Structure based drug design

Purple: A lead for the inhibition of PDE4D;

Yellow: the end result from SBDD.


- Structural work, principally involving MX and NMR, allows the elaboration of a lead to a candidate drug.

- In this step, the chemical space to be explored is much smaller, and the synthetic efforts required much reduced compared to the first step of lead generation.

- Lead generation is thus a bottleneck in drug discovery and warrants much attention.
Structure based drug design

Historical method of drug discovery

Trial-and-error testing of chemical substances on cultured cells or animals, and matching the apparent effects to treatments.

Rational drug design

Knowledge of specific chemical responses in the body or target organism, and tailoring combinations of these to fit a treatment profile.

A particular example of rational drug design is a structure-based drug design.

The first unequivocal example of the application of structure-based drug design leading in 1995 to an approved drug is the carbonic anhydrase inhibitor dorzolamide (Trusopt) - an anti-glaucoma agent applied in the form of eye drops.
The task of ligand fitting

The data is modelled in terms of a protein ...
The task of ligand fitting

…but there is excess density at the end not accounted for…
The task of ligand fitting

…which is where the ligand should go.
The task of ligand fitting

Challenges to be addressed...

- Different resolutions and data quality
- Different ligand complexity/topology
- Partial disorder of a ligand
- Different ligands at the same site?
Automatic binding site identification

The difference density map at a low contour threshold.
Automatic binding site identification

The difference density map at a medium contour threshold.
Automatic binding site identification

The difference density map at a high contour threshold.
Automatic binding site identification

Capturing the dependence of the difference density map on changing the contour thresholds: *Fragmentation tree.*
Automatic binding site identification

Capturing the dependence of the difference density map on changing the contour thresholds: *Fragmentation tree.*
Automatic binding site identification

Capturing the dependence of the difference density map on changing the contour thresholds: *Fragmentation tree.*
The workflow of ligand fitting in ARP/wARP

Organised as a pipeline of core modules for specific sub-tasks following an intuitive approach:

- **Prepare**
  - Identify binding site and/or ligand.
  - Sparse density map and generate ligand topology.

Sparse grid representation
Constructing the ligand: Graph search

Uses the sparse grid representation of the electron density at the chosen binding site & topology of known ligand.

Knowledge about ligand

A) Connectivity
B) Distances
C) Angles
D) Chirality

Knowledge about grid

A) Connectivity
B) Distances
C) Density
D) NO IDENTITIES!

FAD

Grid
Constructing the ligand: Graph search

‘Label swapping’
- The ligand is expanded on the sparse density preserving connectivities.
- Every dummy ‘atom’ of the sparse grid is tried as a start point.
- An exhaustive graph search is performed.
- Models are scored by their fit to density and expected stereochemical features.
Constructing the ligand: Graph search

Index of starting grid point

No of ligand atoms placed

Start atom

Surviving start points

Dead branches

Start atom

Surviving start points
Constructing the ligand: Metropolis search

Perform a random walk in parameter space biased towards the optimum of a score function.

Parameter space: Position, Orientation and Conformation.
Score function: ‘Pseudo’ map correlation.
Advantage: Less degrees of freedom.

Rigid groups

Rotatable bonds
Constructing the ligand: Metropolis search

Perform a random walk in parameter space biased towards the optimum of a score function.

*Parameter space*: Position, Orientation and Conformation.

*Score function*: ‘Pseudo’ map correlation.

*Advantage*: Less degrees of freedom.

---

**ligand size vs Rotatable bonds**

![Plot showing ligand size vs Rotatable bonds](image-url)
The workflow of ligand fitting in ARP/wARP

Organised as a pipeline of core modules for specific sub-tasks following an intuitive approach:

**Prepare**
- Identify binding site and/or ligand.
- Sparse density map and generate ligand topology.

**Construct**
- Construct ensemble of ligand models in plausible conformations to fit sparsed density.
The workflow of ligand fitting in ARP/wARP

Organised as a pipeline of core modules for specific sub-tasks following an intuitive approach:

Prepare
- Identify binding site and/or ligand.
- Sparse density map and generate ligand topology.

Construct
- Construct ensemble of ligand models in plausible conformations to fit sparsed density.

Refine
- Refine ligand coordinates to satisfy geometric constraints and to maximise fit to density.
- Choose best model.
The protocol of a modeling task

Fundamental initial knowledge determines the protocol.

<table>
<thead>
<tr>
<th>Ligand identity</th>
<th>Binding site</th>
<th>1 density cluster</th>
<th>1 ligand</th>
</tr>
</thead>
<tbody>
<tr>
<td>known</td>
<td>known</td>
<td></td>
<td></td>
</tr>
<tr>
<td>known</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>unknown</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>unknown</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The protocol of a modeling task

Fundamental initial knowledge determines the protocol.

\[ N \text{ density clusters} \]

<table>
<thead>
<tr>
<th>Ligand identity</th>
<th>Binding site</th>
</tr>
</thead>
<tbody>
<tr>
<td>known</td>
<td>unknown</td>
</tr>
<tr>
<td>known</td>
<td>X</td>
</tr>
<tr>
<td>unknown</td>
<td></td>
</tr>
</tbody>
</table>

1 ligand
The protocol of a modeling task

Fundamental initial knowledge determines the protocol.

