Protein Data Bank Contents Guide:

Atomic Coordinate Entry Format Description

Version 3.30

Document Published by the wwPDB

This format complies with the PDB Exchange Dictionary (PDBx)
http://mmcif.pdb.org/dictionaries/mmcif_pdbx.dic/Index/index.html.

5. Secondary Structure Section

The secondary structure section of a PDB formatted file describes helices and sheets found in protein and polypeptide structures.

HELIX

Overview

HELIX records are used to identify the position of helices in the molecule. Helices are named, numbered, and classified by type. The residues where the helix begins and ends are noted, as well as the total length.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"HELIX "	
8 - 10	Integer	serNum	Serial number of the helix. This starts at 1 and increases incrementally.
12 - 14	LString(3)	helixID	Helix identifier. In addition to a serial number, each helix is given an alphanumeric character helix identifier.
16 - 18	Residue name	initResName	Name of the initial residue.
20	Character	initChainID	Chain identifier for the chain containing this helix.
22 - 25	Integer	initSeqNum	Sequence number of the initial residue.
26	AChar	initICode	Insertion code of the initial residue.
28 - 30	Residue name	endResName	Name of the terminal residue of the helix.
32	Character	endChainID	Chain identifier for the chain containing this helix.
34 - 37	Integer	endSeqNum	Sequence number of the terminal residue.
38	AChar	endICode	Insertion code of the terminal residue.
39 - 40	Integer	helixClass	Helix class (see below).
41 - 70	String	comment	Comment about this helix.
72 - 76	Integer	length	Length of this helix.

Details

* Additional HELIX records with different serial numbers and identifiers occur if more than one helix is present.

* The initial residue of the helix is the N-terminal residue.

* Helices are classified as follows:

TYPE OF HELIX	CLASS NUMBER (COLUMNS 39 - 40)
Right-handed alpha (default)	1
Right-handed omega	2
Right-handed pi	3
Right-handed gamma	4
Right-handed 310	5
Left-handed alpha	6
Left-handed omega	7
Left-handed gamma	8
27 ribbon/helix	9
Polyproline	10

Relationships to Other Record Types

There may be related information in the REMARKs.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
HELIX    1  HA GLY A   86  GLY A   94  1                                        9
HELIX    2  HB GLY B   86  GLY B   94  1                                        9

