

Molecular Re-Placement of a structurally similar search model

Re-placed partial model: Phases

Actual structure: Data, i.e. amplitudes

Either in one shot - 6-dimensional stochastic/evolutionary search

or by successive 3-d rotation and 3-d translation search using Patterson methods

MSC Chapter 11 - Copyrighted Material © Rupp 2010

24 / 39

Web compiler

[illegible]

Evolutionary 6-D multi-solution search: EPMR



Figure 11-19 Molecular replacement by evolutionary search. Evolutionary algorithms are an efficient approximation technique implemented in multi-dimensional searches. A search model is placed at random into a unit cell, and evaluated in a stochastic tournament. The winners are kept, and a new generation with perturbations around the initial population regenerate base populations (inner loop). After 10 evolution cycles, the best model is kept, and a new run begins (outer loop). The best solution from all runs is finally used to generate a bias minimized map for model building.

```

graph TD
    A[Protein sequence] --> B[Select search model]
    A --> C[Model structure factors in box]
    B --> D[Prepare search model(s)]
    D --> E[Generate random search probes]
    E --> F[Interpolate probe structure factors]
    F --> G[Score solutions by correlation coefficient]
    G --> H[Stochastic tournament]
    H --> I[Regenerate population from survivors]
    I --> J[Bias minimized initial electron density]
    J --> K[Interpretable map: Model building]
    C --> L[Cell and space group Diffraction data]
    L --> G
    G --> M[Next generation]
    M --> N[Next evolution run]
    N --> O[Score coefficient]
    O --> K
  
```

F-based, depends on superfast structure factor calculation (interpolation). Slow but with large radius of convergence - often yielding a flat multi-solution landscape

EMR Center 11 • Copyrighted Material • Run 2010

26 / 39

Web compiler

Typical multi-solution landscape

Figure 11-20 Molecular replacement solution of one independent

crystal. Multiple independent solutions independently or sequentially molecular replacement search are not required. The correlation between the native oligomer and the model is shown in the

Figure 11-21 Molecular replacement search using the evolutionary KARS search program (P4M). The solution landscape of the search for the first molecule (top panel) is quite flat with the correct solution ($C.C. = 0.288$) appearing here in trials 2 and 17. In contrast, the solution for the second copy of the cytochrome dimer is quite sharp and well discriminated from the average solution ($C.C. = 0.444$), but it appears late in the trials.

Figure 11-22 Molecular replacement search using the evolutionary KARS search program (P4M). The solution landscape of the search for the first molecule (top panel) is quite flat with the correct solution ($C.C. = 0.288$) appearing here in trials 2 and 17. In contrast, the solution for the second copy of the cytochrome dimer is quite sharp and well discriminated from the average solution ($C.C. = 0.444$), but it appears late in the trials.

Figure 11-23 Molecular replacement search using the evolutionary KARS search program (P4M). The solution landscape of the search for the first molecule (top panel) is quite flat with the correct solution ($C.C. = 0.288$) appearing here in trials 2 and 17. In contrast, the solution for the second copy of the cytochrome dimer is quite sharp and well discriminated from the average solution ($C.C. = 0.444$), but it appears late in the trials.

Typical multi-solution landscape

Trial number	C.C.
1	0.28
3	0.25
5	0.22
7	0.28
9	0.25
11	0.28
13	0.25
15	0.22
17	0.28
19	0.25

Trial number	C.C.
1	0.28
3	0.28
5	0.28
7	0.28
9	0.28
11	0.28
13	0.28
15	0.28
17	0.44
19	0.28

Figure 11-21 Solution scores of sequential 6-dimensional molecular replacement search. Both panels show linear correlation coefficients of 20 independent runs each of a molecular replacement search using the evolutionary KARS search program (P4M).

The solution landscape of the search for the first molecule (top panel) is quite flat with the correct solution ($C.C. = 0.288$) appearing here in trials 2 and 17. In contrast, the solution for the second copy of the cytochrome dimer is quite sharp and well discriminated from the average solution ($C.C. = 0.444$), but it appears late in the trials.

Run enough trials!

Web emulator

MBC Chapter 11 - Copyrighted Material B. Burn 2010

27 / 30



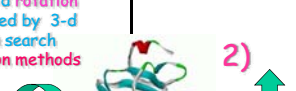
University of Bonn

Plan B: break up into rotation and translation part, use Patterson function



University of Bonn

Successive 3-d rotation search followed by 3-d translation search using Patterson methods



1) 2)

Has to work with reflection data only (no phases yet!) → done in interatomic distance space using Patterson methods

HSC Chapter 11 Copyrighted Material © Peter 2010

28 / 30

With permission

Figure 11-23 Patterson cross-rotation search. Panel A shows a search molecule positioned in a 1° box that is selected to be large enough to contain only intramolecular vectors. The Patterson map containing the intramolecular distance vectors of the search molecule is shown in panel B. The target molecule is in an as-yet-unknown orientation and location in its unit cell (C), and the corresponding vectors are shown in the next map (D). Intermolecular vectors to neighboring molecules are omitted. Note that the intramolecular vector map is independent from the location of the target molecule, but it depends on the orientation and is correspondingly rotated. By successively rotating the search map (blue dots) and calculating the overlap function at each orientation, we can find the correct rotation angle(s), in this case about 45° (panels E and F).

$$\mathfrak{R}(\mathbf{R}) = \frac{1}{V} \int_{\text{unit cell}} P(\mathbf{r}) \cdot P^*(\mathbf{R}^T \mathbf{r}) d\mathbf{r}$$