<table>
<thead>
<tr>
<th>Ligand identity</th>
<th>Binding site</th>
<th>1 density cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>known</td>
<td>known</td>
<td>N ligand candidates</td>
</tr>
<tr>
<td>known</td>
<td>unknown</td>
<td></td>
</tr>
<tr>
<td>unknown</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

G. Langer: Ligand building with ARP/wARP
The protocol of a modeling task

Fundamental initial knowledge determines the protocol.

<table>
<thead>
<tr>
<th>Ligand identity</th>
<th>Binding site</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>known</td>
<td>known</td>
<td></td>
<td></td>
</tr>
<tr>
<td>unknown</td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$N$ density clusters $N$ ligand candidates
Shape features to measure similarity

To assign a likelihood to a site knowing the ligand.

1) Surface to volume ratio
2) Bounding box limits
3) Moments of inertia
4) Rotation match score
5) Eigenvalues
6) Distance histogram
7) Geodesic distance histogram

Find the matching cluster for this ligand!

Decision is based on feature vectors.
Shape features to measure similarity

To assign a likelihood to a ligand knowing the binding site.

Final ligand ranking
High throughput ligand identification

Density map (or otherwise surface of protein pocket)

Segmentation

Projection/normalisation

Calculation of shape descriptors e.g. Zernike moments

Ligand

Ranking

Comparison

Database

G. Langer: Ligand building with ARP/wARP
Identifying Ligands

A skeleton based on a electrostatic simulation
Identifying Ligands

Reconstructions from Zernike-Moments of orders 6, 8, 10, 12, 15, 18
Success rates

Tested on > 20k PDB entries from EDS

X-ray resolution limits: 1.0 to 3.0Å,
Building the largest fully occupied ligand.
Correctness criterion: rmsd < 1.0Å

(Current Version 7.1)

5…6 atoms, any map corr.
7…100 atoms, any map corr.
10…40 atoms, map corr. > 80%
20…40 atoms, map corr. > 80%
Success criterion: r.m.s.d. < 1.0Å

Correct

Correct

In yellow: the deposited ligand

Incorrect

Incorrect

rmsd=0.4Å

rmsd=1.0Å

rmsd=1.8Å

rmsd=1.5Å
Partial disorder, partially occupied ligands

Deposited in PDB

Built automatically with ARP/wARP when the full ligand is given

SAM in 1v2x at 1.5Å

Artificial cocktail:

Compound chosen in cocktail case
Running a task through the CCP4 interface

The ligand building as part of ARP/wARP has its own GUI.

Jobs can also be run from the command line! Series of jobs can be run this way.
Running a task from the command line

Series of jobs can be run this way easily.

datafile [either miffile or mapfile]
protein {starting_PDB_file_without_ligand}
ligand {PDB file with ligand to fit}
[workdir {FULLPATH_WORKING_DIRECTORY}]
[ligandfileout {output_PDB_file}]
[fp {fp_label}][sigfp {sigfp_label}][freer {freer_label}]
ligandcycles {number_of_ligandbuild_cycles [default is 2]}
[search_model {PDB_file_with_model_at_expected_ligand_site}]
[search_position {[x y z]}]
[search_radius {radius_in_angstroms}]
[reflist {textfile_withFULLPATHnames_of_fitted_ligands_for_comparison}]
extrailib {user_defined_library_for_Refrac5}
[parfile {parfilename_if_only_parfile_is_to_be_created}]

- Optional command line arguments are given in square parentheses
- All input files are assumed to be located in working directory
  unless they are given with full path
- If workdir is not given, the current directory will be assumed
- All output files will be written into workdir/subdirectory
Ligand building with arpnavigator
Ligand building with arpnavigator

Select ‘Fit A Ligand’
And fill the form
Ligand building with arpnavigator

Load the density map and the protein. Load the ligand you want to build. Mark a density blob, detach and fit the ligand right there. Easy!
Acknowledgements

Developers

EMBL Hamburg: Ciaran Carolan, Philipp Heuser, Victor Lamzin, Gerrit Langer, Tim Wiegels

NKI Amsterdam: Krista Joosten, Robbie Joosten, Tassos Perrakis

Collaborators

EMBL Hamburg: Gleb Bourenkov, Santosh Panjikar

York University: Garib Murshudov’s group

Daresbury Laboratory: CCP4 team

Former members

Richard Morris, Peter Zwart, Francisco Fernandez, Matheos Kakaris, Olga Kirillova, Wijnand Mooij, Diederick De Vries, Marouane Ben Jelloul, Johan Hattne, Tilo Strutz, Guillaume Evrard, Serge Cohen, Venkat Parthasarathy, Babu Pothineni, Helene Doerksen

Funding