Protein Data Bank Contents Guide:

Atomic Coordinate Entry Format Description

Version 3.20

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.

Table of Contents

1. Introduction

The Protein Data Bank (PDB) is an archive of experimentally determined three-dimensional structures of biological macromolecules that serves a global community of researchers, educators, and students. The data contained in the archive include atomic coordinates, crystallographic structure factors and NMR experimental data. Aside from coordinates, each deposition also includes the names of molecules, primary and secondary structure information, sequence database references, where appropriate, and ligand and biological assembly information, details about data collection and structure solution, and bibliographic citations.

This comprehensive guide describes the "PDB format" used by the members of the worldwide Protein Data Bank (wwPDB; Berman, H.M., Henrick, K. and Nakamura, H. Announcing the worldwide Protein Data Bank. *Nat Struct Biol* **10**, 980 (2003)). Questions should be sent to info@wwpdb.org

Information about file formats and data dictionaries can be found at http://wwpdb.org.

Version History:

Version 2.3: The format in which structures were released from 1998 to July 2007.

Version 3.0: Major update from Version 2.3; incorporates all of the revisions used by the wwPDB to integrate uniformity and remediation data into a single set of archival data files including IUPAC nomenclature. See http://www.wwpdb.org/docs.html for more details.

Version 3.1: Minor addenda to Version 3.0, introducing a small number of changes and extensions supporting the annotation practices adopted by the wwPDB beginning in August 2007 including chain ID standardization and biological assembly .

Version 3.15: Minor addenda to Version 3.20, introducing a small number of changes and extensions supporting the annotation practices adopted by the wwPDB beginning in October 2008 including DBREF, taxonomy and citation information.

Version 3.20: Current version, minor addenda to Version 3.1, introducing a small number of changes and extensions supporting the annotation practices adopted by the wwPDB beginning in December 2008 including DBREF, taxonomy and citation information.
September 15 2008, initial version 3.20.
November 15 2008, add examples for Refmac template and coordinate with alternate conformation.
December 24 2008, update REMARK 3 templates/examples, add Norine database in DBREF, update REMARK 500 on chiral center.
February 12 2009, update example in REMARK 210 and record format in NUMMDL
July 6 2009, update description for REVDAT, DBREF2, MASTER and extend number of columns for AUTHOR, JRNL, CAVEAT, KEYWDS, etc.
December 22, 2009, update CAVEAT and REMARK 265.

Basic Notions of the Format Description

Character Set

Only non-control ASCII characters, as well as the space and end-of-line indicator, appear in a PDB coordinate entry file. Namely:

`abcdefghijklmnopqrstuvwxyzABCDEFGHIJKLMNOPQRSTUVWXYZ`

`1234567890`

``` ` - = [ ] \ ; ' , . / ~ ! @ # $ % ^ & * ( ) _ + { } | : " < > ? ```

The use of punctuation characters in the place of alphanumeric characters is discouraged.

The space, and end-of-line:. The end-of-line indicator is system-specific character; some systems may use a carriage return followed by a line feed, others only a line-feed character.

Special Characters

Greek letters are spelled out, i.e., `alpha`, `beta`, `gamma`, etc.

Bullets are represented as `(DOT)`.

Right arrow is represented as `-->`.

Left arrow is represented as `<--`.

If "=" is surrounded by at least one space on each side, then it is assumed to be an equal sign, e.g., `2 + 4 = 6`.

Commas, colons, and semi-colons are used as list delimiters in records that have one of the following data types:

List

SList

Specification List

Specification

If a comma, colon, or semi-colon is used in any context other than as a delimiting character, then the character must be escaped, i.e., immediately preceded by a backslash, "\".
```

Example - Use of “\” character:

```
COMPND          MOL_ID: 1;
COMPND        2 MOLECULE: GLUTATHIONE SYNTHETASE;
COMPND        3 CHAIN: A;
COMPND        4 SYNONYM: GAMMA-L-GLUTAMYL-L-CYSTEINE\:GLYCINE LIGASE
COMPND        5 (ADP-FORMING);
COMPND        6 EC: 6.3.2.3;
COMPND        7 ENGINEERED: YES
```

```
COMPND          MOL_ID: 1;
COMPND        2 MOLECULE: S-ADENOSYLMETHIONINE SYNTHETASE;
COMPND        3 CHAIN: A, B;
COMPND        4 SYNONYM: MAT, ATP\:L-METHIONINE S-ADENOSYLTRANSFERASE;
COMPND        5 EC: 2.5.1.6;
COMPND        6 ENGINEERED: YES;
COMPND        7 BIOLOGICAL_UNIT: TETRAMER;
COMPND        8 OTHER_DETAILS: TETRAGONAL MODIFICATION
```

Record Format

Every PDB file is presented in a number of lines. Each line in the PDB entry file consists of 80 columns. The last character in each PDB entry should be an end-of- line indicator.

Each line in the PDB file is self-identifying. The first six columns of every line contains a record name, that is left-justified and separated by a blank. The record name must be an exact match to one of the stated record names in this format guide.

The PDB file may also be viewed as a collection of record types. Each record type consists of one or more lines.

Each record type is further divided into fields.

Each record type is detailed in this document. The description of each record type includes the following sections:

- **Overview**
- **Record Format**
- **Details**
- **Verification/Validation/Value Authority Control**
- **Relationship to Other Record Types**
- **Examples**
- **Known Problems**

For records that are fully described in fixed column format, columns not assigned to fields must be left blank.

Types of Records

It is possible to group records into categories based upon how often the record type appears in an entry.

One time, single line: There are records that may only appear one time and without continuations in a file. Listed alphabetically, these are:

RECORD TYPE	DESCRIPTION
CRYST1	Unit cell parameters, space group, and Z.
END	Last record in the file.
HEADER	First line of the entry, contains PDB ID code, classification, and date of deposition.
NUMMDL	Number of models.
MASTER	Control record for bookkeeping.
ORIGXn	Transformation from orthogonal coordinates to the submitted coordinates (n = 1, 2, or 3).
SCALEn	Transformation from orthogonal coordinates to fractional crystallographic coordinates (n = 1, 2, or 3).

It is an error for a duplicate of any of these records to appear in an entry.

One time, multiple lines: There are records that conceptually exist only once in an entry, but the information content may exceed the number of columns available. These records are therefore continued on subsequent lines. Listed alphabetically, these are:

RECORD TYPE	DESCRIPTION
AUTHOR	List of contributors.
CAVEAT	Severe error indicator.
COMPND	Description of macromolecular contents of the entry.
EXPDTA	Experimental technique used for the structure determination.
MDLTYP	Contains additional annotation pertinent to the coordinates presented in the entry.
KEYWDS	List of keywords describing the macromolecule.
OBSLTE	Statement that the entry has been removed from distribution and list of the ID code(s) which replaced it.
SOURCE	Biological source of macromolecules in the entry.
SPLIT	List of PDB entries that compose a larger macromolecular complexes.
SPRSDE	List of entries obsoleted from public release and replaced by current entry.

TITLE	Description of the experiment represented in the entry.

The second and subsequent lines contain a continuation field, which is a right-justified integer. This number increments by one for each additional line of the record, and is followed by a blank character.

<u>Multiple times, one line:</u> Most record types appear multiple times, often in groups where the information is not logically concatenated but is presented in the form of a list. Many of these record types have a custom serialization that may be used not only to order the records, but also to connect to other record types. Listed alphabetically, these are:

RECORD TYPE	DESCRIPTION
ANISOU	Anisotropic temperature factors.
ATOM	Atomic coordinate records for standard groups.
CISPEP	Identification of peptide residues in cis conformation.
CONECT	Connectivity records.
DBREF	Reference to the entry in the sequence database(s).
HELIX	Identification of helical substructures.
HET	Identification of non-standard groups heterogens).
HETATM	Atomic coordinate records for heterogens.
LINK	Identification of inter-residue bonds.
MODRES	Identification of modifications to standard residues.
MTRIXn	Transformations expressing non-crystallographic symmetry (n = 1, 2, or 3). There may be multiple sets of these records.
REVDAT	Revision date and related information.
SEQADV	Identification of conflicts between PDB and the named sequence database.
SHEET	Identification of sheet substructures.
SSBOND	Identification of disulfide bonds.

Multiple times, multiple lines: There are records that conceptually exist multiple times in an entry, but the information content may exceed the number of columns available. These records are therefore continued on subsequent lines. Listed alphabetically, these are:

RECORD TYPE	DESCRIPTION
FORMUL	Chemical formula of non-standard groups.
HETNAM	Compound name of the heterogens.
HETSYN	Synonymous compound names for heterogens.
SEQRES	Primary sequence of backbone residues.
SITE	Identification of groups comprising important entity sites.

The second and subsequent lines contain a continuation field which is a right-justified integer. This number increments by one for each additional line of the record, and is followed by a blank character.

Grouping: There are three record types used to group other records.
Listed alphabetically, these are:

RECORD TYPE	DESCRIPTION
ENDMDL	End-of-model record for multiple structures in a single coordinate entry.
MODEL	Specification of model number for multiple structures in a single coordinate entry.
TER	Chain terminator.

The MODEL/ENDMDL records surround groups of ATOM, HETATM, ANISOU, and TER records. TER records indicate the end of a chain.

Other: The remaining record types have a detailed inner structure.
Listed alphabetically, these are:

RECORD TYPE	DESCRIPTION
JRNL	Literature citation that defines the coordinate set.
REMARK	General remarks; they can be structured or free form.

PDB Format Change Policy

The wwPDB will use the following protocol in making changes to the way PDB coordinate entries are represented and archived. The purpose of the policy is to allow ample time for everyone to understand these changes and to assess their impact on existing programs. PDB format modifications are necessary to address the changing needs of PDB users as well as the changing nature of the data that is archived.

1. Comments and suggestions will be solicited from the community on specific problems and data representation issues as they arise.

2. Proposed format changes will be disseminated through pdb-l@rcsb.org and wwpdb.org.

3. A 60-day discussion period will follow the announcement of proposed changes. Comments and suggestions must be received within this time period. Major changes that are not upwardly compatible will be allotted up to twice the standard amount of discussion time.

4. The wwPDB will then work in consultation with the wwPDB Advisory Committee and the equivalent partner Scientific Advisory Committees to evaluate and reconcile all suggestions. The final decision will be officially announced via pdb-l@rcsb.org and wwpdb.org.

5. Implementation will follow official announcement of the format change. Major changes will not appear in PDB files earlier than 60 days after the announcement, allowing sufficient time to modify files and programs.

Order of Records

All records in a PDB coordinate entry must appear in a defined order. Mandatory record types are present in all entries. When mandatory data are not provided, the record name must appear in the entry with a NULL indicator. Optional items become mandatory when certain conditions exist. Record order and existence are described in the following table:

RECORD TYPE	EXISTENCE	CONDITIONS IF OPTIONAL
HEADER	Mandatory	
OBSLTE	Optional	Mandatory in entries that have been replaced by a newer entry.
TITLE	Mandatory	
SPLIT	Optional	Mandatory when large macromolecular complexes are split into multiple PDB entries.
CAVEAT	Optional	Mandatory when there are outstanding errors such as chirality.
COMPND	Mandatory	
SOURCE	Mandatory	
KEYWDS	Mandatory	
EXPDTA	Mandatory	
NUMMDL	Optional	Mandatory for NMR ensemble entries.
MDLTYP	Optional	Mandatory for NMR minimized average Structures or when the entire polymer chain contains C alpha or P atoms only.
AUTHOR	Mandatory	
REVDAT	Mandatory	
SPRSDE	Optional	Mandatory for a replacement entry.
JRNL	Optional	Mandatory for a publication describes the experiment.
REMARK 0	Optional	Mandatory for a re-refined structure
REMARK 1	Optional	
REMARK 2	Mandatory	
REMARK 3	Mandatory	
REMARK N	Optional	Mandatory under certain conditions.
DBREF	Optional	Mandatory for all polymers.
DBREF1/DBREF2	Optional	Mandatory when certain sequence database accession and/or sequence numbering does not fit preceding DBREF format.

SEQADV	Optional	Mandatory if sequence conflict exists.
SEQRES	Mandatory	Mandatory if ATOM records exist.
MODRES	Optional	Mandatory if modified group exists in the coordinates.
HET	Optional	Mandatory if a non-standard group other than water appears in the coordinates.
HETNAM	Optional	Mandatory if a non-standard group other than water appears in the coordinates.
HETSYN	Optional	
FORMUL	Optional	Mandatory if a non-standard group or water appears in the coordinates.
HELIX	Optional	
SHEET	Optional	
SSBOND	Optional	Mandatory if a disulfide bond is present.
LINK	Optional	Mandatory if non-standard residues appear in a polymer
CISPEP	Optional	
SITE	Optional	
CRYST1	Mandatory	
ORIGX1 ORIGX2 ORIGX3	Mandatory	
SCALE1 SCALE2 SCALE3	Mandatory	
MTRIX1 MTRIX2 MTRIX3	Optional	Mandatory if the complete asymmetric unit must be generated from the given coordinates using non-crystallographic symmetry.
MODEL	Optional	Mandatory if more than one model is present in the entry.
ATOM	Optional	Mandatory if standard residues exist.
ANISOU	Optional	
TER	Optional	Mandatory if ATOM records exist.
HETATM	Optional	Mandatory if non-standard group exists.
ENDMDL	Optional	Mandatory if MODEL appears.
CONECT	Optional	Mandatory if non-standard group appears and if LINK or SSBOND records exist.
MASTER	Mandatory	
END	Mandatory	

Sections of an Entry

The following table lists the various sections of a PDB entry (version 3.2) and the records within it:

SECTION	DESCRIPTION	RECORD TYPE
Title	Summary descriptive remarks	HEADER, OBSLTE, TITLE, SPLIT, CAVEAT, COMPND, SOURCE, KEYWDS,EXPDTA, NUMMDL, MDLTYP, AUTHOR, REVDAT, SPRSDE, JRNL
Remark	Various comments about entry annotations in more depth than standard records	REMARKs 0-999
Primary structure	Peptide and/or nucleotide sequence and the relationship between the PDB sequence and that found in the sequence database(s)	DBREF, SEQADV, SEQRES MODRES
Heterogen	Description of non-standard groups	HET, HETNAM, HETSYN, FORMUL
Secondary structure	Description of secondary structure	HELIX, SHEET
Connectivity annotation	Chemical connectivity	SSBOND, LINK, CISPEP
Miscellaneous features	Features within the macromolecule	SITE
Crystallographic	Description of the crystallographic cell	CRYST1
Coordinate transformation	Coordinate transformation operators	ORIGXn, SCALEn, MTRIXn,
Coordinate	Atomic coordinate data	MODEL, ATOM, ANISOU, TER, HETATM, ENDMDL
Connectivity	Chemical connectivity	CONECT
Bookkeeping	Summary information, end-of-file marker	MASTER, END

Field Formats and Data Types

Each record type is presented in a table which contains the division of the records into fields by column number, defined data type, field name or a quoted string which must appear in the field, and field definition. Any column not specified must be left blank.

Each field contains an identified data type that can be validated by a program. These are:

DATA TYPE	DESCRIPTION
AChar	An alphabetic character (A-Z, a-z).
Atom	Atom name.
Character	Any non-control character in the ASCII character set or a space.
Continuation	A two-character field that is either blank (for the first record of a set) or contains a two digit number right-justified and blank-filled which counts continuation records starting with 2. The continuation number must be followed by a blank.
Date	A 9 character string in the form DD-MMM-YY where DD is the day of the month, zero-filled on the left (e.g., 04); MMM is the common English 3-letter abbreviation of the month; and YY is the last two digits of the year. This must represent a valid date.
IDcode	A PDB identification code which consists of 4 characters, the first of which is a digit in the range 0 - 9; the remaining 3 are alpha-numeric, and letters are upper case only. Entries with a 0 as the first character do not contain coordinate data.
Integer	Right-justified blank-filled integer value.
Token	A sequence of non-space characters followed by a colon and a space.
List	A String that is composed of text separated with commas.
LString	A literal string of characters. All spacing is significant and must be preserved.
LString(n)	An LString with exactly n characters.
Real(n,m)	Real (floating point) number in the FORTRAN format Fn.m.
Record name	The name of the record: 6 characters, left-justified and blank-filled.
Residue name	One of the standard amino acid or nucleic acids, as listed below, or the non-standard group designation as defined in the HET dictionary. Field is right-justified.
SList	A String that is composed of text separated with semi-colons.
Specification	A String composed of a token and its associated value

	separated by a colon.
Specification List	A sequence of Specifications, separated by semi-colons.
String	A sequence of characters. These characters may have arbitrary spacing, but should be interpreted as directed below.
String(n)	A String with exactly n characters.
SymOP	An integer field of from 4 to 6 digits, right-justified, of the form nnnMMM where nnn is the symmetry operator number and MMM is the translation vector.

To interpret a String, concatenate the contents of all continued fields together, collapse all sequences of multiple blanks to a single blank, and remove any leading and trailing blanks. This permits very long strings to be properly reconstructed.

2. Title Section

This section contains records used to describe the experiment and the biological macromolecules present in the entry: HEADER, OBSLTE, TITLE, SPLIT, CAVEAT, COMPND, SOURCE, KEYWDS, EXPDTA, AUTHOR, REVDAT, SPRSDE, JRNL, and REMARK records.

HEADER

Overview

The HEADER record uniquely identifies a PDB entry through the idCode field. This record also provides a classification for the entry. Finally, it contains the date when the coordinates were deposited to the PDB archive.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"HEADER"	
11 - 50	String(40)	classification	Classifies the molecule(s).
51 - 59	Date	depDate	Deposition date. This is the date the coordinates were received at the PDB.
63 - 66	IDcode	idCode	This identifier is unique within the PDB.

Details

* The classification string is left-justified and exactly matches one of a collection of strings. A class list is available from the current wwPDB Annotation Documentation Appendices (http://www.wwpdb.org/docs.html). In the case of macromolecular complexes, the classification field must present a class for each macromolecule present. Due to the limited length of the classification field, strings must sometimes be abbreviated. In these cases, the full terms are given in KEYWDS.

* Classification may be based on function, metabolic role, molecule type, cellular location, etc. This record can describe dual functions of a molecules, and when applicable, separated by a comma “,”. Entries with multiple molecules in a complex will list the classifications of each macromolecule separated by slash “/”.

Verification/Validation/Value Authority Control

The verification program checks that the deposition date is a legitimate date and that the ID code is well-formed.

PDB coordinate entry ID codes do not begin with 0. “No coordinates”, or NOC files, given as 0xxx codes, contained no structural information and were bibliographic only. These entries were subsequently removed from PDB archive.

Relationships to Other Record Types

The classification found in HEADER also appears in KEYWDS, unabbreviated and in no strict order.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
HEADER    PHOTOSYNTHESIS                          28-MAR-07   2UXK

HEADER    TRANSFERASE/TRANSFERASE INHIBITOR       17-SEP-04   1XH6

HEADER    MEMBRANE PROTEIN, TRANSPORT PROTEIN     20-JUL-06   2HRT
```

OBSLTE

Overview

OBSLTE appears in entries that have been removed from public distribution.

This record acts as a flag in an entry that has been removed (“obsoleted”) from the PDB's full release. It indicates which, if any, new entries have replaced the entry that was obsoleted. The format allows for the case of multiple new entries replacing one existing entry.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"OBSLTE"	
9 - 10	Continuation	continuation	Allows concatenation of multiple records
12 - 20	Date	repDate	Date that this entry was replaced.
22 - 25	IDcode	idCode	ID code of this entry.
32 - 35	IDcode	rIdCode	ID code of entry that replaced this one.
37 - 40	IDcode	rIdCode	ID code of entry that replaced this one.
42 - 45	IDcode	rIdCode	ID code of entry that replaced this one.
47 - 50	IDcode	rIdCode	ID code of entry that replaced this one.
52 - 55	IDcode	rIdCode	ID code of entry that replaced this one.
57 - 60	IDcode	rIdCode	ID code of entry that replaced this one.
62 - 65	IDcode	rIdCode	ID code of entry that replaced this one.
67 - 70	IDcode	rIdCode	ID code of entry that replaced this one.
72 - 75	IDcode	rIdCode	ID code of entry that replaced this one.

Details

* It is PDB policy that only the principal investigator and/or the primary author who submitted an entry has the authority to obsolete it. All OBSLTE entries are available from the PDB archive (ftp://ftp.wwpdb.org/pub/pdb/data/structures/obsolete).

* Though the obsolete entry is removed from the public archive, the initial citation that reported the structure is carried over to the superseding entry.

Verification/Validation/Value Authority Control

wwPDB staff adds this record at the time an entry is removed from release.

Relationships to Other Record Types

None.

Example

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

```
OBSLTE     31-JAN-94 1MBP      2MBP
```

TITLE

Overview

The TITLE record contains a title for the experiment or analysis that is represented in the entry. It should identify an entry in the same way that a citation title identifies a publication.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"TITLE "	
9 - 10	Continuation	continuation	Allows concatenation of multiple records.
11 - 80	String	title	Title of the experiment.

Details

* The title of the entry is free text and should describe the contents of the entry and any procedures or conditions that distinguish this entry from similar entries. It presents an opportunity for the depositor to emphasize the underlying purpose of this particular experiment.

* Some items that may be included in TITLE are:

- Experiment type.
- Description of the mutation.
- The fact that only alpha carbon coordinates have been provided in the entry.

Verification/Validation/Value Authority Control

This record is free text so no verification of format is required. The title is supplied by the depositor, but staff may exercise editorial judgment in consultation with depositors in assigning the title.

Relationships to Other Record Types

COMPND, SOURCE, EXPDTA, and REMARKs provide information that may also be found in TITLE. You may think of the title as describing the experiment, and the compound record as describing the molecule(s).

Examples

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
TITLE     RHIZOPUSPEPSIN COMPLEXED WITH REDUCED PEPTIDE INHIBITOR

TITLE     STRUCTURE OF THE TRANSFORMED MONOCLINIC LYSOZYME BY
TITLE    2 CONTROLLED DEHYDRATION
```

```
TITLE     NMR STUDY OF OXIDIZED THIOREDOXIN MUTANT (C62A,C69A,C73A)
TITLE    2 MINIMIZED AVERAGE STRUCTURE
```

SPLIT (added)

Overview

The SPLIT record is used in instances where a specific entry composes part of a large macromolecular complex. It will identify the PDB entries that are required to reconstitute a complete complex.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"SPLIT "	
9 - 10	Continuation	continuation	Allows concatenation of multiple records.
12 - 15	IDcode	idCode	ID code of related entry.
17 - 20	IDcode	idCode	ID code of related entry.
22 - 25	IDcode	idCode	ID code of related entry.
27 – 30	IDcode	idCode	ID code of related entry.
32 - 35	IDcode	idCode	ID code of related entry.
37 - 40	IDcode	idCode	ID code of related entry.
42 - 45	IDcode	idCode	ID code of related entry.
47 - 50	IDcode	idCode	ID code of related entry.
52 - 55	IDcode	idCode	ID code of related entry.
57 - 60	IDcode	idCode	ID code of related entry.
62 - 65	IDcode	idCode	ID code of related entry.
67 - 70	IDcode	idCode	ID code of related entry.
72 - 75	IDcode	idCode	ID code of related entry.
77 - 80	IDcode	idCode	ID code of related entry.

Details

* The SPLIT record can be continued on multiple lines, so that all related PDB entries are cataloged.

Verification/Validation/Value Authority Control

This record will be generated at the time of processing the component PDB files of the large macromolecular complex when all complex constituents are deposited.

Relationships to Other Record Types

REMARK 350 will contain an amended statement to reflect the entire complex.

Examples

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
SPLIT      1VOQ 1VOR 1VOS 1VOU 1VOV 1VOW 1VOX 1VOY 1VP0 1VOZ
```

CAVEAT

Overview

CAVEAT warns of errors and unresolved issues in the entry. Use caution when using an entry containing this record.

Record Format

```
COLUMNS       DATA TYPE     FIELD          DEFINITION
---------------------------------------------------------------------------------------
 1 - 6        Record name   "CAVEAT"

 9 - 10       Continuation  continuation   Allows concatenation of multiple records.

12 - 15       IDcode        idCode         PDB ID code of this entry.