HELIX   21  21 PRO J  385  LEU J  388  5                                        4
HELIX   22  22 PHE J  397  PHE J  402  5                                        6
```

SHEET

Overview

SHEET records are used to identify the position of sheets in the molecule. Sheets are both named and numbered. The residues where the sheet begins and ends are noted.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"SHEET "	
8 - 10	Integer	strand	Strand number which starts at 1 for each strand within a sheet and increases by one.
12 - 14	LString(3)	sheetID	Sheet identifier.
15 - 16	Integer	numStrands	Number of strands in sheet.
18 - 20	Residue name	initResName	Residue name of initial residue.
22	Character	initChainID	Chain identifier of initial residue in strand.
23 - 26	Integer	initSeqNum	Sequence number of initial residue in strand.
27	AChar	initICode	Insertion code of initial residue in strand.
29 - 31	Residue name	endResName	Residue name of terminal residue.
33	Character	endChainID	Chain identifier of terminal residue.
34 - 37	Integer	endSeqNum	Sequence number of terminal residue.
38	AChar	endICode	Insertion code of terminal residue.
39 - 40	Integer	sense	Sense of strand with respect to previous strand in the sheet. 0 if first strand, 1 if parallel,and -1 if anti-parallel.
42 - 45	Atom	curAtom	Registration. Atom name in current strand.
46 - 48	Residue name	curResName	Registration. Residue name in current strand
50	Character	curChainId	Registration. Chain identifier in current strand.
51 - 54	Integer	curResSeq	Registration. Residue sequence number in current strand.
55	AChar	curICode	Registration. Insertion code in current strand.
57 - 60	Atom	prevAtom	Registration. Atom name in previous strand.

61 - 63	Residue name	prevResName	Registration. Residue name in previous strand.
65	Character	prevChainId	Registration. Chain identifier in previous strand.
66 - 69	Integer	prevResSeq	Registration. Residue sequence number in previous strand.
70	AChar	prevICode	Registration. Insertion code in previous strand.

Details

* The initial residue for a strand is its N-terminus. Strand registration information is provided in columns 39 - 70. Strands are listed starting with one edge of the sheet and continuing to the spatially adjacent strand.

* The sense in columns 39 - 40 indicates whether strand n is parallel (sense = 1) or anti-parallel (sense = -1) to strand n-1. Sense is equal to zero (0) for the first strand of a sheet.

* The registration (columns 42 - 70) of strand n to strand n-1 may be specified by one hydrogen bond between each such pair of strands. This is done by providing the hydrogen bonding between the current and previous strands. No register information should be provided for the first strand.

* Split strands, or strands with two or more runs of residues from discontinuous parts of the amino acid sequence, are explicitly listed. Detail description can be included in the REMARK 700 .

Relationships to Other Record Types

If the entry contains bifurcated sheets or beta-barrels, the relevant REMARK 700 records must be provided. See the REMARK section for details.

Examples

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
SHEET    1   A 5 THR A 107  ARG A 110  0
SHEET    2   A 5 ILE A  96  THR A  99 -1  N  LYS A  98   O  THR A 107
SHEET    3   A 5 ARG A  87  SER A  91 -1  N  LEU A  89   O  TYR A  97
SHEET    4   A 5 TRP A  71  ASP A  75 -1  N  ALA A  74   O  ILE A  88
SHEET    5   A 5 GLY A  52  PHE A  56 -1  N  PHE A  56   O  TRP A  71
SHEET    1   B 5 THR B 107  ARG B 110  0
SHEET    2   B 5 ILE B  96  THR B  99 -1  N  LYS B  98   O  THR B 107
SHEET    3   B 5 ARG B  87  SER B  91 -1  N  LEU B  89   O  TYR B  97
SHEET    4   B 5 TRP B  71  ASP B  75 -1  N  ALA B  74   O  ILE B  88
SHEET    5   B 5 GLY B  52  ILE B  55 -1  N  ASP B  54   O  GLU B  73
```

The sheet presented as BS1 below is an eight-stranded beta-barrel. This is represented by a nine-stranded sheet in which the first and last strands are identical.

```
SHEET    1 BS1 9 VAL    13  ILE    17  0
SHEET    2 BS1 9 ALA    70  ILE    73  1  O  TRP    72   N  ILE    17
SHEET    3 BS1 9 LYS   127  PHE   132  1  O  ILE   129   N  ILE    73
SHEET    4 BS1 9 GLY   221  ASP   225  1  O  GLY   221   N  ILE   130
SHEET    5 BS1 9 VAL   248  GLU   253  1  O  PHE   249   N  ILE   222
```

```
SHEET    6 BS1 9 LEU   276  ASP   278  1  N  LEU   277   O  GLY   252
SHEET    7 BS1 9 TYR   310  THR   318  1  O  VAL   317   N  ASP   278
SHEET    8 BS1 9 VAL   351  TYR   356  1  O  VAL   351   N  THR   318
SHEET    9 BS1 9 VAL    13  ILE    17  1  N  VAL    14   O  PRO   352
```

The sheet structure of this example is bifurcated. In order to represent this feature, two sheets are defined. Strands 2 and 3 of BS7 and BS8 are identical.

```
SHEET    1 BS7 3 HIS   662  THR   665  0
SHEET    2 BS7 3 LYS   639  LYS   648 -1  N  PHE   643   O  HIS   662
SHEET    3 BS7 3 ASN   596  VAL   600 -1  N  TYR   598   O  ILE   646
SHEET    1 BS8 3 ASN   653  TRP   656  0
SHEET    2 BS8 3 LYS   639  LYS   648 -1  N  LYS   647   O  THR   655
SHEET    3 BS8 3 ASN   596  VAL   600 -1  N  TYR   598   O  ILE   646
```

ATOM

Overview

The ATOM records present the atomic coordinates for standard amino acids and nucleotides. They also present the occupancy and temperature factor for each atom. Non-polymer chemical coordinates use the HETATM record type. The element symbol is always present on each ATOM record; charge is optional.

