molecular replacement

Airlie McCoy

molecular replacement

1. prepare native crystals
2. collect and process data
3. model preparation

4. molecular replacement

5. model building, refinement and validation

molecular replacement

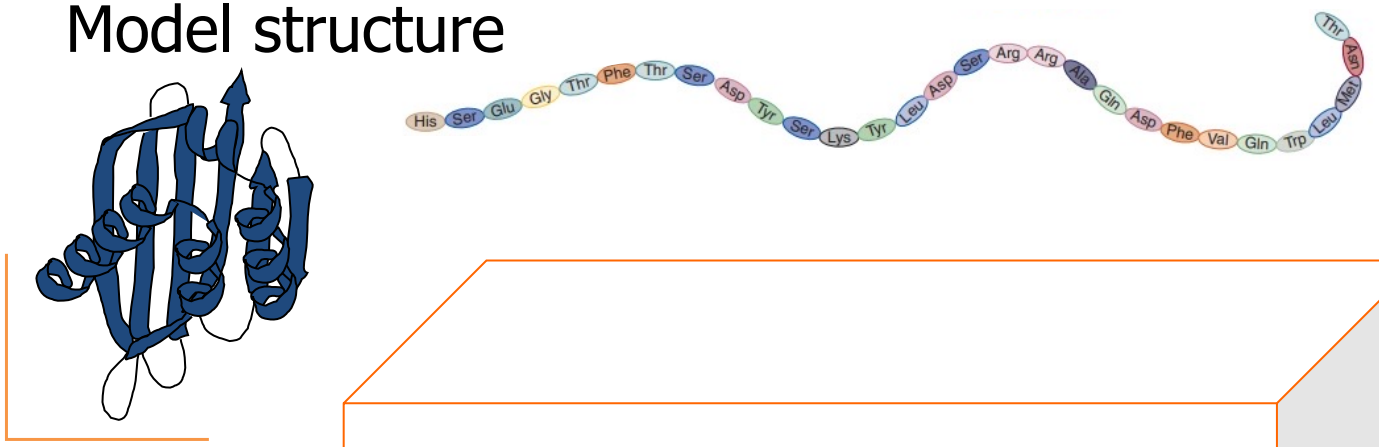
1. prepare native crystals
2. collect and process data
3. model preparation

4. molecular replacement – with Phaser
   - some things will be general to molecular replacement and
   - some things specific to molecular replacement with Phaser
5. model building, refinement and validation

phasing by molecular replacement

molecular replacement

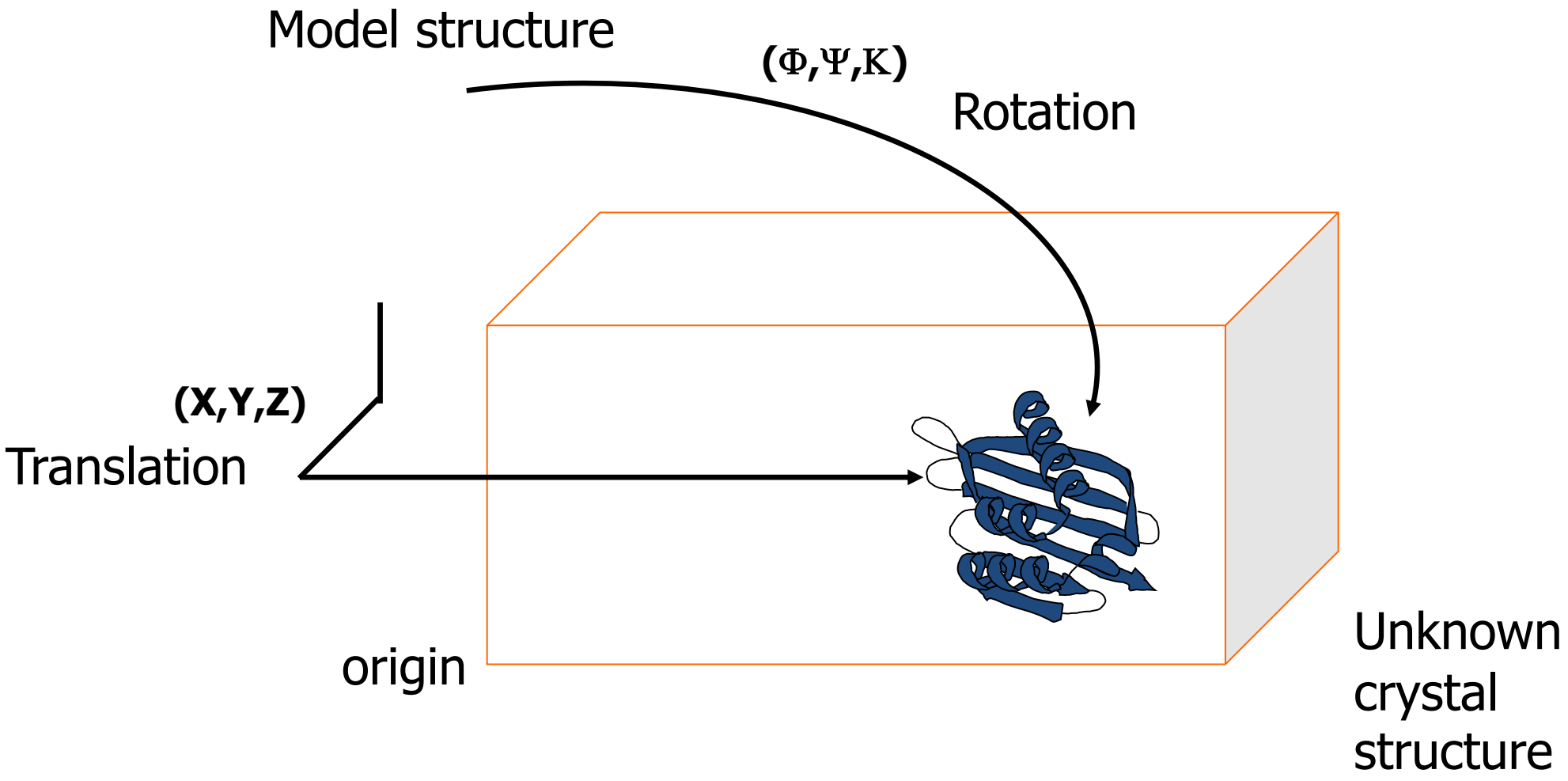
Model structure



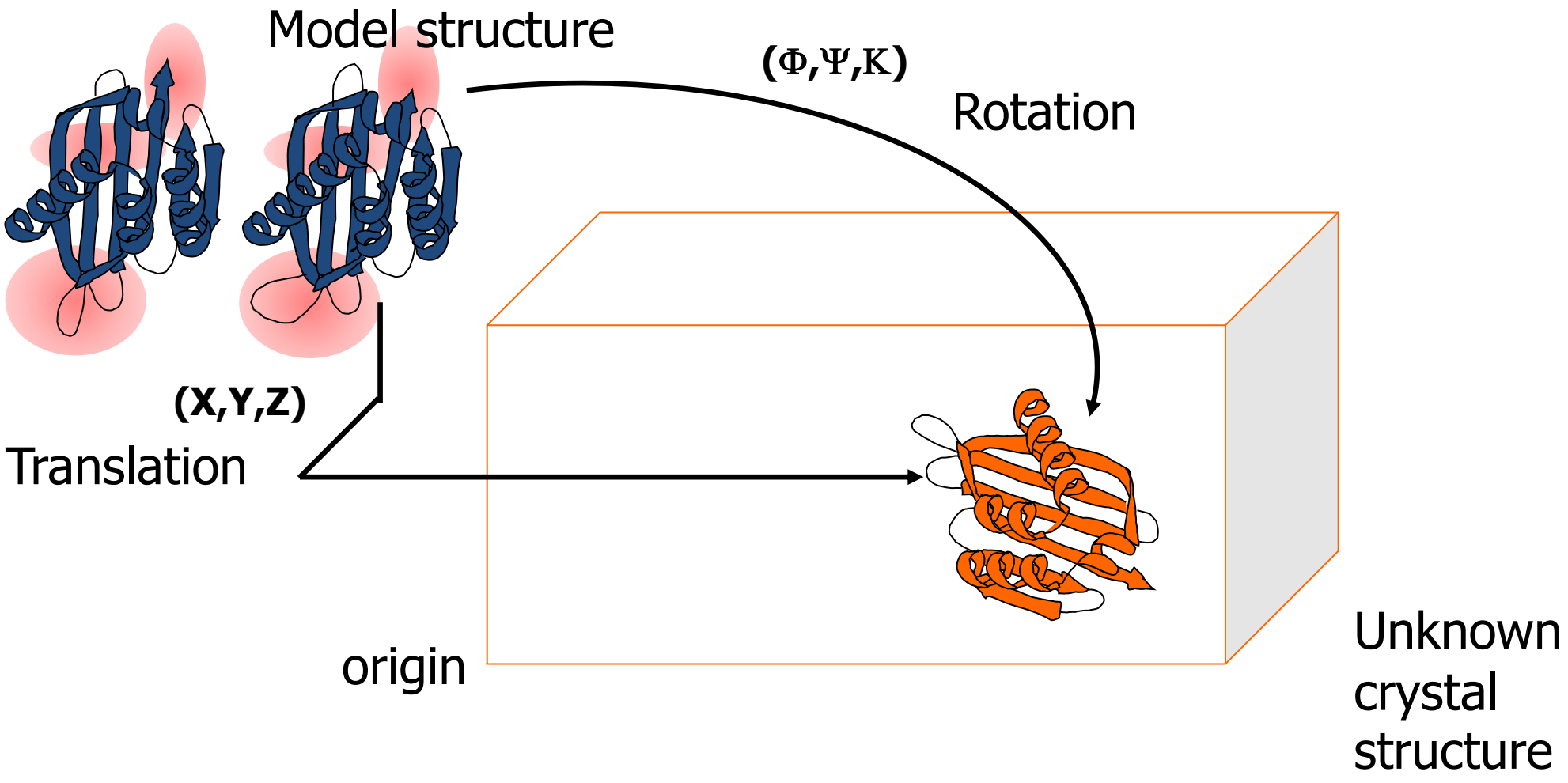
origin

Unknown crystal structure

molecular replacement

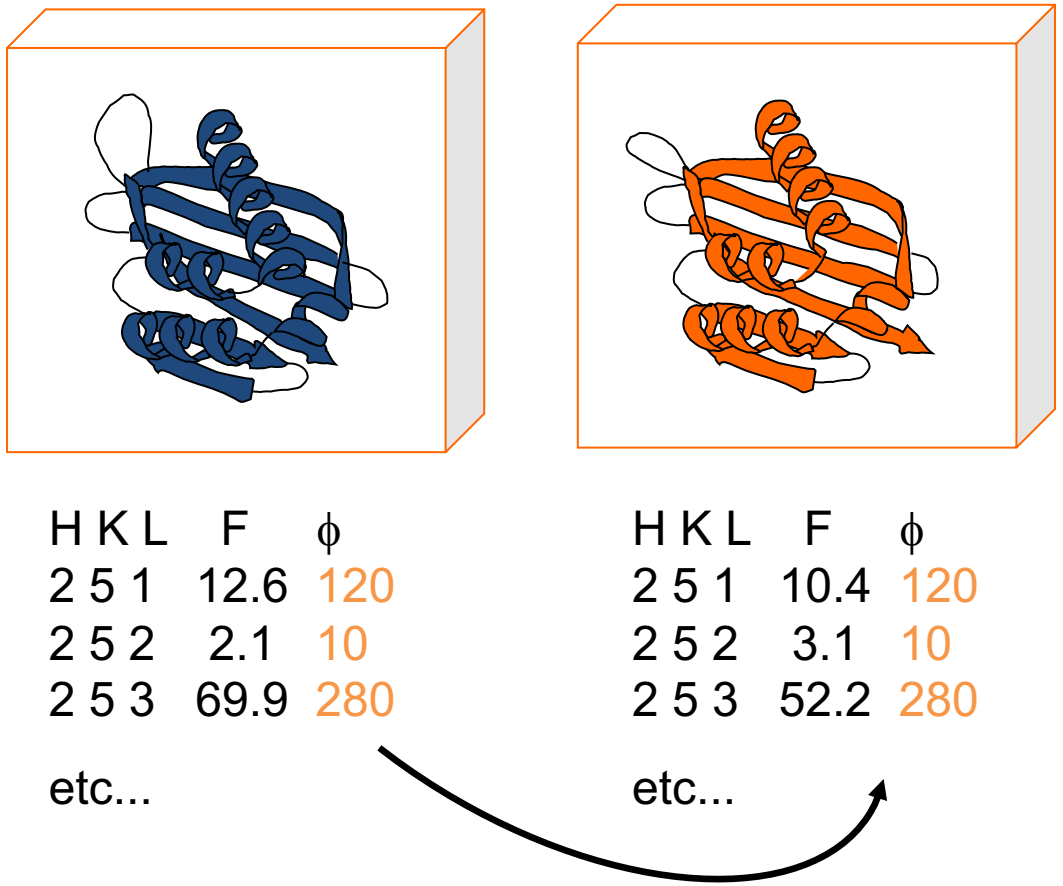


molecular replacement



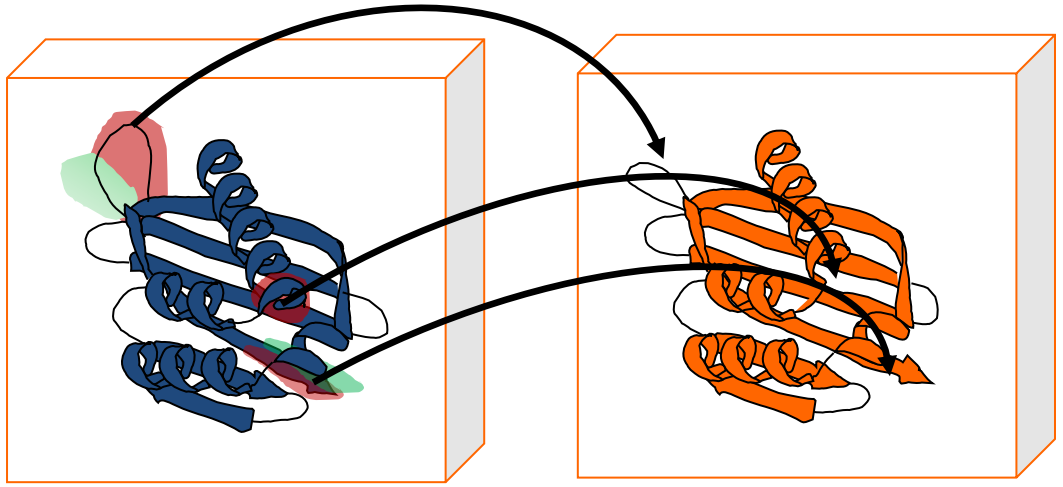
molecular replacement

- Find orientation and position where model overlies the target structure
- Borrow the phases
- Then it becomes a refinement problem
- The phases will change during refinement!



model building and refinement

- After molecular replacement the electron density maps can be inspected to see where the model is wrong or incomplete
- ‘difference density’ shows where atoms need to be deleted from or added to the model

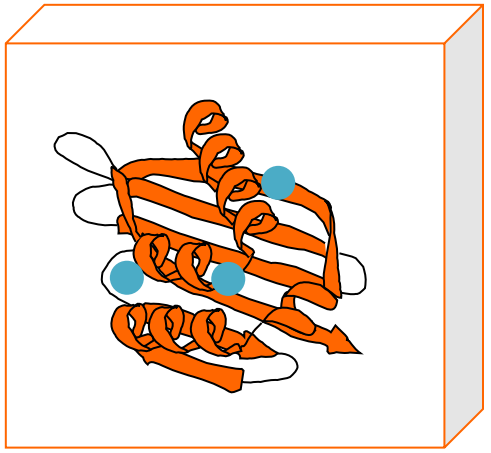


H K L	F	ϕ
2 5 1	12.6	120
2 5 2	2.1	10
2 5 3	69.9	280
etc...		

