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

Rotation

Translation

Model structure

Unknown crystal structure

origin

(X,Y,Z)

(F,Y,K)
molecular replacement

Model structure

Translation

Rotation

Origin

Unknown crystal structure

F, Y, K

(X, Y, Z)
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.

<table>
<thead>
<tr>
<th>H</th>
<th>K</th>
<th>L</th>
<th>F</th>
<th>φ</th>
</tr>
</thead>
<tbody>
<tr>
<td>251</td>
<td>12.6</td>
<td>120</td>
<td></td>
<td></td>
</tr>
<tr>
<td>252</td>
<td>2.1</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>253</td>
<td>69.9</td>
<td>280</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>H</th>
<th>K</th>
<th>L</th>
<th>F</th>
<th>φ</th>
</tr>
</thead>
<tbody>
<tr>
<td>251</td>
<td>10.4</td>
<td>142</td>
<td></td>
<td></td>
</tr>
<tr>
<td>252</td>
<td>3.1</td>
<td>34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>253</td>
<td>52.2</td>
<td>250</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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
symmetry and molecular replacement
molecular replacement

Model structure

Translation

Rotation

(Φ,Ψ,Κ)

(X,Y,Z)

Origin

Symmetry

Unknown crystal structure
molecular replacement

Model structure

Translation

Rotation

(\Phi, \Psi, \kappa)

(X, Y, Z)

origin

symmetry

Unknown crystal structure
molecular replacement

Model structure

origin

Unknown crystal structure
molecular replacement

Translation

Model structure

Rotation

(\Phi, \Psi, \kappa)

Model structure

(X, Y, Z)

symmetry

Unknown crystal structure

origin
molecular replacement

Model structure

Rotation

(Φ,Υ,Κ)

Translation

(X,Y,Z)

unknown crystal structure

origin

symmetry
molecular replacement

Model structure

Translation

(X,Y,Z)

Rotation

(Φ,Ψ,Κ)

Unknown crystal structure

origin

symmetry
molecular replacement

- Model structure
- Rotation
- Translation
- (X,Y,Z)
- Origin
- Symmetry
- Unknown crystal structure
- Biological oligomer placement

Diagram:
- Model structure
- Rotation
- Translation
- (X,Y,Z)
- Origin
- Symmetry

Equation:
\[(\Phi, \Psi, \kappa)\]
molecular replacement

Model structure

Translation

(Φ,Ψ,K)

Rotation

Non-biological oligomer placement

Unknown crystal structure

symmetry

origin

(X,Y,Z)
molecular replacement

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

Model structure

Rotation

Translation

(Φ,Ψ,Κ)

(X,Y,Z)

origin

symmetry

Unknown crystal structure
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 is duck

unit cell

two-fold symmetry axis (out of the plane)
asymmetric unit contents

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

A brace of ducks
No point group symmetry

A raft of ducks
with point group symmetry

A raft of ducks
non-crystallographic symmetry

A leash of ducks
with partial 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

- 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
Do I have a solution?
Will I get a solution?
log-likelihood gain for solutions

Database of over 23000 MR problems

Plot of LLG versus success in structure solution

R.D. Oeffner
When is a model correctly placed?

<table>
<thead>
<tr>
<th>TF Z-score</th>
<th>LLG score</th>
<th>Solved?</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 5</td>
<td>&lt; 25</td>
<td>no</td>
</tr>
<tr>
<td>5 - 6</td>
<td>25 - 36</td>
<td>unlikely</td>
</tr>
<tr>
<td>6 - 7</td>
<td>36 - 49</td>
<td>possibly</td>
</tr>
<tr>
<td>7 - 8</td>
<td>49 - 64</td>
<td>probably</td>
</tr>
<tr>
<td>&gt; 8</td>
<td>&gt; 64</td>
<td>definitely</td>
</tr>
</tbody>
</table>
space group alternatives
(or: why your space group might change)
space group determination

- Space groups that come in enantiomorphous pairs (e.g. P4₁, P4₃) 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  
$0 \leq x \leq 1; \ 0 \leq y \leq 1; \ 0 \leq z \leq 1$

Symmetry operations

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

<table>
<thead>
<tr>
<th>Positions</th>
<th>Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiplicity, Wyckoff letter, Site symmetry</td>
<td>2 $a$ 1 $(1) \ x, y, z$ $(2) \ \bar{x}, y+\frac{1}{2}, \bar{z}$</td>
</tr>
</tbody>
</table>
P2₁

- Origin anchored at symmetry operations
- Symmetry operations \((x,y,z), (-x, y+\frac{1}{2}, -z)\)
• 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