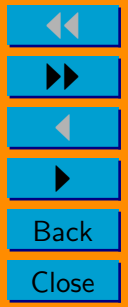
http://www.globalphasing.com

SMILES Strings and Ligand Refinement in BUSTER/TNT

Eric Blanc and Gérard Bricogne

Global Phasing Ltd.
Sheraton House, Castle Park, Cambridge CB3 0AX, England
email: blanc@globalphasing.com

Euro MUG, Cambridge 26 September 2002



Structure Refinement Process

1. The protein structure without ligand is known (or the structure of one close relative is)
2. The structure of the drug candidate is approximately known (up to free dihedral angles)
3. The protein structure is fitted against crystallographic data
4. The ligand is placed in the electron density map, and protein plus ligand are refined against crystallographic data




Back
Close

ARL2-GTP complexed with PDEδ

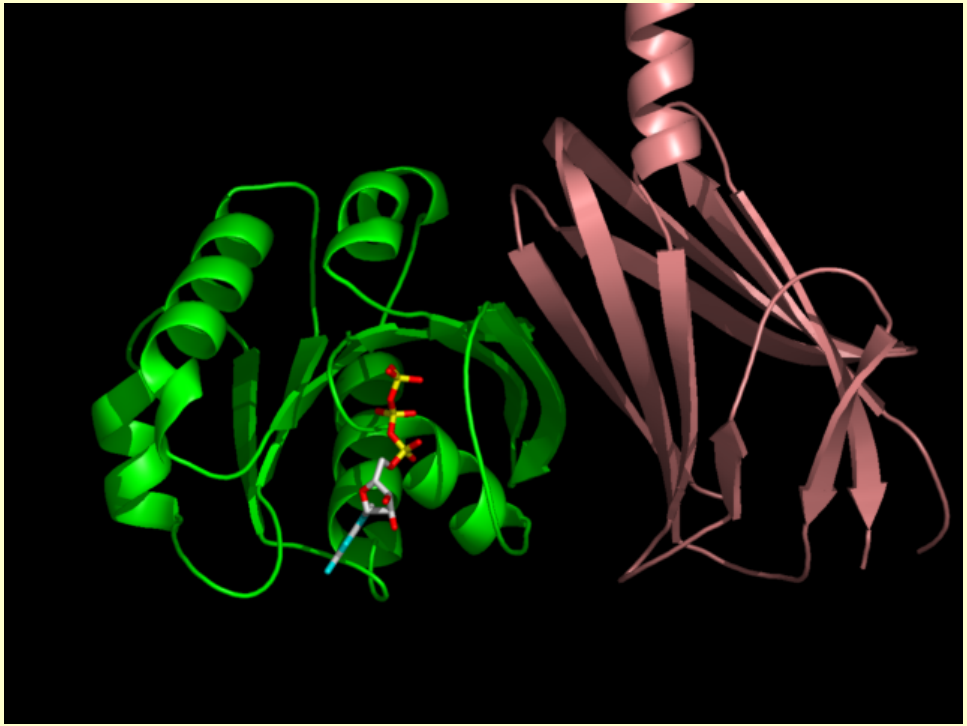
Although not directly related to drug discovery process, this example shows protein-ligand interaction

- ARL2 (Arf-like) regulates PDEδ activity, upon GTP binding and hydrolysis
- PDEδ releases membrane-anchored proteins into the cytosol
- ARL2 and PDEδ are approximately 21.0 and 17.5 kD resp., while GTP is only 0.5 kD