H K L	F	ϕ
2 5 1	10.4	142
2 5 2	3.1	34
2 5 3	52.2	250
etc...		

MR-SAD

- SAD: single-wavelength anomalous dispersion
  - Phasing method
- After molecular replacement any anomalous scattering from the crystal can be used to find anomalous scatterers such as metal ions or sulphur or selenium from seleno-methionine
- Can be used to help phase if MR solution is poor



H	K	L	F+
2	5	1	10.4
2	5	2	3.1
2	5	3	52.2

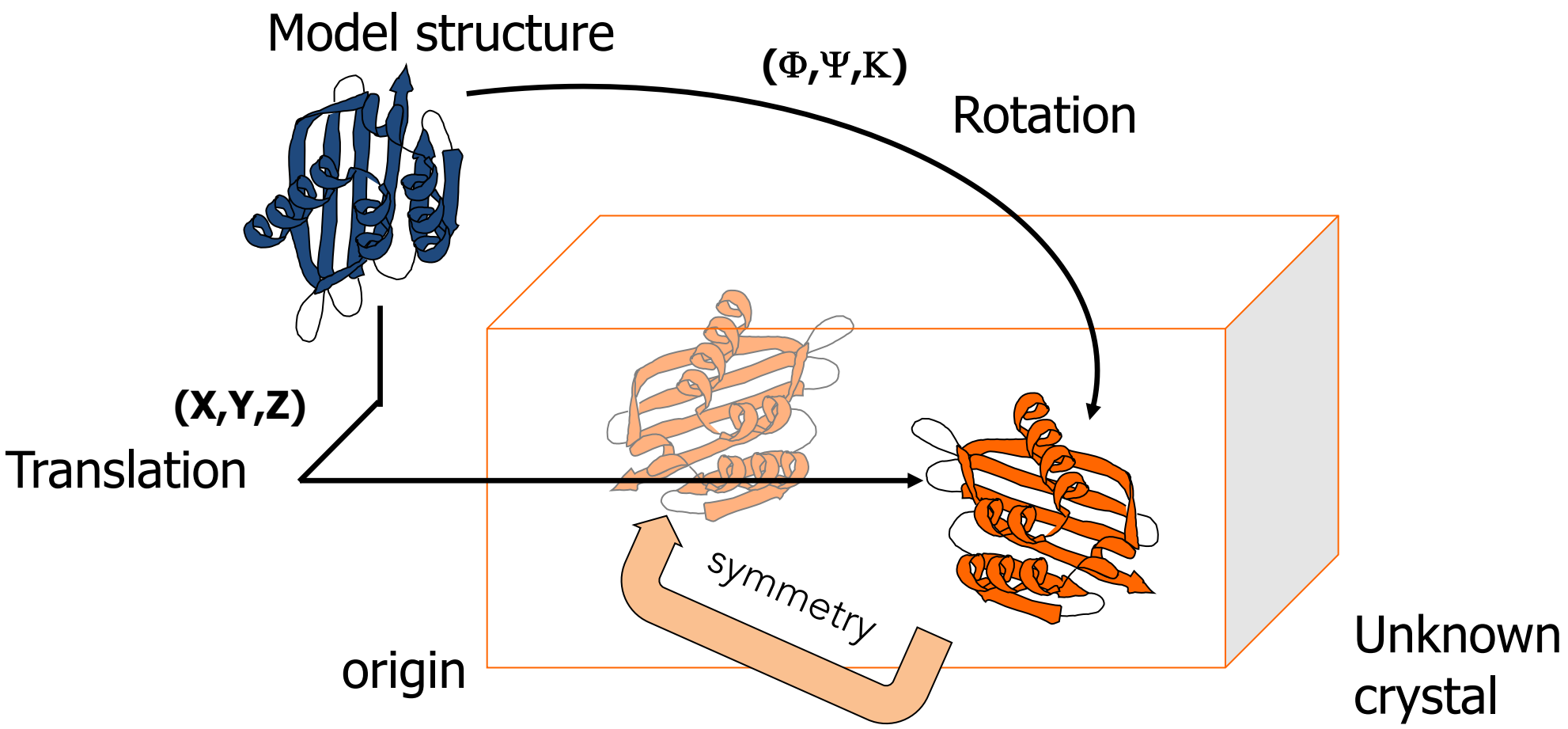
etc...

H	K	L	F-
2	5	1	10.1
2	5	2	3.8
2	5	3	53.6

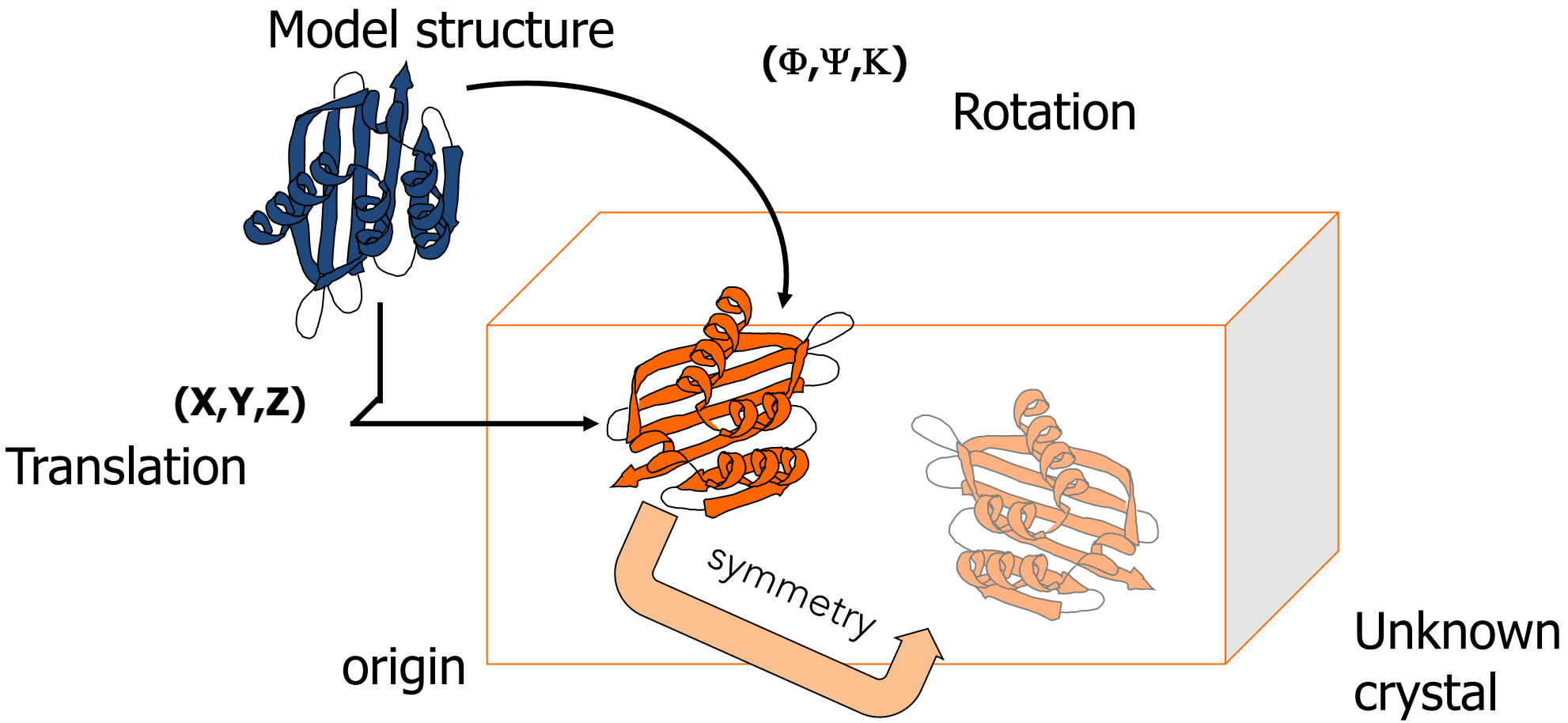
etc...

