NAME

InfoPDBFiles.pl - List information about PDB files

SYNOPSIS

InfoPDBFiles.pl PDBFile(s) PDB(s)...

InfoPDBFiles.pl [-a, --all] [-b, --BoundingBox] [-c, --count "RecordType, [RecordType,...]" | All] [--chains] [-d, --detail infolevel] [-e, --experiment] [-f, --frequency] [-h, --help] [--header] [m, --MasterCheck] [--residues] [--ResiduesMode InChains | All | Both] [--ResidueNumbers] [-w, --WorkingDir dirname] PDBFile(s)...

DESCRIPTION

List information about contents of *PDBFile(s)*: number of each record type, number of chains, count and percent distribution of residues in each chain, bounding box and so on. Multiple PDBFile names are separated by spaces. The valid file extension is *.pdb*. All other file name extensions are ignored during the wild card expansion. All the PDB files in a current directory can be specified either by **.pdb* or the current directory name.

In PDB files containing data for multiple models, all ATOM/HETAM records for chains after the first model are ignored.

OPTIONS

-a, --all

List all the available information.

-b, --BoundingBox

List min/max XYZ coordiates of ATOM/HETATM records.

-c, --count *RecordType,[RecordType,...]|All*

Types of PDB records to count in *PDBFile(s)*. You can specify a list of any valid PDB record type or count all record types found in the files. Possible values: Comma delimited list of valid *RecordTypes* or *All*. Default: *ATOM,HETATM*. And this is also default behavior.

The list of valid PDB record types includes: *HEADER, OBSLTE, TITLE, CAVEAT, COMPND, SOURCE, KEYWDS, EXPDTA, AUTHOR, REVDAT, SPRSDE, JRN, REMARK, DBRE, SEQADV, SEQRES, MODRES, HET, HETNAM, HETSYN, FORMUL, HELIX, SHEET, TURN, SSBOND, LINK, HYDBND, SLTBRG, CISPEP, SITE, CRYST1, ORIGX1, ORIGX2, ORIGX3, SCALE1, SCALE2, SCALE3, MTRIX1 MTRIX2 MTRIX3, TVECT, MODEL, ATOM, SIGATM, ANISOU, SIGUIJ, TER, HETATM, ENDMDL, CONECT, MASTER, END*.

--chains

Count number of chains.

-d, --detail *infolevel*

Level of information to print about PDB during various options. Default: *1*. Possible values: *1, 2 or 3*.

-e, --experiment

List experimental technique information along with any applicable resolution.

-f, --frequency

List distribution of residues: report count and percent of residues in individual chains and across all the chains, or for all the residues in the file. The value of option --residuesmode determines how residues are counted and what is listed. The list is sorted by frequency in descending order. By default, only residue count values are reported. To list percent distribution of residues, specify -d, --detail value of *2* or higher.

-h, --help

Print this help message.

--header

List header information.

m, --MasterCheck

Check master record by explicitly counting the number of REMARK, HET, HELIX, SHEET, TURN, SITE, ORIGX, SCALE, MTRIX, ATOM, HETATM, TER, CONECT and SEQRES records and comparing their values against contents of master record.

--residues

Count residues in *PDBFile(s)*. This is also default behavior.

By default, only residue count values are reported. To list percent distribution of residues, specify -d, --detail value of *2* or higher.

--ResiduesMode <InChains | All | Both>

Specify how to count residues in *PDBFile(s)*: Count residue in each chain and across all the chains, list count iof all the residues in the file, or list both. Possible values: *InChains, All, or Both*. Default: *Both*.

--ResidueNumbers

List information about ATOM residue numbers in each chain before TER record: start and end residue number; gaps in residue numbers corresponding to non-sequential residue numbers; residue numbers not in ascending order.

-w, --WorkingDir *dirname*

Location of working directory. Default: current directory.

EXAMPLES

To list total number of records and number of chain(s) residues in PDB files, type:

```
% InfoPDBFiles.pl Sample1.pdb
% InfoPDBFiles.pl Sample2.pdb
```

To list all available information for PDB file Sample2.pdb, type:

```
% InfoPDBFiles.pl -a Sample2.pdb
```

To list all available information for PDB file Sample2.pdb with all available details, type:

```
% InfoPDBFiles.pl -a -d Sample2.pdb
```

To count ATOM and HETATM records in Sample2.pdb file, type:

```
% InfoPDBFiles.pl -c "ATOM,HETATM" Sample2.pdb
```

To list distribution of residues in chains across the whole PDB file Sample2.pdb along with percent distribution, type

```
% InfoPDBFiles.pl --frequency -d 2 Sample2.pdb
```

To list distribution of residues only across chains in PDB file Sample2.pdb along with percent distribution, type

```
% InfoPDBFiles.pl --frequency -d 2 --ResiduesMode InChains Sample2.pdb
```

To list min/max coordinates of the bounding box which encompasses the structure in Sample1.pdb file, type:

```
% InfoPDBFiles.pl -b Sample1.pdb
```

AUTHOR

Manish Sud <msud@san.rr.com>

SEE ALSO

ExtractFromPDBFiles.pl, InfoAminoAcids.pl, InfoNucleicAcids.pl, InfoSequenceFiles.pl, ModifyPDBFiles.pl

COPYRIGHT

Copyright (C) 2026 Manish Sud. All rights reserved.

This file is part of MayaChemTools.

MayaChemTools is free software; you can redistribute it and/or modify it under the terms of the GNU Lesser General Public License as published by the Free Software Foundation; either version 3 of the License, or (at your option) any later version.