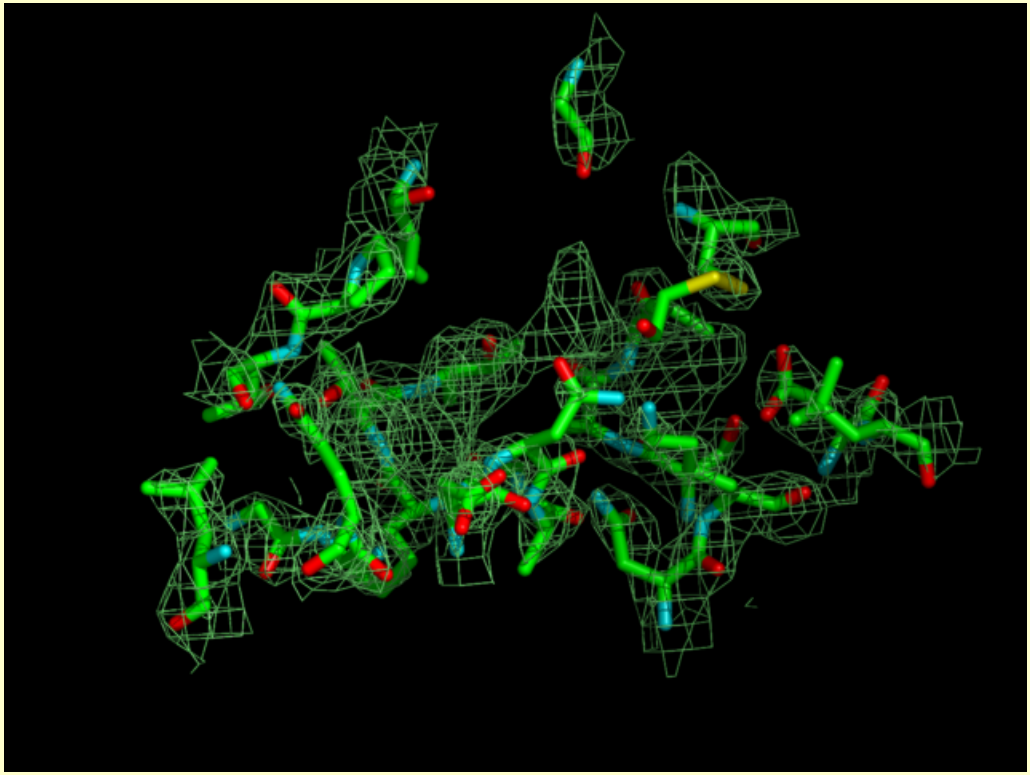



Back
Close

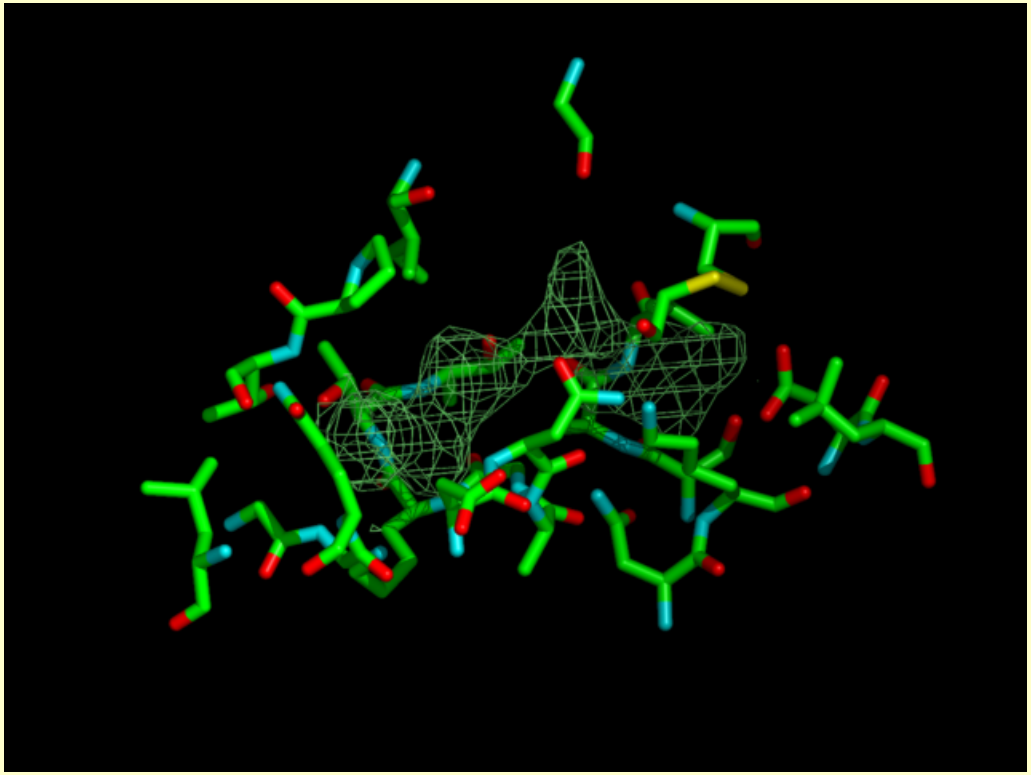
ARL2-GTP complexed with PDEδ



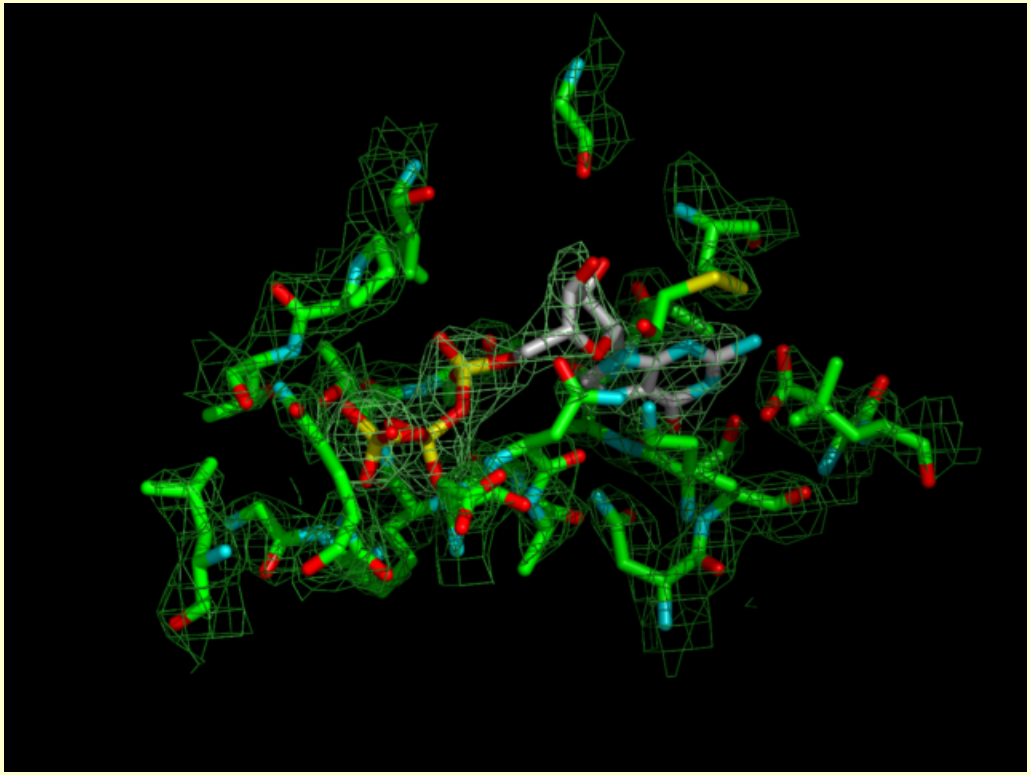
GTP Binding Site



GTP Binding Site



GTP Binding Site



Software Description

The software converts chemical information in a format suitable for crystallographic refinement

- It must be able to generate a reasonable 3D model from the SMILES string
- It must generate stereochemical restraints for the ligand to allow for further structural refinement against the crystallographic data, by matching the SMILES string describing the ligand with a set of stereochemical rules associated to SMARTS patterns

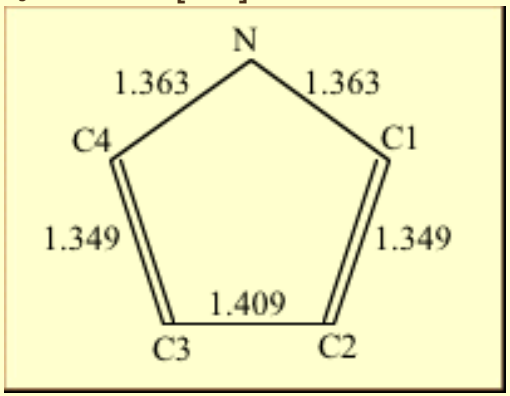
The programs are written is JAVA2 1.3, under script control for easy inclusion in the BUSTER/TNT macromolecular software package