symmetry and molecular replacement

molecular replacement

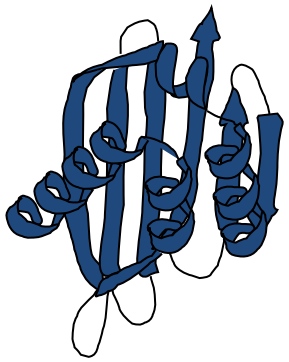


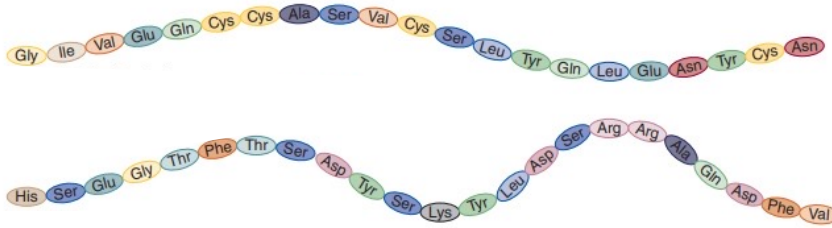
molecular replacement

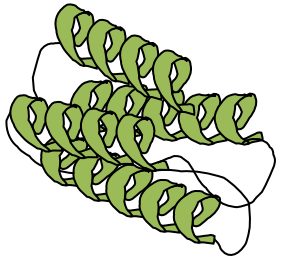


molecular replacement

Model structure



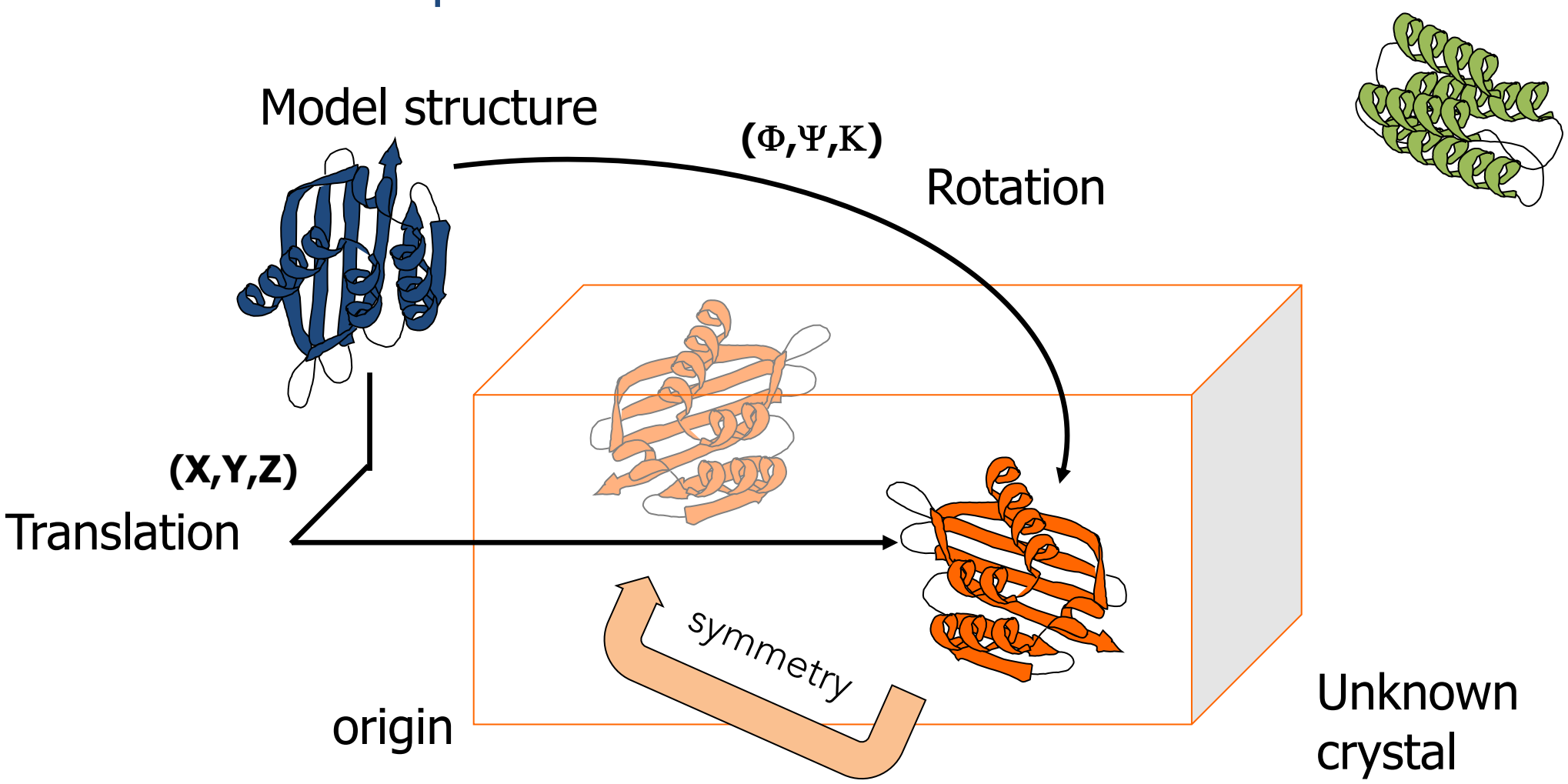




origin

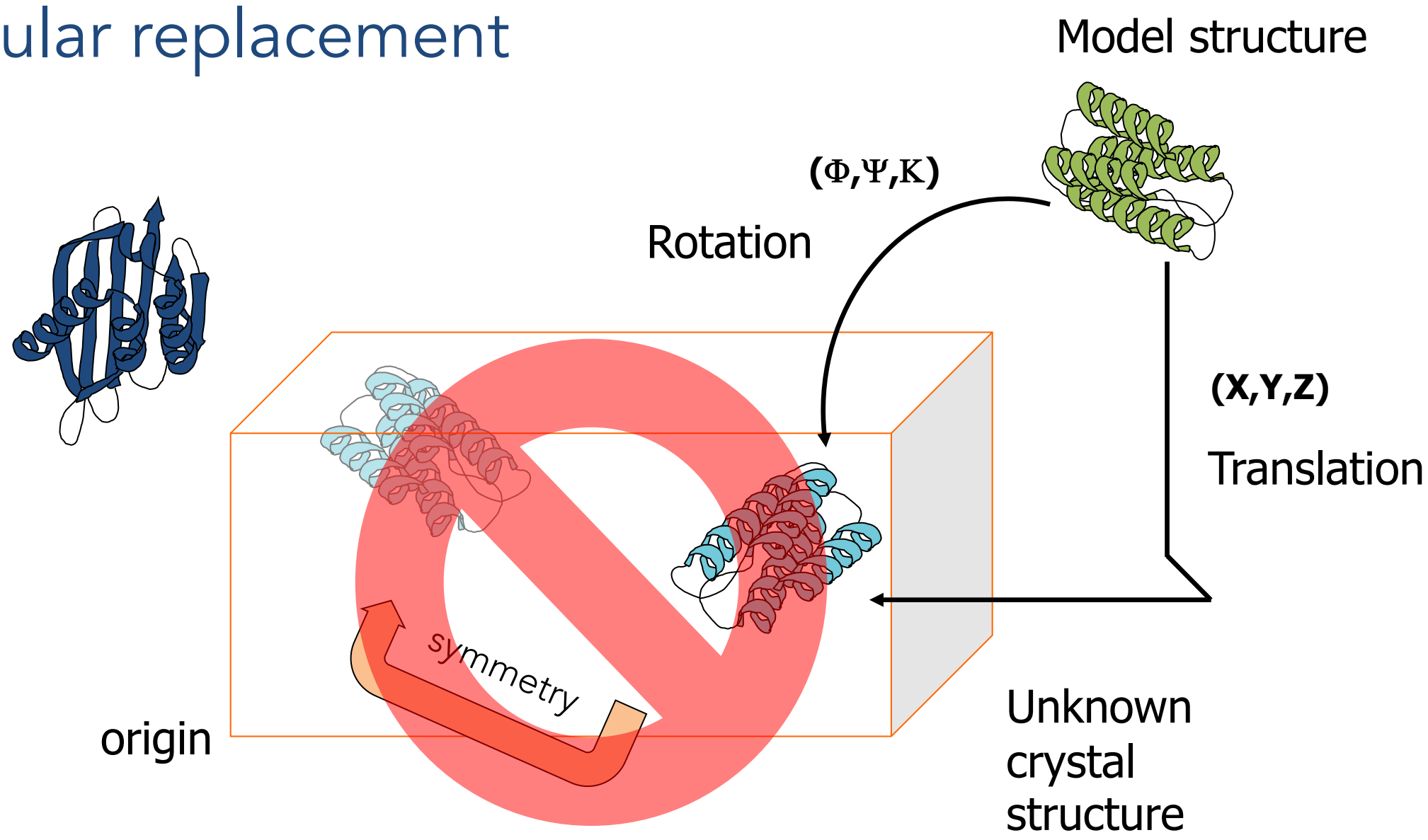
Unknown crystal structure

molecular replacement

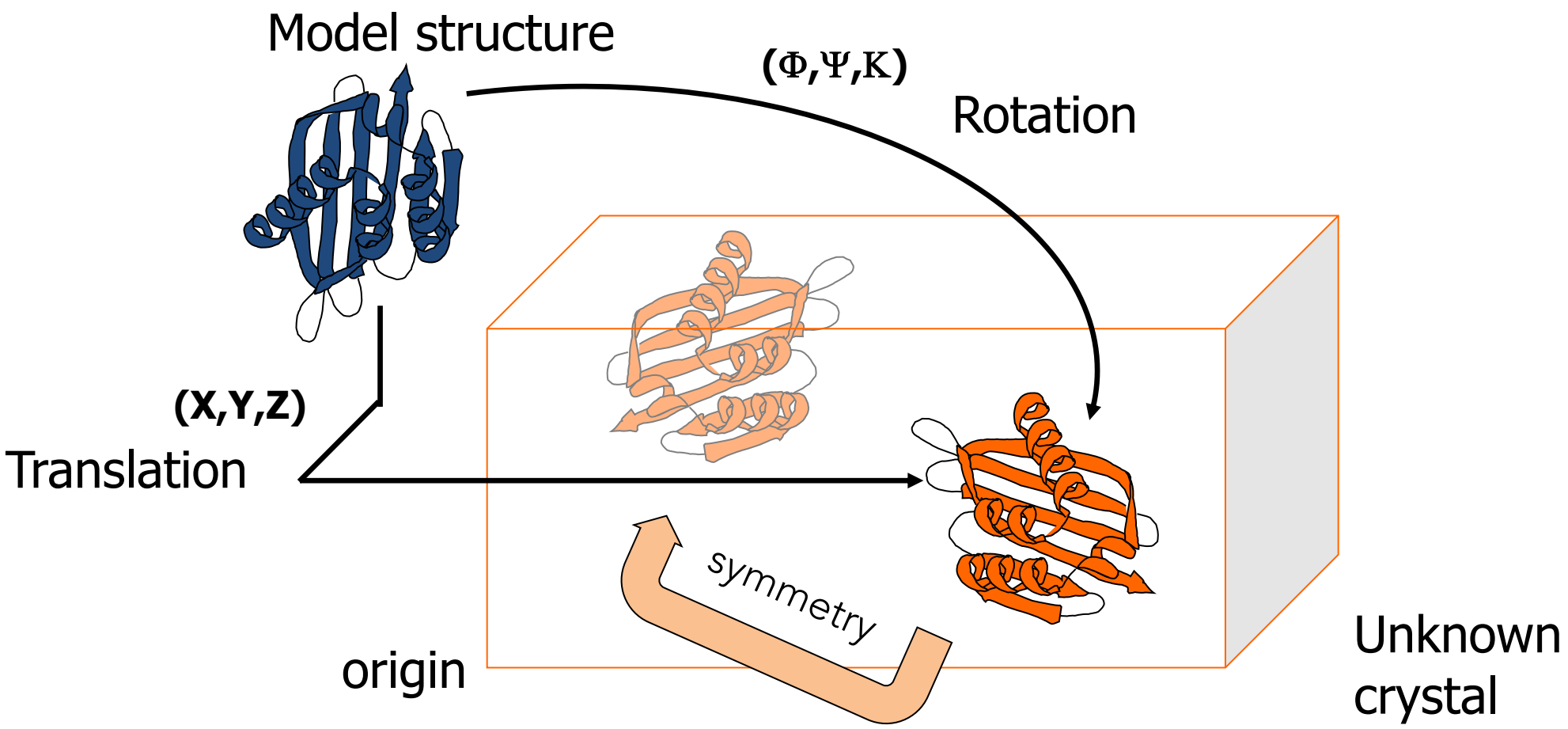


Model structure

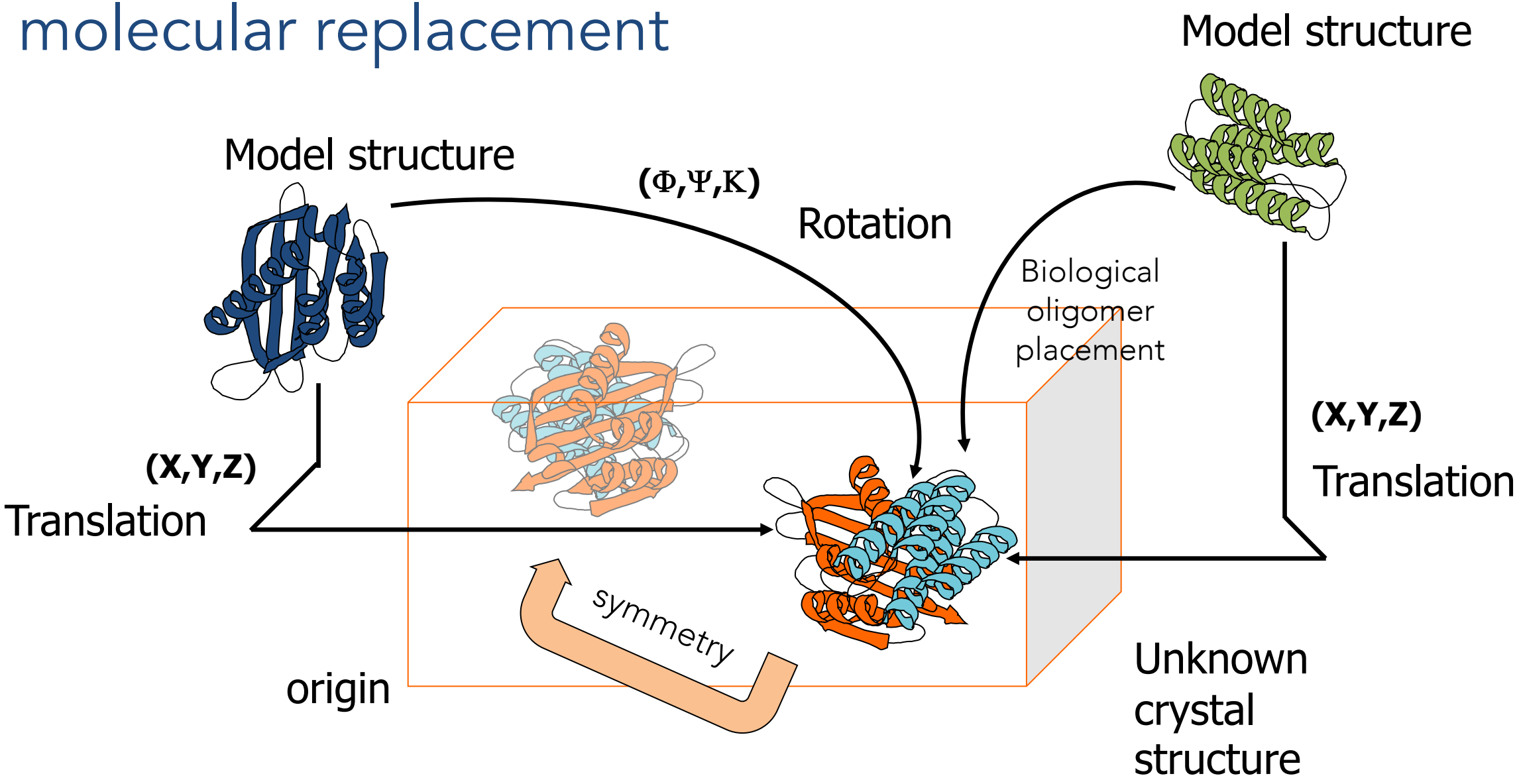
molecular replacement



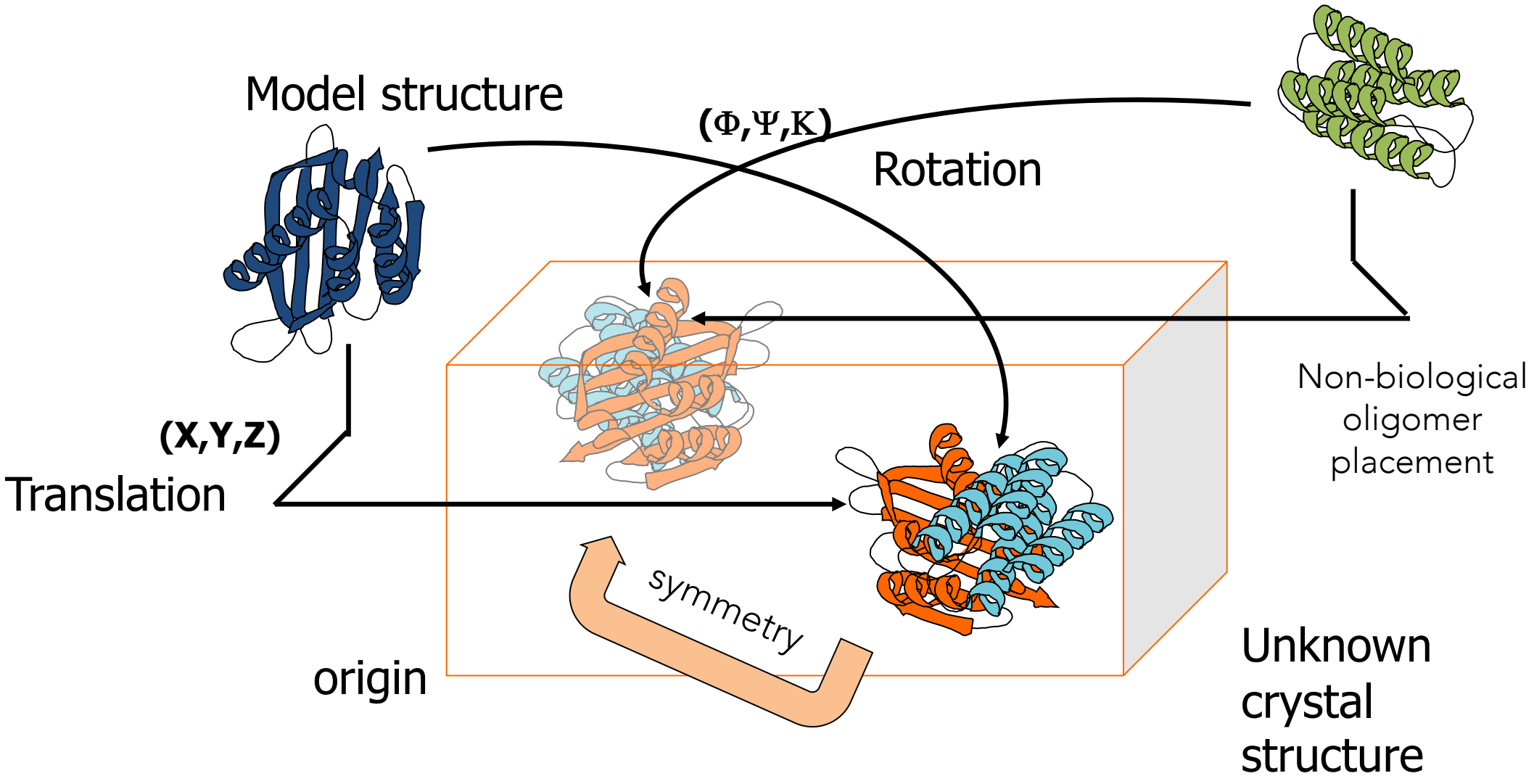
molecular replacement



molecular replacement

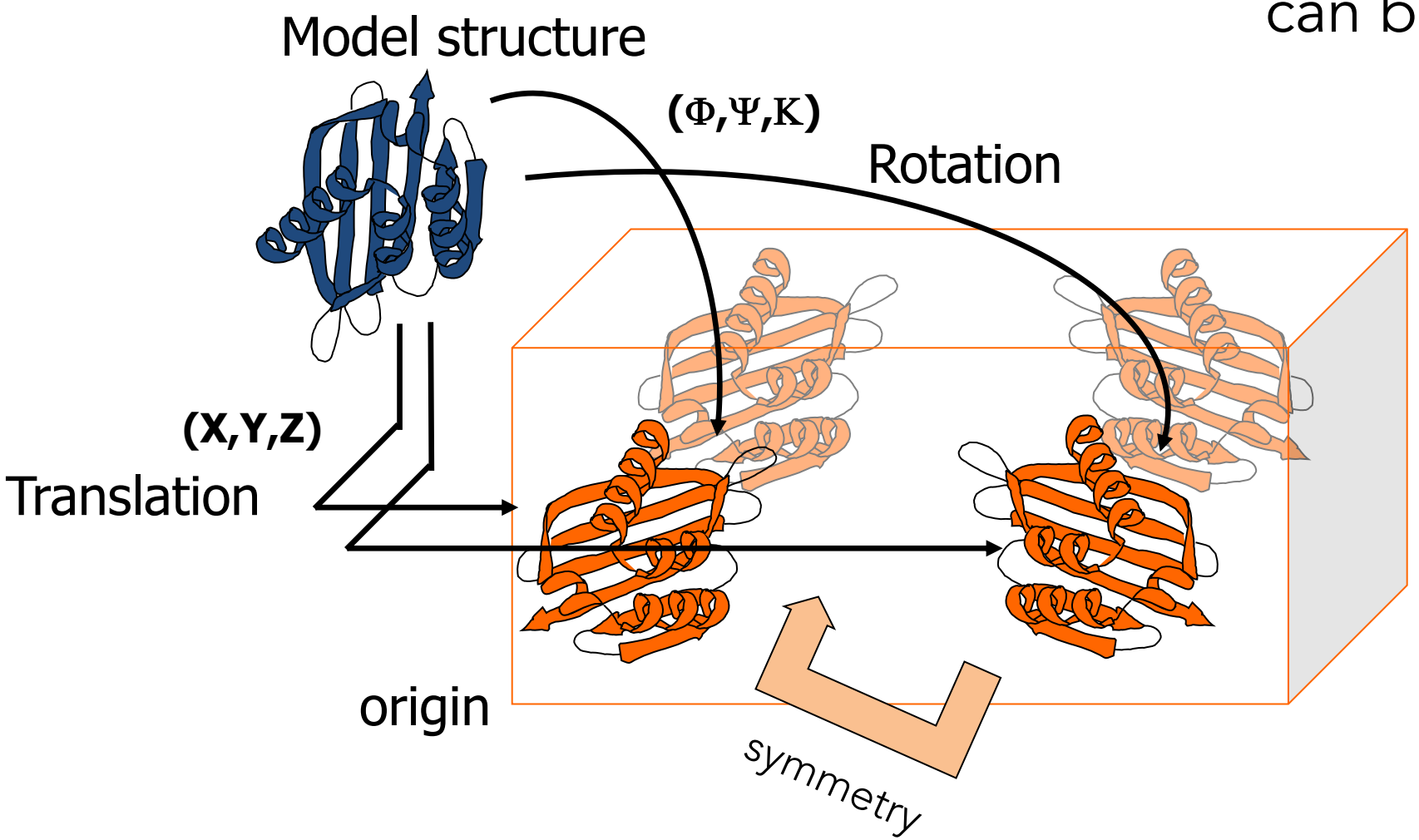


molecular replacement



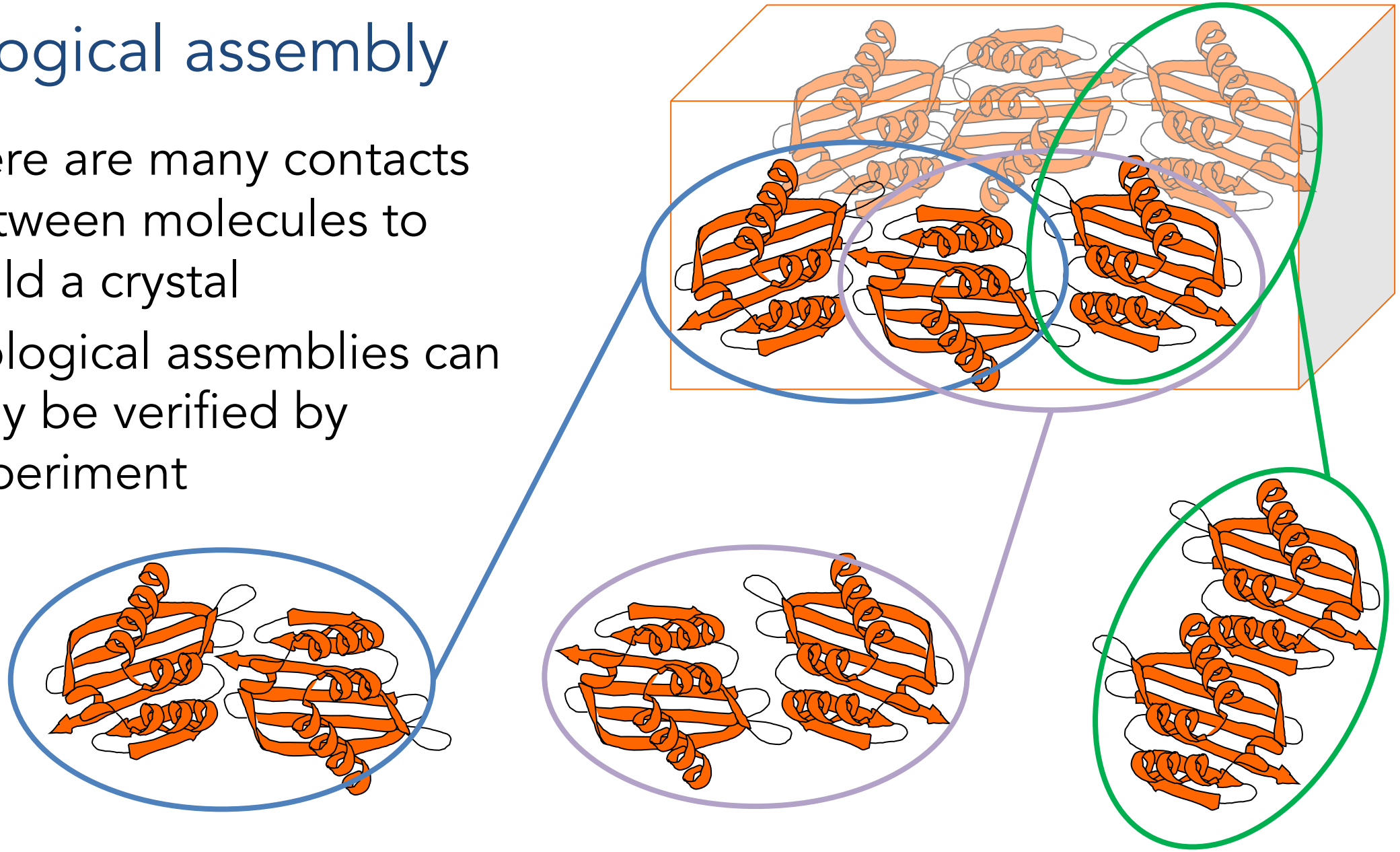
molecular replacement

number of components in the asymmetric unit can be more than one



biological assembly

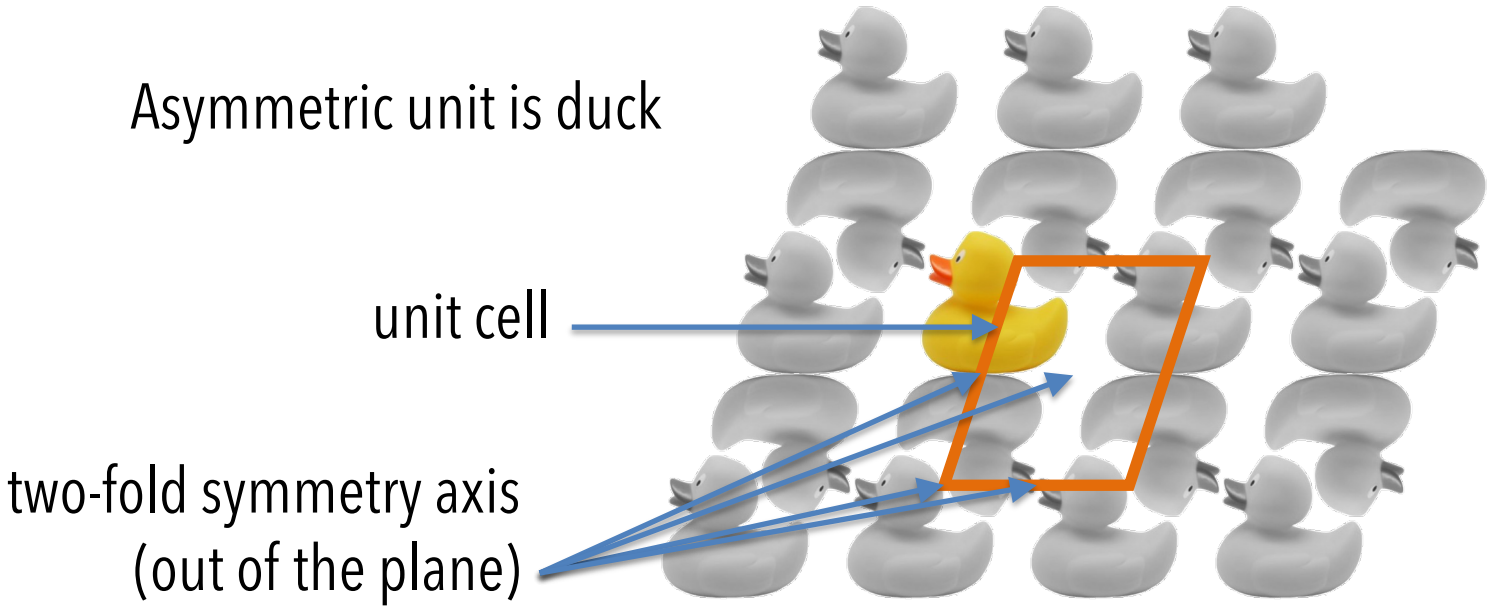