Changes in ATOM/HETATM records result from the standardization atom and residue nomenclature. This nomenclature is described in the Chemical Component Dictionary (ftp://ftp.wwpdb.org/pub/pdb/data/monomers).

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"ATOM "	
7 - 11	Integer	serial	Atom serial number.
13 - 16	Atom	name	Atom name.
17	Character	altLoc	Alternate location indicator.
18 - 20	Residue name	resName	Residue name.
22	Character	chainID	Chain identifier.
23 - 26	Integer	resSeq	Residue sequence number.
27	AChar	iCode	Code for insertion of residues.
31 - 38	Real(8.3)	x	Orthogonal coordinates for X in Angstroms.
39 - 46	Real(8.3)	y	Orthogonal coordinates for Y in Angstroms.
47 - 54	Real(8.3)	z	Orthogonal coordinates for Z in Angstroms.
55 - 60	Real(6.2)	occupancy	Occupancy.
61 - 66	Real(6.2)	tempFactor	Temperature factor.
77 - 78	LString(2)	element	Element symbol, right-justified.
79 - 80	LString(2)	charge	Charge on the atom.

Details

* ATOM records for proteins are listed from amino to carboxyl terminus.

* Nucleic acid residues are listed from the 5' → 3' terminus.
* Alignment of one-letter atom name such as C starts at column 14, while two-letter atom name such as FE starts at column 13.

* Atom nomenclature begins with atom type.
* No ordering is specified for polysaccharides.
* Non-blank alphanumerical character is used for chain identifier.
* The list of ATOM records in a chain is terminated by a TER record.
* If more than one model is present in the entry, each model is delimited by MODEL and ENDMDL records.

* AltLoc is the place holder to indicate alternate conformation. The alternate conformation can be in the entire polymer chain, or several residues or partial residue (several atoms within one residue). If an atom is provided in more than one position, then a non-blank alternate location indicator must be used for each of the atomic positions. Within a residue, all atoms that are associated with each other in a given conformation are assigned the same alternate position indicator. There are two ways of representing alternate conformation- either at atom level or at residue level (see examples).

* For atoms that are in alternate sites indicated by the alternate site indicator, sorting of atoms in the ATOM/HETATM list uses the following general rules:
  - In the simple case that involves a few atoms or a few residues with alternate sites, the coordinates occur one after the other in the entry.
  - In the case of a large heterogen groups which are disordered, the atoms for each conformer are listed together.

* Alphabet letters are commonly used for insertion code. The insertion code is used when two residues have the same numbering. The combination of residue numbering and insertion code defines the unique residue.

* If the depositor provides the data, then the isotropic B value is given for the temperature factor.

* If there are neither isotropic B values from the depositor, nor anisotropic temperature factors in ANISOU, then the default value of 0.0 is used for the temperature factor.

* Columns 79 - 80 indicate any charge on the atom, e.g., 2+, 1-. In most cases, these are blank.

Verification/Validation/Value Authority Control

The ATOM/HETATM records are checked for PDB file format, sequence information, and packing.

Relationships to Other Record Types

The ATOM records are compared to the corresponding sequence database. Sequence discrepancies appear in the SEQADV record. Missing atoms are annotated in the remarks. HETATM records are formatted in the same way as ATOM records. The sequence implied by ATOM records must be identical to that given in SEQRES, with the exception that residues that have no coordinates, e.g., due to disorder, must appear in SEQRES.