20 - 79       String        comment        Free text giving the reason for the CAVEAT.
```

Details

* The CAVEAT will also be included in cases where the wwPDB is unable to verify the transformation of the coordinates back to the crystallographic cell. In these cases, the molecular structure may still be correct.

Verification/Validation/Value Authority Control

CAVEAT will be added to entries known to be incorrect.

COMPND (updated)

Overview

The COMPND record describes the macromolecular contents of an entry. Some cases where the entry contains a standalone drug or inhibitor, the name of the non-polymeric molecule will appear in this record. Each macromolecule found in the entry is described by a set of token: value pairs, and is referred to as a COMPND record component. Since the concept of a molecule is difficult to specify exactly, staff may exercise editorial judgment in consultation with depositors in assigning these names.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"COMPND"	
8 - 10	Continuation	continuation	Allows concatenation of multiple records.
11 - 80	Specification list	compound	Description of the molecular components.

Details

* The compound record is a Specification list. The specifications, or tokens, that may be used are listed below:

TOKEN	VALUE DEFINITION
MOL_ID	Numbers each component; also used in SOURCE to associate the information.
MOLECULE	Name of the macromolecule.
CHAIN	Comma-separated list of chain identifier(s).
FRAGMENT	Specifies a domain or region of the molecule.
SYNONYM	Comma-separated list of synonyms for the MOLECULE.
EC	The Enzyme Commission number associated with the molecule. If there is more than one EC number, they are presented as a comma-separated list.
ENGINEERED	Indicates that the molecule was produced using recombinant technology or by purely chemical synthesis.
MUTATION	Indicates if there is a mutation.
OTHER_DETAILS	Additional comments.

* In the case of synthetic molecules, the depositor will provide the description.

* For chimeric proteins, the protein name is comma-separated and may refer to the presence of a linker (protein_1, linker, protein_2).

* Asterisks in nucleic acid names (in MOLECULE) are for ease of reading.
* No specific rules apply to the ordering of the tokens, except that the occurrence of MOL_ID or FRAGMENT indicates that the subsequent tokens are related to that specific molecule or fragment of the molecule.

* When insertion codes are given as part of the residue name, they must be given within square brackets, i.e., H57[A]N. This might occur when listing residues in FRAGMENT or OTHER_DETAILS.

* For multi-chain molecules, e.g., the hemoglobin tetramer, a comma-separated list of CHAIN identifiers is used.

Verification/Validation/Value Authority Control

CHAIN must match the chain identifiers(s) of the molecule(s). EC numbers are also checked.

Relationships to Other Record Types

In the case of mutations, the SEQADV records will present differences from the reference molecule. REMARK records may further describe the contents of the entry. Also see verification above.

Examples

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
COMPND    MOL_ID: 1;
COMPND   2 MOLECULE: HEMOGLOBIN ALPHA CHAIN;
COMPND   3 CHAIN: A, C;
COMPND   4 SYNONYM: DEOXYHEMOGLOBIN ALPHA CHAIN;
COMPND   5 ENGINEERED: YES;
COMPND   6 MUTATION: YES;
COMPND   7 MOL_ID: 2;
COMPND   8 MOLECULE: HEMOGLOBIN BETA CHAIN;
COMPND   9 CHAIN: B, D;
COMPND  10 SYNONYM: DEOXYHEMOGLOBIN BETA CHAIN;
COMPND  11 ENGINEERED: YES;
COMPND  12 MUTATION: YES

COMPND    MOL_ID: 1;
COMPND   2 MOLECULE: COWPEA CHLOROTIC MOTTLE VIRUS;
COMPND   3 CHAIN: A, B, C;
COMPND   4 SYNONYM: CCMV;
COMPND   5 MOL_ID: 2;
COMPND   6 MOLECULE: RNA (5'-(*AP*UP*AP*U)-3');
COMPND   7 CHAIN: D, F;
COMPND   8 ENGINEERED: YES;
COMPND   9 MOL_ID: 3;
COMPND  10 MOLECULE: RNA (5'-(*AP*U)-3');
COMPND  11 CHAIN: E;
COMPND  12 ENGINEERED: YES

COMPND    MOL_ID: 1;
COMPND   2 MOLECULE: HEVAMINE A;
```

```
COMPND   3 CHAIN: A;
COMPND   4 EC: 3.2.1.14, 3.2.1.17;
COMPND   5 OTHER_DETAILS: PLANT ENDOCHITINASE/LYSOZYME
```

SOURCE (updated)

Overview

The SOURCE record specifies the biological and/or chemical source of each biological molecule in the entry. Some cases where the entry contains a standalone drug or inhibitor, the source information of this molecule will appear in this record. Sources are described by both the common name and the scientific name, e.g., genus and species. Strain and/or cell-line for immortalized cells are given when they help to uniquely identify the biological entity studied.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"SOURCE"	
8 - 10	Continuation	continuation	Allows concatenation of multiple records.
11 - 79	Specification List	srcName	Identifies the source of the macromolecule in a token: value format.

Details

TOKEN	VALUE DEFINITION
MOL_ID	Numbers each molecule. Same as appears in COMPND.
SYNTHETIC	Indicates a chemically-synthesized source.
FRAGMENT	A domain or fragment of the molecule may be specified.
ORGANISM_SCIENTIFIC	Scientific name of the organism.
ORGANISM_COMMON	Common name of the organism.
ORGANISM_TAXID	NCBI Taxonomy ID number of the organism.
STRAIN	Identifies the strain.
VARIANT	Identifies the variant.
CELL_LINE	The specific line of cells used in the experiment.
ATCC	American Type Culture Collection tissue culture number.
ORGAN	Organized group of tissues that carries on

	a specialized function.
TISSUE	Organized group of cells with a common function and structure.
CELL	Identifies the particular cell type.
ORGANELLE	Organized structure within a cell.
SECRETION	Identifies the secretion, such as saliva, urine, or venom, from which the molecule was isolated.
CELLULAR_LOCATION	Identifies the location inside/outside the cell.
PLASMID	Identifies the plasmid containing the gene.
GENE	Identifies the gene.
EXPRESSION_SYSTEM	Scientific name of the organism in which the molecule was expressed.
EXPRESSION_SYSTEM_COMMON	Common name of the organism in which the molecule was expressed.
EXPRESSION_SYSTEM_TAXID	NCBI Taxonomy ID of the organism used as the expression system.
EXPRESSION_SYSTEM_STRAIN	Strain of the organism in which the molecule was expressed.
EXPRESSION_SYSTEM_VARIANT	Variant of the organism used as the expression system.
EXPRESSION_SYSTEM_CELL_LINE	The specific line of cells used as the expression system.
EXPRESSION_SYSTEM_ATCC_NUMBER	Identifies the ATCC number of the expression System.
EXPRESSION_SYSTEM_ORGAN	Specific organ which expressed the molecule.
EXPRESSION_SYSTEM_TISSUE	Specific tissue which expressed the molecule.
EXPRESSION_SYSTEM_CELL	Specific cell type which expressed the molecule.
EXPRESSION_SYSTEM_ORGANELLE	Specific organelle which expressed the molecule.
EXPRESSION_SYSTEM_CELLULAR_LOCATION	Identifies the location inside or outside the cell which expressed the molecule.
EXPRESSION_SYSTEM_VECTOR_TYPE	Identifies the type of vector used, i.e., plasmid, virus, or cosmid.
EXPRESSION_SYSTEM_VECTOR	Identifies the vector used.
EXPRESSION_SYSTEM_PLASMID	Plasmid used in the recombinant experiment.
EXPRESSION_SYSTEM_GENE	Name of the gene used in recombinant experiment.
OTHER_DETAILS	Used to present information on the source which is not given elsewhere.

* The srcName is a list of tokens: value pairs describing each biological component of the entry.

* As in COMPND, the order is not specified except that MOL_ID or FRAGMENT indicates subsequent specifications are related to that molecule or fragment of the molecule.

* Only the relevant tokens need to appear in an entry.

* Molecules prepared by purely chemical synthetic methods are described by the specification SYNTHETIC followed by "YES" or an optional value, such as NON-BIOLOGICAL SOURCE or BASED ON THE NATURAL SEQUENCE. ENGINEERED must appear in the COMPND record.

* In the case of a chemically synthesized molecule using a biologically functional sequence (nucleic or amino acid), SOURCE reflects the biological origin of the sequence and COMPND reflects its synthetic nature by inclusion of the token ENGINEERED. The token SYNTHETIC appears in SOURCE.

* If made from a synthetic gene, ENGINEERED appears in COMPND and the expression system is described in SOURCE (SYNTHETIC does NOT appear in SOURCE).

* If the molecule was made using recombinant techniques, ENGINEERED appears in COMPND and the system is described in SOURCE.

* When multiple macromolecules appear in the entry, each MOL_ID, as given in the COMPND record, must be repeated in the SOURCE record along with the source information for the corresponding molecule.

* Hybrid molecules prepared by fusion of genes are treated as multi-molecular systems for the purpose of specifying the source. The token FRAGMENT is used to associate the source with its corresponding fragment.

- When necessary to fully describe hybrid molecules, tokens may appear more than once for a given MOL_ID.

- All relevant token: value pairs that taken together fully describe each fragment are grouped following the appropriate FRAGMENT.

- Descriptors relative to the full system appear before the FRAGMENT (see third example below).

* ORGANISM_SCIENTIFIC provides the Latin genus and species. Virus names are listed as the scientific name.

* Cellular origin is described by giving cellular compartment, organelle, cell, tissue, organ, or body part from which the molecule was isolated.

* CELLULAR_LOCATION may be used to indicate where in the organism the compound was found. Examples are: extracellular, periplasmic, cytosol.

* Entries containing molecules prepared by recombinant techniques are described as follows:

- The expression system is described.
- The organism and cell location given are for the source of the gene used in the cloning experiment.

* Transgenic organisms, such as mouse producing human proteins, are treated as expression systems.

* New tokens may be added by the wwPDB.

Verification/Validation/Value Authority Control

The biological source is compared to that found in the sequence databases. The Tax ID is identified and the corresponding scientific and common names for the organism is matched to a standard taxonomy database (such as NCBI).

Relationships to Other Record Types

Each macromolecule listed in COMPND must have a corresponding source.

Examples

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
SOURCE    MOL_ID: 1;
SOURCE   2 ORGANISM_SCIENTIFIC: AVIAN SARCOMA VIRUS;
SOURCE   3 ORGANISM_TAXID: 11876
SOURCE   4 STRAIN: SCHMIDT-RUPPIN B;
SOURCE   5 EXPRESSION_SYSTEM: ESCHERICHIA COLI;
SOURCE   6 EXPRESSION_SYSTEM_TAXID: 562
SOURCE   7 EXPRESSION_SYSTEM_PLASMID: PRC23IN

SOURCE    MOL_ID: 1;
SOURCE   2 ORGANISM_SCIENTIFIC: GALLUS GALLUS;
SOURCE   3 ORGANISM_COMMON: CHICKEN;
SOURCE   3 ORGANISM_TAXID: 9031
SOURCE   4 ORGAN: HEART;
SOURCE   5 TISSUE: MUSCLE
```

For a Chimera protein:

```
SOURCE    MOL_ID: 1;
SOURCE   2 ORGANISM_SCIENTIFIC: MUS MUSCULUS, HOMO SAPIENS;
SOURCE   3 ORGANISM_COMMON: MOUSE, HUMAN;
SOURCE   3 ORGANISM_TAXID: 10090, 9606
SOURCE   5 EXPRESSION_SYSTEM: ESCHERICHIA COLI;
SOURCE   6 EXPRESSION_SYSTEM_TAXID: 344601
SOURCE   6 EXPRESSION_SYSTEM_STRAIN: B171;
```

```
SOURCE    7 EXPRESSION_SYSTEM_VECTOR_TYPE: PLASMID;
SOURCE    8 EXPRESSION_SYSTEM_PLASMID: P4XH-M13;
```

KEYWDS

Overview

The KEYWDS record contains a set of terms relevant to the entry. Terms in the KEYWDS record provide a simple means of categorizing entries and may be used to generate index files. This record addresses some of the limitations found in the classification field of the HEADER record. It provides the opportunity to add further annotation to the entry in a concise and computer-searchable fashion.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"KEYWDS"	
9 - 10	Continuation	continuation	Allows concatenation of records if necessary.
11 - 79	List	keywds	Comma-separated list of keywords relevant to the entry.

Details

* The KEYWDS record contains a list of terms relevant to the entry, similar to that found in journal articles. A phrase may be used if it presents a single concept (e.g., reaction center). Terms provided in this record may include those that describe the following:

- Functional classification.
- Metabolic role.
- Known biological or chemical activity.
- Structural classification.

*Other classifying terms may be used. No particular ordering is required. A number of PDB entries contain complexes of macromolecules. In these cases, all terms applicable to each molecule should be provided separated by a comma.

*Note that the terms in the KEYWDS record duplicate those found in the classification field of the HEADER record. Terms abbreviated in the HEADER record are unabbreviated in KEYWDS.

Verification/Validation/Value Authority Control

Terms used in the KEYWDS record are subject to scientific and editorial review. A list of terms, definitions, and synonyms will be maintained by the wwPDB. Every attempt will be made to provide some level of consistency with keywords used in other biological databases.

Relationships to Other Record Types

HEADER records contain a classification term which must also appear in KEYWDS. Scientific judgment will dictate when terms used in one entry to describe a molecule should be included in other entries with the same or similar molecules.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
KEYWDS    LYASE, TRICARBOXYLIC ACID CYCLE, MITOCHONDRION, OXIDATIVE
KEYWDS   2 METABOLISM
```

EXPDTA (updated)

Overview

The EXPDTA record presents information about the experiment.

The EXPDTA record identifies the experimental technique used. This may refer to the type of radiation and sample, or include the spectroscopic or modeling technique. Permitted values include:

X-RAY DIFFRACTION
FIBER DIFFRACTION
NEUTRON DIFFRACTION
ELECTRON CRYSTALLOGRAPHY
ELECTRON MICROSCOPY
SOLID-STATE NMR
SOLUTION NMR
SOLUTION SCATTERING

**Note:Since October 15, 2006, theoretical models are no longer accepted for deposition. Any theoretical models deposited prior to this date are archived at* ftp://ftp.wwpdb.org/pub/pdb/data/structures/models.
Please see the documentation from previous versions for the related file format description.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"EXPDTA"	
9 - 10	Continuation	continuation	Allows concatenation of multiple records.
11 - 79	SList	technique	The experimental technique(s) with optional comment describing the sample or experiment.

Details

* EXPDTA is mandatory and appears in all entries. The technique must match one of the permitted values. See above.

* If more than one technique was used for the structure determination and is being represented in the entry, EXPDTA presents the techniques as a semi-colon separated list.

Verification/Validation/Value Authority Control

The verification program checks that the EXPDTA record appears in the entry and that the technique matches one of the allowed values. It also checks that the relevant standard REMARK is added, as in the cases of NMR or electron microscopy studies, that the appropriate CRYST1 and SCALE values are used.

Relationships to Other Record Types

If the experiment is an NMR or electron microscopy study, this may be stated in the TITLE, and the appropriate EXPDTA and REMARK records should appear. Specific details of the data collection and experiment appear in the REMARKs.

In the case of a polycrystalline fiber diffraction study, CRYST1 and SCALE contain the normal unit cell data.

Examples

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
EXPDTA    X-RAY DIFFRACTION

EXPDTA    NEUTRON DIFFRACTION; X-RAY DIFFRACTION

EXPDTA    SOLUTION NMR

EXPDTA    ELECTRON MICROSCOPY
```

NUMMDL (added)

Overview

The NUMMDL record indicates total number of models in a PDB entry.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"NUMMDL"	
11 - 14	Integer	modelNumber	Number of models.

Details

* The modelNumber field lists total number of models in a PDB entry and is left justified.

* If more than one model appears in the entry, the number of models included must be stated.

* NUMMDL is mandatory if a PDB entry contains more than one models.

Verification/Validation/Value Authority Control

The verification program checks that the modelNumber field is correctly formatted.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
NUMMDL    20
```

MDLTYP (added)

Overview

The MDLTYP record contains additional annotation pertinent to the coordinates presented in the entry.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"MDLTYP"	
9 - 10	Continuation	continuation	Allows concatenation of multiple records.
11 - 80	SList	comment	Free Text providing additional structural annotation.

Details

* The MDLTYP record will be used by the wwPDB to highlight certain features of the deposited coordinates as described below.

* For entries that are determined by NMR methods and the coordinates deposited are either a minimized average or regularized mean structure, this record will contain the tag "MINIMIZED AVERAGE" to highlight the nature of the deposited coordinates in the entry.

* Where the entry contains entire polymer chains that have only either C-alpha (for proteins) or P atoms (for nucleotides), the MDLTYP record will be used to describe the contents of such chains along with the chain identifier. For these polymeric chains, REMARK 470 (Missing Atoms) will be omitted.

* If multiple features need to be described in this record, they will be separated by a ";" delineator.

* Where an entry has multiple features requiring description in this record including MINIMIZED AVERAGE, the MINIMIZED AVERAGE value will precede all other annotation.

* New descriptors may be added by the wwPDB.

Verification/Validation/Value Authority Control

The chain_identifiers described in this record must be present in the COMPND, SEQRES and the coordinate section of the entry.

Example

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

```
1234567890123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890
MDLTYP    MINIMIZED AVERAGE

MDLTYP    CA ATOMS ONLY, CHAIN A, B, C, D, E, F, G, H, I, J, K ; P ATOMS ONLY,
MDLTYP   2 CHAIN X, Y, Z

MDLTYP    MINIMIZED AVERAGE; CA ATOMS ONLY, CHAIN A, B
```

AUTHOR

Overview

The AUTHOR record contains the names of the people responsible for the contents of the entry.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"AUTHOR"	
9 - 10	Continuation	continuation	Allows concatenation of multiple records.
11 - 79	List	authorList	List of the author names, separated by commas.

Details

* The authorList field lists author names separated by commas with no subsequent spaces.

* Representation of personal names:

- First and middle names are indicated by initials, each followed by a period, and precede the surname.
- Only the surname (family or last name) of the author is given in full.
- Hyphens can be used if they are part of the author's name.
- Apostrophes are allowed in surnames.
- Umlauts and other character modifiers are not given.

* Structure of personal names:

- There is no space after any initial and its following period.
- Blank spaces are used in a name only if properly part of the surname (e.g., J.VAN DORN), or between surname and Jr., II, or III

  Abbreviations that are part of a surname, such as Jr., St. or Ste., are followed by a period and a space before the next part of the surname.

* Representation of corporate, organization or university names:

- Group names used for one or all of the authors should be spelled out in full.

- The name of the larger group comes before the name of a subdivision, e.g., University of Somewhere, Department of Chemistry.

* Structure of list:

- Line breaks between multiple lines in the authorList occur only after a comma.
- Personal names are not split across two lines.

* Special cases:

- Names are given in English if there is an accepted English version; otherwise in the native language, transliterated if necessary.

Verification/Validation/Value Authority Control

The verification program checks that the authorList field is correctly formatted. It does not perform any spelling checks or name verification.

Relationships to Other Record Types

The format of the names in the AUTHOR record is the same as in JRNL and REMARK 1 references.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
AUTHOR    M.B.BERRY,B.MEADOR,T.BILDERBACK,P.LIANG,M.GLASER,
AUTHOR   2 G.N.PHILLIPS JR.,T.L.ST. STEVENS
```

REVDAT (updated)

Overview

REVDAT records contain a history of the modifications made to an entry since its release.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"REVDAT"	
8 - 10	Integer	modNum	Modification number.
11 - 12	Continuation	continuation	Allows concatenation of multiple records.
14 - 22	Date	modDate	Date of modification (or release for new entries) in DD-MMM-YY format. This is not repeated on continued lines.
24 - 27	IDCode	modId	ID code of this entry. This is not repeated on continuation lines.
32	Integer	modType	An integer identifying the type of modification. For all revisions, the modification type is listed as 1
40 - 45	LString(6)	record	Modification detail.
47 - 52	LString(6)	record	Modification detail.
54 - 59	LString(6)	record	Modification detail.
61 - 66	LString(6)	record	Modification detail.

Details

* Each time revisions are made to the entry, a modification number is assigned in increasing (by 1) numerical order. REVDAT records appear in descending order (most recent modification appears first). New entries have a REVDAT record with modNum equal to 1 and modType equal to 0. Allowed modTypes are:

0	Initial released entry.
1	Other modification.

* Each revision may have more than one REVDAT record, and each revision has a separate continuation field.

* Modification details are typically PDB record names such as JRNL, SOURCE, TITLE, or COMPND. A special modification detail VERSN indicates that the file has undergone a change in version. The current version will be specified in REMARK 4.

Verification/Validation/Value Authority Control

The modType must be one of the defined types, and the given record type must be valid. If modType is 0, the modId must match the entry's ID code in the HEADER record.

Relationships to Other Record Types

In the case of a version revision, the current will be specified in REMARK 4.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REVDAT   2   15-OCT-99 1ABC    1       REMARK
REVDAT   1   09-JAN-89 1ABC    0
```

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REVDAT   2   11-MAR-08 2ABC    1       JRNL   VERSN
REVDAT   1   09-DEC-03 2ABC    0
```

SPRSDE

Overview

The SPRSDE records contain a list of the ID codes of entries that were made obsolete by the given coordinate entry and removed from the PDB release set. One entry may replace many.

It is wwPDB policy that only the principal investigator of a structure has the authority to obsolete it.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"SPRSDE"	
9 - 10	Continuation	continuation	Allows for multiple ID codes.
12 - 20	Date	sprsdeDate	Date this entry superseded the listed entries. This field is not copied on continuations.
22 - 25	IDcode	idCode	ID code of this entry. This field is not copied on continuations.
32 - 35	IDcode	sIdCode	ID code of a superseded entry.
37 - 40	IDcode	sIdCode	ID code of a superseded entry.
42 - 45	IDcode	sIdCode	ID code of a superseded entry.
47 - 50	IDcode	sIdCode	ID code of a superseded entry.
52 - 55	IDcode	sIdCode	ID code of a superseded entry.
57 - 60	IDcode	sIdCode	ID code of a superseded entry.
62 - 65	IDcode	sIdCode	ID code of a superseded entry.
67 - 70	IDcode	sIdCode	ID code of a superseded entry.
72 - 75	IDcode	sIdCode	ID code of a superseded entry.

Details

* The ID code list is terminated by the first blank sIdCode field.

Verification/Validation/Value Authority Control

wwPDB checks that the superseded entries have actually been removed from release.

Relationships to Other Record Types

The sprsdeDate is usually the date the entry is released, and therefore matches the date in the REVDAT 1 record. The ID code found in the idCode field must be the same as one found in the idCode field of the HEADER record.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
SPRSDE     17-JUL-84 4HHB      1HHB
```

```
SPRSDE     27-FEB-95 1GDJ      1LH4 2LH4
```

JRNL (updated)

Overview

The JRNL record contains the primary literature citation that describes the experiment which resulted in the deposited coordinate set. There is at most one JRNL reference per entry. If there is no primary reference, then there is no JRNL reference. Other references are given in REMARK 1.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"JRNL "	
13 - 79	LString	text	See Details below.

Details

* The following tables are used to describe the sub-record types of the JRNL record.

* The AUTH sub-record is mandatory in JRNL. This is followed by TITL, EDIT, REF, PUBL, REFN, PMID and DOI sub- record types. REF and REFN are also mandatory in JRNL. EDIT and PUBL may appear only if the reference is to a non-journal.

1. AUTH

* AUTH contains the list of authors associated with the cited article or contribution to a larger work (i.e., AUTH is not used for the editor of a book).

* The author list is formatted similarly to the AUTHOR record. It is a comma-separated list of names. Spaces at the end of a sub-record are not significant; all other spaces are significant. See the AUTHOR record for full details.

* The authorList field of continuation sub-records in JRNL differs from that in AUTHOR by leaving no leading blank in column 20 of any continuation lines.

* One author's name, consisting of the initials and family name, cannot be split across two lines. If there are continuation sub-records, then all but the last sub-record must end in a comma.

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"REMARK"	
10	LString(1)	"1"	
13 - 16	LString(4)	"AUTH"	Appears on all continuation records.
17 - 18	Continuation	continuation	Allows a long list of authors.
20 - 79	List	authorList	List of the authors.

2. TITL

* TITL specifies the title of the reference. This is used for the title of a journal article, chapter, or part of a book. The TITL line is omitted if the author(s) listed in authorList wrote the entire book (or other work) listed in REF and no section of the book is being cited.

* If an article is in a language other than English and is printed with an alternate title in English, the English language title is given, followed by a space and then the name of the language (in its English form, in square brackets) in which the article is written.

* If the title of an article is in a non-Roman alphabet the title is transliterated.

* The actual title cited is reconstructed in a manner identical to other continued records, i.e., trailing blanks are discarded and the continuation line is concatenated with a space inserted.

* A line cannot end with a hyphen. A compound term (two elements connected by a hyphen) or chemical names which include a hyphen must appear on a single line, unless they are too long to fit on one line, in which case the split is made at a normally-occurring hyphen. An individual word cannot be hyphenated at the end of a line and put on two lines. An exception is when there is a repeating compound term where the second element is omitted, e.g., "DOUBLE- AND TRIPLE-RESONANCE". In such a case the non-completed word "DOUBLE-" could end a line and not alter reconstruction of the title.

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"REMARK"	
10	LString(1)	"1"	
13 - 16	LString(4)	"TITL"	Appears on all continuation records.
17 - 18	Continuation	continuation	Permits long titles.
20 - 79	LString	title	Title of the article.

3. EDIT

* EDIT appears if editors are associated with a non-journal reference. The editor list is formatted and concatenated in the same way that author lists are.

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"REMARK"	
10	LString(1)	"1"	
13 - 16	LString(4)	"TITL"	Appears on all continuation records.
17 - 18	Continuation	continuation	Permits long titles.
20 - 79	LString	title	Title of the article.

4. REF

* REF is a group of fields that contain either the publication status or the name of the publication (and any supplement and/or report information), volume, page, and year. There are two forms of this sub-record group, depending upon the citation's publication status.

4a. If the reference has not been published yet, the sub-record type group has the form:

```
COLUMNS        DATA TYPE      FIELD          DEFINITION
------------------------------------------------------------------------------------
 1 -  6        Record name    "JRNL  "

13 - 16        LString(3)     "REF"

20 - 34        LString(15)    "TO BE PUBLISHED"
```

* Publication name (first item in pubName field):

If the publication is a serial (i.e., a journal, an annual, or other non-book or non-monographic item issued in parts and intended to be continued indefinitely), use the abbreviated name of the publication as listed in PubMed with periods.

If the publication is a book, monograph, or other non-serial item, use its full name according to the Anglo-American Cataloguing Rules, 2nd Revised Edition; (AACR2R). (Non-serial items include theses, videos, computer programs, and anything that is complete in one or a finite number of parts.) If there is a sub-title, verifiable in an online catalog, it will be included using the same punctuation as in the source of verification. Preference will be given to verification using cataloging of the Library of Congress, the National Library of Medicine, and the British Library, in that order.

If a book is part of a monographic series: the full name of the book (according to the AACR2R) is listed first, followed by the name of the series in which it was published. The series information is given within parentheses and the series name is preceded by "IN:" and a space. The series name should be listed in full unless the series has an accepted ISO abbreviation. If applicable, the series name should be followed, after a comma and a space, by a volume (V.) and/or number (NO.) and/or part (PT.) indicator and its number and/or letter in the series.

* Supplement (follows publication name in pubName field):

If a reference is in a supplement to the volume listed, or if information about a "part" is needed to distinguish multiple parts with the same page numbering, such information should be put in the REF sub-record.

A supplement indication should follow the name of the publication and should be preceded by a comma and a space. Supplement should be abbreviated as "SUPPL." If there is a supplement number or letter, it should follow "SUPPL." without an intervening space. A part indication should also follow the name of the publication and be preceded by a comma and a space. A part should be abbreviated as "PT.", and the number or letter should follow without an intervening space.

If there is both a supplement and a part, their order should reflect the order printed on the work itself.

* Report (follows publication name and any supplement or part information in pubName field):

If a book has a report designation, the report information should follow the title and precede series information. The name and number of the report is given in parentheses, and the name is preceded by "REPORT:" and a space.

* Reconstruction of publication name:

The name of the publication is reconstructed by removing any trailing blanks in the pubName field, and concatenating all of the pubName fields from the continuation lines with an intervening space. There are two conditions where no intervening space is added between lines: when the pubName field on a line ends with a hyphen or a period, or when the line ends with a hyphen (-). When the line ends with a period (.), add a space if this is the only period in the entire pubName field; do not add a space if there are two or more periods throughout the pubName field, excluding any periods after the designations "SUPPL", "V", "NO", or "PT".

* Volume, page, and year (volume, first page, year fields respectively):

The REF sub-record type group also contains information about volume, page, and year when applicable.

In the case of a monograph with multiple volumes which is also in a numbered series, the number in the volume field represents the number of the book, not the series. (The volume number of the series is in parentheses with the name of the series, as described above under publication name.)

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"JRNL "	
13 - 16	LString(3)	"REF "	
17 - 18	Continuation	continuation	Allows long publication names.
20 - 47	LString	pubName	Name of the publication including section or series designation. This is the only field of this sub-record which may be continued on successive sub-records.
50 - 51	LString(2)	"V."	Appears in the first sub-record only, and only if column 55 is non-blank.
52 - 55	String	volume	Right-justified blank-filled volume information; appears in the first sub-record only.
57 - 61	String	page	First page of the article; appears in the first sub-record only.
63 - 66	Integer	year	Year of publication; first sub-record only.

5. PUBL

* PUBL contains the name of the publisher and place of publication if the reference is to a book or other non-journal publication. If the non-journal has not yet been published or released, this sub-record is absent.

* The place of publication is listed first, followed by a space, a colon, another space, and then the name of the publisher/issuer. This arrangement is based on the ISBD(M) International Standard Bibliographic Description for Monographic Publications (Rev.Ed., 1987) and the AACR2R, and is used in public online catalogs in libraries. Details on the contents of PUBL are given below.

* Place of publication:

Give the place of publication. If the name of the country, state, province, etc. is considered necessary to distinguish the place of publication from others of the same name, or for identification, then follow the city with a comma, a space, and the name of the larger geographic area.

If there is more than one place of publication, only the first listed will be used. If an online catalog record is used to verify the item, the first place listed there will be used, omitting any brackets. Preference will be given to the cataloging done by the Library of Congress, the National Library of Medicine, and the British Library, in that order.

* Publisher's name (or name of other issuing entity):

Give the name of the publisher in the shortest form in which it can be understood and identified internationally, according to AACR2R rule 1.4D.

If there is more than one publisher listed in the publication, only the first will be used in the PDB file. If an online catalog record is used to verify the item, the first place listed there will be used for the name of the publisher. Preference will be given to the cataloging of the Library of Congress, the National Library of Medicine, and the British Library, in that order.

* Ph.D. and other theses:

Theses are presented in the PUBL record if the degree has been granted and the thesis made available for public consultation by the degree-granting institution.
The name of the degree-granting institution (the issuing agency) is followed by a space and "(THESIS)".

* Reconstruction of place and publisher:

The PUBL sub-record type can be reconstructed by removing all trailing blanks in the pub field and concatenating all of the pub fields from the continuation lines with an intervening space. Continued lines do not begin with a space.

```
COLUMNS       DATA TYPE     FIELD         DEFINITION
------------------------------------------------------------------------------------
 1 -  6       Record name   "JRNL  "
```

13 - 16	LString(4)	"PUBL"	
17 - 18	Continuation	continuation	Allows long publisher and place names.
20 - 70	LString	pub	City of publication and name of the publisher/institution.

6. REFN (changed)

* REFN is a group of fields that contain encoded references to the citation. No continuation lines are possible. Each piece of coded information has a designated field.

* There are two forms of this sub-record type group, depending upon the publication status.

6a. This form of the REFN sub-record type group is used if the citation has not been published.

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"JRNL "	
13 - 16	LString(4)	"REFN"	

6b. This form of the REFN sub-record type group is used if the citation has been published.

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"JRNL "	
13 - 16	LString(4)	"REFN"	
36 - 39	LString(4)	"ISSN" or "ESSN"	International Standard Serial Number or Electronic Standard Serial Number.
41 - 65	LString	issn	ISSN number (final digit may be a letter and may contain one or more dashes).

7. PMID (added)

* PMID lists the PubMed unique accession number of the publication related to the entry.

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"JRNL "	
13 - 16	LString(4)	"PMID"	
20 – 79	Integer	continuation	unique PubMed identifier number assigned to the publication describing the experiment. Allows for a long PubMed ID number.

8. DOI (added)

* DOI is the Digital Object Identifier for the related electronic publication (“e-pub”), if applicable.

* Every DOI consists of a publisher prefix, a fore-slash (“/”), and then a suffix which can be any length and may include a combination of numbers and alphabets.
For example: 10.1073/PNAS.0712393105

```
COLUMNS        DATA TYPE     FIELD          DEFINITION
--------------------------------------------------------------------------------
 1 -  6        Record name   "JRNL  "

13 - 16        LString(4)    "DOI "

20 – 79        LString       continuation   Unique DOI assigned to the publication
                                            describing the experiment.
                                            Allows for a long DOI string.
```

Verification/Validation/Value Authority Control

wwPDB verifies that this record is correctly formatted.

Citations appearing in JRNL may not also appear in REMARK 1.

Relationships to Other Record Types

The publication cited as the JRNL record may not be repeated in REMARK 1.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
JRNL        AUTH   G.FERMI,M.F.PERUTZ,B.SHAANAN,R.FOURME
JRNL        TITL   THE CRYSTAL STRUCTURE OF HUMAN DEOXYHAEMOGLOBIN AT
JRNL        TITL 2 1.74 A RESOLUTION
JRNL        REF    J.MOL.BIOL.                   V. 175   159 1984
JRNL        REFN                    ISSN 0022-2836
JRNL        PMID   6726807
JRNL        DOI    10.1016/0022-2836(84)90472-8
```

Known Problems

* Interchange of bibliographic information and linking with other databases is hampered by the lack of labels or specific locations for certain types of information or by more than one type of information being in a particular location. This is most likely to occur with books, series, and reports. Some of the points below provide details about the variations and/or blending of information.

* Titles of the publications that require more than 28 characters on the REF line must be continued on subsequent lines. There is some awkwardness due to volume, page, and year appearing on the first REF line, thereby splitting up the title.

* Information about a supplement and its number/letter is presented in the publication's title field (on

the REF lines in columns 20 - 47).

* When series information for a book is presented, it is added to the REF line. The number of REF lines can become large in some cases because of the 28-column limit for title information in REF.

* Books that are issued in more than one series are not accommodated.

* Pagination is limited to the beginning page.

REMARK

Overview

REMARK records present experimental details, annotations, comments, and information not included in other records. In a number of cases, REMARKs are used to expand the contents of other record types. A new level of structure is being used for some REMARK records. This is expected to facilitate searching and will assist in the conversion to a relational database.

The very first line of every set of REMARK records is used as a spacer to aid in reading.

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"REMARK"	
8 - 10	Integer	remarkNum	Remark number. It is not an error for remark n to exist in an entry when remark n-1 does not.
12 - 70	LString	empty	Left as white space in first line of each new remark.

REMARKs 0-4

REMARK 0, 1, 2, 3, and 4 detailed below, are specific for re-refinement, references, resolution, final refinement, and PDB File Format version, respectively.

REMARK 0 (added), Re-refinement notice

REMARK 0 identifies entries in which a re-refinement has been performed using the data from an existing entry.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK   0
REMARK   0 THIS ENTRY yyyy REFLECTS AN ALTERNATIVE MODELING OF THE
REMARK   0 ORIGINAL STRUCTURAL DATA (RxxxxSF or xxxx.MR) DETERMINED BY
REMARK   0 AUTHORS OF THE PDB ENTRY xxxx:
REMARK   0 AUTHOR INITIALS, AUTHOR LAST NAME
```

Note: In entries where REMARK 0 is included as described above, remarks REMARK 1 and REMARK 900 will also reflect the reuse of existing experimental data.

```
REMARK 200 REMARK: AUTHOR USED THE SF(MR) DATA FROM ENTRY xxxx.
```

NOTE: the rest of REMARKs 200 and 280 are blank, since the re-refinement author did not collect original data.

REMARK 1 (updated), Related publications

REMARK 1 lists important publications related to the structure presented in the entry. These citations are chosen by the depositor. They are listed in reverse-chronological order. Citations are not repeated from the JRNL records. After the first blank record and the REFERENCE sub-record, the sub-record types for REMARK 1 are the same as in the JRNL sub-record types. For details, see the JRNL section.

Record Format and Details

As with all other remarks, the first line is empty and is used as a spacer.

The following tables are used to describe the sub-record types of REMARK 1.

1. REFERENCE

Each reference is preceded by a line indicating the reference number in the entry.

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"REMARK"	
10	LString(1)	"1"	
12 - 20	LString(9)	"REFERENCE"	
22 - 70	Integer	refNum	Reference number. Starts with 1 and increments by 1.

2. AUTH

AUTH contains the list of authors of the reference.

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"REMARK"	
10	LString(1)	"1"	
13 - 16	LString(4)	"AUTH"	Appears on all continuation records.
17 - 18	Continuation	continuation	Allows a long list of authors.
20 - 79	List	authorList	List of the authors.

See JRNL AUTH for details.

3. TITL

TITL specifies the title of the reference.

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"REMARK"	
10	LString(1)	"1"	
13 - 16	LString(4)	"TITL"	Appears on all continuation records.
17 - 18	Continuation	continuation	Permits long titles.
20 - 79	LString	title	Title of the article.

See JRNL TITL for details.

4. EDIT

EDIT appears if editors are associated with a non-journal reference.

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"REMARK"	
10	LString(1)	"1"	
13 - 16	LString(4)	"TITL"	Appears on all continuation records.
17 - 18	Continuation	continuation	Permits long list of editors.
20 - 79	LString	editorList	List of the editors.

See JRNL EDIT for details.

5. REF

REF is a group of fields which contains the name of the publication.

5a. If it has not been published yet, the REF sub-record type has the form:

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"REMARK"	
10	LString(1)	"1"	
13 - 16	LString(3)	"REF"	
20 - 34	LString(15)	"TO BE PUBLISHED"	

At the present time, there is no formal mechanism in place for monitoring the subsequent publication of referenced papers. wwPDB relies upon the depositor to provide reference update information since preliminary information can change by the time of actual publication.

5b. If the reference has been published, then the REF sub-record type group contains information about the name of the publication, supplement, report, volume, page, and year, in the appropriate fields.

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"REMARK"	
10	LString(1)	"1"	
13 - 16	LString(3)	"REF"	
17 - 18	Continuation	continuation	Permits long publication names.
20 - 47	LString	pubName	Name of the publication including section or series designation. This is the only field of this record which may be continued on successive records.
50 - 51	LString(2)	"V."	Appears in the first record only, and only if column 55 is filled in.
52 - 55	String	volume	Right-justified blank-filled volume information; appears in the first sub-record only.
57 - 61	String	page	First page of the article; appears in the first sub-record only.
63 - 66	Integer	year	First record year of publication.

See JRNL REF for details.

6. PUBL

PUBL contains the name of the publisher and place of publication if the reference is to a book or other non-journal publication. If the reference has not yet been published or released, this sub-record is absent.

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"REMARK"	
10	LString(1)	"1"	
13 - 16	LString(4)	"PUBL"	
17 - 18	Continuation	continuation	Permits long publisher and city information.
20 - 70	LString	pub	Name of the publisher and city of publication.

See JRNL PUBL for details.

7. REFN (changed)

REFN is a group of fields which contains encoded references to the citation.

7a. If the citation has not been published, this form of the REFN sub-record type group is used.

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"REMARK"	
10	LString(1)	"1"	
13 - 16	LString(4)	"REFN"	

7b. If the citation has been published, this form of the REFN sub-record type group is used.

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"REMARK"	
10	LString(1)	"1"	
13 - 16	LString(4)	"REFN"	
36 - 39	LString(4)	"ISSN" or "ESSN"	International Standard Serial Number or Electronic Standard Serial Number.
41 - 65	LString	issn	ISSN number.

See JRNL REFN for details.

8. PMID (added)

* PMID lists the PubMed unique accession number of the publication related to the entry.

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"JRNL "	
13 - 16	LString(4)	"PMID"	
20 – 79	Integer	continuation	unique PubMed identifier number assigned to the publication describing the experiment. Allows for a long pubmed id number.

9. DOI (added)

* DOI is the Digital Object Identifier for the related electronic publication (“e-pub”), if applicable.

* Every DOI consists of a publisher prefix, a fore-slash (“/”), and then a suffix which can be any length and may include a combination of numbers and alphabets. For example:
10.1073/PNAS.0712393105

COLUMNS	DATA TYPE	FIELD	DEFINITION

1 - 6	Record name	"JRNL "	
13 - 16	LString(4)	"DOI "	
20 - 79	LString	continuation	Unique DOI assigned to the Publication describing the experiment. Allows for a long DOI string.

Verification/Validation/Value Authority Control

wwPDB verifies that this record is correctly formatted.

Relationships to Other Record Types

Citations appearing in REMARK 1 may not appear in JRNL.

Examples

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK   1
REMARK   1 REFERENCE 1
REMARK   1  AUTH   J.N.BREG,J.H.J.VAN OPHEUSDEN,M.J.M.BURGERING,
REMARK   1  AUTH 2 R.BOELENS,R.KAPTEIN
REMARK   1  TITL   STRUCTURE OF ARC REPRESSOR IN SOLUTION: EVIDENCE
REMARK   1  TITL 2 FOR A FAMILY OF B-SHEET DNA-BINDING PROTEIN
REMARK   1  REF    NATURE                        V. 346   586 1990
REMARK   1  REFN                   ISSN 0028-0836
REMARK   1  PMID   2377232
REMARK   1  DOI    10.1038/346586a0
REMARK   1 REFERENCE 2
REMARK   1  AUTH   J.N.BREG,R.BOELENS,A.V.E.GEORGE,R.KAPTEIN
REMARK   1  TITL   SEQUENCE-SPECIFIC 1H NMR ASSIGNMENT AND SECONDARY
REMARK   1  TITL 2 STRUCTURE OF THE ARC REPRESSOR OF BACTERIOPHAGE
REMARK   1  TITL 3 P22 AS DETERMINED BY 2D 1H NMR SPECTROSCOPY
REMARK   1  REF    BIOCHEMISTRY                  V.  28  9826 1989
REMARK   1  REFN                   ISSN 0006-2960
REMARK   1  PMID   2611268

REMARK   1
REMARK   1 REFERENCE 1
REMARK   1  AUTH   J.MAREK,J.VEVODOVA,I.SMATANOVA,Y.NAGATA,
REMARK   1  AUTH 2 L.A.SVENSSON,J.NEWMAN,M.TAKAGI,J.DAMBORSKY
REMARK   1  TITL   CRYSTAL STRUCTURE OF THE HALOALKANE DEHALOGENASE
REMARK   1  TITL 2 FROM SPHINGOMONAS PAUCIMOBILIS UT26
REMARK   1  REF    BIOCHEMISTRY                  V.  39 14082 2000
REMARK   1  REFN                   ISSN 0006-2960
REMARK   1  PMID   11087355
REMARK   1  DOI    10.1021/bi001539c
```

Known Problems

See JRNL for a listing of problems associated with references.

REMARK 2 (updated), Resolution

REMARK 2 states the highest resolution, in Angstroms, that was used in building the model. As with all the remarks, the first REMARK 2 record is empty and is used as a spacer.

Record Format and Details

* The second REMARK 2 record has one of two formats. The first is used for diffraction studies, the second for other types of experiments in which resolution is not relevant, e.g., NMR.

* For diffraction experiments,

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"REMARK"	
10	LString(1)	"2"	
12 - 22	LString(11)	"RESOLUTION."	
24 - 30	Real(7.2)	resolution	Resolution.
32 - 41	LString(10)	"ANGSTROMS."	

* REMARK 2 when not a diffraction experiment:

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"REMARK"	
10	LString(1)	"2"	
12 - 38	LString(28)	"RESOLUTION. NOT APPLICABLE."	
41 - 70	String	comment	Comment in free text.

* Additional explanatory text may be included starting with the third line of the REMARK 2 record. For example, depositors may wish to qualify the resolution value provided due to unusual experimental conditions.

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"REMARK"	
10	LString(1)	"2"	
12 - 22	LString(11)	"RESOLUTION."	
24 - 70	String	comment	Comment in free text.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK   2
REMARK   2 RESOLUTION.    1.74 ANGSTROMS.

REMARK   2
REMARK   2 RESOLUTION. NOT APPLICABLE.

REMARK   2
REMARK   2 RESOLUTION.    7.50 ANGSTROMS.
```

REMARK 3 (updated), Final refinement information

Overview

REMARK 3 presents information on refinement program(s) used and related statistics. For non-diffraction studies, REMARK 3 is used to describe any refinement done, but its format is mostly free text.

Details

* The value "NULL" is given when there is no data available for a particular token.

* If more than one refinement package was used, they may be named in "OTHER REFINEMENT REMARKS". However, REMARK 3 statistics are given for the final refinement run.

The format of this remark changes with the evolution of refinement software. Selected representative templates or examples are provided here.

Refinement using X-PLOR

Template/example

```
REMARK   3
REMARK   3 REFINEMENT.
REMARK   3   PROGRAM     : X-PLOR 3.851
REMARK   3   AUTHORS     : BRUNGER
REMARK   3
REMARK   3  DATA USED IN REFINEMENT.
REMARK   3   RESOLUTION RANGE HIGH (ANGSTROMS) : 2.47
REMARK   3   RESOLUTION RANGE LOW  (ANGSTROMS) : 34.50
REMARK   3   DATA CUTOFF            (SIGMA(F)) : 0.000
REMARK   3   DATA CUTOFF HIGH         (ABS(F)) : NULL
REMARK   3   DATA CUTOFF LOW          (ABS(F)) : NULL
REMARK   3   COMPLETENESS (WORKING+TEST)   (%) : NULL
REMARK   3   NUMBER OF REFLECTIONS             : 28372
REMARK   3
REMARK   3  FIT TO DATA USED IN REFINEMENT.
REMARK   3   CROSS-VALIDATION METHOD          : THROUGHOUT
REMARK   3   FREE R VALUE TEST SET SELECTION  : RANDOM
REMARK   3   R VALUE            (WORKING SET) : 0.174
REMARK   3   FREE R VALUE                     : 0.244
REMARK   3   FREE R VALUE TEST SET SIZE   (%) : NULL
REMARK   3   FREE R VALUE TEST SET COUNT      : 2819
REMARK   3   ESTIMATED ERROR OF FREE R VALUE  : NULL
REMARK   3
REMARK   3  FIT IN THE HIGHEST RESOLUTION BIN.
REMARK   3   TOTAL NUMBER OF BINS USED           : NULL
REMARK   3   BIN RESOLUTION RANGE HIGH       (A) : NULL
REMARK   3   BIN RESOLUTION RANGE LOW        (A) : NULL
REMARK   3   BIN COMPLETENESS (WORKING+TEST) (%) : NULL
REMARK   3   REFLECTIONS IN BIN    (WORKING SET) : NULL
REMARK   3   BIN R VALUE           (WORKING SET) : NULL
REMARK   3   BIN FREE R VALUE                    : NULL
REMARK   3   BIN FREE R VALUE TEST SET SIZE  (%) : NULL
REMARK   3   BIN FREE R VALUE TEST SET COUNT     : NULL
REMARK   3   ESTIMATED ERROR OF BIN FREE R VALUE : NULL
REMARK   3
REMARK   3  NUMBER OF NON-HYDROGEN ATOMS USED IN REFINEMENT.
REMARK   3   PROTEIN ATOMS            : 5711
REMARK   3   NUCLEIC ACID ATOMS       : 0
REMARK   3   HETEROGEN ATOMS          : 0
REMARK   3   SOLVENT ATOMS            : 132
REMARK   3
REMARK   3  B VALUES.
REMARK   3   FROM WILSON PLOT           (A**2) : 31.94
REMARK   3   MEAN B VALUE      (OVERALL, A**2) : 27.18
REMARK   3   OVERALL ANISOTROPIC B VALUE.
REMARK   3    B11 (A**2) : NULL
REMARK   3    B22 (A**2) : NULL
REMARK   3    B33 (A**2) : NULL
REMARK   3    B12 (A**2) : NULL
REMARK   3    B13 (A**2) : NULL
REMARK   3    B23 (A**2) : NULL
REMARK   3
REMARK   3  ESTIMATED COORDINATE ERROR.
REMARK   3   ESD FROM LUZZATI PLOT        (A) : 0.24
REMARK   3   ESD FROM SIGMAA              (A) : 0.25
REMARK   3   LOW RESOLUTION CUTOFF        (A) : NULL
REMARK   3
REMARK   3  CROSS-VALIDATED ESTIMATED COORDINATE ERROR.
```

```
REMARK   3   ESD FROM C-V LUZZATI PLOT    (A) : NULL
REMARK   3   ESD FROM C-V SIGMAA          (A) : NULL
REMARK   3
REMARK   3  RMS DEVIATIONS FROM IDEAL VALUES.
REMARK   3   BOND LENGTHS                 (A) : 0.006
REMARK   3   BOND ANGLES            (DEGREES) : 1.18
REMARK   3   DIHEDRAL ANGLES        (DEGREES) : 27.95
REMARK   3   IMPROPER ANGLES        (DEGREES) : NULL
REMARK   3
REMARK   3  ISOTROPIC THERMAL MODEL : GROUPED ISOTROPIC B-FACTORS, 2 B-
REMARK   3                            VALUES/RESIDUE
REMARK   3
REMARK   3  ISOTROPIC THERMAL FACTOR RESTRAINTS.    RMS    SIGMA
REMARK   3   MAIN-CHAIN BOND              (A**2) : NULL  ; NULL
REMARK   3   MAIN-CHAIN ANGLE             (A**2) : NULL  ; NULL
REMARK   3   SIDE-CHAIN BOND              (A**2) : NULL  ; NULL
REMARK   3   SIDE-CHAIN ANGLE             (A**2) : NULL  ; NULL
REMARK   3
REMARK   3  NCS MODEL : NULL
REMARK   3
REMARK   3  NCS RESTRAINTS.                         RMS   SIGMA/WEIGHT
REMARK   3   GROUP  1  POSITIONAL            (A) : NULL  ; NULL
REMARK   3   GROUP  1  B-FACTOR           (A**2) : NULL  ; NULL
REMARK   3
REMARK   3  PARAMETER FILE  1  : NULL
REMARK   3  TOPOLOGY FILE  1   : NULL
REMARK   3
REMARK   3  OTHER REFINEMENT REMARKS: NULL
```

Refinement using CNS

Template/example

```
REMARK   3
REMARK   3 REFINEMENT.
REMARK   3   PROGRAM     : CNS 1.2
REMARK   3   AUTHORS     : BRUNGER,ADAMS,CLORE,DELANO,GROS,GROSSE-
REMARK   3               : KUNSTLEVE,JIANG,KUSZEWSKI,NILGES, PANNU,
REMARK   3               : READ,RICE,SIMONSON,WARREN
REMARK   3
REMARK   3  REFINEMENT TARGET : ENGH & HUBER
REMARK   3
REMARK   3  DATA USED IN REFINEMENT.
REMARK   3   RESOLUTION RANGE HIGH (ANGSTROMS) : 2.20
REMARK   3   RESOLUTION RANGE LOW  (ANGSTROMS) : 19.87
REMARK   3   DATA CUTOFF            (SIGMA(F)) : 2.0
REMARK   3   DATA CUTOFF HIGH         (ABS(F)) : 89190.68
REMARK   3   DATA CUTOFF LOW          (ABS(F)) : 0.0000
REMARK   3   COMPLETENESS (WORKING+TEST)   (%) : 91.1
REMARK   3   NUMBER OF REFLECTIONS             : 32745
REMARK   3
REMARK   3  FIT TO DATA USED IN REFINEMENT.
REMARK   3   CROSS-VALIDATION METHOD          : THROUGHOUT
REMARK   3   FREE R VALUE TEST SET SELECTION  : RANDOM
REMARK   3   R VALUE            (WORKING SET) : 0.203
REMARK   3   FREE R VALUE                     : 0.237
REMARK   3   FREE R VALUE TEST SET SIZE   (%) : 5.0
REMARK   3   FREE R VALUE TEST SET COUNT      : 1633
REMARK   3   ESTIMATED ERROR OF FREE R VALUE  : 0.006
REMARK   3
REMARK   3  FIT IN THE HIGHEST RESOLUTION BIN.
REMARK   3   TOTAL NUMBER OF BINS USED           : 6
REMARK   3   BIN RESOLUTION RANGE HIGH       (A) : 2.00
REMARK   3   BIN RESOLUTION RANGE LOW        (A) : 2.13
REMARK   3   BIN COMPLETENESS (WORKING+TEST) (%) : NULL
REMARK   3   REFLECTIONS IN BIN    (WORKING SET) : 0
REMARK   3   BIN R VALUE           (WORKING SET) : 0.237
REMARK   3   FREE R VALUE                     : NULL
REMARK   3   FREE R VALUE TEST SET SIZE   (%) : NULL
REMARK   3   FREE R VALUE TEST SET COUNT      : NULL
REMARK   3   ESTIMATED ERROR OF FREE R VALUE  : NULL
REMARK   3
REMARK   3  NUMBER OF NON-HYDROGEN ATOMS USED IN REFINEMENT.
REMARK   3   PROTEIN ATOMS            : 2172
REMARK   3   NUCLEIC ACID ATOMS       : 0
REMARK   3   HETEROGEN ATOMS          : 0
REMARK   3   SOLVENT ATOMS            : 127
REMARK   3
REMARK   3  B VALUES.
REMARK   3   FROM WILSON PLOT           (A**2) : 11.20
REMARK   3   MEAN B VALUE      (OVERALL, A**2) : 25.20
REMARK   3   OVERALL ANISOTROPIC B VALUE.
REMARK   3    B11 (A**2) : 2.38000
REMARK   3    B22 (A**2) : 2.38000
REMARK   3    B33 (A**2) : -4.76000
REMARK   3    B12 (A**2) : 0.00000
REMARK   3    B13 (A**2) : 0.00000
REMARK   3    B23 (A**2) : 0.00000
REMARK   3
REMARK   3  ESTIMATED COORDINATE ERROR.
REMARK   3   ESD FROM LUZZATI PLOT        (A) : 0.22
```

```
REMARK   3   ESD FROM SIGMAA               (A) : 0.07
REMARK   3   LOW RESOLUTION CUTOFF         (A) : 5.00
REMARK   3
REMARK   3  CROSS-VALIDATED ESTIMATED COORDINATE ERROR.
REMARK   3   ESD FROM C-V LUZZATI PLOT    (A) : 0.26
REMARK   3   ESD FROM C-V SIGMAA          (A) : 0.18
REMARK   3
REMARK   3  RMS DEVIATIONS FROM IDEAL VALUES.
REMARK   3   BOND LENGTHS                 (A) : 0.006
REMARK   3   BOND ANGLES            (DEGREES) : 1.30
REMARK   3   DIHEDRAL ANGLES        (DEGREES) : 24.30
REMARK   3   IMPROPER ANGLES        (DEGREES) : 0.82
REMARK   3
REMARK   3  ISOTROPIC THERMAL MODEL : RESTRAINED
REMARK   3
REMARK   3  ISOTROPIC THERMAL FACTOR RESTRAINTS.    RMS    SIGMA
REMARK   3   MAIN-CHAIN BOND              (A**2) : NULL  ; NULL
REMARK   3   MAIN-CHAIN ANGLE             (A**2) : NULL  ; NULL
REMARK   3   SIDE-CHAIN BOND              (A**2) : NULL  ; NULL
REMARK   3   SIDE-CHAIN ANGLE             (A**2) : NULL  ; NULL
REMARK   3
REMARK   3  BULK SOLVENT MODELING.
REMARK   3   METHOD USED : FLAT MODEL
REMARK   3   KSOL        : 0.45
REMARK   3   BSOL        : 64.83
REMARK   3
REMARK   3  NCS MODEL : NULL
REMARK   3
REMARK   3  NCS RESTRAINTS.                         RMS   SIGMA/WEIGHT
REMARK   3   GROUP  1  POSITIONAL            (A) : NULL  ; NULL
REMARK   3   GROUP  1  B-FACTOR           (A**2) : NULL  ; NULL
REMARK   3
REMARK   3  PARAMETER FILE  1  : PROTEIN_REP.PARAM
REMARK   3  PARAMETER FILE  2  : WATER_REP.PARAM
REMARK   3  PARAMETER FILE  3  : ION.PARAM
REMARK   3  PARAMETER FILE  4  : NULL
REMARK   3  TOPOLOGY FILE  1   : PROTEIN.TOP
REMARK   3  TOPOLOGY FILE  2   : WATER.TOP
REMARK   3  TOPOLOGY FILE  3   : ION.TOP
REMARK   3  TOPOLOGY FILE  4   : NULL
REMARK   3
REMARK   3  OTHER REFINEMENT REMARKS: BULK SOLVENT MODEL USED
```

Refinement using CNX

Template/example

```
REMARK   3
REMARK   3 REFINEMENT.
REMARK   3   PROGRAM     : CNX
REMARK   3   AUTHORS     : BRUNGER,ADAMS,CLORE,DELANO,GROS,GROSSE-
REMARK   3               : KUNSTLEVE,JIANG,KUSZEWSKI,NILGES, PANNU,
REMARK   3               : READ,RICE,SIMONSON,WARREN
REMARK   3
REMARK   3  DATA USED IN REFINEMENT.
REMARK   3   RESOLUTION RANGE HIGH (ANGSTROMS) : 3.00
REMARK   3   RESOLUTION RANGE LOW  (ANGSTROMS) : 50.00
REMARK   3   DATA CUTOFF            (SIGMA(F)) : 0.000
REMARK   3   DATA CUTOFF HIGH         (ABS(F)) : 1000.000
REMARK   3   DATA CUTOFF LOW          (ABS(F)) : 0.0000
REMARK   3   COMPLETENESS (WORKING+TEST)   (%) : 94.0
REMARK   3   NUMBER OF REFLECTIONS             : 20693
REMARK   3
REMARK   3  FIT TO DATA USED IN REFINEMENT.
REMARK   3   CROSS-VALIDATION METHOD          : NULL
REMARK   3   FREE R VALUE TEST SET SELECTION  : NULL
REMARK   3   R VALUE     (WORKING + TEST SET) : NULL
REMARK   3   R VALUE            (WORKING SET) : 0.219
REMARK   3   FREE R VALUE                     : 0.319
REMARK   3   FREE R VALUE TEST SET SIZE   (%) : 7.500
REMARK   3   FREE R VALUE TEST SET COUNT      : 1643
REMARK   3   ESTIMATED ERROR OF FREE R VALUE  : NULL
REMARK   3
REMARK   3  FIT/AGREEMENT OF MODEL WITH ALL DATA.
REMARK   3   R VALUE     (WORKING + TEST SET, NO CUTOFF) : NULL
REMARK   3   R VALUE            (WORKING SET, NO CUTOFF) : NULL
REMARK   3   FREE R VALUE                   (NO CUTOFF) : NULL
REMARK   3   FREE R VALUE TEST SET SIZE   (%, NO CUTOFF) : NULL
REMARK   3   FREE R VALUE TEST SET COUNT    (NO CUTOFF) : NULL
REMARK   3   ESTIMATED ERROR OF FREE R VALUE (NO CUTOFF) : NULL
REMARK   3   TOTAL NUMBER OF REFLECTIONS    (NO CUTOFF) : 20693
REMARK   3
REMARK   3  FIT IN THE HIGHEST RESOLUTION BIN.
REMARK   3   TOTAL NUMBER OF BINS USED           : NULL
REMARK   3   BIN RESOLUTION RANGE HIGH       (A) : NULL
REMARK   3   BIN RESOLUTION RANGE LOW        (A) : NULL
REMARK   3   BIN COMPLETENESS (WORKING+TEST) (%) : NULL
REMARK   3   REFLECTIONS IN BIN    (WORKING SET) : NULL
REMARK   3   BIN R VALUE           (WORKING SET) : NULL
REMARK   3   BIN FREE R VALUE                    : NULL
REMARK   3   BIN FREE R VALUE TEST SET SIZE  (%) : NULL
REMARK   3   BIN FREE R VALUE TEST SET COUNT     : NULL
REMARK   3   ESTIMATED ERROR OF BIN FREE R VALUE : NULL
REMARK   3
REMARK   3  NUMBER OF NON-HYDROGEN ATOMS USED IN REFINEMENT.
REMARK   3   PROTEIN ATOMS            : 7895
REMARK   3   NUCLEIC ACID ATOMS       : 0
REMARK   3   HETEROGEN ATOMS          : 276
REMARK   3   SOLVENT ATOMS            : 0
REMARK   3
REMARK   3  B VALUES.
REMARK   3   FROM WILSON PLOT           (A**2) : NULL
REMARK   3   MEAN B VALUE      (OVERALL, A**2) : 43.13
REMARK   3   OVERALL ANISOTROPIC B VALUE.
REMARK   3    B11 (A**2) : -6.46200
```

```
REMARK   3    B22 (A**2) : 0.93900
REMARK   3    B33 (A**2) : 5.52300
REMARK   3    B12 (A**2) : 0.00000
REMARK   3    B13 (A**2) : 0.00000
REMARK   3    B23 (A**2) : 0.00000
REMARK   3
REMARK   3  ESTIMATED COORDINATE ERROR.
REMARK   3   ESD FROM LUZZATI PLOT        (A) : NULL
REMARK   3   ESD FROM SIGMAA              (A) : NULL
REMARK   3   LOW RESOLUTION CUTOFF        (A) : NULL
REMARK   3
REMARK   3  CROSS-VALIDATED ESTIMATED COORDINATE ERROR.
REMARK   3   ESD FROM C-V LUZZATI PLOT    (A) : NULL
REMARK   3   ESD FROM C-V SIGMAA          (A) : NULL
REMARK   3
REMARK   3  RMS DEVIATIONS FROM IDEAL VALUES.
REMARK   3   BOND LENGTHS                 (A) : 0.010
REMARK   3   BOND ANGLES            (DEGREES) : 1.85
REMARK   3   DIHEDRAL ANGLES        (DEGREES) : NULL
REMARK   3   IMPROPER ANGLES        (DEGREES) : NULL
REMARK   3
REMARK   3  ISOTROPIC THERMAL MODEL : NULL
REMARK   3
REMARK   3  ISOTROPIC THERMAL FACTOR RESTRAINTS.    RMS    SIGMA
REMARK   3   MAIN-CHAIN BOND              (A**2) : 1.284 ; 1.500
REMARK   3   MAIN-CHAIN ANGLE             (A**2) : 2.239 ; 2.000
REMARK   3   SIDE-CHAIN BOND              (A**2) : 1.709 ; 2.000
REMARK   3   SIDE-CHAIN ANGLE             (A**2) : 2.698 ; 2.500
REMARK   3
REMARK   3  BULK SOLVENT MODELING.
REMARK   3   METHOD USED : NULL
REMARK   3   KSOL        : NULL
REMARK   3   BSOL        : NULL
REMARK   3
REMARK   3  NCS MODEL : NULL
REMARK   3
REMARK   3  PARAMETER FILE  1  : PROTEIN_REP.PARAM
REMARK   3  PARAMETER FILE  2  : DNA-RNA_REP.PARAM
REMARK   3  PARAMETER FILE  3  : WATER_REP.PARAM
REMARK   3  PARAMETER FILE  4  : N1234.XPRM
REMARK   3  PARAMETER FILE  5  : LIG12AB.XPRM
REMARK   3  TOPOLOGY FILE  1   : PROTEIN.TOP
REMARK   3  TOPOLOGY FILE  2   : DNA-RNA.TOP
REMARK   3  TOPOLOGY FILE  3   : WATER.TOP
REMARK   3  TOPOLOGY FILE  4   : ION.TOP
REMARK   3  TOPOLOGY FILE  5   : NULL
REMARK   3
REMARK   3  OTHER REFINEMENT REMARKS: NULL
```

Refinement using REFMAC

Template/example 1

```
REMARK   3
REMARK   3 REFINEMENT.
REMARK   3   PROGRAM     : REFMAC 5.3.0017
REMARK   3   AUTHORS     : MURSHUDOV,VAGIN,DODSON
REMARK   3
REMARK   3    REFINEMENT TARGET : MAXIMUM LIKELIHOOD
REMARK   3
REMARK   3  DATA USED IN REFINEMENT.
REMARK   3   RESOLUTION RANGE HIGH (ANGSTROMS) : 2.20
REMARK   3   RESOLUTION RANGE LOW  (ANGSTROMS) : 50.00
REMARK   3   DATA CUTOFF            (SIGMA(F)) : 0.000
REMARK   3   COMPLETENESS FOR RANGE        (%) : 99.7
REMARK   3   NUMBER OF REFLECTIONS             : 41377
REMARK   3
REMARK   3  FIT TO DATA USED IN REFINEMENT.
REMARK   3   CROSS-VALIDATION METHOD          : THROUGHOUT
REMARK   3   FREE R VALUE TEST SET SELECTION  : RANDOM
REMARK   3   R VALUE     (WORKING + TEST SET) : 0.228
REMARK   3   R VALUE            (WORKING SET) : 0.225
REMARK   3   FREE R VALUE                     : 0.283
REMARK   3   FREE R VALUE TEST SET SIZE   (%) : 5.200
REMARK   3   FREE R VALUE TEST SET COUNT      : 2256
REMARK   3
REMARK   3  FIT IN THE HIGHEST RESOLUTION BIN.
REMARK   3   TOTAL NUMBER OF BINS USED           : 20
REMARK   3   BIN RESOLUTION RANGE HIGH       (A) : 2.20
REMARK   3   BIN RESOLUTION RANGE LOW        (A) : 2.26
REMARK   3   REFLECTION IN BIN     (WORKING SET) : 2978
REMARK   3   BIN COMPLETENESS (WORKING+TEST) (%) : 98.65
REMARK   3   BIN R VALUE           (WORKING SET) : 0.2840
REMARK   3   BIN FREE R VALUE SET COUNT          : 161
REMARK   3   BIN FREE R VALUE                    : 0.3680
REMARK   3
REMARK   3  NUMBER OF NON-HYDROGEN ATOMS USED IN REFINEMENT.
REMARK   3   PROTEIN ATOMS            : 2043
REMARK   3   NUCLEIC ACID ATOMS       : 0
REMARK   3   HETEROGEN ATOMS          : 12
REMARK   3   SOLVENT ATOMS            : 88
REMARK   3
REMARK   3  B VALUES.
REMARK   3   FROM WILSON PLOT           (A**2) : 41.60
REMARK   3   MEAN B VALUE      (OVERALL, A**2) : 27.81
REMARK   3   OVERALL ANISOTROPIC B VALUE.
REMARK   3    B11 (A**2) : 4.03000
REMARK   3    B22 (A**2) : -3.59000
REMARK   3    B33 (A**2) : -0.44000
REMARK   3    B12 (A**2) : 0.00000
REMARK   3    B13 (A**2) : -0.05000
REMARK   3    B23 (A**2) : 0.00000
REMARK   3
REMARK   3  ESTIMATED OVERALL COORDINATE ERROR.
REMARK   3   ESU BASED ON R VALUE                            (A): 0.345
REMARK   3   ESU BASED ON FREE R VALUE                       (A): 0.256
REMARK   3   ESU BASED ON MAXIMUM LIKELIHOOD                 (A): 0.236
REMARK   3   ESU FOR B VALUES BASED ON MAXIMUM LIKELIHOOD (A**2): 17.897
REMARK   3
REMARK   3 CORRELATION COEFFICIENTS.
REMARK   3   CORRELATION COEFFICIENT FO-FC      : 0.948
```

```
REMARK   3   CORRELATION COEFFICIENT FO-FC FREE : 0.922
REMARK   3
REMARK   3  RMS DEVIATIONS FROM IDEAL VALUES        COUNT    RMS    WEIGHT
REMARK   3   BOND LENGTHS REFINED ATOMS        (A):  6752 ; 0.012 ; 0.022
REMARK   3   BOND LENGTHS OTHERS               (A):  4487 ; 0.002 ; 0.020
REMARK   3   BOND ANGLES REFINED ATOMS   (DEGREES):  9197 ; 1.419 ; 1.960
REMARK   3   BOND ANGLES OTHERS          (DEGREES): 10983 ; 0.883 ; 3.004
REMARK   3   TORSION ANGLES, PERIOD 1    (DEGREES):   856 ; 6.794 ; 5.000
REMARK   3   TORSION ANGLES, PERIOD 2    (DEGREES):   254 ;35.063 ;24.724
REMARK   3   TORSION ANGLES, PERIOD 3    (DEGREES):  1071 ;16.530 ;15.000
REMARK   3   TORSION ANGLES, PERIOD 4    (DEGREES):    18 ;20.218 ;15.000
REMARK   3   CHIRAL-CENTER RESTRAINTS       (A**3):  1031 ; 0.082 ; 0.200
REMARK   3   GENERAL PLANES REFINED ATOMS      (A):  7482 ; 0.005 ; 0.020
REMARK   3   GENERAL PLANES OTHERS             (A):  1286 ; 0.001 ; 0.020
REMARK   3   NON-BONDED CONTACTS REFINED ATOMS (A):  1064 ; 0.196 ; 0.200
REMARK   3   NON-BONDED CONTACTS OTHERS        (A):  4210 ; 0.199 ; 0.200
REMARK   3   NON-BONDED TORSION REFINED ATOMS  (A):  3041 ; 0.179 ; 0.200
REMARK   3   NON-BONDED TORSION OTHERS         (A):  3676 ; 0.087 ; 0.200
REMARK   3   H-BOND (X...Y) REFINED ATOMS      (A):   245 ; 0.159 ; 0.200
REMARK   3   H-BOND (X...Y) OTHERS             (A):  NULL ;  NULL ;  NULL
REMARK   3   POTENTIAL METAL-ION REFINED ATOMS (A):  NULL ;  NULL ;  NULL
REMARK   3   POTENTIAL METAL-ION OTHERS        (A):  NULL ;  NULL ;  NULL
REMARK   3   SYMMETRY VDW REFINED ATOMS        (A):     7 ; 0.185 ; 0.200
REMARK   3   SYMMETRY VDW OTHERS               (A):    30 ; 0.167 ; 0.200
REMARK   3   SYMMETRY H-BOND REFINED ATOMS     (A):     5 ; 0.189 ; 0.200
REMARK   3   SYMMETRY H-BOND OTHERS            (A):  NULL ;  NULL ;  NULL
REMARK   3   SYMMETRY METAL-ION REFINED ATOMS  (A):  NULL ;  NULL ;  NULL
REMARK   3   SYMMETRY METAL-ION OTHERS         (A):  NULL ;  NULL ;  NULL
REMARK   3
REMARK   3  ISOTROPIC THERMAL FACTOR RESTRAINTS.     COUNT   RMS    WEIGHT
REMARK   3   MAIN-CHAIN BOND REFINED ATOMS  (A**2):  5474 ; 0.659 ; 1.500
REMARK   3   MAIN-CHAIN BOND OTHER ATOMS    (A**2):  1741 ; 0.125 ; 1.500
REMARK   3   MAIN-CHAIN ANGLE REFINED ATOMS (A**2):  6943 ; 0.782 ; 2.000
REMARK   3   SIDE-CHAIN BOND REFINED ATOMS  (A**2):  2998 ; 1.339 ; 3.000
REMARK   3   SIDE-CHAIN ANGLE REFINED ATOMS (A**2):  2254 ; 1.913 ; 4.500
REMARK   3
REMARK   3 ANISOTROPIC THERMAL FACTOR RESTRAINTS.    COUNT   RMS   WEIGHT
REMARK   3   RIGID-BOND RESTRAINTS          (A**2):  NULL ;  NULL ;  NULL
REMARK   3   SPHERICITY; FREE ATOMS         (A**2):  NULL ;  NULL ;  NULL
REMARK   3   SPHERICITY; BONDED ATOMS       (A**2):  NULL ;  NULL ;  NULL
REMARK   3
REMARK   3  NCS RESTRAINTS STATISTICS
REMARK   3   NUMBER OF DIFFERENT NCS GROUPS : 5
REMARK   3
REMARK   3  NCS GROUP NUMBER               : 1
REMARK   3     CHAIN NAMES                    : L A
REMARK   3     NUMBER OF COMPONENTS NCS GROUP : 1
REMARK   3       COMPONENT C  SSSEQI  TO  C   SSSEQI   CODE
REMARK   3           1     L      1       L     107      6
REMARK   3           1     A      1       A     107      6
REMARK   3                   GROUP CHAIN        COUNT   RMS     WEIGHT
REMARK   3   LOOSE POSITIONAL   1         (A):   1461 ;  0.25 ;  5.00
REMARK   3   LOOSE THERMAL      1      (A**2):   1461 ;  0.73 ; 10.00
REMARK   3
REMARK   3  NCS GROUP NUMBER               : 2
REMARK   3     CHAIN NAMES                    : L A
REMARK   3     NUMBER OF COMPONENTS NCS GROUP : 1
REMARK   3       COMPONENT C  SSSEQI  TO  C   SSSEQI   CODE
REMARK   3           1     L    108       L     211      6
REMARK   3           1     A    108       A     211      6
REMARK   3                   GROUP CHAIN        COUNT   RMS     WEIGHT
REMARK   3   LOOSE POSITIONAL   2         (A):   1359 ;  0.23 ;  5.00
REMARK   3   LOOSE THERMAL      2      (A**2):   1359 ;  1.29 ; 10.00
REMARK   3
```

```
REMARK   3  NCS GROUP NUMBER               : 3
REMARK   3     CHAIN NAMES                    : H B
REMARK   3     NUMBER OF COMPONENTS NCS GROUP : 1
REMARK   3       COMPONENT C  SSSEQI  TO  C   SSSEQI   CODE
REMARK   3           1     H      1       H     113      6
REMARK   3           1     B      1       B     113      6
REMARK   3                   GROUP CHAIN        COUNT   RMS     WEIGHT
REMARK   3   LOOSE POSITIONAL   3          (A):   1547 ;  0.11 ;  5.00
REMARK   3   LOOSE THERMAL      3       (A**2):   1547 ;  1.00 ; 10.00
REMARK   3
REMARK   3  NCS GROUP NUMBER               : 4
REMARK   3     CHAIN NAMES                    : H B
REMARK   3     NUMBER OF COMPONENTS NCS GROUP : 1
REMARK   3       COMPONENT C  SSSEQI  TO  C   SSSEQI   CODE
REMARK   3           1     H    114       H     126      6
REMARK   3           1     B    114       B     126      6
REMARK   3                   GROUP CHAIN        COUNT   RMS     WEIGHT
REMARK   3   LOOSE POSITIONAL   4          (A):    155 ;  0.38 ;  5.00
REMARK   3   LOOSE THERMAL      4       (A**2):    155 ;  1.95 ; 10.00
REMARK   3
REMARK   3  NCS GROUP NUMBER               : 5
REMARK   3     CHAIN NAMES                    : H B
REMARK   3     NUMBER OF COMPONENTS NCS GROUP : 1
REMARK   3       COMPONENT C  SSSEQI  TO  C   SSSEQI   CODE
REMARK   3           1     H    136       H     227      6
REMARK   3           1     B    136       B     227      6
REMARK   3                   GROUP CHAIN        COUNT   RMS     WEIGHT
REMARK   3   LOOSE POSITIONAL   5          (A):    973 ;  0.23 ;  5.00
REMARK   3   LOOSE THERMAL      5       (A**2):    973 ;  0.82 ; 10.00
REMARK   3
REMARK   3  TLS DETAILS
REMARK   3   NUMBER OF TLS GROUPS  : 4
REMARK   3
REMARK   3   TLS GROUP : 1
REMARK   3    NUMBER OF COMPONENTS GROUP : 1
REMARK   3    COMPONENTS        C SSSEQI   TO  C SSSEQI
REMARK   3    RESIDUE RANGE :   A      1        A   221
REMARK   3    ORIGIN FOR THE GROUP (A):  38.5186   9.2498  17.0299
REMARK   3    T TENSOR
REMARK   3      T11:   0.2639 T22:   0.1856
REMARK   3      T33:   0.0412 T12:   0.0129
REMARK   3      T13:  -0.0229 T23:   0.0075
REMARK   3    L TENSOR
REMARK   3      L11:   1.2476 L22:  18.8186
REMARK   3      L33:   0.7358 L12:  -0.9182
REMARK   3      L13:  -0.4633 L23:  -2.8572
REMARK   3    S TENSOR
REMARK   3      S11:  -0.1230 S12:  -0.1350 S13:   0.1070
REMARK   3      S21:   0.1833 S22:   0.1989 S23:  -0.0673
REMARK   3      S31:   0.2988 S32:   0.3017 S33:  -0.0759
REMARK   3
REMARK   3   TLS GROUP : 2
REMARK   3    NUMBER OF COMPONENTS GROUP : 1
REMARK   3    COMPONENTS        C SSSEQI   TO  C SSSEQI
REMARK   3    RESIDUE RANGE :   B      1        B   227
REMARK   3    ORIGIN FOR THE GROUP (A):  18.6717  -2.2091  -2.3508
REMARK   3    T TENSOR
REMARK   3      T11:   0.3169 T22:   0.0830
REMARK   3      T33:   0.0521 T12:   0.0175
REMARK   3      T13:  -0.0382 T23:   0.0060
REMARK   3    L TENSOR
REMARK   3      L11:   2.8160 L22:   1.2951
REMARK   3      L33:   2.1804 L12:   0.8548
REMARK   3      L13:  -2.1037 L23:  -1.0227
```

```
REMARK   3    S TENSOR
REMARK   3      S11:   0.1656 S12:   0.1951 S13:   0.1602
REMARK   3      S21:  -0.3132 S22:   0.0276 S23:   0.3597
REMARK   3      S31:  -0.0658 S32:  -0.1993 S33:  -0.1933
REMARK   3
REMARK   3   TLS GROUP : 3
REMARK   3    NUMBER OF COMPONENTS GROUP : 1
REMARK   3    COMPONENTS        C SSSEQI   TO  C SSSEQI
REMARK   3    RESIDUE RANGE :   H     1         H   227
REMARK   3    ORIGIN FOR THE GROUP (A):  17.9538  -1.3383  43.2238
REMARK   3    T TENSOR
REMARK   3      T11:   0.3108 T22:   0.1076
REMARK   3      T33:   0.0689 T12:  -0.0386
REMARK   3      T13:   0.0466 T23:   0.0367
REMARK   3    L TENSOR
REMARK   3      L11:   0.7004 L22:   1.2871
REMARK   3      L33:   2.0590 L12:  -0.2673
REMARK   3      L13:   1.1558 L23:  -0.0172
REMARK   3    S TENSOR
REMARK   3      S11:   0.0754 S12:  -0.1826 S13:  -0.0304
REMARK   3      S21:   0.2142 S22:   0.1332 S23:   0.2512
REMARK   3      S31:   0.1977 S32:  -0.2560 S33:  -0.2086
REMARK   3
REMARK   3   TLS GROUP : 4
REMARK   3    NUMBER OF COMPONENTS GROUP : 1
REMARK   3    COMPONENTS        C SSSEQI   TO  C SSSEQI
REMARK   3    RESIDUE RANGE :   L     1         L   221
REMARK   3    ORIGIN FOR THE GROUP (A):  36.2584  -4.5702  24.8879
REMARK   3    T TENSOR
REMARK   3      T11:   0.2779 T22:   0.1946
REMARK   3      T33:   0.0079 T12:   0.0203
REMARK   3      T13:  -0.0113 T23:   0.0538
REMARK   3    L TENSOR
REMARK   3      L11:   2.9390 L22:  17.8781
REMARK   3      L33:   6.7012 L12:   4.8729
REMARK   3      L13:   1.9743 L23:   1.1500
REMARK   3    S TENSOR
REMARK   3      S11:  -0.0794 S12:   0.0598 S13:   0.3426
REMARK   3      S21:   0.2222 S22:   0.0581 S23:   0.7020
REMARK   3      S31:   0.0016 S32:   0.1934 S33:   0.0213
REMARK   3
REMARK   3  BULK SOLVENT MODELLING.
REMARK   3   METHOD USED : MASK
REMARK   3   PARAMETERS FOR MASK CALCULATION
REMARK   3   VDW PROBE RADIUS   : 1.20
REMARK   3   ION PROBE RADIUS   : 0.80
REMARK   3   SHRINKAGE RADIUS   : 0.80
REMARK   3
REMARK   3  OTHER REFINEMENT REMARKS: HYDROGENS HAVE BEEN ADDED IN THE
REMARK   3  RIDING POSITIONS
```

Template/example 2

```
REMARK   3
REMARK   3 REFINEMENT.
REMARK   3   PROGRAM     : REFMAC 5.5.0057
REMARK   3   AUTHORS     : MURSHUDOV,VAGIN,DODSON
REMARK   3
REMARK   3    REFINEMENT TARGET : MAXIMUM LIKELIHOOD
```

```
REMARK   3
REMARK   3  DATA USED IN REFINEMENT.
REMARK   3   RESOLUTION RANGE HIGH (ANGSTROMS) : 1.40
REMARK   3   RESOLUTION RANGE LOW  (ANGSTROMS) : 38.27
REMARK   3   DATA CUTOFF            (SIGMA(F)) : NULL
REMARK   3   COMPLETENESS FOR RANGE        (%) : 99.8
REMARK   3   NUMBER OF REFLECTIONS             : 102082
REMARK   3
REMARK   3  FIT TO DATA USED IN REFINEMENT.
REMARK   3   CROSS-VALIDATION METHOD          : THROUGHOUT
REMARK   3   FREE R VALUE TEST SET SELECTION  : RANDOM
REMARK   3   R VALUE     (WORKING + TEST SET) : 0.179
REMARK   3   R VALUE            (WORKING SET) : 0.179
REMARK   3   FREE R VALUE                     : 0.193
REMARK   3   FREE R VALUE TEST SET SIZE   (%) : 5.000
REMARK   3   FREE R VALUE TEST SET COUNT      : 5384
REMARK   3
REMARK   3  FIT IN THE HIGHEST RESOLUTION BIN.
REMARK   3   TOTAL NUMBER OF BINS USED           : 20
REMARK   3   BIN RESOLUTION RANGE HIGH       (A) : 1.40
REMARK   3   BIN RESOLUTION RANGE LOW        (A) : 1.44
REMARK   3   REFLECTION IN BIN     (WORKING SET) : 7804
REMARK   3   BIN COMPLETENESS (WORKING+TEST) (%) : 103.34
REMARK   3   BIN R VALUE           (WORKING SET) : 0.3830
REMARK   3   BIN FREE R VALUE SET COUNT          : 404
REMARK   3   BIN FREE R VALUE                    : 0.3790
REMARK   3
REMARK   3  NUMBER OF NON-HYDROGEN ATOMS USED IN REFINEMENT.
REMARK   3   PROTEIN ATOMS            : 1043
REMARK   3   NUCLEIC ACID ATOMS       : 0
REMARK   3   HETEROGEN ATOMS          : 0
REMARK   3   SOLVENT ATOMS            : 65
REMARK   3
REMARK   3  B VALUES.
REMARK   3   FROM WILSON PLOT           (A**2) : NULL
REMARK   3   MEAN B VALUE      (OVERALL, A**2) : 14.92
REMARK   3   OVERALL ANISOTROPIC B VALUE.
REMARK   3    B11 (A**2) : -7.82000
REMARK   3    B22 (A**2) : -6.91000
REMARK   3    B33 (A**2) : 14.73000
REMARK   3    B12 (A**2) : 0.00000
REMARK   3    B13 (A**2) : 7.83000
REMARK   3    B23 (A**2) : 0.00000
REMARK   3
REMARK   3  ESTIMATED OVERALL COORDINATE ERROR.
REMARK   3   ESU BASED ON R VALUE                            (A): 0.013
REMARK   3   ESU BASED ON FREE R VALUE                       (A): 0.012
REMARK   3   ESU BASED ON MAXIMUM LIKELIHOOD                 (A): 0.037
REMARK   3   ESU FOR B VALUES BASED ON MAXIMUM LIKELIHOOD (A**2): 1.010
REMARK   3
REMARK   3 CORRELATION COEFFICIENTS.
REMARK   3   CORRELATION COEFFICIENT FO-FC      : 0.969
REMARK   3   CORRELATION COEFFICIENT FO-FC FREE : 0.965
REMARK   3
REMARK   3  RMS DEVIATIONS FROM IDEAL VALUES        COUNT    RMS    WEIGHT
REMARK   3   BOND LENGTHS REFINED ATOMS        (A):  4192 ; 0.012 ; 0.022
REMARK   3   BOND LENGTHS OTHERS               (A):  2871 ; 0.001 ; 0.020
REMARK   3   BOND ANGLES REFINED ATOMS   (DEGREES):  5694 ; 1.405 ; 1.992
REMARK   3   BOND ANGLES OTHERS          (DEGREES):  7056 ; 0.928 ; 3.000
REMARK   3   TORSION ANGLES, PERIOD 1    (DEGREES):   558 ; 5.807 ; 5.000
REMARK   3   TORSION ANGLES, PERIOD 2    (DEGREES):   172 ;35.581 ;23.953
REMARK   3   TORSION ANGLES, PERIOD 3    (DEGREES):   773 ;12.566 ;15.000
REMARK   3   TORSION ANGLES, PERIOD 4    (DEGREES):    29 ;12.738 ;15.000
REMARK   3   CHIRAL-CENTER RESTRAINTS       (A**3):   659 ; 0.087 ; 0.200
```

```
REMARK   3   GENERAL PLANES REFINED ATOMS      (A):  4650 ; 0.006 ; 0.020
REMARK   3   GENERAL PLANES OTHERS             (A):   817 ; 0.001 ; 0.020
REMARK   3   NON-BONDED CONTACTS REFINED ATOMS (A):  NULL ;  NULL ;  NULL
REMARK   3   NON-BONDED CONTACTS OTHERS        (A):  NULL ;  NULL ;  NULL
REMARK   3   NON-BONDED TORSION REFINED ATOMS  (A):  NULL ;  NULL ;  NULL
REMARK   3   NON-BONDED TORSION OTHERS         (A):  NULL ;  NULL ;  NULL
REMARK   3   H-BOND (X...Y) REFINED ATOMS      (A):  NULL ;  NULL ;  NULL
REMARK   3   H-BOND (X...Y) OTHERS             (A):  NULL ;  NULL ;  NULL
REMARK   3   POTENTIAL METAL-ION REFINED ATOMS (A):  NULL ;  NULL ;  NULL
REMARK   3   POTENTIAL METAL-ION OTHERS        (A):  NULL ;  NULL ;  NULL
REMARK   3   SYMMETRY VDW REFINED ATOMS        (A):  NULL ;  NULL ;  NULL
REMARK   3   SYMMETRY VDW OTHERS               (A):  NULL ;  NULL ;  NULL
REMARK   3   SYMMETRY H-BOND REFINED ATOMS     (A):  NULL ;  NULL ;  NULL
REMARK   3   SYMMETRY H-BOND OTHERS            (A):  NULL ;  NULL ;  NULL
REMARK   3   SYMMETRY METAL-ION REFINED ATOMS  (A):  NULL ;  NULL ;  NULL
REMARK   3   SYMMETRY METAL-ION OTHERS         (A):  NULL ;  NULL ;  NULL
REMARK   3
REMARK   3  ISOTROPIC THERMAL FACTOR RESTRAINTS.     COUNT   RMS    WEIGHT
REMARK   3   MAIN-CHAIN BOND REFINED ATOMS  (A**2):  2610 ; 0.801 ; 1.500
REMARK   3   MAIN-CHAIN BOND OTHER ATOMS    (A**2):  1080 ; 0.208 ; 1.500
REMARK   3   MAIN-CHAIN ANGLE REFINED ATOMS (A**2):  4200 ; 1.425 ; 2.000
REMARK   3   SIDE-CHAIN BOND REFINED ATOMS  (A**2):  1582 ; 2.325 ; 3.000
REMARK   3   SIDE-CHAIN ANGLE REFINED ATOMS (A**2):  1470 ; 3.654 ; 4.500
REMARK   3
REMARK   3 ANISOTROPIC THERMAL FACTOR RESTRAINTS.    COUNT   RMS   WEIGHT
REMARK   3   RIGID-BOND RESTRAINTS          (A**2):  NULL ;  NULL ;  NULL
REMARK   3   SPHERICITY; FREE ATOMS         (A**2):  NULL ;  NULL ;  NULL
REMARK   3   SPHERICITY; BONDED ATOMS       (A**2):  NULL ;  NULL ;  NULL
REMARK   3
REMARK   3  NCS RESTRAINTS STATISTICS
REMARK   3   NUMBER OF DIFFERENT NCS GROUPS : 1
REMARK   3
REMARK   3  NCS GROUP NUMBER               : 1
REMARK   3     CHAIN NAMES                    : A B C D
REMARK   3     NUMBER OF COMPONENTS NCS GROUP : 1
REMARK   3       COMPONENT C  SSSEQI  TO  C   SSSEQI   CODE
REMARK   3           1     A      3       A     300      6
REMARK   3           1     B      3       B     300      6
REMARK   3           1     C      3       C     300      6
REMARK   3           1     D      3       D     300      6
REMARK   3                   GROUP CHAIN        COUNT   RMS     WEIGHT
REMARK   3   LOOSE POSITIONAL   1    A    (A):   1265 ; 0.730 ; 5.000
REMARK   3   LOOSE POSITIONAL   1    B    (A):   1265 ; 0.550 ; 5.000
REMARK   3   LOOSE POSITIONAL   1    C    (A):   1265 ; 0.670 ; 5.000
REMARK   3   LOOSE POSITIONAL   1    D    (A):   1265 ; 0.640 ; 5.000
REMARK   3   LOOSE THERMAL      1    A (A**2):   1265 ; 5.080 ;10.000
REMARK   3   LOOSE THERMAL      1    B (A**2):   1265 ; 2.980 ;10.000
REMARK   3   LOOSE THERMAL      1    C (A**2):   1265 ; 4.360 ;10.000
REMARK   3   LOOSE THERMAL      1    D (A**2):   1265 ; 3.660 ;10.000
REMARK   3
REMARK   3  TWIN DETAILS
REMARK   3   NUMBER OF TWIN DOMAINS  : 2
REMARK   3      TWIN DOMAIN   : 1
REMARK   3      TWIN OPERATOR : H,K,L
REMARK   3      TWIN FRACTION : 0.867
REMARK   3      TWIN DOMAIN   : 2
REMARK   3      TWIN OPERATOR : L,-K,H
REMARK   3      TWIN FRACTION : 0.133
REMARK   3
REMARK   3  TLS DETAILS
REMARK   3   NUMBER OF TLS GROUPS  : 4
REMARK   3
REMARK   3   TLS GROUP : 1
REMARK   3    NUMBER OF COMPONENTS GROUP : 1
```

```
REMARK   3    COMPONENTS        C SSSEQI   TO  C SSSEQI
REMARK   3    RESIDUE RANGE :   A     0        A   300
REMARK   3    ORIGIN FOR THE GROUP (A):  18.1228  50.5084  22.0082
REMARK   3    T TENSOR
REMARK   3      T11:  -0.1434 T22:  -0.0225
REMARK   3      T33:  -0.0349 T12:  -0.0039
REMARK   3      T13:  -0.0011 T23:  -0.0142
REMARK   3    L TENSOR
REMARK   3      L11:   0.2499 L22:   2.6544
REMARK   3      L33:   3.1537 L12:  -0.0787
REMARK   3      L13:  -0.1230 L23:  -0.1410
REMARK   3    S TENSOR
REMARK   3      S11:  -0.0939 S12:  -0.0135 S13:   0.0635
REMARK   3      S21:   0.4552 S22:   0.0720 S23:   0.3304
REMARK   3      S31:  -0.0916 S32:  -0.2462 S33:   0.0219
REMARK   3
REMARK   3   TLS GROUP : 2
REMARK   3    NUMBER OF COMPONENTS GROUP : 1
REMARK   3    COMPONENTS        C SSSEQI   TO  C SSSEQI
REMARK   3    RESIDUE RANGE :   B     2        B   300
REMARK   3    ORIGIN FOR THE GROUP (A):  26.3124  33.3641  14.1724
REMARK   3    T TENSOR
REMARK   3      T11:   0.0332 T22:  -0.0920
REMARK   3      T33:  -0.0634 T12:  -0.0201
REMARK   3      T13:   0.0037 T23:   0.0130
REMARK   3    L TENSOR
REMARK   3      L11:   0.8985 L22:   2.2480
REMARK   3      L33:   0.4623 L12:  -1.0984
REMARK   3      L13:  -0.0114 L23:   0.6608
REMARK   3    S TENSOR
REMARK   3      S11:  -0.0856 S12:  -0.0481 S13:  -0.1305
REMARK   3      S21:   0.2230 S22:   0.0179 S23:   0.0879
REMARK   3      S31:   0.1878 S32:   0.0617 S33:   0.0677
REMARK   3
REMARK   3   TLS GROUP : 3
REMARK   3    NUMBER OF COMPONENTS GROUP : 1
REMARK   3    COMPONENTS        C SSSEQI   TO  C SSSEQI
REMARK   3    RESIDUE RANGE :   C     3        C   300
REMARK   3    ORIGIN FOR THE GROUP (A):  31.0030  33.2958  50.0967
REMARK   3    T TENSOR
REMARK   3      T11:  -0.1785 T22:  -0.0337
REMARK   3      T33:  -0.0199 T12:   0.0202
REMARK   3      T13:  -0.0342 T23:  -0.0065
REMARK   3    L TENSOR
REMARK   3      L11:   1.1097 L22:   4.1071
REMARK   3      L33:   2.4647 L12:   0.0878
REMARK   3      L13:   0.3839 L23:   0.9425
REMARK   3    S TENSOR
REMARK   3      S11:  -0.0357 S12:   0.1026 S13:  -0.1055
REMARK   3      S21:  -0.2843 S22:   0.0934 S23:   0.3463
REMARK   3      S31:   0.4808 S32:  -0.0374 S33:  -0.0577
REMARK   3
REMARK   3   TLS GROUP : 4
REMARK   3    NUMBER OF COMPONENTS GROUP : 1
REMARK   3    COMPONENTS        C SSSEQI   TO  C SSSEQI
REMARK   3    RESIDUE RANGE :   D     2        D   300
REMARK   3    ORIGIN FOR THE GROUP (A):  36.9044  51.7770  58.1371
REMARK   3    T TENSOR
REMARK   3      T11:   0.0198 T22:  -0.0873
REMARK   3      T33:  -0.0907 T12:  -0.0053
REMARK   3      T13:   0.0045 T23:   0.0206
REMARK   3    L TENSOR
REMARK   3      L11:   0.6326 L22:   2.2127
REMARK   3      L33:   1.0826 L12:   0.2556
```

```
REMARK   3      L13:   0.1927 L23:   0.6195
REMARK   3    S TENSOR
REMARK   3      S11:  -0.0674 S12:   0.0446 S13:   0.0174
REMARK   3      S21:  -0.1848 S22:   0.0310 S23:   0.0198
REMARK   3      S31:  -0.2108 S32:   0.0752 S33:   0.0364
REMARK   3
REMARK   3  BULK SOLVENT MODELLING.
REMARK   3   METHOD USED : BABINET MODEL WITH MASK
REMARK   3   PARAMETERS FOR MASK CALCULATION
REMARK   3   VDW PROBE RADIUS   : 1.40
REMARK   3   ION PROBE RADIUS   : 0.80
REMARK   3   SHRINKAGE RADIUS   : 0.80
REMARK   3
REMARK   3  OTHER REFINEMENT REMARKS: HYDROGENS HAVE BEEN ADDED IN THE
REMARK   3  RIDING POSITIONS
```

Refinement using NUCLSQ

Template

```
REMARK   3
REMARK   3 REFINEMENT.
REMARK   3   PROGRAM     : NUCLSQ
REMARK   3   AUTHORS     : WESTHOF,DUMAS,MORAS
REMARK   3
REMARK   3  DATA USED IN REFINEMENT.
REMARK   3   RESOLUTION RANGE HIGH (ANGSTROMS) :
REMARK   3   RESOLUTION RANGE LOW  (ANGSTROMS) :
REMARK   3   DATA CUTOFF            (SIGMA(F)) :
REMARK   3   COMPLETENESS FOR RANGE        (%) :
REMARK   3   NUMBER OF REFLECTIONS             :
REMARK   3
REMARK   3  FIT TO DATA USED IN REFINEMENT.
REMARK   3   CROSS-VALIDATION METHOD          :
REMARK   3   FREE R VALUE TEST SET SELECTION  :
REMARK   3   R VALUE     (WORKING + TEST SET) :
REMARK   3   R VALUE            (WORKING SET) :
REMARK   3   FREE R VALUE                     :
REMARK   3   FREE R VALUE TEST SET SIZE   (%) :
REMARK   3   FREE R VALUE TEST SET COUNT      :
REMARK   3
REMARK   3  FIT/AGREEMENT OF MODEL WITH ALL DATA.
REMARK   3   R VALUE   (WORKING + TEST SET, NO CUTOFF) :
REMARK   3   R VALUE          (WORKING SET, NO CUTOFF) :
REMARK   3   FREE R VALUE                  (NO CUTOFF) :
REMARK   3   FREE R VALUE TEST SET SIZE (%, NO CUTOFF) :
REMARK   3   FREE R VALUE TEST SET COUNT   (NO CUTOFF) :
REMARK   3   TOTAL NUMBER OF REFLECTIONS   (NO CUTOFF) :
REMARK   3
REMARK   3  NUMBER OF NON-HYDROGEN ATOMS USED IN REFINEMENT.
REMARK   3   PROTEIN ATOMS            :
REMARK   3   NUCLEIC ACID ATOMS       :
REMARK   3   HETEROGEN ATOMS          :
REMARK   3   SOLVENT ATOMS            :
REMARK   3
REMARK   3  B VALUES.
REMARK   3   FROM WILSON PLOT           (A**2) :
REMARK   3   MEAN B VALUE      (OVERALL, A**2) :
REMARK   3   OVERALL ANISOTROPIC B VALUE.
REMARK   3    B11 (A**2) :
REMARK   3    B22 (A**2) :
REMARK   3    B33 (A**2) :
REMARK   3    B12 (A**2) :
REMARK   3    B13 (A**2) :
REMARK   3    B23 (A**2) :
REMARK   3
REMARK   3  ESTIMATED COORDINATE ERROR.
REMARK   3   ESD FROM LUZZATI PLOT        (A) :
REMARK   3   ESD FROM SIGMAA              (A) :
REMARK   3   LOW RESOLUTION CUTOFF        (A) :
REMARK   3
REMARK   3  RMS DEVIATIONS FROM IDEAL VALUES.
REMARK   3   DISTANCE RESTRAINTS.                    RMS     SIGMA
REMARK   3    SUGAR-BASE BOND DISTANCE        (A) :       ;
REMARK   3    SUGAR-BASE BOND ANGLE DISTANCE  (A) :       ;
REMARK   3    PHOSPHATE BONDS DISTANCE        (A) :       ;
REMARK   3    PHOSPHATE BOND ANGLE, H-BOND    (A) :       ;
REMARK   3
```

```
REMARK   3   PLANE RESTRAINT                  (A) :       ;
REMARK   3   CHIRAL-CENTER RESTRAINT       (A**3) :       ;
REMARK   3
REMARK   3   NON-BONDED CONTACT RESTRAINTS.
REMARK   3    SINGLE TORSION CONTACT          (A) :       ;
REMARK   3    MULTIPLE TORSION CONTACT        (A) :       ;
REMARK   3
REMARK   3  ISOTROPIC THERMAL FACTOR RESTRAINTS.    RMS    SIGMA
REMARK   3   SUGAR-BASE BONDS             (A**2) :       ;
REMARK   3   SUGAR-BASE ANGLES            (A**2) :       ;
REMARK   3   PHOSPHATE BONDS              (A**2) :       ;
REMARK   3   PHOSPHATE BOND ANGLE, H-BOND (A**2) :       ;
REMARK   3
REMARK   3  OTHER REFINEMENT REMARKS:
```