- there are many contacts between molecules to build a crystal
- biological assemblies can only be verified by experiment

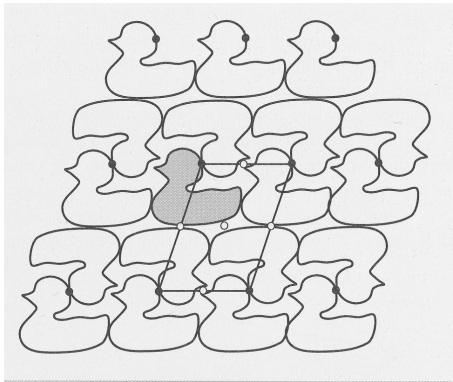


searching for multiple copies

asymmetric unit

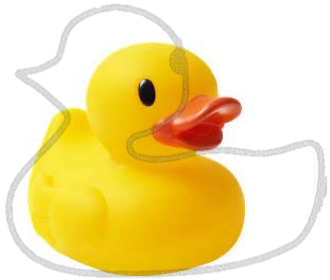
- The asymmetric unit is the smallest unit of structure that can generate the whole crystal after application of the crystal symmetry



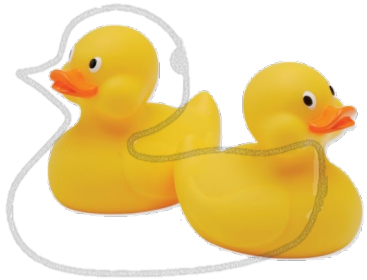


asymmetric unit contents

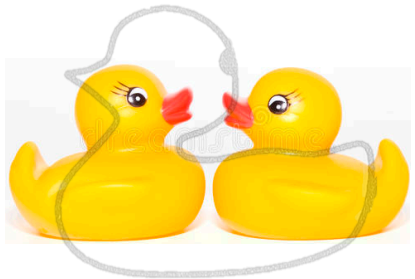
- duplication: non-crystallographic symmetry
- does not propagate through crystal



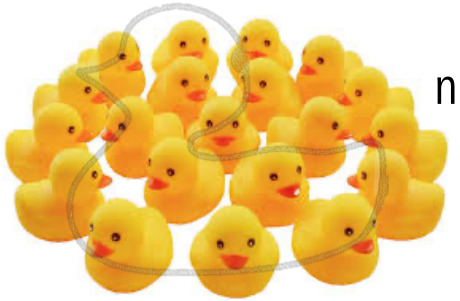
A duck



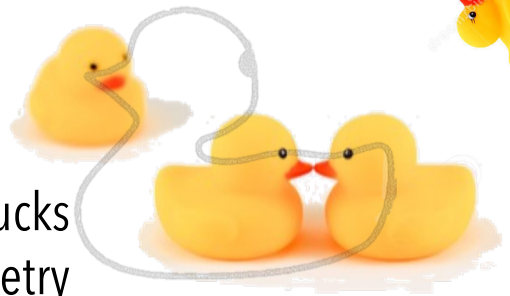
A brace of ducks
No point group symmetry



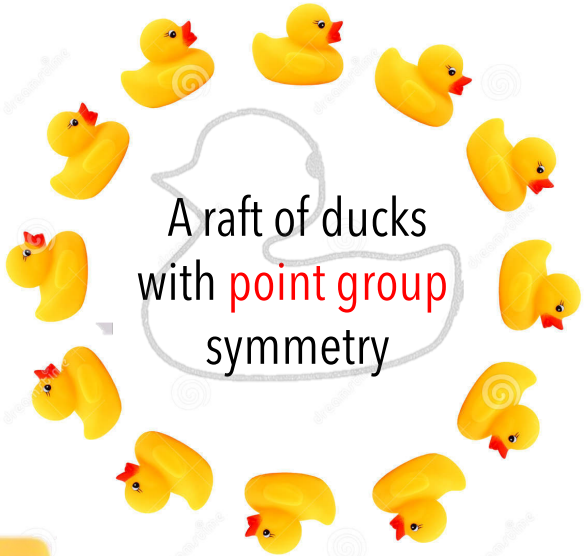
A brace of ducks
with point group
symmetry



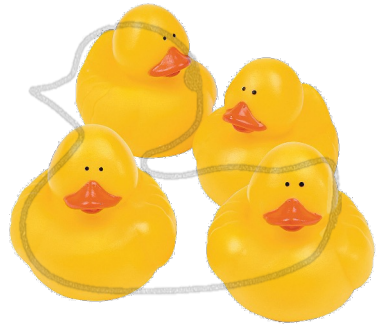
A raft of ducks
non-crystallographic
symmetry