Examples

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
```

```
ATOM     32  N  AARG A  -3      11.281  86.699  94.383  0.50 35.88           N
ATOM     33  N  BARG A  -3      11.296  86.721  94.521  0.50 35.60           N
ATOM     34  CA AARG A  -3      12.353  85.696  94.456  0.50 36.67           C
ATOM     35  CA BARG A  -3      12.333  85.862  95.041  0.50 36.42           C
ATOM     36  C  AARG A  -3      13.559  86.257  95.222  0.50 37.37           C
ATOM     37  C  BARG A  -3      12.759  86.530  96.365  0.50 36.39           C
ATOM     38  O  AARG A  -3      13.753  87.471  95.270  0.50 37.74           O
ATOM     39  O  BARG A  -3      12.924  87.757  96.420  0.50 37.26           O
ATOM     40  CB AARG A  -3      12.774  85.306  93.039  0.50 37.25           C
ATOM     41  CB BARG A  -3      13.428  85.746  93.980  0.50 36.60           C
ATOM     42  CG AARG A  -3      11.754  84.432  92.321  0.50 38.44           C
ATOM     43  CG BARG A  -3      12.866  85.172  92.651  0.50 37.31           C
ATOM     44  CD AARG A  -3      11.698  84.678  90.815  0.50 38.51           C
ATOM     45  CD BARG A  -3      13.374  85.886  91.406  0.50 37.66           C
ATOM     46  NE AARG A  -3      12.984  84.447  90.163  0.50 39.94           N
ATOM     47  NE BARG A  -3      12.644  85.487  90.195  0.50 38.24           N
ATOM     48  CZ AARG A  -3      13.202  84.534  88.850  0.50 40.03           C
ATOM     49  CZ BARG A  -3      13.114  85.582  88.947  0.50 39.55           C
ATOM     50  NH1AARG A  -3      12.218  84.840  88.007  0.50 40.76           N
ATOM     51  NH1BARG A  -3      14.338  86.056  88.706  0.50 40.23           N
ATOM     52  NH2AARG A  -3      14.421  84.308  88.373  0.50 40.45           N

         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
ATOM     32  N  AARG A  -3      11.281  86.699  94.383  0.50 35.88           N
ATOM     33  CA AARG A  -3      12.353  85.696  94.456  0.50 36.67           C
ATOM     34  C  AARG A  -3      13.559  86.257  95.222  0.50 37.37           C
ATOM     35  O  AARG A  -3      13.753  87.471  95.270  0.50 37.74           O
ATOM     36  CB AARG A  -3      12.774  85.306  93.039  0.50 37.25           C
ATOM     37  CG AARG A  -3      11.754  84.432  92.321  0.50 38.44           C
ATOM     38  CD AARG A  -3      11.698  84.678  90.815  0.50 38.51           C
ATOM     39  NE AARG A  -3      12.984  84.447  90.163  0.50 39.94           N
ATOM     40  CZ AARG A  -3      13.202  84.534  88.850  0.50 40.03           C
ATOM     41  NH1AARG A  -3      12.218  84.840  88.007  0.50 40.76           N
ATOM     42  NH2AARG A  -3      14.421  84.308  88.373  0.50 40.45           N
ATOM     43  N  BARG A  -3      11.296  86.721  94.521  0.50 35.60           N
ATOM     44  CA BARG A  -3      12.333  85.862  95.041  0.50 36.42           C
ATOM     45  C  BARG A  -3      12.759  86.530  96.365  0.50 36.39           C
ATOM     46  O  BARG A  -3      12.924  87.757  96.420  0.50 37.26           O
ATOM     47  CB BARG A  -3      13.428  85.746  93.980  0.50 36.60           C
ATOM     48  CG BARG A  -3      12.866  85.172  92.651  0.50 37.31           C
ATOM     49  CD BARG A  -3      13.374  85.886  91.406  0.50 37.66           C
ATOM     50  NE BARG A  -3      12.644  85.487  90.195  0.50 38.24           N
ATOM     51  CZ BARG A  -3      13.114  85.582  88.947  0.50 39.55           C
ATOM     52  NH1BARG A  -3      14.338  86.056  88.706  0.50 40.23           N
```

ANISOU

Overview

The ANISOU records present the anisotropic temperature factors.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION