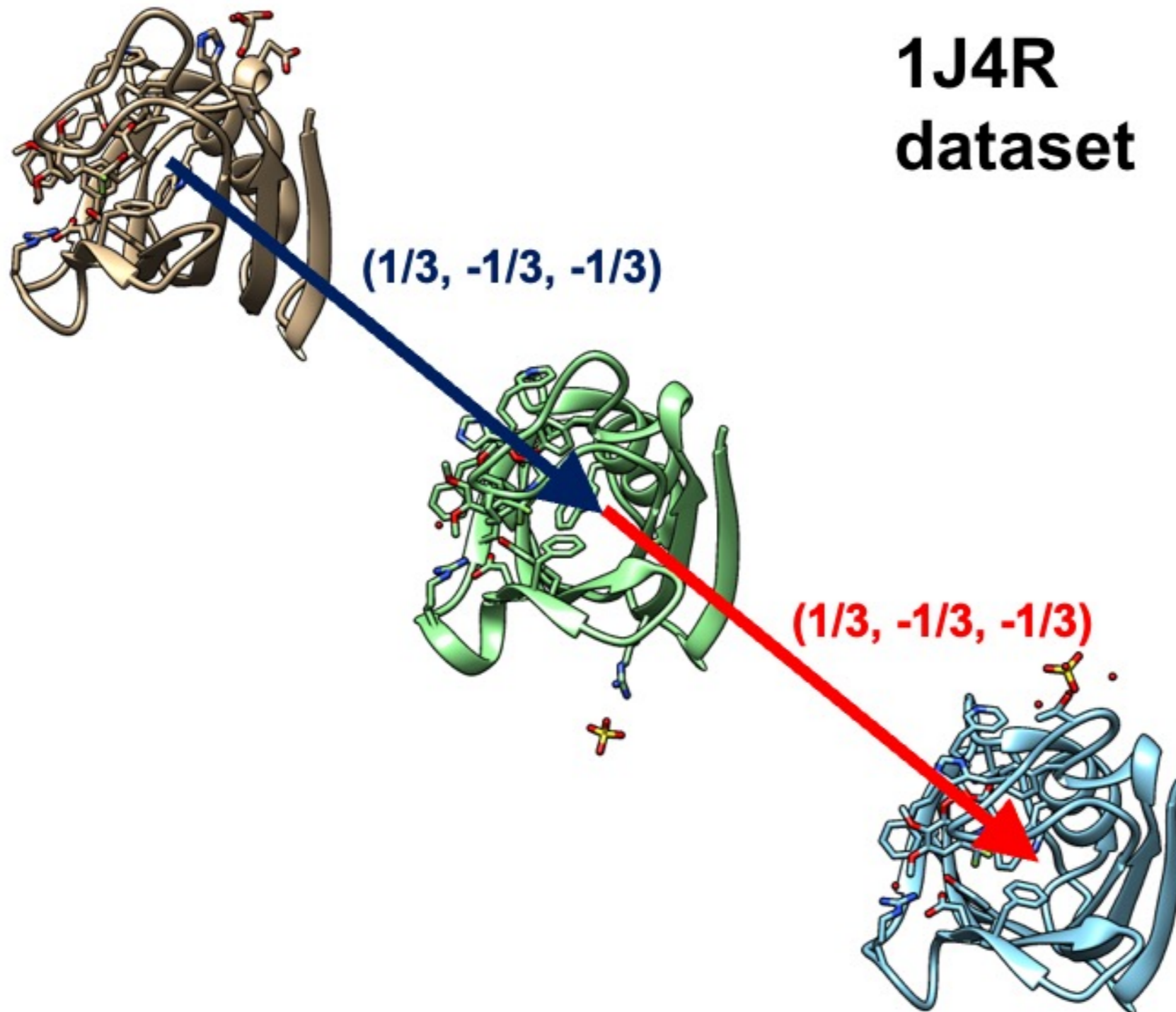
- Are there ice rings?
- Is translational NCS present?
- Is twinning present?