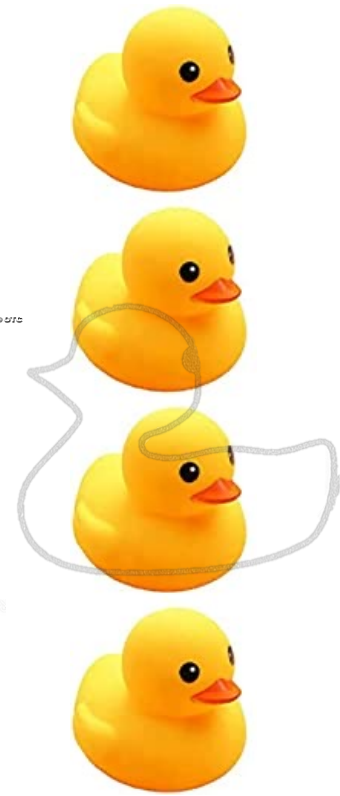
A leash of ducks
with partial point group symmetry



A raft of ducks
with point group
symmetry



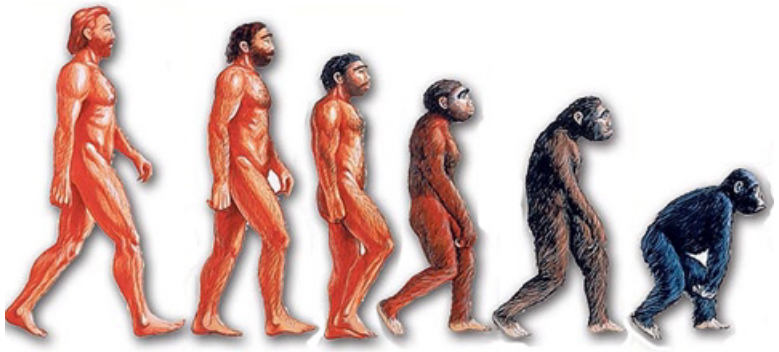
Four ducks
without translational
symmetry

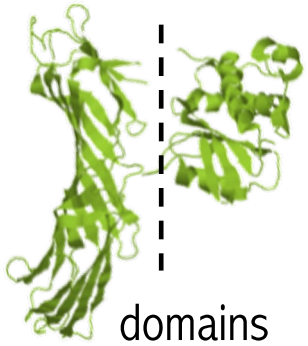


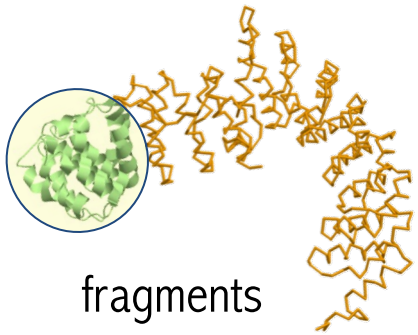
Four ducks
with translational
symmetry

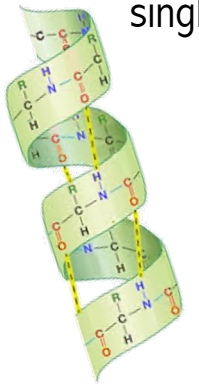
asymmetric unit contents

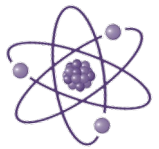
- A crystal structure is not solved by molecular replacement until ALL the components in the asymmetric unit have been found
- For molecular replacement each component of the asymmetric unit is a collection of atoms with correct local structure (disposition of atoms)











Matthew's coefficient

- First calculated by Brian Matthews in 1968 (over 3500 citations)
- Most crystals are 50% protein by volume
- Can be used to estimate the contents of the asymmetric unit

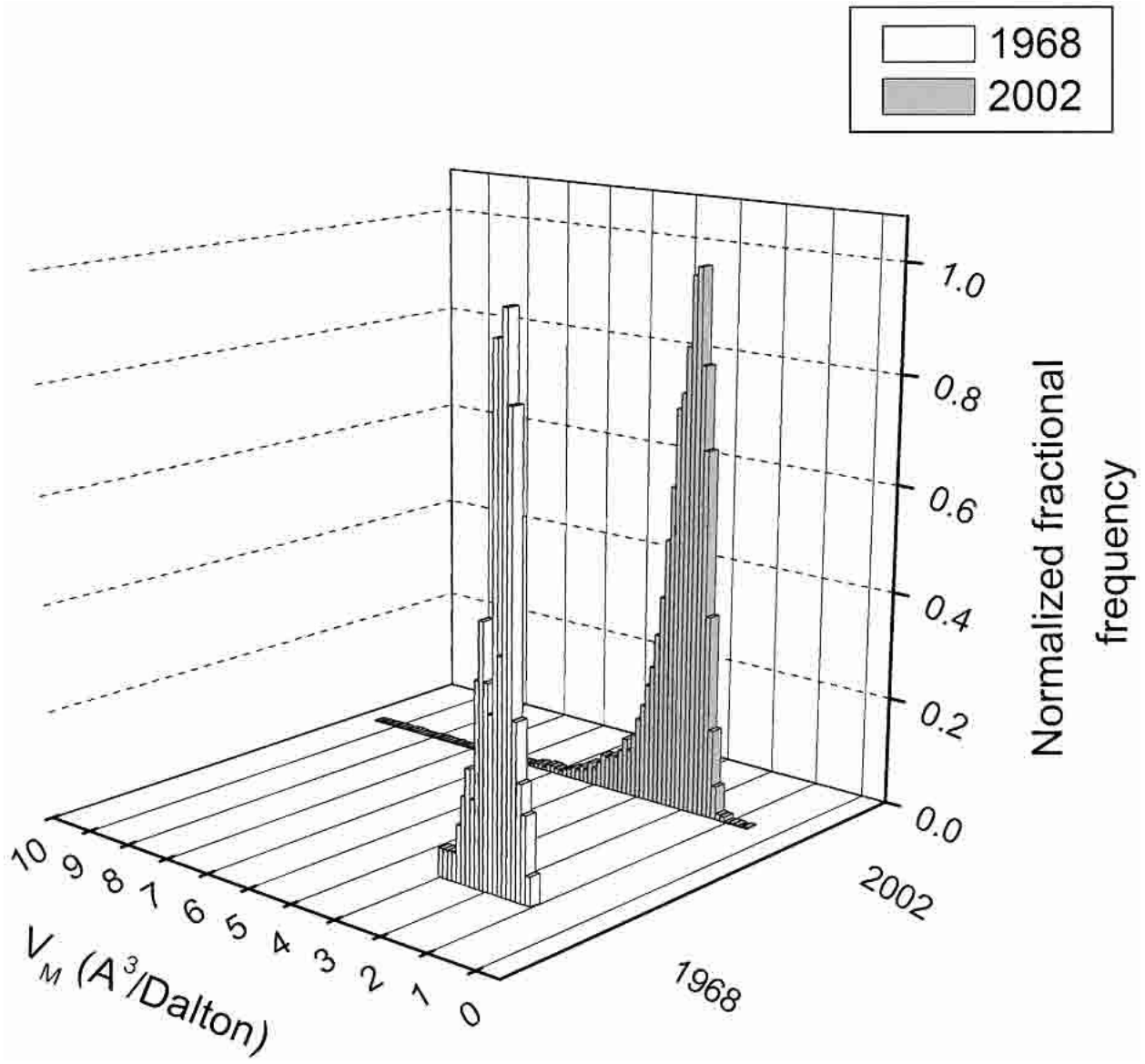
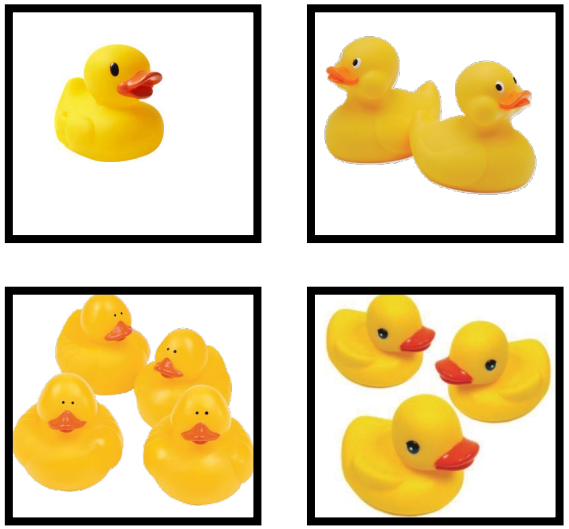


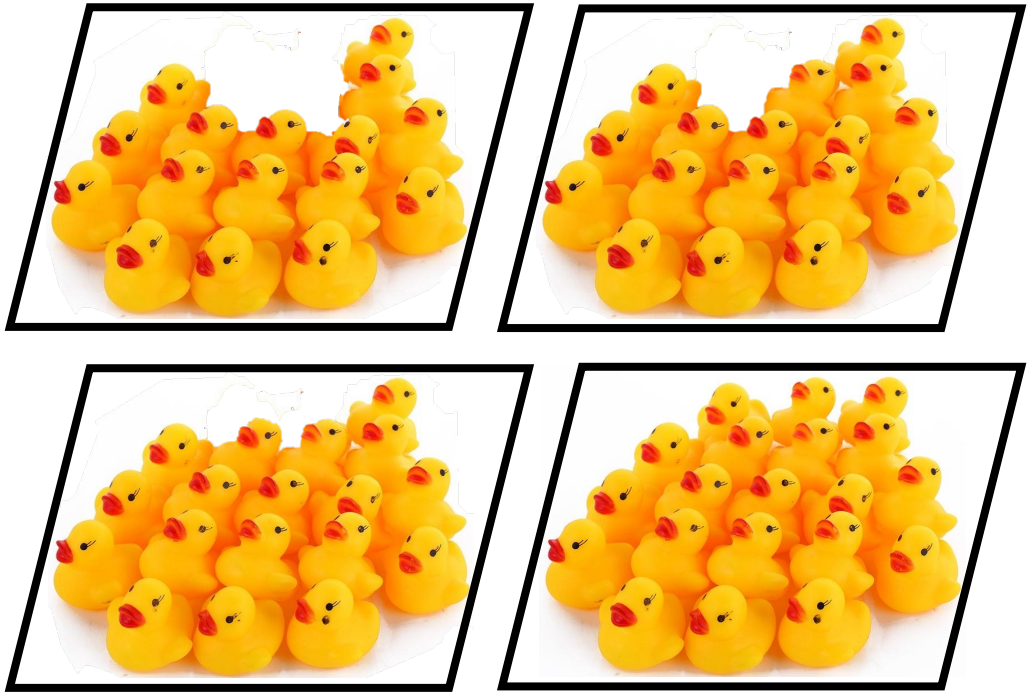
Figure 1: Kantardjieff and Rupp (2003)