Refinement using CCP4, PROFFT, GPRLSA, and related programs

Template/example

```
REMARK   3
REMARK   3 REFINEMENT.
REMARK   3   PROGRAM     : PROFFT
REMARK   3   AUTHORS     : KONNERT,HENDRICKSON,FINZEL
REMARK   3
REMARK   3  DATA USED IN REFINEMENT.
REMARK   3   RESOLUTION RANGE HIGH (ANGSTROMS) : 1.65
REMARK   3   RESOLUTION RANGE LOW  (ANGSTROMS) : 5.00
REMARK   3   DATA CUTOFF            (SIGMA(F)) : 2.000
REMARK   3   COMPLETENESS FOR RANGE        (%) : NULL
REMARK   3   NUMBER OF REFLECTIONS             : 10699
REMARK   3
REMARK   3  FIT TO DATA USED IN REFINEMENT.
REMARK   3   CROSS-VALIDATION METHOD          : NONE
REMARK   3   FREE R VALUE TEST SET SELECTION  : NULL
REMARK   3   R VALUE     (WORKING + TEST SET) : 0.180
REMARK   3   R VALUE            (WORKING SET) : NULL
REMARK   3   FREE R VALUE                     : NULL
REMARK   3   FREE R VALUE TEST SET SIZE   (%) : NULL
REMARK   3   FREE R VALUE TEST SET COUNT      : NULL
REMARK   3
REMARK   3  FIT/AGREEMENT OF MODEL WITH ALL DATA.
REMARK   3   R VALUE   (WORKING + TEST SET, NO CUTOFF) : NULL
REMARK   3   R VALUE          (WORKING SET, NO CUTOFF) : NULL
REMARK   3   FREE R VALUE                  (NO CUTOFF) : NULL
REMARK   3   FREE R VALUE TEST SET SIZE (%, NO CUTOFF) : NULL
REMARK   3   FREE R VALUE TEST SET COUNT   (NO CUTOFF) : NULL
REMARK   3   TOTAL NUMBER OF REFLECTIONS   (NO CUTOFF) : NULL
REMARK   3
REMARK   3  NUMBER OF NON-HYDROGEN ATOMS USED IN REFINEMENT.
REMARK   3   PROTEIN ATOMS            : 843
REMARK   3   NUCLEIC ACID ATOMS       : 0
REMARK   3   HETEROGEN ATOMS          : 6
REMARK   3   SOLVENT ATOMS            : 85
REMARK   3
REMARK   3  B VALUES.
REMARK   3   FROM WILSON PLOT           (A**2) : NULL
REMARK   3   MEAN B VALUE      (OVERALL, A**2) : 17.28
REMARK   3   OVERALL ANISOTROPIC B VALUE.
REMARK   3    B11 (A**2) : NULL
REMARK   3    B22 (A**2) : NULL
REMARK   3    B33 (A**2) : NULL
REMARK   3    B12 (A**2) : NULL
REMARK   3    B13 (A**2) : NULL
REMARK   3    B23 (A**2) : NULL
REMARK   3
REMARK   3  ESTIMATED COORDINATE ERROR.
REMARK   3   ESD FROM LUZZATI PLOT        (A) : 0.20
REMARK   3   ESD FROM SIGMAA              (A) : NULL
REMARK   3   LOW RESOLUTION CUTOFF        (A) : NULL
REMARK   3
REMARK   3  RMS DEVIATIONS FROM IDEAL VALUES.
REMARK   3   DISTANCE RESTRAINTS.                    RMS    SIGMA
REMARK   3    BOND LENGTH                     (A) : 0.012 ; NULL
REMARK   3    ANGLE DISTANCE                  (A) : NULL  ; NULL
REMARK   3    INTRAPLANAR 1-4 DISTANCE        (A) : NULL  ; NULL
REMARK   3    H-BOND OR METAL COORDINATION    (A) : NULL  ; NULL
REMARK   3
REMARK   3   PLANE RESTRAINT                  (A) : NULL  ; NULL
```

```
REMARK   3   CHIRAL-CENTER RESTRAINT       (A**3) : NULL  ; NULL
REMARK   3
REMARK   3   NON-BONDED CONTACT RESTRAINTS.
REMARK   3    SINGLE TORSION                  (A) : NULL  ; NULL
REMARK   3    MULTIPLE TORSION                (A) : NULL  ; NULL
REMARK   3    H-BOND (X...Y)                  (A) : NULL  ; NULL
REMARK   3    H-BOND (X-H...Y)                (A) : NULL  ; NULL
REMARK   3
REMARK   3   CONFORMATIONAL TORSION ANGLE RESTRAINTS.
REMARK   3    SPECIFIED                 (DEGREES) : NULL  ; NULL
REMARK   3    PLANAR                    (DEGREES) : NULL  ; NULL
REMARK   3    STAGGERED                 (DEGREES) : NULL  ; NULL
REMARK   3    TRANSVERSE                (DEGREES) : NULL  ; NULL
REMARK   3
REMARK   3  ISOTROPIC THERMAL FACTOR RESTRAINTS.    RMS    SIGMA
REMARK   3   MAIN-CHAIN BOND               (A**2) : NULL  ; NULL
REMARK   3   MAIN-CHAIN ANGLE              (A**2) : NULL  ; NULL
REMARK   3   SIDE-CHAIN BOND               (A**2) : NULL  ; NULL
REMARK   3   SIDE-CHAIN ANGLE              (A**2) : NULL  ; NULL
REMARK   3
REMARK   3  OTHER REFINEMENT REMARKS: NULL
```

Refinement using SHELXL

Template

```
REMARK   3
REMARK   3   PROGRAM     : SHELXL-97
REMARK   3   AUTHORS     : G.M.SHELDRICK
REMARK   3
REMARK   3  DATA USED IN REFINEMENT.
REMARK   3   RESOLUTION RANGE HIGH (ANGSTROMS) : 1.15
REMARK   3   RESOLUTION RANGE LOW  (ANGSTROMS) : 30.00
REMARK   3   DATA CUTOFF            (SIGMA(F)) : 0.000
REMARK   3   COMPLETENESS FOR RANGE        (%) : 99.8
REMARK   3   CROSS-VALIDATION METHOD           : FREE R
REMARK   3   FREE R VALUE TEST SET SELECTION   : RANDOM
REMARK   3
REMARK   3  FIT TO DATA USED IN REFINEMENT (NO CUTOFF).
REMARK   3   R VALUE   (WORKING + TEST SET, NO CUTOFF) : 0.116
REMARK   3   R VALUE          (WORKING SET, NO CUTOFF) : 0.116
REMARK   3   FREE R VALUE                 (NO CUTOFF) : 0.145
REMARK   3   FREE R VALUE TEST SET SIZE (%, NO CUTOFF) : 5.000
REMARK   3   FREE R VALUE TEST SET COUNT   (NO CUTOFF) : 4279
REMARK   3   TOTAL NUMBER OF REFLECTIONS   (NO CUTOFF) : 85756
REMARK   3
REMARK   3  FIT/AGREEMENT OF MODEL FOR DATA WITH F>4SIG(F).
REMARK   3   R VALUE   (WORKING + TEST SET, F>4SIG(F)) : 0.010
REMARK   3   R VALUE          (WORKING SET, F>4SIG(F)) : 0.010
REMARK   3   FREE R VALUE                 (F>4SIG(F)) : 0.136
REMARK   3   FREE R VALUE TEST SET SIZE (%, F>4SIG(F)) : 5.000
REMARK   3   FREE R VALUE TEST SET COUNT   (F>4SIG(F)) : 3859
REMARK   3   TOTAL NUMBER OF REFLECTIONS   (F>4SIG(F)) : 77074
REMARK   3
REMARK   3  NUMBER OF NON-HYDROGEN ATOMS USED IN REFINEMENT.
REMARK   3   PROTEIN ATOMS      : 1439
REMARK   3   NUCLEIC ACID ATOMS : 0
REMARK   3   HETEROGEN ATOMS    : 0
REMARK   3   SOLVENT ATOMS      : 288
REMARK   3
REMARK   3  MODEL REFINEMENT.
REMARK   3   OCCUPANCY SUM OF NON-HYDROGEN ATOMS      : 1638.00
REMARK   3   OCCUPANCY SUM OF HYDROGEN ATOMS          : 1406.00
REMARK   3   NUMBER OF DISCRETELY DISORDERED RESIDUES : 7
REMARK   3   NUMBER OF LEAST-SQUARES PARAMETERS       : 15553
REMARK   3   NUMBER OF RESTRAINTS                     : 19134
REMARK   3
REMARK   3  RMS DEVIATIONS FROM RESTRAINT TARGET VALUES.
REMARK   3   BOND LENGTHS                         (A) : 0.015
REMARK   3   ANGLE DISTANCES                      (A) : 0.030
REMARK   3   SIMILAR DISTANCES (NO TARGET VALUES) (A) : 0.000
REMARK   3   DISTANCES FROM RESTRAINT PLANES      (A) : 0.030
REMARK   3   ZERO CHIRAL VOLUMES               (A**3) : 0.086
REMARK   3   NON-ZERO CHIRAL VOLUMES           (A**3) : 0.095
REMARK   3   ANTI-BUMPING DISTANCE RESTRAINTS     (A) : 0.032
REMARK   3   RIGID-BOND ADP COMPONENTS         (A**2) : 0.005
REMARK   3   SIMILAR ADP COMPONENTS            (A**2) : 0.048
REMARK   3   APPROXIMATELY ISOTROPIC ADPS      (A**2) : 0.107
REMARK   3
REMARK   3  BULK SOLVENT MODELING.
REMARK   3   METHOD USED: MOEWS & KRETSINGER, J.MOL.BIOL.91(1973)201-228
REMARK   3
REMARK   3  STEREOCHEMISTRY TARGET VALUES : ENGH & HUBER
REMARK   3   SPECIAL CASE: NULL
```

```
REMARK   3
REMARK   3  OTHER REFINEMENT REMARKS: NULL
```

Refinement using TNT/BUSTER

Template/example 1

```
REMARK   3
REMARK   3 REFINEMENT.
REMARK   3   PROGRAM     : TNT
REMARK   3   AUTHORS     : TRONRUD,TEN EYCK,MATTHEWS
REMARK   3
REMARK   3  DATA USED IN REFINEMENT.
REMARK   3   RESOLUTION RANGE HIGH (ANGSTROMS) : 1.60
REMARK   3   RESOLUTION RANGE LOW  (ANGSTROMS) : 30.00
REMARK   3   DATA CUTOFF            (SIGMA(F)) : 0.000
REMARK   3   COMPLETENESS FOR RANGE        (%) : 93.6
REMARK   3   NUMBER OF REFLECTIONS             : 80952
REMARK   3
REMARK   3  USING DATA ABOVE SIGMA CUTOFF.
REMARK   3   CROSS-VALIDATION METHOD          : THROUGHOUT
REMARK   3   FREE R VALUE TEST SET SELECTION  : RANDOM
REMARK   3   R VALUE     (WORKING + TEST SET) : 0.160
REMARK   3   R VALUE            (WORKING SET) : 0.158
REMARK   3   FREE R VALUE                     : 0.215
REMARK   3   FREE R VALUE TEST SET SIZE   (%) : NULL
REMARK   3   FREE R VALUE TEST SET COUNT      : 6164
REMARK   3
REMARK   3  USING ALL DATA, NO SIGMA CUTOFF.
REMARK   3   R VALUE   (WORKING + TEST SET, NO CUTOFF) : 0.1600
REMARK   3   R VALUE          (WORKING SET, NO CUTOFF) : NULL
REMARK   3   FREE R VALUE                  (NO CUTOFF) : NULL
REMARK   3   FREE R VALUE TEST SET SIZE (%, NO CUTOFF) : NULL
REMARK   3   FREE R VALUE TEST SET COUNT   (NO CUTOFF) : NULL
REMARK   3   TOTAL NUMBER OF REFLECTIONS   (NO CUTOFF) : 80952
REMARK   3
REMARK   3  NUMBER OF NON-HYDROGEN ATOMS USED IN REFINEMENT.
REMARK   3   PROTEIN ATOMS            : 5555
REMARK   3   NUCLEIC ACID ATOMS       : 0
REMARK   3   HETEROGEN ATOMS          : 69
REMARK   3   SOLVENT ATOMS            : 681
REMARK   3
REMARK   3  WILSON B VALUE (FROM FCALC, A**2) : NULL
REMARK   3
REMARK   3  RMS DEVIATIONS FROM IDEAL VALUES.    RMS    WEIGHT  COUNT
REMARK   3   BOND LENGTHS                 (A) : NULL  ; NULL  ; NULL
REMARK   3   BOND ANGLES            (DEGREES) : NULL  ; NULL  ; NULL
REMARK   3   TORSION ANGLES         (DEGREES) : NULL  ; NULL  ; NULL
REMARK   3   PSEUDOROTATION ANGLES  (DEGREES) : NULL  ; NULL  ; NULL
REMARK   3   TRIGONAL CARBON PLANES       (A) : NULL  ; NULL  ; NULL
REMARK   3   GENERAL PLANES               (A) : NULL  ; NULL  ; NULL
REMARK   3   ISOTROPIC THERMAL FACTORS (A**2) : NULL  ; NULL  ; NULL
REMARK   3   NON-BONDED CONTACTS          (A) : NULL  ; NULL  ; NULL
REMARK   3
REMARK   3  INCORRECT CHIRAL-CENTERS (COUNT) : NULL
REMARK   3
REMARK   3  BULK SOLVENT MODELING.
REMARK   3   METHOD USED : NULL
REMARK   3   KSOL        : NULL
REMARK   3   BSOL        : NULL
REMARK   3
REMARK   3  RESTRAINT LIBRARIES.
REMARK   3   STEREOCHEMISTRY : ENGH & HUBER
REMARK   3   ISOTROPIC THERMAL FACTOR RESTRAINTS : NULL
REMARK   3
```

```
REMARK   3  OTHER REFINEMENT REMARKS: NULL
```

Template/example 2

```
REMARK   3
REMARK   3 REFINEMENT.
REMARK   3   PROGRAM     : BUSTER-TNT 2.1.1
REMARK   3   AUTHORS     : BLANC,ROVERSI,VONRHEIN,BRICOGNE,TRONRUD,
REMARK   3               : TEN EYCK,MATTHEWS
REMARK   3
REMARK   3  DATA USED IN REFINEMENT.
REMARK   3   RESOLUTION RANGE HIGH (ANGSTROMS) : 1.75
REMARK   3   RESOLUTION RANGE LOW  (ANGSTROMS) : 34.65
REMARK   3   DATA CUTOFF            (SIGMA(F)) : 0.000
REMARK   3   COMPLETENESS FOR RANGE        (%) : 97.4
REMARK   3   NUMBER OF REFLECTIONS             : 53863
REMARK   3
REMARK   3  USING DATA ABOVE SIGMA CUTOFF.
REMARK   3   CROSS-VALIDATION METHOD          : THROUGHOUT
REMARK   3   FREE R VALUE TEST SET SELECTION  : RANDOM
REMARK   3   R VALUE     (WORKING + TEST SET) : 0.182
REMARK   3   R VALUE            (WORKING SET) : 0.180
REMARK   3   FREE R VALUE                     : 0.218
REMARK   3   FREE R VALUE TEST SET SIZE   (%) : 5.000
REMARK   3   FREE R VALUE TEST SET COUNT      : 2691
REMARK   3
REMARK   3  FIT IN THE HIGHEST RESOLUTION BIN.
REMARK   3   TOTAL NUMBER OF BINS USED               : 9
REMARK   3   BIN RESOLUTION RANGE HIGH   (ANGSTROMS) : 1.75
REMARK   3   BIN RESOLUTION RANGE LOW    (ANGSTROMS) : 1.86
REMARK   3   BIN COMPLETENESS (WORKING+TEST)     (%) : 97.43
REMARK   3   REFLECTIONS IN BIN (WORKING + TEST SET) : 7826
REMARK   3   BIN R VALUE        (WORKING + TEST SET) : 0.2400
REMARK   3   REFLECTIONS IN BIN        (WORKING SET) : 7447
REMARK   3   BIN R VALUE               (WORKING SET) : 0.2376
REMARK   3   BIN FREE R VALUE                        : 0.2873
REMARK   3   BIN FREE R VALUE TEST SET SIZE      (%) : 4.84
REMARK   3   BIN FREE R VALUE TEST SET COUNT         : 379
REMARK   3   ESTIMATED ERROR OF FREE R VALUE         : NULL
REMARK   3
REMARK   3  NUMBER OF NON-HYDROGEN ATOMS USED IN REFINEMENT.
REMARK   3   PROTEIN ATOMS            : 3944
REMARK   3   NUCLEIC ACID ATOMS       : 0
REMARK   3   HETEROGEN ATOMS          : 44
REMARK   3   SOLVENT ATOMS            : 257
REMARK   3
REMARK   3  B VALUES.
REMARK   3   FROM WILSON PLOT           (A**2) : 26.46
REMARK   3   MEAN B VALUE      (OVERALL, A**2) : 28.30
REMARK   3   OVERALL ANISOTROPIC B VALUE.
REMARK   3    B11 (A**2) : 0.83777
REMARK   3    B22 (A**2) : 6.21317
REMARK   3    B33 (A**2) : -7.05095
REMARK   3    B12 (A**2) : 0.00000
REMARK   3    B13 (A**2) : -7.10165
REMARK   3    B23 (A**2) : 0.00000
REMARK   3
REMARK   3  ESTIMATED COORDINATE ERROR.
REMARK   3   ESD FROM LUZZATI PLOT        (A) : NULL
REMARK   3
REMARK   3 CORRELATION COEFFICIENTS.
REMARK   3   CORRELATION COEFFICIENT FO-FC      : NULL
REMARK   3   CORRELATION COEFFICIENT FO-FC FREE : NULL
REMARK   3
```

```
REMARK   3  RMS DEVIATIONS FROM IDEAL VALUES        RMS    WEIGHT   COUNT
REMARK   3   BOND LENGTHS                 (A) : 0.011  ; 2.000 ; 4060
REMARK   3   BOND ANGLES            (DEGREES) : 1.186  ; 2.000 ; 5477
REMARK   3   TORSION ANGLES         (DEGREES) : 16.973 ; 0.000 ; 827
REMARK   3   PSEUDOROTATION ANGLES  (DEGREES) : NULL   ; NULL  ; NULL
REMARK   3   TRIGONAL CARBON PLANES       (A) : 0.012  ; 2.000 ; 101
REMARK   3   GENERAL PLANES               (A) : 0.018  ; 5.000 ; 608
REMARK   3   ISOTROPIC THERMAL FACTORS (A**2) : 1.634  ; 20.000; 4060
REMARK   3   NON-BONDED CONTACTS          (A) : 0.044  ; 5.000 ; 34
REMARK   3
REMARK   3  INCORRECT CHIRAL-CENTERS (COUNT) : NULL
REMARK   3
REMARK   3  OTHER REFINEMENT REMARKS: NULL
```

Refinement using PHENIX

Template/example 1

```
REMARK   3
REMARK   3 REFINEMENT.
REMARK   3   PROGRAM     : PHENIX (PHENIX.REFINE)
REMARK   3   AUTHORS     : PAUL ADAMS,PAVEL AFONINE,VICENT CHEN,IAN
REMARK   3               : DAVIS,KRESHNA GOPAL,RALF GROSSE-
REMARK   3               : KUNSTLEVE,LI-WEI HUNG,ROBERT IMMORMINO,
REMARK   3               : TOM IOERGER,AIRLIE MCCOY,ERIK MCKEE,NIGEL
REMARK   3               : MORIARTY,REETAL PAI,RANDY READ,JANE
REMARK   3               : RICHARDSON,DAVID RICHARDSON,TOD ROMO,JIM
REMARK   3               : SACCHETTINI,NICHOLAS SAUTER,JACOB SMITH,
REMARK   3               : LAURENT STORONI,TOM TERWILLIGER,PETER
REMARK   3               : ZWART
REMARK   3
REMARK   3    REFINEMENT TARGET : ML
REMARK   3
REMARK   3  DATA USED IN REFINEMENT.
REMARK   3   RESOLUTION RANGE HIGH (ANGSTROMS) : 2.99
REMARK   3   RESOLUTION RANGE LOW  (ANGSTROMS) : 40.07
REMARK   3   MIN(FOBS/SIGMA_FOBS)              : 0.000
REMARK   3   COMPLETENESS FOR RANGE        (%) : 96.7
REMARK   3   NUMBER OF REFLECTIONS             : 242645
REMARK   3
REMARK   3  FIT TO DATA USED IN REFINEMENT.
REMARK   3   R VALUE     (WORKING + TEST SET) : 0.293
REMARK   3   R VALUE            (WORKING SET) : 0.291
REMARK   3   FREE R VALUE                     : 0.335
REMARK   3   FREE R VALUE TEST SET SIZE   (%) : 4.980
REMARK   3   FREE R VALUE TEST SET COUNT      : 12081
REMARK   3
REMARK   3  FIT TO DATA USED IN REFINEMENT (IN BINS).
REMARK   3   BIN  RESOLUTION RANGE  COMPL.    NWORK NFREE   RWORK  RFREE
REMARK   3     1 40.0700 -  9.2600    0.98     8197   419  0.1970 0.2050
REMARK   3     2  9.2600 -  7.3700    0.98     7994   409  0.1560 0.1990
REMARK   3     3  7.3700 -  6.4400    0.99     7965   413  0.2060 0.2470
REMARK   3     4  6.4400 -  5.8500    0.99     7924   426  0.2330 0.2740
REMARK   3     5  5.8500 -  5.4300    0.98     7833   444  0.2550 0.3160
REMARK   3     6  5.4300 -  5.1200    0.98     7811   408  0.2530 0.3110
REMARK   3     7  5.1200 -  4.8600    0.97     7819   387  0.2550 0.3210
REMARK   3     8  4.8600 -  4.6500    0.97     7693   423  0.2690 0.3260
REMARK   3     9  4.6500 -  4.4700    0.97     7737   394  0.2790 0.2920
REMARK   3    10  4.4700 -  4.3200    0.97     7691   403  0.2690 0.3280
REMARK   3    11  4.3200 -  4.1800    0.97     7731   402  0.2560 0.3040
REMARK   3    12  4.1800 -  4.0600    0.98     7760   407  0.2610 0.3170
REMARK   3    13  4.0600 -  3.9500    0.97     7685   398  0.2710 0.3070
REMARK   3    14  3.9500 -  3.8600    0.98     7758   403  0.2970 0.3650
REMARK   3    15  3.8600 -  3.7700    0.98     7713   431  0.2890 0.3260
REMARK   3    16  3.7700 -  3.6900    0.98     7737   386  0.2870 0.3520
REMARK   3    17  3.6900 -  3.6200    0.98     7719   410  0.2910 0.3230
REMARK   3    18  3.6200 -  3.5500    0.98     7683   426  0.2770 0.3200
REMARK   3    19  3.5500 -  3.4800    0.98     7756   375  0.2950 0.3480
REMARK   3    20  3.4800 -  3.4300    0.98     7720   414  0.3110 0.3780
REMARK   3    21  3.4300 -  3.3700    0.98     7742   372  0.3200 0.3760
REMARK   3    22  3.3700 -  3.3200    0.98     7667   411  0.3440 0.4360
REMARK   3    23  3.3200 -  3.2700    0.98     7700   414  0.3410 0.3840
REMARK   3    24  3.2700 -  3.2200    0.97     7667   411  0.3350 0.3870
REMARK   3    25  3.2200 -  3.1800    0.97     7541   419  0.3400 0.3790
REMARK   3    26  3.1800 -  3.1400    0.96     7637   402  0.3460 0.4220
REMARK   3    27  3.1400 -  3.1000    0.96     7613   381  0.3580 0.3940
```

```
REMARK   3    28  3.1000 -  3.0600    0.96      7538   427  0.3790 0.4290
REMARK   3    29  3.0600 -  3.0300    0.95      7440   376  0.3760 0.4350
REMARK   3    30  3.0300 -  2.9900    0.77      6093   290  0.3950 0.4490
REMARK   3
REMARK   3  BULK SOLVENT MODELLING.
REMARK   3   METHOD USED        : FLAT BULK SOLVENT MODEL
REMARK   3   SOLVENT RADIUS     : 1.11
REMARK   3   SHRINKAGE RADIUS   : 0.90
REMARK   3   K_SOL              : 0.30
REMARK   3   B_SOL              : 56.99
REMARK   3
REMARK   3  ERROR ESTIMATES.
REMARK   3   COORDINATE ERROR (MAXIMUM-LIKELIHOOD BASED)     : 0.510
REMARK   3   PHASE ERROR (DEGREES, MAXIMUM-LIKELIHOOD BASED) : 36.180
REMARK   3
REMARK   3  B VALUES.
REMARK   3   FROM WILSON PLOT           (A**2) : 50.24
REMARK   3   MEAN B VALUE      (OVERALL, A**2) : 62.67
REMARK   3   OVERALL ANISOTROPIC B VALUE.
REMARK   3    B11 (A**2) : -12.34000
REMARK   3    B22 (A**2) : -11.49000
REMARK   3    B33 (A**2) : 23.84000
REMARK   3    B12 (A**2) : 0.00000
REMARK   3    B13 (A**2) : 0.00000
REMARK   3    B23 (A**2) : 0.00000
REMARK   3
REMARK   3  TWINNING INFORMATION.
REMARK   3   FRACTION: 0.2950
REMARK   3   OPERATOR: -H,-K,L
REMARK   3
REMARK   3  DEVIATIONS FROM IDEAL VALUES.
REMARK   3                RMSD          COUNT
REMARK   3   BOND      :  0.059          59703
REMARK   3   ANGLE     :  3.995          80640
REMARK   3   CHIRALITY :  0.243           9800
REMARK   3   PLANARITY :  0.013          10535
REMARK   3   DIHEDRAL  : 24.960          22449
REMARK   3
REMARK   3  TLS DETAILS
REMARK   3   NUMBER OF TLS GROUPS  : 4
REMARK   3   TLS GROUP : 1
REMARK   3    SELECTION: CHAIN A
REMARK   3    ORIGIN FOR THE GROUP (A):  34.3280 -44.3362 -33.2464
REMARK   3    T TENSOR
REMARK   3      T11:  -0.3752 T22:  -0.2836
REMARK   3      T33:  -0.1972 T12:  -0.0686
REMARK   3      T13:   0.0888 T23:  -0.1454
REMARK   3    L TENSOR
REMARK   3      L11:  -0.0328 L22:  -0.0264
REMARK   3      L33:  -0.0458 L12:  -0.0047
REMARK   3      L13:   0.0289 L23:  -0.0435
REMARK   3    S TENSOR
REMARK   3      S11:  -0.0220 S12:   0.7030 S13:   0.0888
REMARK   3      S21:  -0.7542 S22:  -0.0140 S23:  -0.0265
REMARK   3      S31:  -0.2584 S32:   0.0315 S33:   0.0000
REMARK   3   TLS GROUP : 2
REMARK   3    SELECTION: CHAIN B
REMARK   3    ORIGIN FOR THE GROUP (A):  45.1940 -82.3594 -33.3841
REMARK   3    T TENSOR
REMARK   3      T11:  -0.0302 T22:  -0.0378
REMARK   3      T33:  -0.0668 T12:  -0.0642
REMARK   3      T13:   0.1450 T23:  -0.0811
REMARK   3    L TENSOR
REMARK   3      L11:   0.0154 L22:   0.0032
```

```
REMARK   3      L33:  -0.0145 L12:   0.0014
REMARK   3      L13:   0.0441 L23:   0.0209
REMARK   3    S TENSOR
REMARK   3      S11:  -0.0023 S12:   0.6259 S13:   0.1176
REMARK   3      S21:  -0.6677 S22:  -0.0510 S23:   0.0868
REMARK   3      S31:  -0.1498 S32:  -0.2360 S33:   0.0000
REMARK   3   TLS GROUP : 3
REMARK   3    SELECTION: CHAIN C
REMARK   3    ORIGIN FOR THE GROUP (A):  81.9697 -97.2090 -33.4842
REMARK   3    T TENSOR
REMARK   3      T11:  -0.2480 T22:  -0.3700
REMARK   3      T33:  -0.1970 T12:   0.0271
REMARK   3      T13:   0.0948 T23:   0.1261
REMARK   3    L TENSOR
REMARK   3      L11:  -0.0431 L22:  -0.0280
REMARK   3      L33:  -0.0440 L12:   0.0175
REMARK   3      L13:   0.0465 L23:   0.0381
REMARK   3    S TENSOR
REMARK   3      S11:  -0.0226 S12:   0.7182 S13:  -0.0052
REMARK   3      S21:  -0.7057 S22:  -0.0069 S23:   0.1273
REMARK   3      S31:   0.0400 S32:  -0.2722 S33:   0.0000
REMARK   3   TLS GROUP : 4
REMARK   3    SELECTION: CHAIN D
REMARK   3    ORIGIN FOR THE GROUP (A): 116.5141 -77.7951 -33.2613
REMARK   3    T TENSOR
REMARK   3      T11:  -0.3864 T22:  -0.1961
REMARK   3      T33:  -0.1626 T12:   0.0163
REMARK   3      T13:   0.0020 T23:   0.1413
REMARK   3    L TENSOR
REMARK   3      L11:  -0.0288 L22:  -0.0282
REMARK   3      L33:  -0.0395 L12:   0.0063
REMARK   3      L13:  -0.0122 L23:   0.0401
REMARK   3    S TENSOR
REMARK   3      S11:  -0.0178 S12:   0.6303 S13:  -0.0642
REMARK   3      S21:  -0.7512 S22:  -0.0110 S23:   0.0796
REMARK   3      S31:   0.1866 S32:  -0.1247 S33:   0.0000
REMARK   3
REMARK   3  NCS DETAILS
REMARK   3   NUMBER OF NCS GROUPS : 3
REMARK   3   NCS GROUP : 1
REMARK   3    NCS OPERATOR : 1
REMARK   3     REFERENCE SELECTION: CHAIN A AND (RESSEQ 2:525 )
REMARK   3     SELECTION          : CHAIN B AND (RESSEQ 2:525 )
REMARK   3     ATOM PAIRS NUMBER  : 3856
REMARK   3     RMSD               : 0.214
REMARK   3    NCS OPERATOR : 2
REMARK   3     REFERENCE SELECTION: CHAIN A AND (RESSEQ 2:525 )
REMARK   3     SELECTION          : CHAIN C AND (RESSEQ 2:525 )
REMARK   3     ATOM PAIRS NUMBER  : 3856
REMARK   3     RMSD               : 0.214
REMARK   3    NCS OPERATOR : 3
REMARK   3     REFERENCE SELECTION: CHAIN A AND (RESSEQ 2:525 )
REMARK   3     SELECTION          : CHAIN D AND (RESSEQ 2:525 )
REMARK   3     ATOM PAIRS NUMBER  : 3856
REMARK   3     RMSD               : 0.186
REMARK   3    NCS OPERATOR : 4
REMARK   3     REFERENCE SELECTION: CHAIN A AND (RESSEQ 2:525 )
REMARK   3     SELECTION          : CHAIN E AND (RESSEQ 2:525 )
REMARK   3     ATOM PAIRS NUMBER  : 3856
REMARK   3     RMSD               : 0.213
REMARK   3    NCS OPERATOR : 5
REMARK   3     REFERENCE SELECTION: CHAIN A AND (RESSEQ 2:525 )
REMARK   3     SELECTION          : CHAIN F AND (RESSEQ 2:525 )
REMARK   3     ATOM PAIRS NUMBER  : 3856
```

```
REMARK   3      RMSD                 : 0.226
REMARK   3    NCS OPERATOR : 6
REMARK   3      REFERENCE SELECTION: CHAIN A AND (RESSEQ 2:525 )
REMARK   3      SELECTION          : CHAIN G AND (RESSEQ 2:525 )
REMARK   3      ATOM PAIRS NUMBER  : 3856
REMARK   3      RMSD               : 0.214
REMARK   3   NCS GROUP : 2
REMARK   3    NCS OPERATOR : 1
REMARK   3      REFERENCE SELECTION: CHAIN H AND (RESSEQ 2:525 )
REMARK   3      SELECTION          : CHAIN I AND (RESSEQ 2:525 )
REMARK   3      ATOM PAIRS NUMBER  : 3856
REMARK   3      RMSD               : 0.224
REMARK   3    NCS OPERATOR : 2
REMARK   3      REFERENCE SELECTION: CHAIN H AND (RESSEQ 2:525 )
REMARK   3      SELECTION          : CHAIN J AND (RESSEQ 2:525 )
REMARK   3      ATOM PAIRS NUMBER  : 3856
REMARK   3      RMSD               : 0.231
REMARK   3    NCS OPERATOR : 3
REMARK   3      REFERENCE SELECTION: CHAIN H AND (RESSEQ 2:525 )
REMARK   3      SELECTION          : CHAIN K AND (RESSEQ 2:525 )
REMARK   3      ATOM PAIRS NUMBER  : 3856
REMARK   3      RMSD               : 0.203
REMARK   3    NCS OPERATOR : 4
REMARK   3      REFERENCE SELECTION: CHAIN H AND (RESSEQ 2:525 )
REMARK   3      SELECTION          : CHAIN L AND (RESSEQ 2:525 )
REMARK   3      ATOM PAIRS NUMBER  : 3856
REMARK   3      RMSD               : 0.215
REMARK   3    NCS OPERATOR : 5
REMARK   3      REFERENCE SELECTION: CHAIN H AND (RESSEQ 2:525 )
REMARK   3      SELECTION          : CHAIN M AND (RESSEQ 2:525 )
REMARK   3      ATOM PAIRS NUMBER  : 3856
REMARK   3      RMSD               : 0.239
REMARK   3    NCS OPERATOR : 6
REMARK   3      REFERENCE SELECTION: CHAIN H AND (RESSEQ 2:525 )
REMARK   3      SELECTION          : CHAIN N AND (RESSEQ 2:525 )
REMARK   3      ATOM PAIRS NUMBER  : 3856
REMARK   3      RMSD               : 0.227
REMARK   3   NCS GROUP : 3
REMARK   3    NCS OPERATOR : 1
REMARK   3      REFERENCE SELECTION: CHAIN O AND (RESSEQ 1:97 )
REMARK   3      SELECTION          : CHAIN P AND (RESSEQ 1:97 )
REMARK   3      ATOM PAIRS NUMBER  : 728
REMARK   3      RMSD               : 0.207
REMARK   3    NCS OPERATOR : 2
REMARK   3      REFERENCE SELECTION: CHAIN O AND (RESSEQ 1:97 )
REMARK   3      SELECTION          : CHAIN Q AND (RESSEQ 1:97 )
REMARK   3      ATOM PAIRS NUMBER  : 728
REMARK   3      RMSD               : 0.211
REMARK   3    NCS OPERATOR : 3
REMARK   3      REFERENCE SELECTION: CHAIN O AND (RESSEQ 1:97 )
REMARK   3      SELECTION          : CHAIN R AND (RESSEQ 1:97 )
REMARK   3      ATOM PAIRS NUMBER  : 728
REMARK   3      RMSD               : 0.196
REMARK   3    NCS OPERATOR : 4
REMARK   3      REFERENCE SELECTION: CHAIN O AND (RESSEQ 1:97 )
REMARK   3      SELECTION          : CHAIN S AND (RESSEQ 1:97 )
REMARK   3      ATOM PAIRS NUMBER  : 728
REMARK   3      RMSD               : 0.200
REMARK   3    NCS OPERATOR : 5
REMARK   3      REFERENCE SELECTION: CHAIN O AND (RESSEQ 1:97 )
REMARK   3      SELECTION          : CHAIN T AND (RESSEQ 1:97 )
REMARK   3      ATOM PAIRS NUMBER  : 728
REMARK   3      RMSD               : 0.214
REMARK   3    NCS OPERATOR : 6
```

```
REMARK   3     REFERENCE SELECTION: CHAIN O AND (RESSEQ 1:97 )
REMARK   3     SELECTION          : CHAIN U AND (RESSEQ 1:97 )
REMARK   3     ATOM PAIRS NUMBER  : 728
REMARK   3     RMSD               : 0.205
REMARK   3
REMARK   3  OTHER REFINEMENT REMARKS: NULL
```

Template/example 2 (Xray/Neutron hybrid)

```
REMARK   3 REFINEMENT.
REMARK   3   PROGRAM     : PHENIX (PHENIX.REFINE)
REMARK   3   AUTHORS     : PAUL ADAMS,PAVEL AFONINE,VICENT CHEN,IAN
REMARK   3               : DAVIS,KRESHNA GOPAL,RALF GROSSE-
REMARK   3               : KUNSTLEVE,LI-WEI HUNG,ROBERT IMMORMINO,
REMARK   3               : TOM IOERGER,AIRLIE MCCOY,ERIK MCKEE,NIGEL
REMARK   3               : MORIARTY,REETAL PAI,RANDY READ,JANE
REMARK   3               : RICHARDSON,DAVID RICHARDSON,TOD ROMO,JIM
REMARK   3               : SACCHETTINI,NICHOLAS SAUTER,JACOB SMITH,
REMARK   3               : LAURENT STORONI,TOM TERWILLIGER,PETER
REMARK   3               : ZWART
REMARK   3
REMARK   3  X-RAY DATA.
REMARK   3
REMARK   3    REFINEMENT TARGET : ML
REMARK   3
REMARK   3  DATA USED IN REFINEMENT.
REMARK   3   RESOLUTION RANGE HIGH (ANGSTROMS) : 1.75
REMARK   3   RESOLUTION RANGE LOW  (ANGSTROMS) : 33.56
REMARK   3   MIN(FOBS/SIGMA_FOBS)              : 1.330
REMARK   3   COMPLETENESS FOR RANGE        (%) : 98.8
REMARK   3   NUMBER OF REFLECTIONS             : 31524
REMARK   3
REMARK   3  FIT TO DATA USED IN REFINEMENT.
REMARK   3   R VALUE     (WORKING + TEST SET) : 0.132
REMARK   3   R VALUE            (WORKING SET) : 0.129
REMARK   3   FREE R VALUE                     : 0.166
REMARK   3   FREE R VALUE TEST SET SIZE   (%) : 9.360
REMARK   3   FREE R VALUE TEST SET COUNT      : 2952
REMARK   3
REMARK   3  FIT TO DATA USED IN REFINEMENT (IN BINS).
REMARK   3   BIN  RESOLUTION RANGE  COMPL.    NWORK NFREE   RWORK  RFREE
REMARK   3     1 33.5691 -  4.8295    0.99     1424   130  14.0900 17.9800
REMARK   3     2  4.8295 -  3.8351    1.00     1420   125  10.3500 11.3600
REMARK   3     3  3.8351 -  3.3508    1.00     1378   129  10.5000 12.4600
REMARK   3     4  3.3508 -  3.0447    1.00     1413   122  11.7500 13.7300
REMARK   3     5  3.0447 -  2.8265    1.00     1385   138  11.9800 16.9100
REMARK   3     6  2.8265 -  2.6600    1.00     1384   131  12.7800 16.5000
REMARK   3     7  2.6600 -  2.5268    1.00     1351   162  12.1800 17.9600
REMARK   3     8  2.5268 -  2.4168    1.00     1394   137  12.3800 15.7200
REMARK   3     9  2.4168 -  2.3238    1.00     1381   130  12.3100 15.4300
REMARK   3    10  2.3238 -  2.2437    1.00     1385   142  12.5900 17.5500
REMARK   3    11  2.2437 -  2.1735    1.00     1368   123  12.0600 18.7300
REMARK   3    12  2.1735 -  2.1114    1.00     1377   143  11.8600 15.5300
REMARK   3    13  2.1114 -  2.0558    1.00     1337   171  11.3000 16.9600
REMARK   3    14  2.0558 -  2.0057    1.00     1353   143  11.9200 15.9400
REMARK   3    15  2.0057 -  1.9601    0.99     1386   144  12.2000 16.0500
REMARK   3    16  1.9601 -  1.9184    1.00     1340   136  12.9800 19.2100
REMARK   3    17  1.9184 -  1.8800    1.00     1380   142  13.7700 21.1700
REMARK   3    18  1.8800 -  1.8445    1.00     1371   141  14.1400 18.2000
REMARK   3    19  1.8445 -  1.8116    1.00     1344   172  15.5900 22.0300
REMARK   3    20  1.8116 -  1.7809    0.99     1321   158  16.7800 22.1500
REMARK   3    21  1.7809 -  1.7522    0.79     1080   133  20.5300 25.6900
```

```
REMARK   3
REMARK   3  BULK SOLVENT MODELLING.
REMARK   3   METHOD USED        : FLAT BULK SOLVENT MODEL
REMARK   3   SOLVENT RADIUS     : 1.11
REMARK   3   SHRINKAGE RADIUS   : 0.90
REMARK   3   K_SOL              : 0.35
REMARK   3   B_SOL              : 29.39
REMARK   3
REMARK   3  ERROR ESTIMATES.
REMARK   3   COORDINATE ERROR (MAXIMUM-LIKELIHOOD BASED)     : 0.190
REMARK   3   PHASE ERROR (DEGREES, MAXIMUM-LIKELIHOOD BASED) : 15.390
REMARK   3
REMARK   3  B VALUES.
REMARK   3   FROM WILSON PLOT           (A**2) : 17.52
REMARK   3   MEAN B VALUE      (OVERALL, A**2) : NULL
REMARK   3   OVERALL ANISOTROPIC B VALUE.
REMARK   3    B11 (A**2) : 5.33780
REMARK   3    B22 (A**2) : 5.45600
REMARK   3    B33 (A**2) : 4.24070
REMARK   3    B12 (A**2) : 0.00000
REMARK   3    B13 (A**2) : 0.43090
REMARK   3    B23 (A**2) : 0.00000
REMARK   3
REMARK   3  TWINNING INFORMATION.
REMARK   3   FRACTION: NULL
REMARK   3   OPERATOR: NULL
REMARK   3
REMARK   3  DEVIATIONS FROM IDEAL VALUES.
REMARK   3                RMSD          COUNT
REMARK   3   BOND      :   NULL           NULL
REMARK   3   ANGLE     :   NULL           NULL
REMARK   3   CHIRALITY :   NULL           NULL
REMARK   3   PLANARITY :   NULL           NULL
REMARK   3   DIHEDRAL  :   NULL           NULL
REMARK   3
REMARK   3  TLS DETAILS
REMARK   3   NUMBER OF TLS GROUPS  : NULL
REMARK   3
REMARK   3  NCS DETAILS
REMARK   3   NUMBER OF NCS GROUPS : NULL
REMARK   3
REMARK   3  OTHER REFINEMENT REMARKS: NULL
REMARK   3
REMARK   3  NEUTRON DATA.
REMARK   3
REMARK   3    REFINEMENT TARGET : ML
REMARK   3
REMARK   3  DATA USED IN REFINEMENT.
REMARK   3   RESOLUTION RANGE HIGH (ANGSTROMS) : 2.19
REMARK   3   RESOLUTION RANGE LOW  (ANGSTROMS) : 40.11
REMARK   3   MIN(FOBS/SIGMA_FOBS)              : 1.530
REMARK   3   COMPLETENESS FOR RANGE        (%) : 72.8
REMARK   3   NUMBER OF REFLECTIONS             : 11884
REMARK   3
REMARK   3  FIT TO DATA USED IN REFINEMENT.
REMARK   3   R VALUE     (WORKING + TEST SET) : 0.260
REMARK   3   R VALUE            (WORKING SET) : 0.257
REMARK   3   FREE R VALUE                     : 0.291
REMARK   3   FREE R VALUE TEST SET SIZE   (%) : 8.350
REMARK   3   FREE R VALUE TEST SET COUNT      : 992
REMARK   3
REMARK   3  FIT TO DATA USED IN REFINEMENT (IN BINS).
REMARK   3   BIN  RESOLUTION RANGE  COMPL.    NWORK NFREE   RWORK  RFREE
REMARK   3     1 40.1164 -  4.1959    0.93     2018   176  19.7800 20.5900
```

```
REMARK   3     2  4.1959 -  3.3309    0.89      1915   175  21.1000 25.3500
REMARK   3     3  3.3309 -  2.9100    0.78      1669   152  25.7100 30.8200
REMARK   3     4  2.9100 -  2.6440    0.70      1475   141  26.7100 30.7500
REMARK   3     5  2.6440 -  2.4545    0.64      1350   133  28.7700 33.5100
REMARK   3     6  2.4545 -  2.3098    0.60      1269   117  30.9500 33.6400
REMARK   3     7  2.3098 -  2.1942    0.56      1196    98  32.3200 36.3200
REMARK   3
REMARK   3  BULK SOLVENT MODELLING.
REMARK   3   METHOD USED        : FLAT BULK SOLVENT MODEL
REMARK   3   SOLVENT RADIUS     : 1.11
REMARK   3   SHRINKAGE RADIUS   : 0.90
REMARK   3   K_SOL              : 0.53
REMARK   3   B_SOL              : 82.90
REMARK   3
REMARK   3  ERROR ESTIMATES.
REMARK   3   COORDINATE ERROR (MAXIMUM-LIKELIHOOD BASED)     : 0.450
REMARK   3   PHASE ERROR (DEGREES, MAXIMUM-LIKELIHOOD BASED) : 26.920
REMARK   3
REMARK   3  B VALUES.
REMARK   3   FROM WILSON PLOT           (A**2) : NULL
REMARK   3   MEAN B VALUE      (OVERALL, A**2) : NULL
REMARK   3   OVERALL ANISOTROPIC B VALUE.
REMARK   3    B11 (A**2) : -0.70340
REMARK   3    B22 (A**2) : -6.62540
REMARK   3    B33 (A**2) : -7.07190
REMARK   3    B12 (A**2) : -0.00000
REMARK   3    B13 (A**2) : -1.25130
REMARK   3    B23 (A**2) : -0.00000
REMARK   3
REMARK   3  TWINNING INFORMATION.
REMARK   3   FRACTION: NULL
REMARK   3   OPERATOR: NULL
REMARK   3
REMARK   3  DEVIATIONS FROM IDEAL VALUES.
REMARK   3                 RMSD          COUNT
REMARK   3   BOND      :  0.018           5216
REMARK   3   ANGLE     :  1.759           9458
REMARK   3   CHIRALITY :  0.160            392
REMARK   3   PLANARITY :  0.011            779
REMARK   3   DIHEDRAL  : 25.100           1363
REMARK   3
REMARK   3  TLS DETAILS
REMARK   3   NUMBER OF TLS GROUPS  : NULL
REMARK   3
REMARK   3  NCS DETAILS
REMARK   3   NUMBER OF NCS GROUPS : NULL
REMARK   3
REMARK   3  OTHER REFINEMENT REMARKS: NULL
```

