

BMP - Brute force method for Molecular Replacement

R. Keegan

September 23, 2005

Contents

1	Introduction	1
1.1	Template model search	1
1.2	Template model preparation	2
1.3	Molecular Replacement and Refinement	2
2	Inatallation Notes	2
3	Installation	3

1 Introduction

BMP is a brute force method for doing the Molecular Replacement step in the determination of the structure of a protein in a protein crystallography experiment. It automates the process of finding a suitable template structure from the set of solved structures contained in the Protein Data Bank (PDB), preparing the template structures and passing them to a Molecular Replacement program. The templates are then refined to ascertain their suitability.

1.1 Template model search

Several methods are used to construct a list of template models that can be used in Molecular Replacement to determine the initial phases for our target structure. Here, we will give a brief outline of these methods.

Sequence-based search

The first step in BMP is the search for suitable homologous structures based on sequence matching against with the target sequence. A fasta search is carried out using either the OCA web-based application provided by the

EBI or using a local installation of fasta34 to perform the search using a local copy of the PDB sequence database. This search produces a list of PDB id's for the structures most closely related to the target structure in terms of their sequence.

Structure-based search (optional)

An additional, optional step may be carried out using the SSM web service provided by the EBI to search for homologous structures based on secondary structure. This is done by giving the SSM search service the top hit from the sequence-based search. If any new PDB id's are thrown up by this search, they are added to the list of PDB id's from the sequence-based search. We refer to this list as the 'match list'.

Domain-based search

Using the match list, a domain-based search is carried out. The SCOP database (parseable file) provides a list of all of the known domains contained in the solved protein structures in the PDB. This database is parsed to find any domains that may exist in the PDBs contained in the match list. If any are found they are extracted from the parent PDB file and added to the match list.

Multimer-based search

Another service provided by the EBI is the Protein Quaternary Service (PQS) which provides a database of all of the known multimeric structures based on the solved protein structures in the PDB. This database is parsed and any multimers based on PDBs contained in the match list are retrieved from the EBI web server and added to the match list.

1.2 Template model preparation

When the match list has been constructed

1.3 Molecular Replacement and Refinement

2 Inatallation Notes

The compulsory dependencies for BMP are as follows:

- CCP4 6.0. Programs used include:
 - Phaser
 - Molrep

- Amore
- Chainsaw
- Pdbcur
- Pdbset
- Coord_format
- Mtzdmp
- python 2.3 or a later
- gnuplot
- MAFFT or Clustalw

Optional dependencies are:

- Fasta34 (for local searches)
- perl (if you wish to use the SSM search facility)

3 Installation

To install BMP on your system carry out the following steps:

- Download the tar-ball `bmp.tar.gz` the ftp site.
- Extract the source code from the tar-ball using the command:

```
gzip -dc bmp.tar.gz | tar xvf -
```

- Edit the ‘`bmp_setup.sh`’ file in the ‘`bmp`’ folder to suit your system