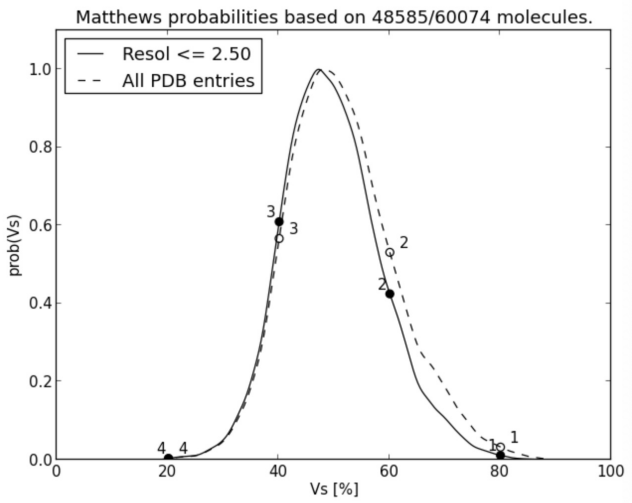
components of asymmetric unit

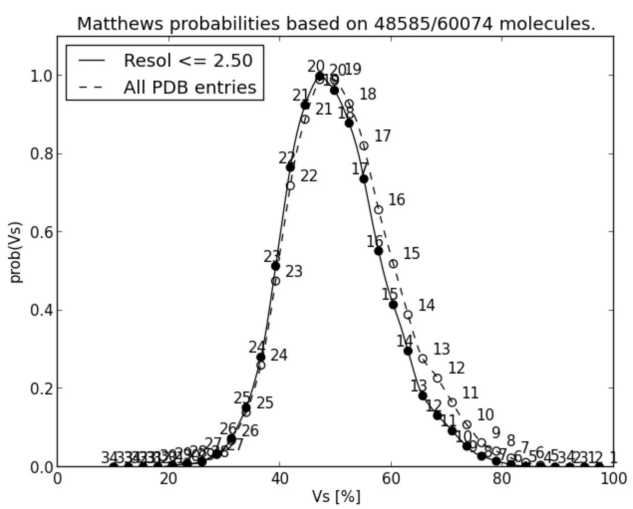


With low numbers of possible copies, options are low

With high numbers of possible copies, options are much greater



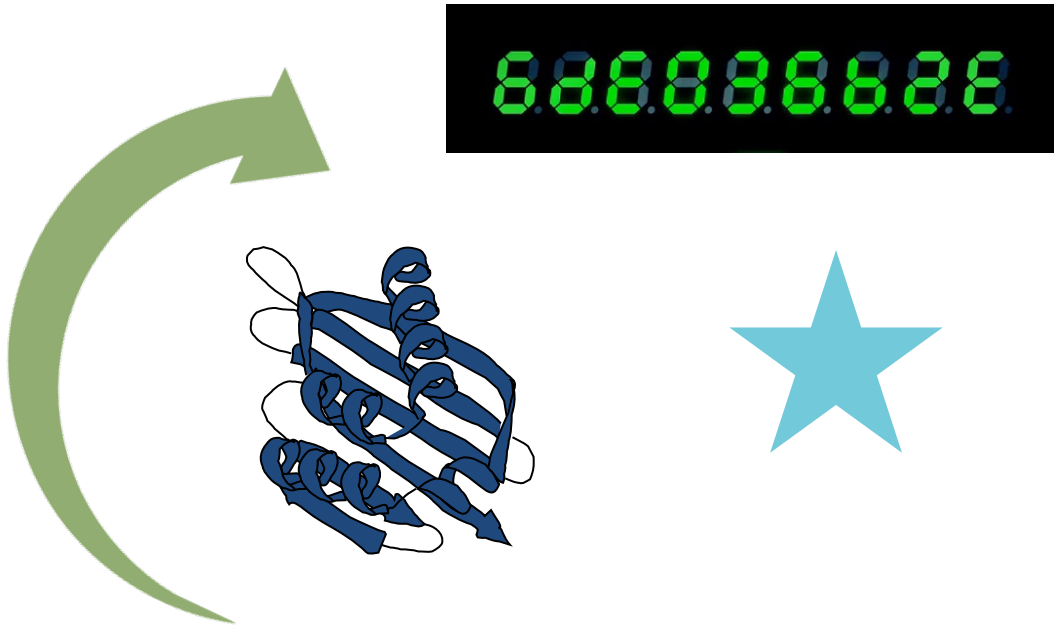




running phaser

rotation function search

- Place model at orientations and calculate probability of each being correct

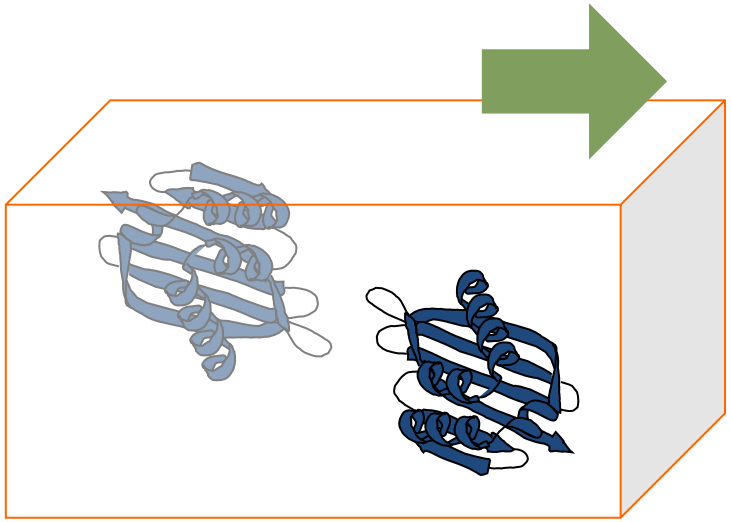


Euler Angles

- The scoring function is the LLG

translation function search

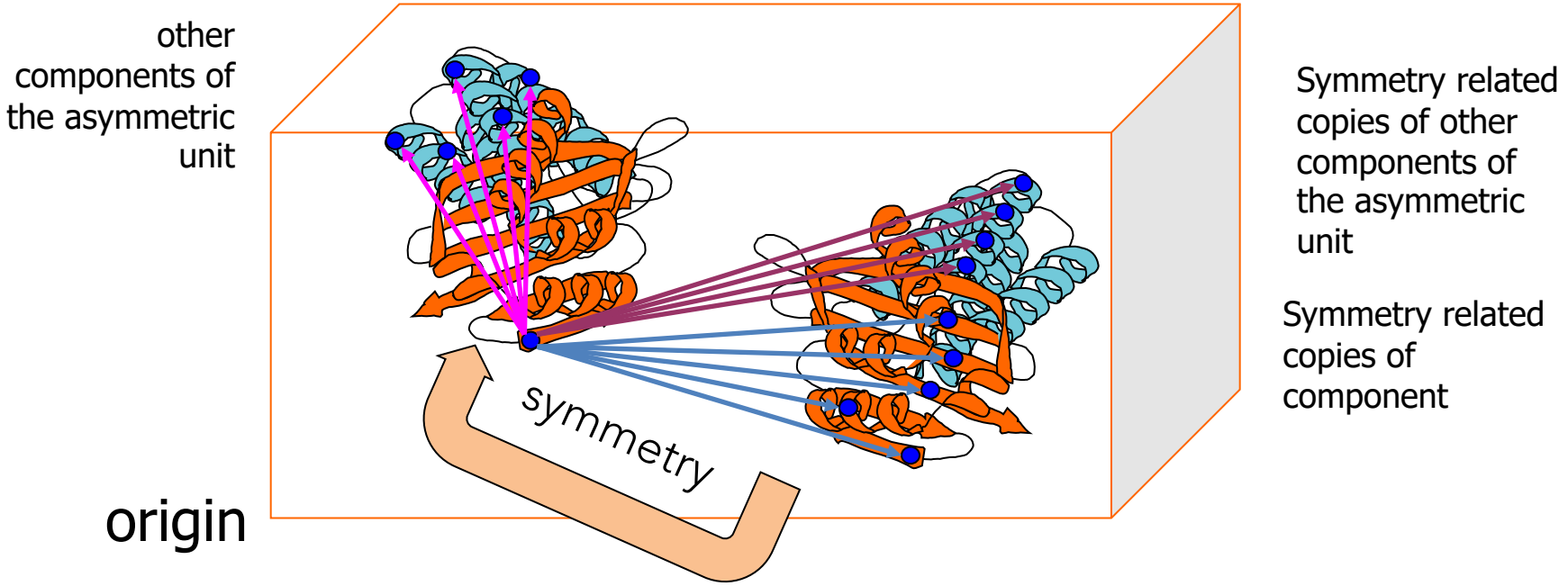
- Place model at points in unit cell and calculate probability that it is in each position



- The scoring function is the LLG

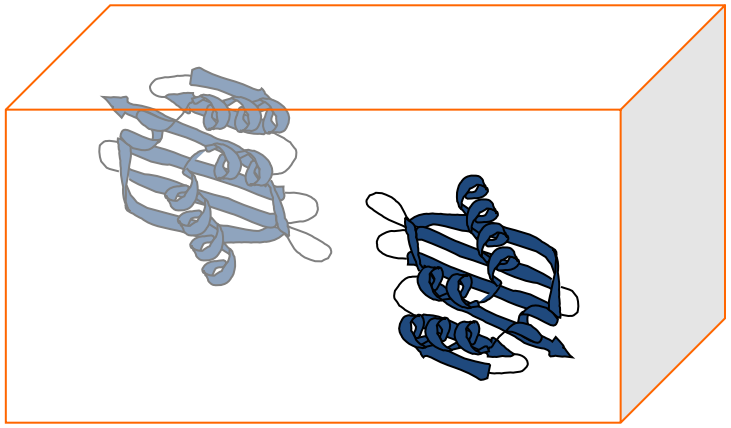
packing analysis

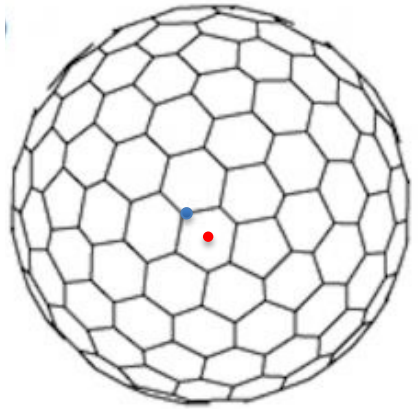
- Cα clash test
- (mostly) independent of likelihood score

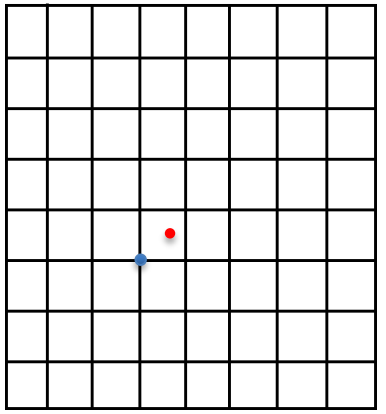


refinement

- Optimize orientation and position away from grid search locations



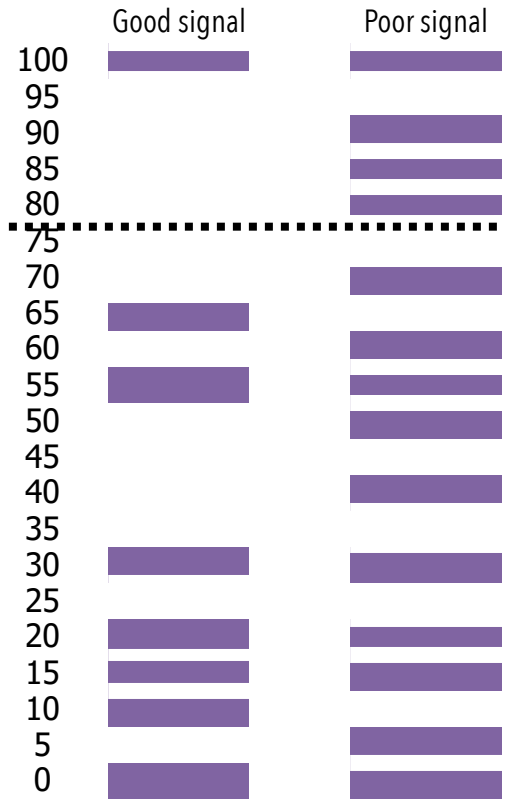


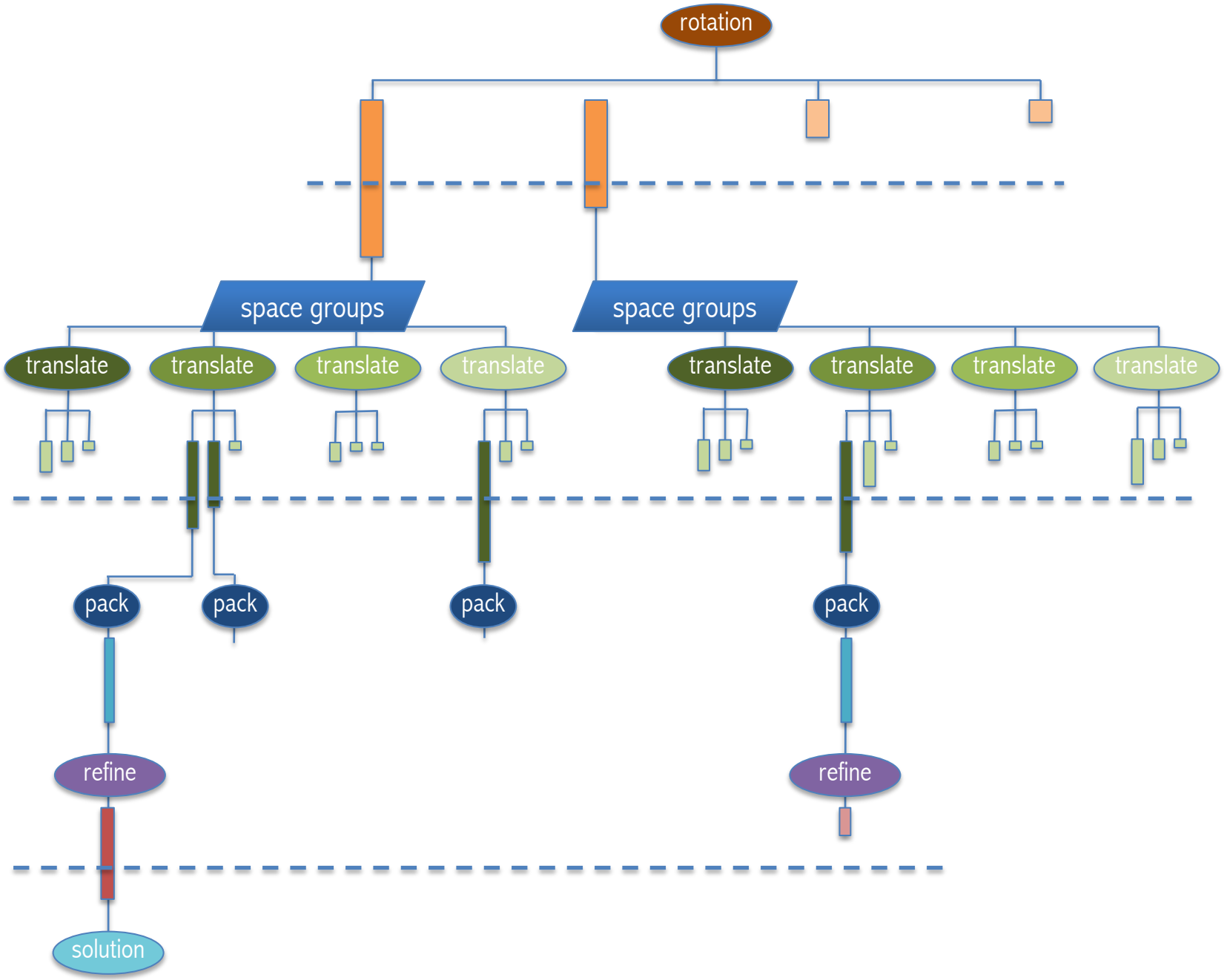


- The scoring function is the LLG

peak selection

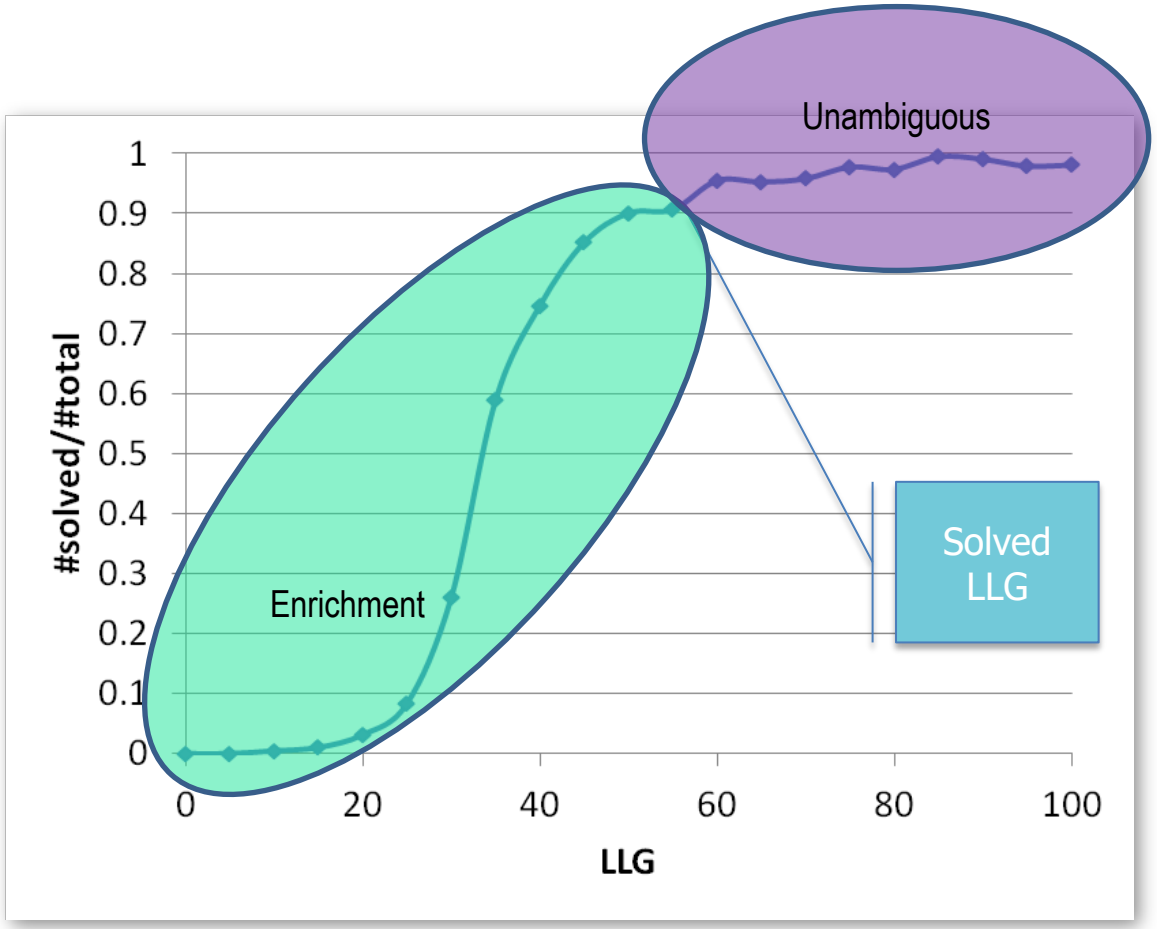
- The scoring function is the LLG
  - Log-likelihood gain
- Must chose a selection criteria to carry potential solutions through to the next step
- By default, solutions over 75% of the difference between the top peak and the mean are selected
  - Good signal, few potential solutions
  - Poor signal, many potential solutions




rotation
space groups
space groups
translate
translate
translate
translate
translate
translate
translate
translate
pack
pack
pack
pack
refine
refine
solution