## Other things to check with Xtrriage

- Can my data be indexed in a different crystal lattice?
- Do the systematic absences in my data match the space group?

# Translational NCS

Two or more copies of the molecule are related by pure translation.



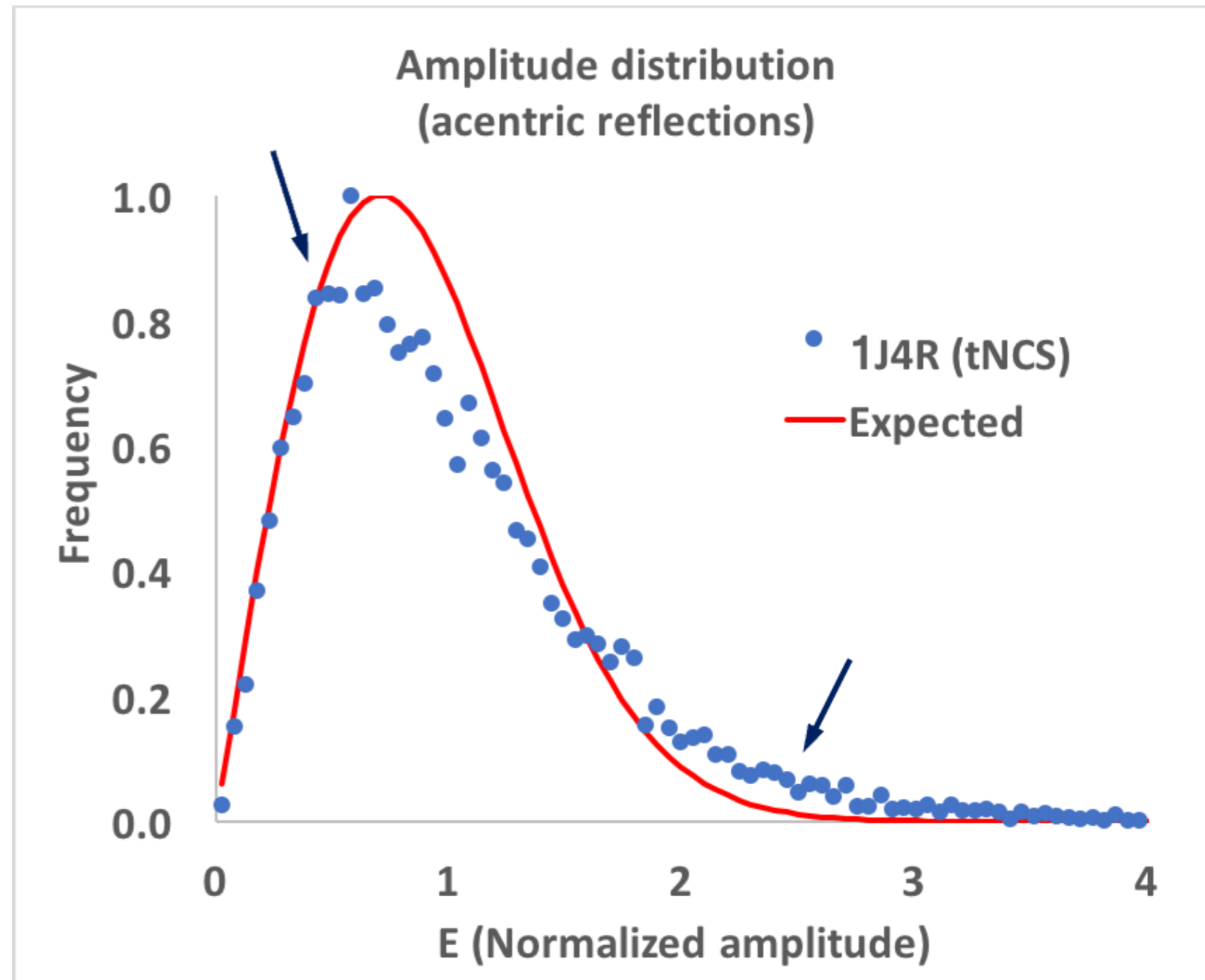
Strong peak in  
Patterson function

In this case at  
 $(1/3, -1/3, -1/3)$

# Translational NCS

## Effect of tNCS:

- Pattern of strong/weak reflections
- Broadening of intensity and amplitude distributions



# If your crystal has tNCS

---

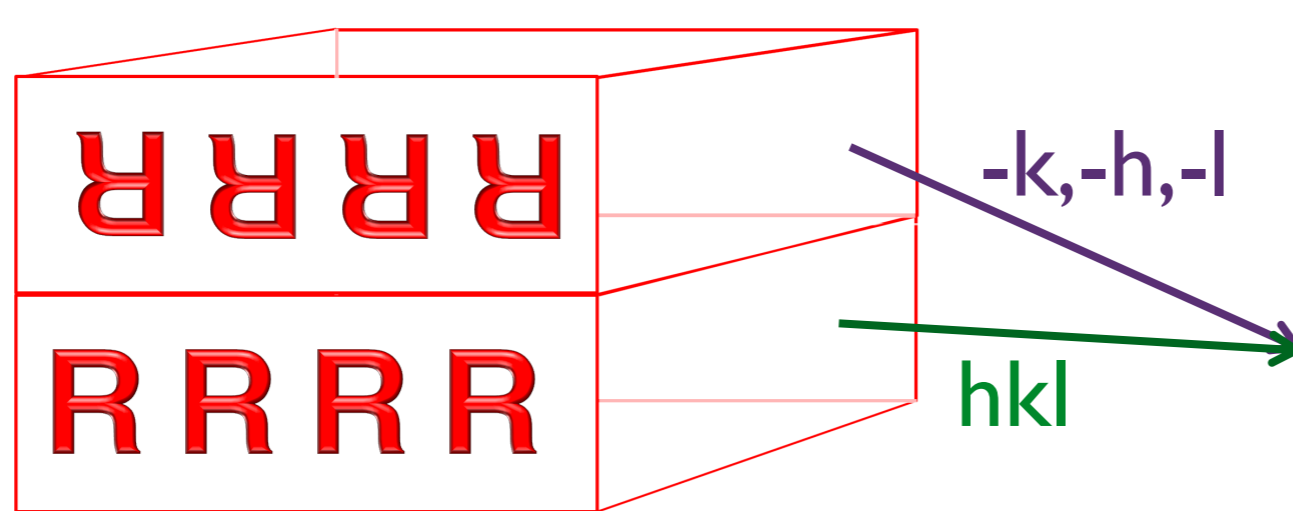
- Use L-test for twinning
- Use tNCS-corrected likelihood target in molecular replacement
- Delete corresponding parts of all copies in omit maps



# Twinning

---

Identical but rotated crystals sandwiched together



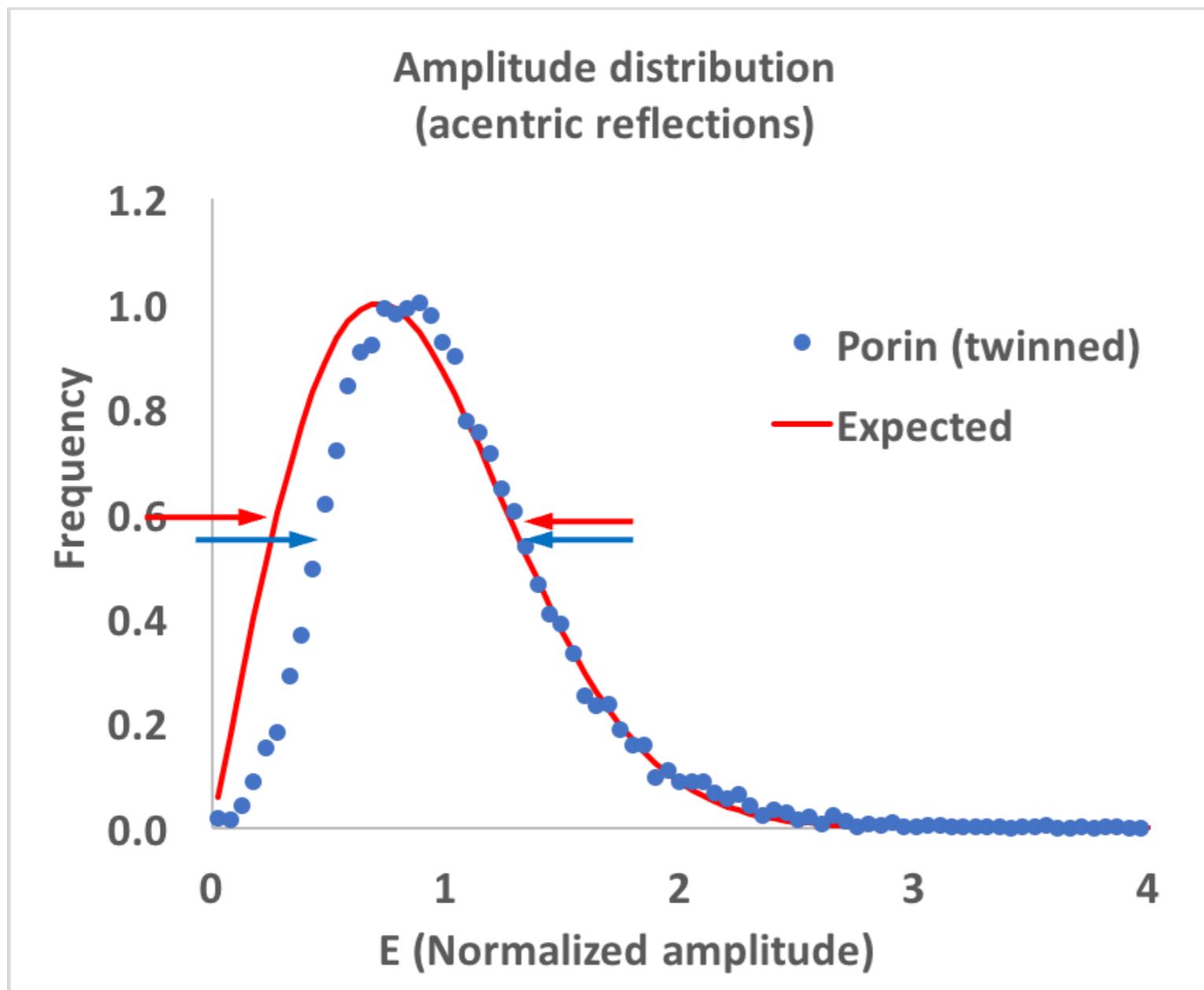
...is mixed with twin-law related reflection (e.g.,  $-k,-h,-l$ )

Diffraction spot for  $(h,k,l)$  reflection...

Possible twin laws depend on your crystal symmetry and cell dimensions

# Effect of twinning

- Mixing of intensities
- Intensities become more average
- Narrowing of intensity and amplitude distributions



tNCS can mask  
twinning effects  
on intensities and  
amplitudes

# Are the data twinned?

---

Is translational NCS present?

No



Check Wilson ratio:  $\langle I^2 \rangle / \langle I \rangle^2$   
2 for untwinned  
1.5 twinned  
(acentric reflections)

Yes



Check L-test\*  
(corrected for NCS)

\*Padilla, J. E. & Yeates, T. O. (2003). *Acta Crystallogr. D Biol. Crystallogr.* **59**, 1124–1130.