Back
Close

Atom Labelling

Pyrrole c1[nH]ccc1



- Atoms must be uniquely labelled to remove ambiguities in the stereochemistry definition
- Care must be taken that the atom labelling in the PDB model is compatible with the labels used for the stereochemistry definition

Software Description: Input

- The ligand described by its SMILES string
- Stereochemical knowledge embedded into SMARTS patterns
- Optionally, a 3D structure of the model in PDB format

Software Description: Output

- A dictionary of stereochemical restraints for the ligand in the TNT format suitable for the macromolecular refinement program BUSTER
- An initial structural (3D) model for the ligand in PDB format, which atom labels synchronised with the labels used fot the stereochemical dictionary
- An assessment of the stereochemical quality of the ligand's 3D model, on a restraint by restraint basis

An Example: GTP

```
% $BDG_home/bin/buster/createTNTDict.sh GTP \
  'Nc1nc2n(cnc2c(=O)[nH]1)C3OC(COP(=O)(O)OP(=O)(O)...' \
  gtp.dat gtp.pdb gtp.log
```

An Example: GTP

```
% $BDG_home/bin/buster/createTNTDict.sh GTP \
  'Nc1nc2n(cnc2c(=O)[nH]1)C3OC(COP(=O)(O)OP(=O)(O)...' \
  gtp.dat gtp.pdb gtp.log
% more gtp.dat
GEOMETRY  GTP  BOND  1.25  0.05  O9  C8
GEOMETRY  GTP  BOND  1.35  0.05  O12 C11
...
GEOMETRY  GTP  ANGLE  120.0  5.0  C1  N10  C8
...
GEOMETRY  GTP  PLANE  5  0.02  C7  C3  N4  C5  N6
%
```




An Example: GTP

```
% $BDG_home/bin/buster/createTNTDict.sh GTP \
  'Nc1nc2n(cnc2c(=O)[nH]1)C3OC(COP(=O)(O)OP(=O)(O)...' \
  gtp.dat gtp.pdb gtp.log
% more gtp.pdb
CRYST1  100.000  100.000  100.000  90.00  90.00  90.00
ATOM      1  N0  GTP     1       4.350  -3.570  -1.040
ATOM      2  C1  GTP     1       4.830  -2.300  -0.810
ATOM      3  N2  GTP     1       3.920  -1.270  -0.310
ATOM      4  C3  GTP     1       4.410   0.030  -0.060
ATOM      5  N4  GTP     1       3.770   1.160   0.430
...
ATOM     32  O31 GTP     1       2.330  -0.690   2.120
%
```

An Example: GTP

```
% $BDG_home/bin/buster/createTNTDict.sh GTP \
  'Nc1nc2n(cnc2c(=O)[nH]1)C3OC(COP(=O)(O)OP(=O)(O)...' \
  gtp.dat gtp.pdb gtp.log
% more gtp.log
       obs          calc          diff  atoms
   1.40000      1.48714    -0.08714   |1:C7    |1:C3
   1.50000      1.57099    -0.07099   |1:P24    |1:O27
   1.50000      1.57054    -0.07054   |1:P16    |1:O18
   1.40000      1.46540    -0.06540   |1:C5    |1:N6
   1.50000      1.56234    -0.06234   |1:P20    |1:O22
...
   1.25000      1.24988     0.00012   |1:O9    |1:C8
%
```

Future Developments

- Improvements in atom labelling:
  Atom labels should allow some intervention from external software
  The user should have control over the final label assignement through a GUI
- Decomposition of the ligand into rigid parts:
  The manual positioning of the ligand into the electron density should be automated
  Rigid parts are found in electron density, then connected with flexible linkers




Back
Close

Acknowledgements

- Richard Morris for discussions and collaborations
- Pietro Roversi provided me with the ARL-PDE complex structure
- Clemens Vonrhein for the interface libraries to BUSTER/TNT
- John Bradshaw for his help with the Daylight Toolkit
- Pictures were prepared with PyMOL by Warren L. DeLano