Do I have a solution?
Will I get a solution?

log-likelihood gain for solutions



Plot of LLG versus success in structure solution

Database of over 23000 MR problems

R.D. Oeffner

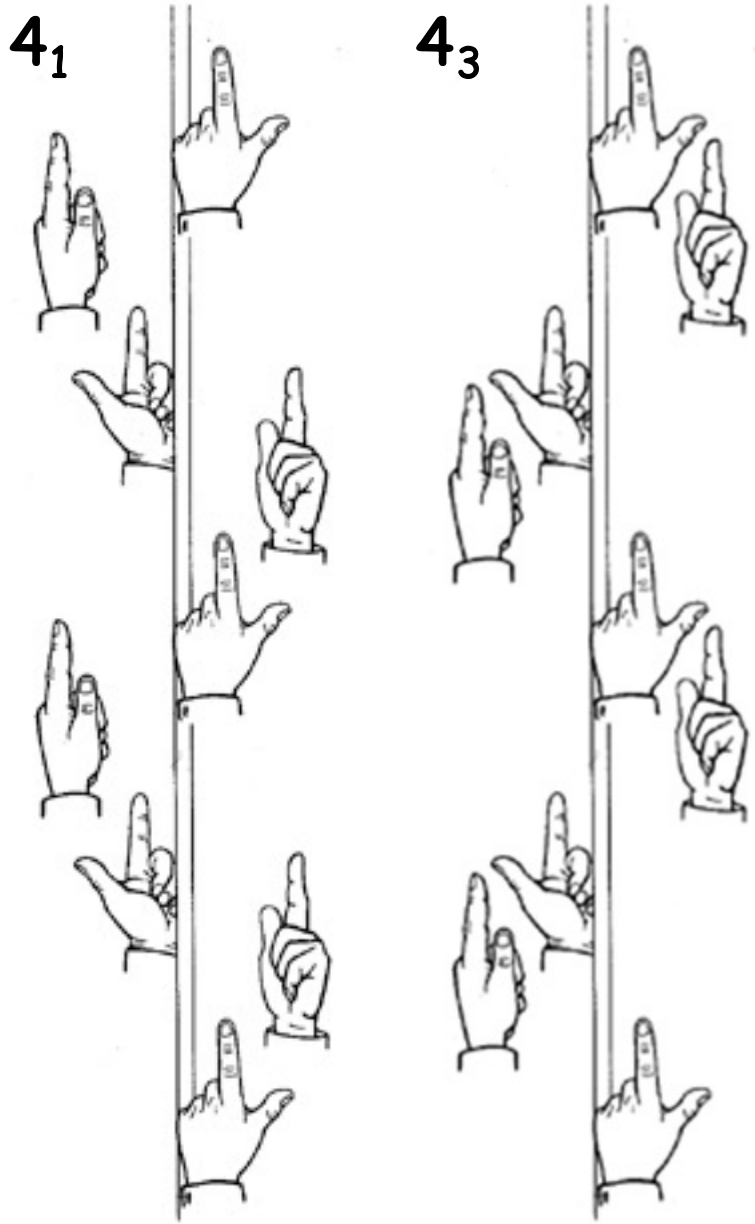
When is a model correctly placed?

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

space group alternatives
(or: why your space group might change)

space group determination

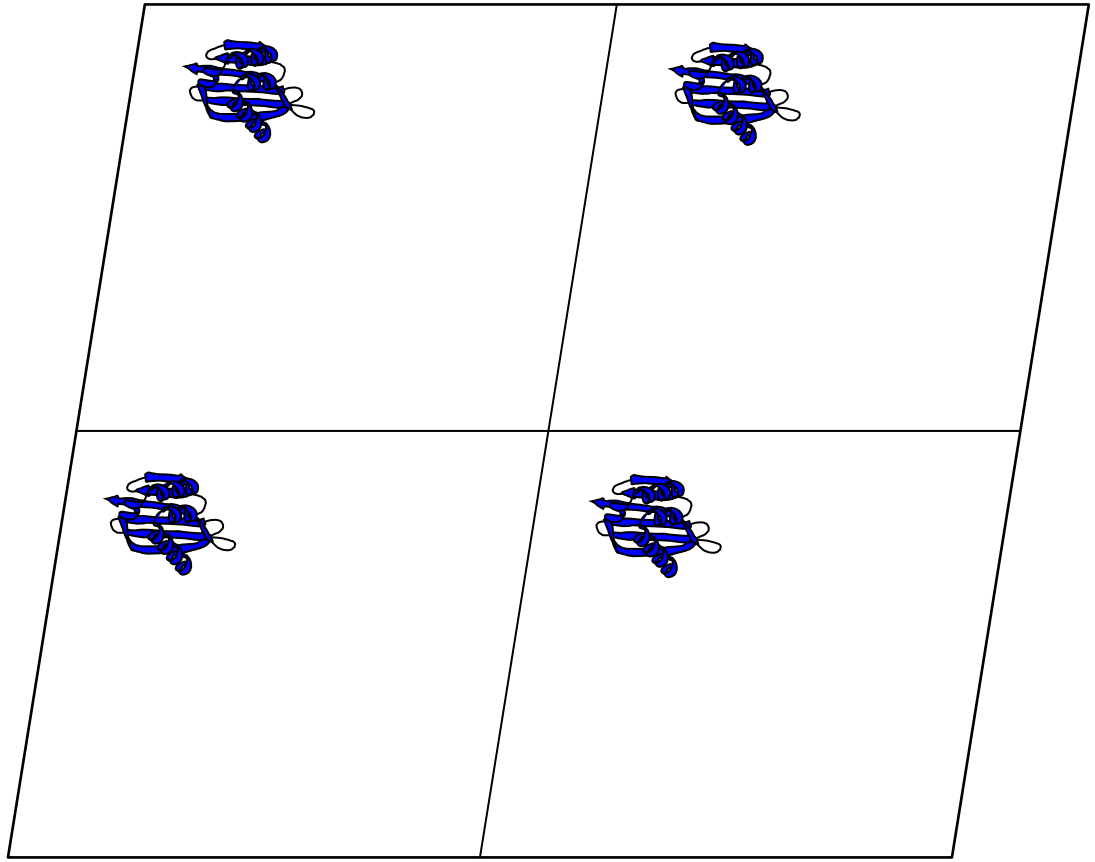
- Space groups that come in enantiomorphic pairs (e.g. $P4_1$, $P4_3$) cannot be distinguished at the data processing stage
- The space group is only confirmed when the structure is solved



alternative origins
(or: why you and your neighbour might get completely different, yet also correct, solutions)

P1

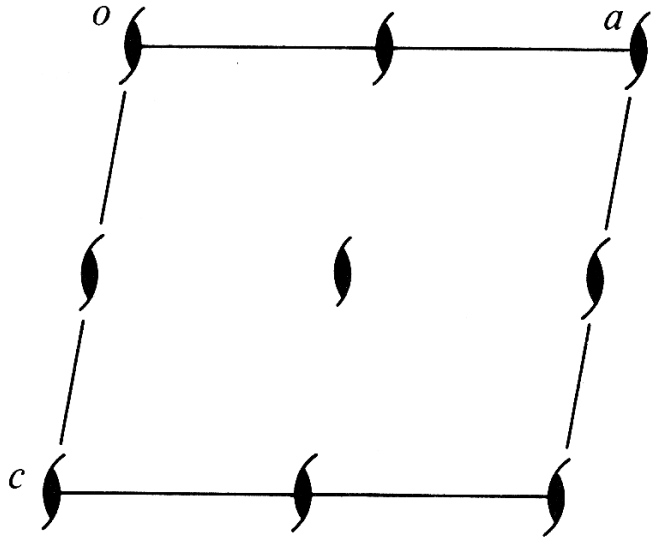
- Origin arbitrary

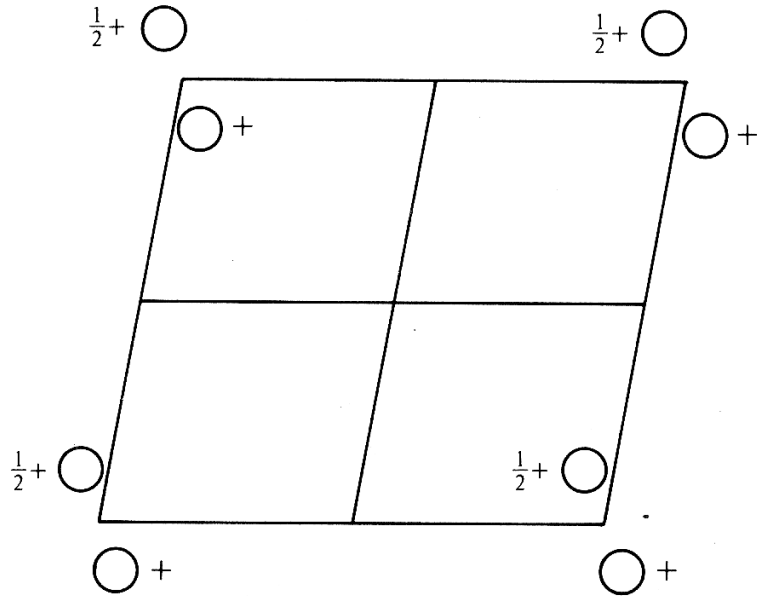


$P\,2_1$

No. 4

UNIQUE AXIS *b*





Origin on 2_1

Asymmetric unit $0 \le x \le 1;\quad 0 \le y \le 1;\quad 0 \le z \le \frac{1}{2}$

Symmetry operations

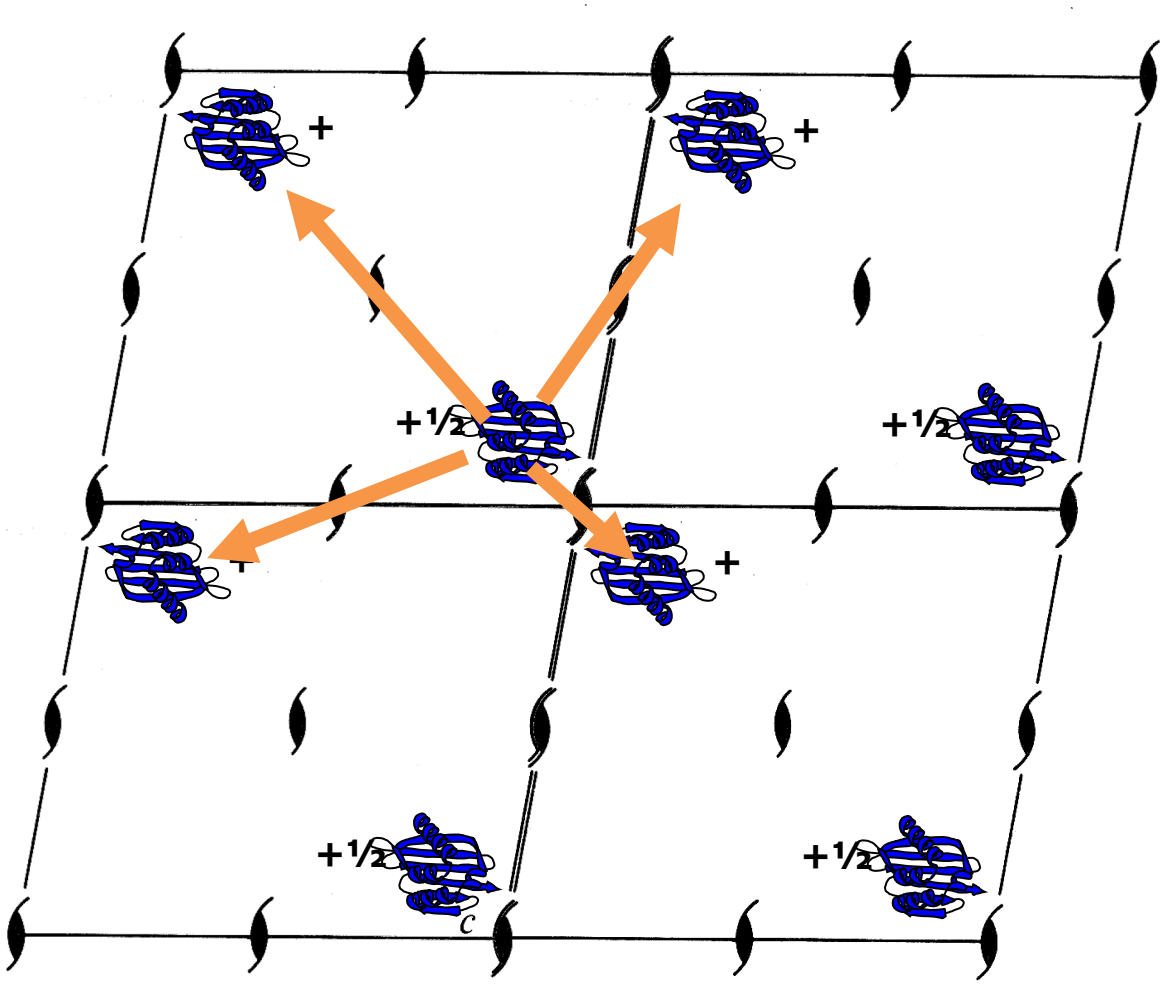
(1) 1 (2) $2(0,\frac{1}{2},0)$ $0,y,0$

Positions

Multiplicity, Wyckoff letter, Site symmetry			Coordinates	
2	*a*	1	(1) x,y,z	(2) $\bar{x},y+\frac{1}{2},\bar{z}$

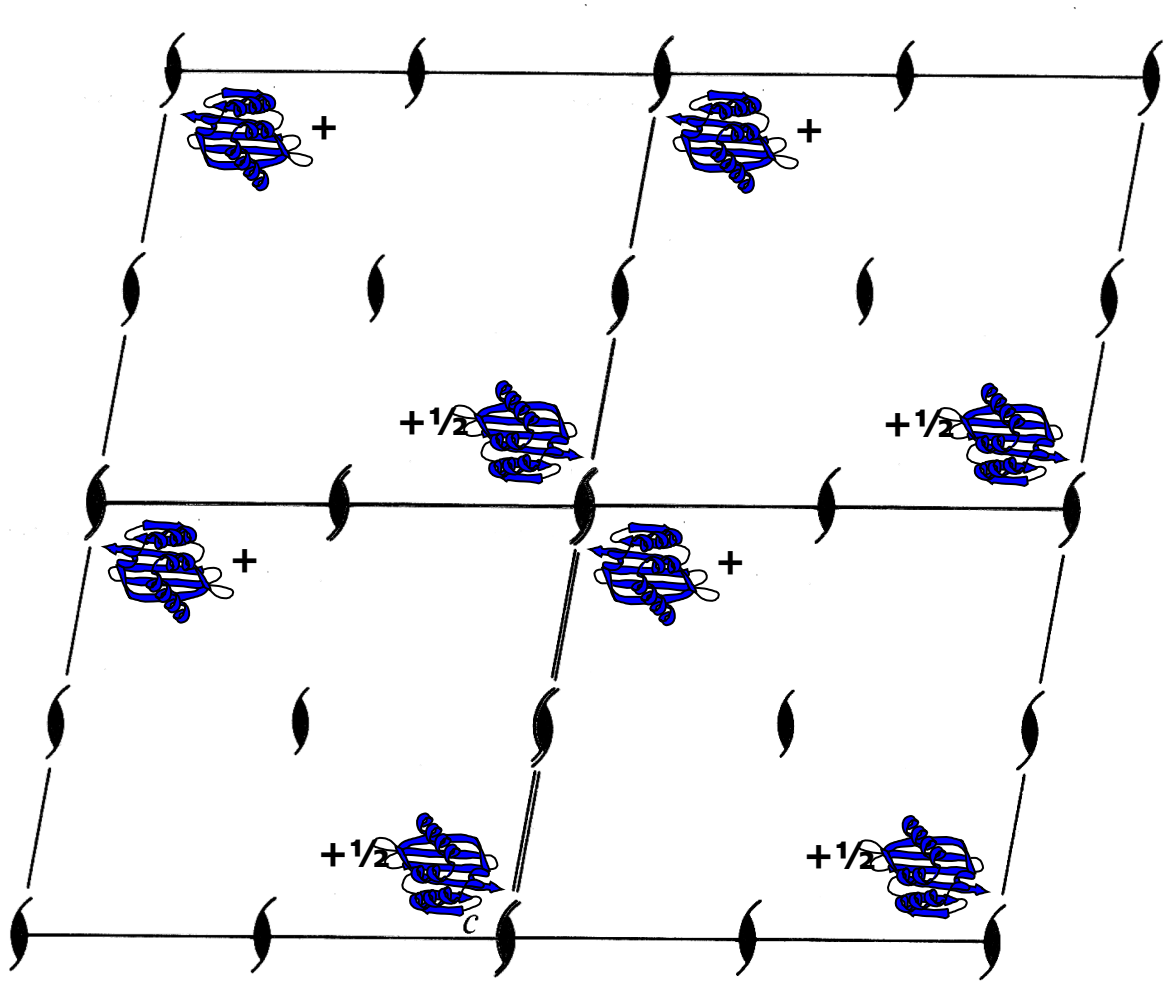
$P2_1$

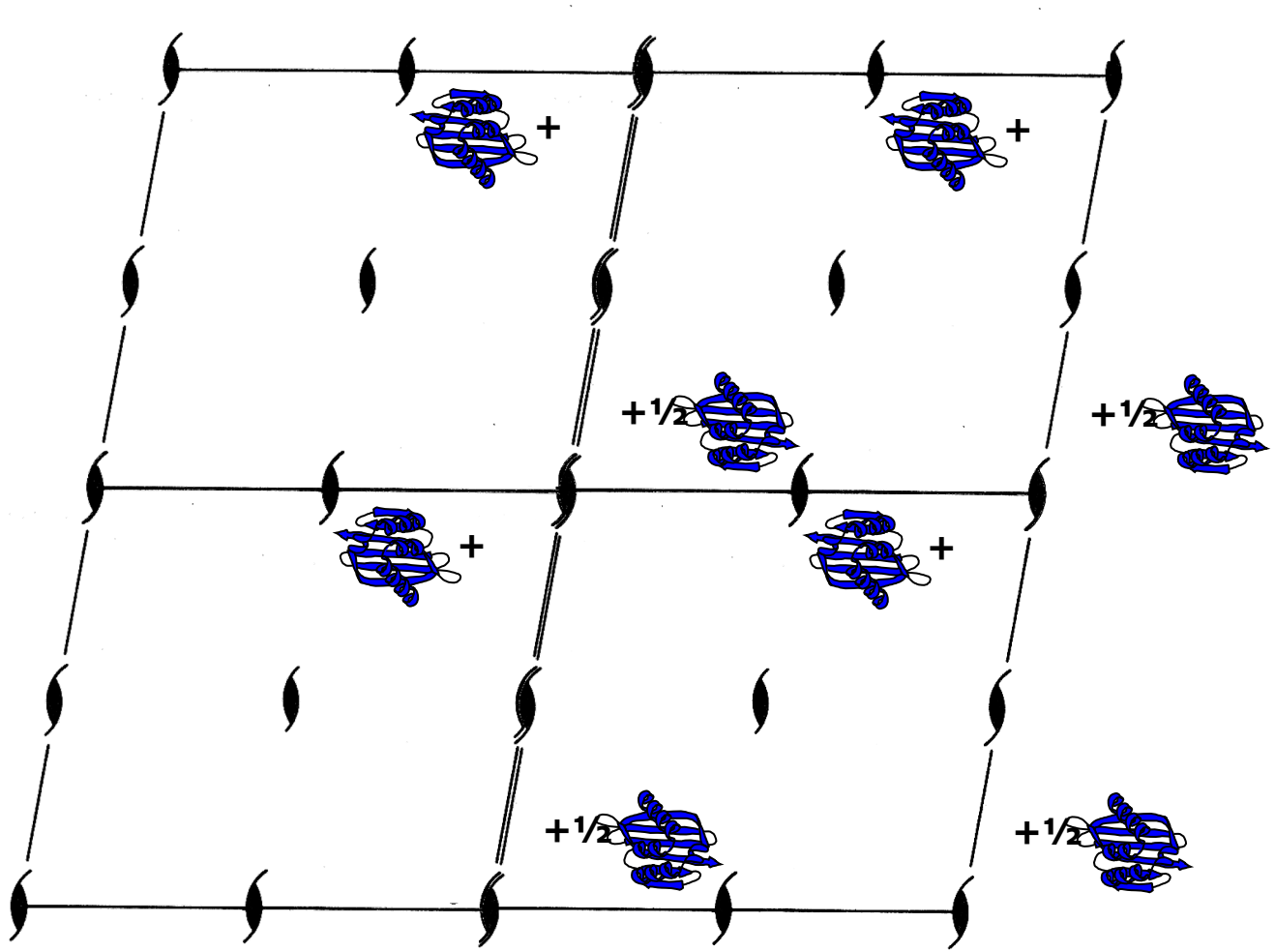
- Origin anchored at symmetry operations
- Symmetry operations (x,y,z), (-x,y+½,-z)

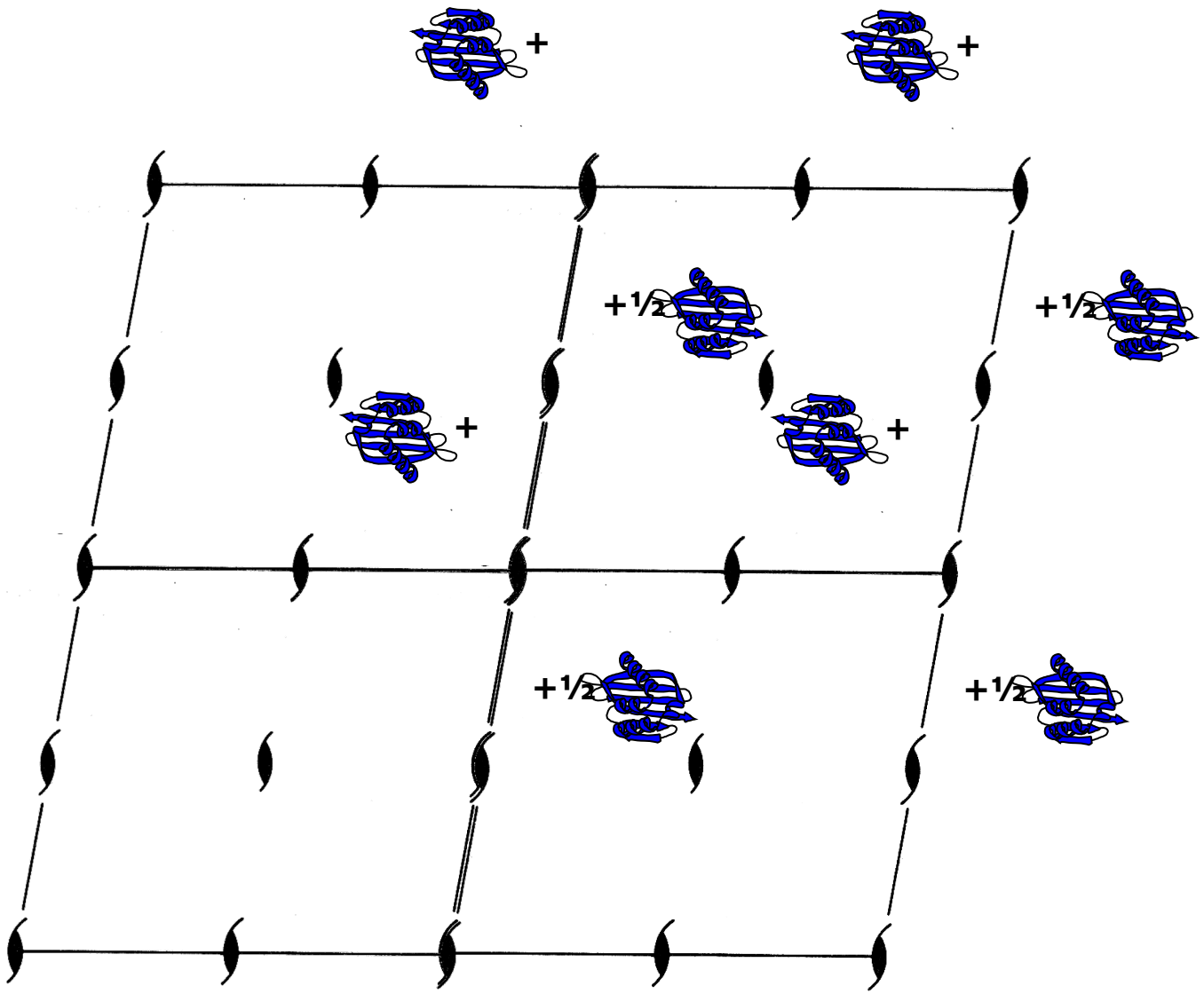


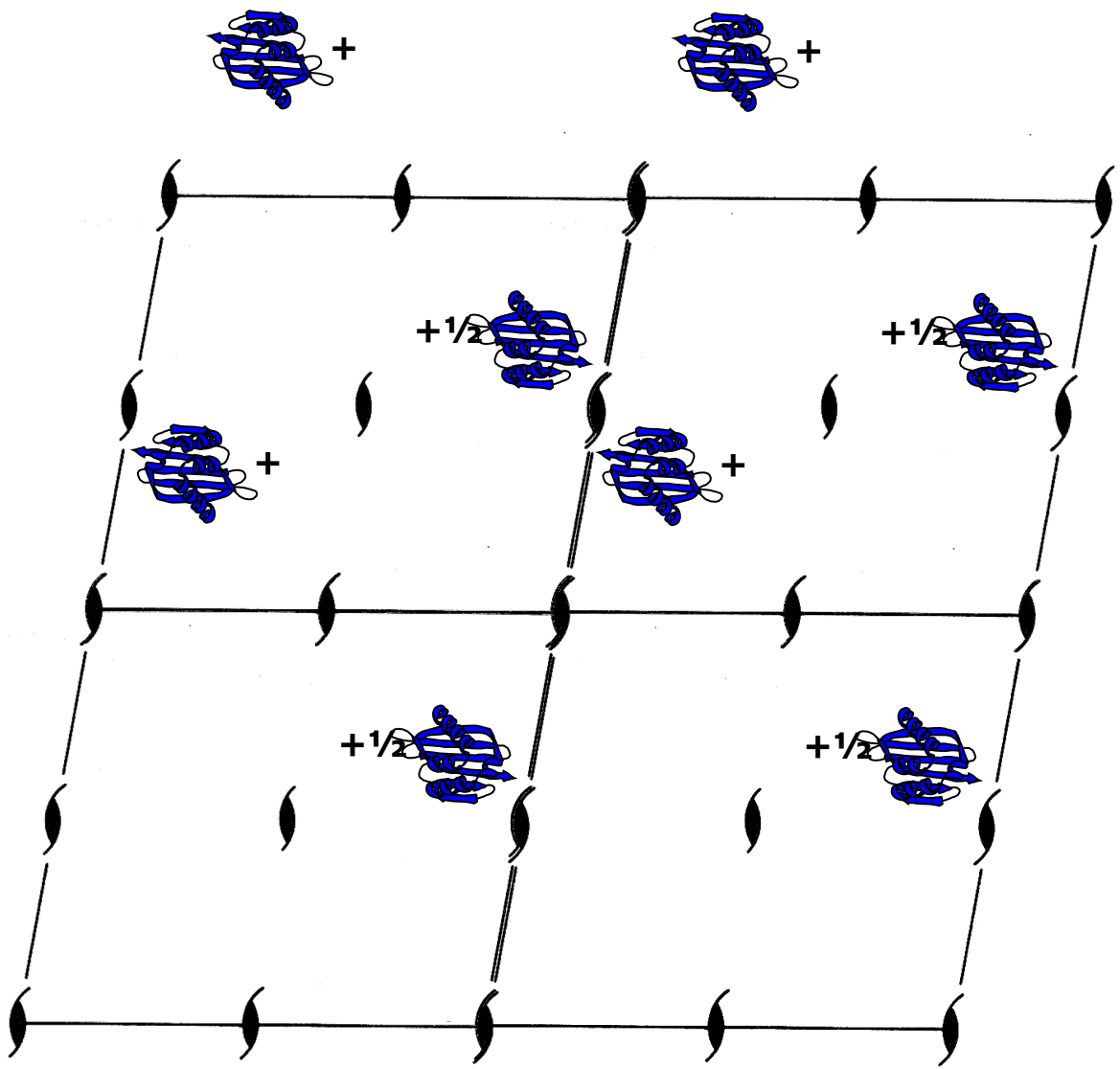
origins

- Different molecular replacement solutions may be on “different origins” and the translation values may be different
- But when you ‘build the crystal’ from the solutions the crystal looks the same


+
+
+½
+½
+
+
+½
+½
c


+
+
+½
+½
+
+
+½
+½


+
+
+½
+½
+
+
+½
+½


+
+
+½
+½
+
+
+½
+½