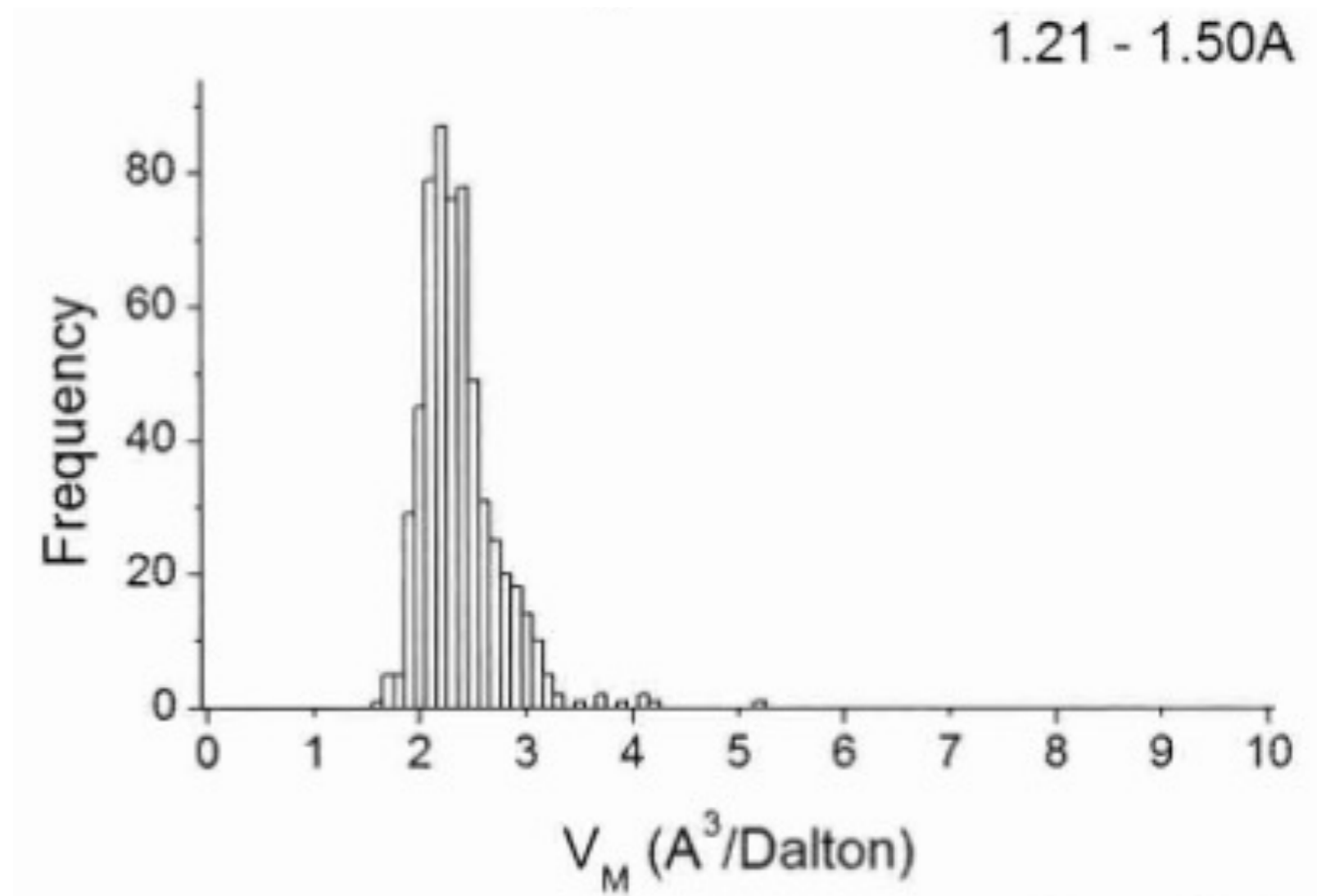
# Matthews coefficient

---

Estimate for the number of molecules in the asymmetric unit  
→ Important for Molecular Replacement

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

Get histograms for  $V_M$   
for known structures  
(from PDB).



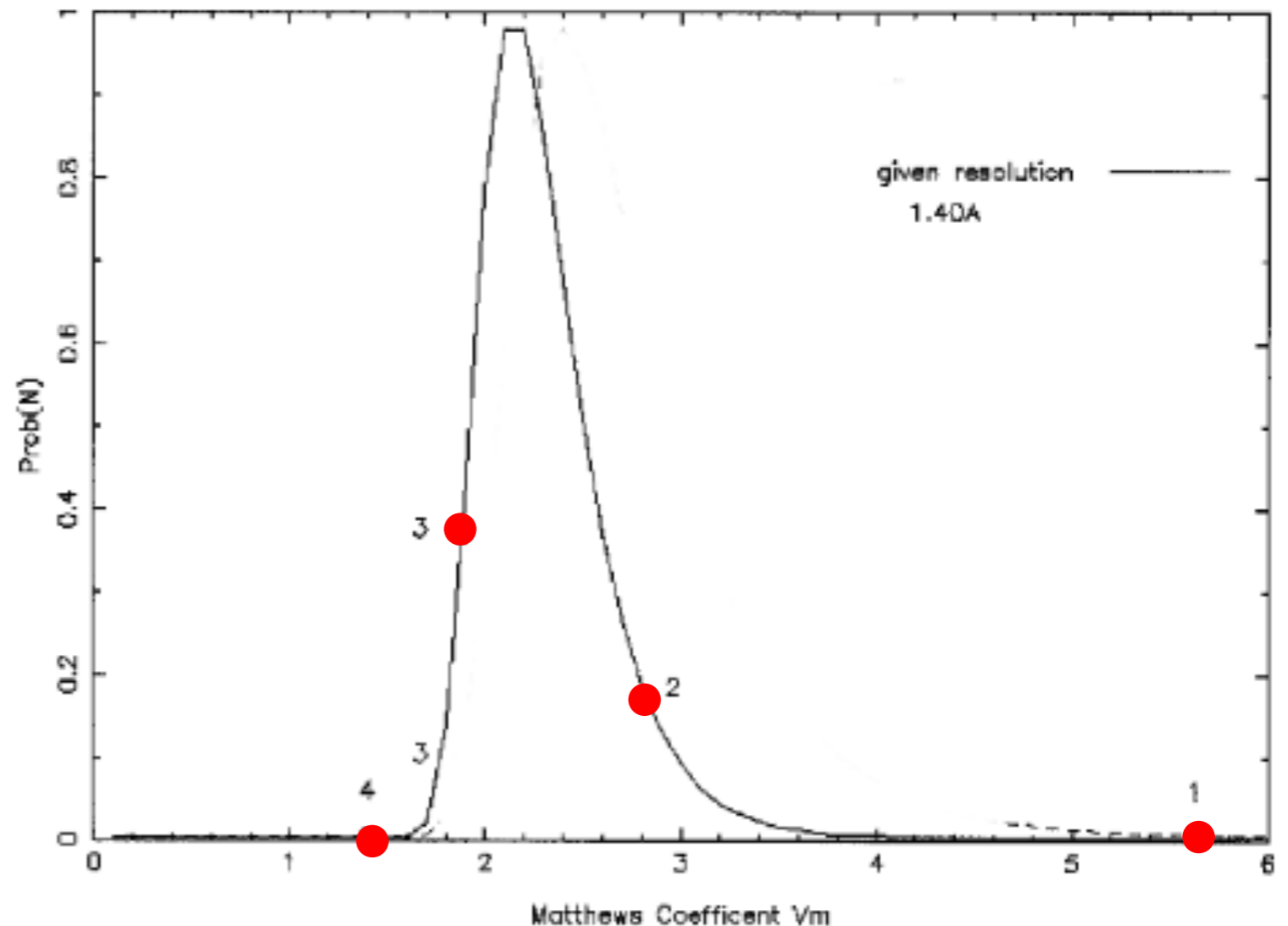
# Matthews coefficient

Estimate for the number of molecules in the asymmetric unit  
→ Important for Molecular Replacement

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

Get an estimate of the number of molecules in the asymmetric unit.

→ Molecular Replacement!



# Summary

---

- Check your diffraction data with Xtrriage
- Look at the Matthews coefficient
- Beware of pathologies