Refinement using Electron Microscopy

Template/Example

```
REMARK   3
REMARK   3 REFINEMENT
REMARK   3   SOFTWARE PACKAGES      : SIMPLEX, PYPFT, EMFIT, O, XPLOR
REMARK   3   RECONSTRUCTION SCHEMA  : ICOSAHEDRAL
REMARK   3
REMARK   3 EM MAP-MODEL FITTING AND REFINEMENT
REMARK   3   PDB ENTRY                    : PDB ID 1HX6
REMARK   3   REFINEMENT SPACE             : RECIPROCAL
REMARK   3   REFINEMENT PROTOCOL          : RIGID BODY REFINEMENT
REMARK   3   REFINEMENT TARGET            : R-FACTOR
REMARK   3   OVERALL ANISOTROPIC B VALUE  : NULL
REMARK   3
REMARK   3 FITTING PROCEDURE : THE CRYSTAL STRUCTURE OF THE MAJOR COAT
REMARK   3 PROTEIN P3 (PDB FILE 1HX6) WAS PLACED INTO THE CRYO-EM
REMARK   3 DENSITY MAP. THE CAPSID PROTEIN WAS FIRST MANUALLY
REMARK   3 POSITIONED INTO THE CRYO-EM DENSITY CORRESPONDING TO
REMARK   3 POSITIONS OF THE FOUR INDEPENDENT TRIMERS IN THE
REMARK   3 ICOSAHEDRAL ASYMMETRIC UNIT. THESE POSITIONS WERE THEN
REMARK   3 REFINED BY RIGID BODY REFINEMENT IN RECIPROCAL SPACE WITH
REMARK   3 THE PROGRAM XPLOR.
REMARK   3 QUALITY OF THE FIT R-FACTOR= 0.339, CROSS-CORRELATI0N
REMARK   3 COEFFICIENT 0.915, ATOMS OUTSIDE DENSITY PER ICOSAHEDRAL
REMARK   3 ASYMMETRIC UNIT 527 (1.5%),ATOM CLASHES PER ICOSAHEDRAL
REMARK   3 ASYMMETRIC UNIT 115 (0.3%)
REMARK   3
REMARK   3 EM IMAGE RECONSTRUCTION STATISTICS
REMARK   3   NOMINAL PIXEL SIZE (ANGSTROMS)    : 3.68
REMARK   3   ACTUAL PIXEL SIZE  (ANGSTROMS)    : 3.44
REMARK   3   EFFECTIVE RESOLUTION (ANGSTROMS)  : 14.0
REMARK   3   NUMBER OF PARTICLES               : 1800
REMARK   3   CTF CORRECTION METHOD             : NULL
REMARK   3
REMARK   3 EM RECONSTRUCTION MAGNIFICATION CALIBRATION: THE PIXEL
REMARK   3  SIZE OF THE CRYO-EM MAP WAS OBTAINED USING THE X-RAY
REMARK   3  STRUCTURE OF THE P3 TRIMER AS A REFERENCE. AFTER AN INITIAL
REMARK   3  FITTING USING THE NOMINAL PIXEL SIZE, THE P3 TRIMERS IN THE
REMARK   3  ICOSAHEDRAL ASYMMETRIC UNIT WERE GRADUALLY TRANSLATED TOWARDS
REMARK   3  THE CENTER OF THE PARTICLE UNTIL THE CRYSTALLOGRAPHIC R-FACTOR
REMARK   3  WAS MINIMISED.
REMARK   3
REMARK   3 OTHER DETAILS: THE ORIENTATIONS WERE REFINED BY THE CROSS
REMARK   3  COMMON LINES LINES METHOD (SIMPLEX) AND THE POLAR FOURIER
REMARK   3  TRANSFORM METHOD. MODEL-BASED, POLAR-FOURIER-TRANSFORM
REMARK   3  (FULLER ET AL. 1996, J.STRUC.BIOL. 116, 48-55; BAKER AND
REMARK   3  CHENG, 1996, J.STRUC.BIOL. 116, 120-130) MODEL-BASED CROSS
REMARK   3  COMMON LINES SEARCH AND REFINEMENT (CROWTHER ET AL. 1970,
REMARK   3  NATURE (LONDON) 226, 421-425; FULLER ET AL. 1996,
REMARK   3  J.STRUC.BIOL. 116, 48-55;FERLENGHI ET AL. 1998, J.MOL.BIOL.
REMARK   3  283, 71-81). THE EFFECTIVE RESOLUTION OF THE FINAL
REMARK   3  RECONSTRUCTED DENSITY WAS DETERMINED TO BE AT LEAST 25
REMARK   3  ANGSTROMS, AS MEASURED BY RANDOMLY SPLITTING THE PARTICLES
REMARK   3  INTO TWO SETS AND CALCULATING THE FOURIER SHELL CORRELATION
REMARK   3  OBTAINED FROM SEPARATE RECONSTRUCTIONS (HARAUZ AND VAN HEEL
REMARK   3  1986, OPTIK 73, 146-156). THE EIGENVALUE SPECTRUM GAVE AN
REMARK   3  INDICATION OF THE RANDOMNESS OF THE DATA THAT WAS INCLUDED
REMARK   3  IN THE RECONSTRUCTION. THE COMPLETENESS OF THE DATA WAS
REMARK   3  VERIFIED IN THAT ALL EIGENVALUES EXCEEDED 100. THE COORDINATES
REMARK   3  ARE IN THE P, Q, R FRAME IN ANGSTROM UNITS AND CORRESPOND
```

```
REMARK   3  TO ICOSAHEDRAL SYMMETRY AXES. THE ORIGIN IS CHOSEN AT THE
REMARK   3  CENTER OF THE VIRUS WITH P, Q AND R ALONG MUTUALLY
REMARK   3  PERPENDICULAR TWO-FOLD AXES OF THE ICOSAHEDRON. THEY SHOULD
REMARK   3  REMAIN IN THAT FRAME FOR THE EASE OF THE USER IN CREATING
REMARK   3  THE BIOLOGICALLY SIGNIFICANT VIRAL COMPLEX PARTICLE USING
REMARK   3  THE 60 ICOSAHEDRAL SYMMETRY OPERATORS. RESIDUES NOT VISIBLE
REMARK   3  IN THE ORIGINAL CRYSTAL STRUCTURES ARE NOT INCLUDED IN THE
REMARK   3  CRYO-EM STRUCTURE MODEL.
```

Example for Solution Scattering

```
REMARK   3
REMARK   3 REFINEMENT.
REMARK   3   PROGRAM     : INSIGHT II 98.0
REMARK   3   AUTHORS     : MSI
REMARK   3
REMARK   3  NUMBER OF NON-HYDROGEN ATOMS USED IN REFINEMENT.
REMARK   3   PROTEIN ATOMS            : 1213
REMARK   3   NUCLEIC ACID ATOMS       : 0
REMARK   3   HETEROGEN ATOMS          : 0
REMARK   3   SOLVENT ATOMS            : 0
REMARK   3
REMARK   3  OTHER REFINEMENT REMARKS: DISCOVER WAS USED FOR ENERGY
REMARK   3   MINIMISATION
```

Non-diffraction studies

Until standard refinement remarks are adopted for non-diffraction studies, refinement details will appear in REMARK 3 formatted in free text, beginning on the sixth line of the remark.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK   3
REMARK   3 REFINEMENT.
REMARK   3   PROGRAM     :
REMARK   3   AUTHORS     :
REMARK   3
REMARK   3 FREE TEXT
```

Example

```
REMARK   3
REMARK   3 REFINEMENT.
REMARK   3   PROGRAM     : CNSSOLVE 1.1, X-PLOR 2.11.2, PROCHECK NMR 3.51,
REMARK   3                 MOLPROBITY 3.01, QUEEN 1.1, PSVS 1.3
REMARK   3   AUTHORS     : BRUNGER, ET. AL. (CNSSOLVE), CLORE ET. AL. (X-
REMARK   3                 PLOR), LASKOWSKI, MACARTHUR (PROCHECK NMR),
REMARK   3                 LOVELL, RICHARDSON ET. AL. (MOLPROBITY),
REMARK   3                 NABUURS, VUISTER (QUEEN), BHATTACHARYA,
REMARK   3                 MONTELIONE (PSVS)
REMARK   3
REMARK   3  OTHER REFINEMENT REMARKS: NOESY ASSIGNMENT MADE WITH ITERATIVE
REMARK   3  METHOD USING CNS, HYPER (DIHEDRAL) AND DYANA FOLLOWED BY NIH-
REMARK   3  XPLOR FOR SIMMULATED ANNEALING MD. CONVERGED STRUCTURES WERE
REMARK   3  FURTHER MINIMIZED USING CNS IN EXPLICIT H2O SHELL (NILGES
REMARK   3  PROTOCOL). FULL LENGTH SEQUENCE WAS CARRIED THROUGH THE
REMARK   3  REFINEMENT PROTOCOL. COORDINATES FROM DISORDERED REGIONS,
REMARK   3  INCLUDING HEXHIS TAG, WERE NOT REPORTED. STRUCTURE IS BASED ON
REMARK   3  439 CONSTRAINTS (216 LONG RANGE), 43 DIHEDRAL AND 20 H-BOND.
```

REMARK 4 (updated), Format

Remark 4 indicates the version of the PDB File Format used to generate the file.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK   4
REMARK   4 XXXX COMPLIES WITH FORMAT V. N.MM, DD-MMM-YY
```

XXXX refers to the ID code of the entry.

N.MM refers to the version number. The current version is 3.20.
DD-MMM-YY refers to the release date of that version of the format. DD is the day of the month (a number 01 through 31), MMM is the English 3-letter abbreviation for the month, and YY is the year.

Example

```
REMARK   4
REMARK   4 1ABC COMPLIES WITH FORMAT V. 3.20, 01-DEC-08
```

REMARKs 5 - 99

Overview

REMARKs following the format REMARK 4 consist of free text annotation, pre-defined templates, and token: value pair-styled templates. Presented here are examples of REMARK sections.

Record Format and Details

* As with all other REMARKs, the first line of each REMARK is empty and is used as a spacer.

* REMARKs 5-99 are no longer for use of free text annotation.

REMARK 100 (updated), Deposition or Processing Site

This REMARK indicates PDB process site: RCSB, PDBe, PDBj or BNL. This remark also contains process date and site id code with exception of BNL entries.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 100 THIS ENTRY HAS BEEN PROCESSED BY XXXX on DD-MMM-YY.
REMARK 100 THE XXXX ID CODE IS vvvvvvvvvv.
```

XXXX is the process site. VVVVVVVVVV is the site id code. The format of the date in this remark is DD-MMM-YY. DD is the day of the month (a number 01 through 31), MMM is the English 3-letter abbreviation for the month, and YY is the year.

Examples

```
REMARK 100
REMARK 100 THIS ENTRY HAS BEEN PROCESSED BY RCSB on 10-MAR-06.
REMARK 100 THE RCSB ID CODE IS RCSB036809.

REMARK 100
REMARK 100 THIS ENTRY HAS BEEN PROCESSED BY PDBE  on 13-FEB-07.
REMARK 100 THE PDBE ID CODE IS EBI-28843.

REMARK 100
REMARK 100 THIS ENTRY HAS BEEN PROCESSED BY PDBJ on 21-MAR-05.
REMARK 100 THE RCSB ID CODE IS RCSB026278.

REMARK 100
REMARK 100 THIS ENTRY HAS BEEN PROCESSED BY BNL.
```

REMARKs 200-265, Experimental Details

REMARKs in this range present the data collection details for the data which resulted in the refinement statistics of REMARK 3. They provide information on the structure determination experiment, which may have been done by diffraction, NMR or some other technique.

The "NULL" value will be used if the data for a token is not supplied by the depositor.

REMARK 200 (updated), X-ray Diffraction Experimental Details

REMARK 200 is mandatory if single crystal, fiber, or polycrystalline X-ray diffraction experiments were performed. The format of date in this remark is DD-MMM-YY. DD is the day of the month (a number 01 through 31), MMM is the English 3-letter abbreviation for the month, and YY is the year.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 200
REMARK 200 EXPERIMENTAL DETAILS
REMARK 200  EXPERIMENT TYPE                : X-RAY DIFFRACTION
REMARK 200  DATE OF DATA COLLECTION        :
REMARK 200  TEMPERATURE           (KELVIN) :
REMARK 200  PH                             :
REMARK 200  NUMBER OF CRYSTALS USED        :
REMARK 200
REMARK 200  SYNCHROTRON              (Y/N) :
REMARK 200  RADIATION SOURCE               :
REMARK 200  BEAMLINE                       :
REMARK 200  X-RAY GENERATOR MODEL          :
REMARK 200  MONOCHROMATIC OR LAUE    (M/L) :
REMARK 200  WAVELENGTH OR RANGE        (A) :
REMARK 200  MONOCHROMATOR                  :
REMARK 200  OPTICS                         :
REMARK 200
REMARK 200  DETECTOR TYPE                  :
REMARK 200  DETECTOR MANUFACTURER          :
REMARK 200  INTENSITY-INTEGRATION SOFTWARE :
REMARK 200  DATA SCALING SOFTWARE          :
REMARK 200
REMARK 200  NUMBER OF UNIQUE REFLECTIONS   :
REMARK 200  RESOLUTION RANGE HIGH      (A) :
REMARK 200  RESOLUTION RANGE LOW       (A) :
REMARK 200  REJECTION CRITERIA  (SIGMA(I)) :
REMARK 200
REMARK 200 OVERALL.
REMARK 200  COMPLETENESS FOR RANGE     (%) :
REMARK 200  DATA REDUNDANCY                :
REMARK 200  R MERGE                    (I) :
REMARK 200  R SYM                      (I) :
REMARK 200  <I/SIGMA(I)> FOR THE DATA SET  :
REMARK 200
REMARK 200 IN THE HIGHEST RESOLUTION SHELL.
REMARK 200  HIGHEST RESOLUTION SHELL, RANGE HIGH (A) :
REMARK 200  HIGHEST RESOLUTION SHELL, RANGE LOW  (A) :
REMARK 200  COMPLETENESS FOR SHELL     (%) :
REMARK 200  DATA REDUNDANCY IN SHELL       :
```

```
REMARK 200  R MERGE FOR SHELL          (I) :
REMARK 200  R SYM FOR SHELL            (I) :
REMARK 200  <I/SIGMA(I)> FOR SHELL         :
REMARK 200
REMARK 200 METHOD USED TO DETERMINE THE STRUCTURE:
REMARK 200 SOFTWARE USED:
REMARK 200 STARTING MODEL:
REMARK 200
REMARK 200 REMARK:
```

Examples

The following **example** illustrates the how REMARK 200 will be used in cases in which multiple data collections are described. In this example, data items corresponding to different data collection sessions are separated by semi-colons. Multiple data values within a single session (e.g. wavelength) are separated by commas.

```
REMARK 200
REMARK 200 EXPERIMENTAL DETAILS
REMARK 200  EXPERIMENT TYPE                : X-RAY DIFFRACTION
REMARK 200  DATE OF DATA COLLECTION        : 17-MAR-02; 17-MAR-02
REMARK 200  TEMPERATURE           (KELVIN) : 100; 100
REMARK 200  PH                             : 8.00
REMARK 200  NUMBER OF CRYSTALS USED        : 2
REMARK 200
REMARK 200  SYNCHROTRON              (Y/N) : Y; Y
REMARK 200  RADIATION SOURCE               : APS ; APS
REMARK 200  BEAMLINE                       : 17ID; 17ID
REMARK 200  X-RAY GENERATOR MODEL          : NULL
REMARK 200  MONOCHROMATIC OR LAUE    (M/L) : M; M
REMARK 200  WAVELENGTH OR RANGE        (A) : 1.5545; 1.0720, 1.0723,
REMARK 200                                   1.0543
REMARK 200  MONOCHROMATOR                  : SI (111); SI (111)
REMARK 200  OPTICS                         : NULL
REMARK 200
REMARK 200  DETECTOR TYPE                  : CCD; CCD
REMARK 200  DETECTOR MANUFACTURER          : ADSC QUANTUM 210; ADSC
REMARK 200                                   QUANTUM 210
REMARK 200  INTENSITY-INTEGRATION SOFTWARE : DENZO
REMARK 200  DATA SCALING SOFTWARE          : HKL
REMARK 200
REMARK 200  NUMBER OF UNIQUE REFLECTIONS   : 29132
REMARK 200  RESOLUTION RANGE HIGH      (A) : 1.900
REMARK 200  RESOLUTION RANGE LOW       (A) : 30.000
REMARK 200  REJECTION CRITERIA  (SIGMA(I)) : 0.000
REMARK 200
REMARK 200 OVERALL.
REMARK 200  COMPLETENESS FOR RANGE     (%) : 98.3
REMARK 200  DATA REDUNDANCY                : 19.800
REMARK 200  R MERGE                    (I) : NULL
REMARK 200  R SYM                      (I) : 0.07500
REMARK 200  <I/SIGMA(I)> FOR THE DATA SET  : 17.0000
REMARK 200
REMARK 200 IN THE HIGHEST RESOLUTION SHELL.
REMARK 200  HIGHEST RESOLUTION SHELL, RANGE HIGH (A) : 1.90
REMARK 200  HIGHEST RESOLUTION SHELL, RANGE LOW  (A) : 1.97
REMARK 200  COMPLETENESS FOR SHELL     (%) : 83.4
REMARK 200  DATA REDUNDANCY IN SHELL       : 3.00
REMARK 200  R MERGE FOR SHELL          (I) : NULL
REMARK 200  R SYM FOR SHELL            (I) : 0.65000
REMARK 200  <I/SIGMA(I)> FOR SHELL         : 1.500
```

```
REMARK 200
REMARK 200 DIFFRACTION PROTOCOL: SINGLE WAVELENGTH; MAD
REMARK 200 METHOD USED TO DETERMINE THE STRUCTURE: MAD
REMARK 200 SOFTWARE USED: SOLVE 2.02
REMARK 200 STARTING MODEL: NULL
REMARK 200
REMARK 200 REMARK: NULL
```

REMARK 205, Fiber Diffraction, Fiber Sample Experiment Details

REMARK 205 is mandatory if data was obtained from a fiber diffraction - non-crystalline sample study

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 205
REMARK 205 THESE COORDINATES WERE GENERATED FROM FIBER DIFFRACTION
REMARK 205 DATA.    PROTEIN DATA BANK CONVENTIONS REQUIRE THAT CRYST1
REMARK 205 AND SCALE RECORDS BE INCLUDED, BUT THE VALUES OF THESE
REMARK 205 RECORDS ARE MEANINGLESS.
```

REMARKs 210 and 215/217, NMR Experiment Details

Remark 210 is mandatory if data was obtained from an NMR experiment.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 210
REMARK 210 EXPERIMENTAL DETAILS
REMARK 210  EXPERIMENT TYPE                :
REMARK 210  TEMPERATURE           (KELVIN) :
REMARK 210  PH                             :
REMARK 210  IONIC STRENGTH                 :
REMARK 210  PRESSURE                       :
REMARK 210  SAMPLE CONTENTS                :
REMARK 210
REMARK 210  NMR EXPERIMENTS CONDUCTED      :
REMARK 210  SPECTROMETER FIELD STRENGTH    :
REMARK 210  SPECTROMETER MODEL             :
REMARK 210  SPECTROMETER MANUFACTURER      :
REMARK 210
REMARK 210      STRUCTURE DETERMINATION.
REMARK 210       SOFTWARE USED             :
REMARK 210       METHOD USED               :
REMARK 210
REMARK 210 CONFORMERS, NUMBER CALCULATED   :
REMARK 210 CONFORMERS, NUMBER SUBMITTED    :
REMARK 210 CONFORMERS, SELECTION CRITERIA  :
REMARK 210
REMARK 210
REMARK 210 BEST REPRESENTATIVE CONFORMER IN THIS ENSEMBLE :
REMARK 210
REMARK 210 REMARK:
```

Example

```
REMARK 210
REMARK 210 EXPERIMENTAL DETAILS
REMARK 210  EXPERIMENT TYPE                : NMR
REMARK 210  TEMPERATURE           (KELVIN) : 293
```

```
REMARK 210  PH                             : 7.0
REMARK 210  IONIC STRENGTH                 : NULL
REMARK 210  PRESSURE                       : AMBIENT
REMARK 210  SAMPLE CONTENTS                : 4.0 MM PHYLLOSEPTIN-2,
REMARK 210                                   TRIFLUOROETHANOL/WATER (60%/
REMARK 210                                   40%)
REMARK 210
REMARK 210  NMR EXPERIMENTS CONDUCTED      : 2D 1H-1H TOCSY, 2D 1H-1H
REMARK 210                                   NOESY, 2D 1H-13C HSQC, 2D 1H-
REMARK 210                                   15N HSQC
REMARK 210  SPECTROMETER FIELD STRENGTH    : 600; 800
REMARK 210  SPECTROMETER MODEL             : DMX600; DRX800
REMARK 210  SPECTROMETER MANUFACTURER      : BRUKER
REMARK 210
REMARK 210  STRUCTURE DETERMINATION.
REMARK 210   SOFTWARE USED                 : XWINNMR, NMRPIPE, NMRVIEW
REMARK 210                                   5.0.4, X-PLOR_NIH 2.17.0,
REMARK 210                                   MOLMOL 2K.2, PROCHECK 3.5.4
REMARK 210   METHOD USED                   : SIMULATED ANNEALING
REMARK 210
REMARK 210 CONFORMERS, NUMBER CALCULATED   : 200
REMARK 210 CONFORMERS, NUMBER SUBMITTED    : 10
REMARK 210 CONFORMERS, SELECTION CRITERIA  : STRUCTURES WITH THE LOWEST
REMARK 210                                   ENERGY
REMARK 210
REMARK 210 BEST REPRESENTATIVE CONFORMER IN THIS ENSEMBLE : 1
REMARK 210
REMARK 210 REMARK: NULL
```

REMARK 215 is necessary if data was obtained from a solution NMR experiment.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 215
REMARK 215 NMR STUDY
REMARK 215 THE COORDINATES IN THIS ENTRY WERE GENERATED FROM SOLUTION
REMARK 215 NMR DATA.        PROTEIN DATA BANK CONVENTIONS REQUIRE THAT
REMARK 215 CRYST1 AND SCALE RECORDS BE INCLUDED, BUT THE VALUES ON
REMARK 215 THESE RECORDS ARE MEANINGLESS.
```

Remark 217 is used in place of REMARK 215 if a Solid State NMR experiment was performed.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 217
REMARK 217 SOLID STATE NMR STUDY
REMARK 217 THE COORDINATES IN THIS ENTRY WERE GENERATED FROM SOLID
REMARK 217 STATE NMR DATA.  PROTEIN DATA BANK CONVENTIONS REQUIRE THAT
REMARK 217 CRYST1 AND SCALE RECORDS BE INCLUDED, BUT THE VALUES ON
REMARK 217 THESE RECORDS ARE MEANINGLESS.
```

REMARK 230, Neutron Diffraction Experiment Details

REMARK 230 is mandatory if data was obtained from a neutron diffraction study. The format of the date in this remark is DD-MMM-YY. DD is the day of the month (a number 01 through 31), MMM is the English 3-letter abbreviation for the month, and YY is the year.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 230
REMARK 230 EXPERIMENTAL DETAILS
REMARK 230  EXPERIMENT TYPE                : NEUTRON DIFFRACTION
REMARK 230  DATE OF DATA COLLECTION        :
REMARK 230  TEMPERATURE           (KELVIN) :
REMARK 230  PH                             :
REMARK 230  NUMBER OF CRYSTALS USED        :
REMARK 230
REMARK 230  NEUTRON SOURCE                 :
REMARK 230  BEAMLINE                       :
REMARK 230  WAVELENGTH OR RANGE        (A) :
REMARK 230  MONOCHROMATOR                  :
REMARK 230  OPTICS                         :
REMARK 230
REMARK 230  DETECTOR TYPE                  :
REMARK 230  DETECTOR MANUFACTURER          :
REMARK 230  INTENSITY-INTEGRATION SOFTWARE :
REMARK 230  DATA SCALING SOFTWARE          :
REMARK 230
REMARK 230  NUMBER OF UNIQUE REFLECTIONS   :
REMARK 230  RESOLUTION RANGE HIGH      (A) :
REMARK 230  RESOLUTION RANGE LOW       (A) :
REMARK 230  REJECTION CRITERIA  (SIGMA(I)) :
REMARK 230
REMARK 230 OVERALL.
REMARK 230  COMPLETENESS FOR RANGE     (%) :
REMARK 230  DATA REDUNDANCY                :
REMARK 230  R MERGE                    (I) :
REMARK 230  R SYM                      (I) :
REMARK 230  <I/SIGMA(I)> FOR THE DATA SET  :
REMARK 230
REMARK 230 IN THE HIGHEST RESOLUTION SHELL.
REMARK 230  HIGHEST RESOLUTION SHELL, RANGE HIGH (A) :
REMARK 230  HIGHEST RESOLUTION SHELL, RANGE LOW  (A) :
REMARK 230  COMPLETENESS FOR SHELL     (%) :
REMARK 230  DATA REDUNDANCY IN SHELL       :
REMARK 230  R MERGE FOR SHELL          (I) :
REMARK 230  R SYM FOR SHELL            (I) :
REMARK 230  <I/SIGMA(I)> FOR SHELL         :
REMARK 230
REMARK 230 METHOD USED TO DETERMINE THE STRUCTURE:
REMARK 230 SOFTWARE USED :
REMARK 230 STARTING MODEL:
REMARK 230
REMARK 230 REMARK:
```

Example

```
REMARK 230
REMARK 230 EXPERIMENTAL DETAILS
REMARK 230  EXPERIMENT TYPE                : NEUTRON DIFFRACTION
REMARK 230  DATE OF DATA COLLECTION        : 05-MAY-03
REMARK 230  TEMPERATURE           (KELVIN) : 293.0
REMARK 230  PH                             : 5.30
REMARK 230  NUMBER OF CRYSTALS USED        : 1
REMARK 230
REMARK 230  NEUTRON SOURCE                 : NULL
REMARK 230  BEAMLINE                       : NULL
REMARK 230  WAVELENGTH OR RANGE        (A) : 2.88
REMARK 230  MONOCHROMATOR                  : ELLASTICALLY BENT SILICON
REMARK 230  OPTICS                         : MONOCHROMATOR
REMARK 230
REMARK 230  DETECTOR TYPE                  : NEUTRON IMAGING PLATE
REMARK 230  DETECTOR MANUFACTURER          : BIX-3
REMARK 230  INTENSITY-INTEGRATION SOFTWARE : DENZO
REMARK 230  DATA SCALING SOFTWARE          : SCALEPACK
REMARK 230
REMARK 230  NUMBER OF UNIQUE REFLECTIONS   : 7001
REMARK 230  RESOLUTION RANGE HIGH      (A) : 2.400
REMARK 230  RESOLUTION RANGE LOW       (A) : 100.000
REMARK 230  REJECTION CRITERIA  (SIGMA(I)) : 1.000
REMARK 230
REMARK 230 OVERALL.
REMARK 230  COMPLETENESS FOR RANGE     (%) : 92.5
REMARK 230  DATA REDUNDANCY                : NULL
REMARK 230  R MERGE                    (I) : 0.14300
REMARK 230  R SYM                      (I) : NULL
REMARK 230  <I/SIGMA(I)> FOR THE DATA SET  : NULL
REMARK 230
REMARK 230 IN THE HIGHEST RESOLUTION SHELL.
REMARK 230  HIGHEST RESOLUTION SHELL, RANGE HIGH (A) : 2.40
REMARK 230  HIGHEST RESOLUTION SHELL, RANGE LOW  (A) : 2.49
REMARK 230  COMPLETENESS FOR SHELL     (%) : 82.1
REMARK 230  DATA REDUNDANCY IN SHELL       : NULL
REMARK 230  R MERGE FOR SHELL          (I) : 0.39500
REMARK 230  R SYM FOR SHELL            (I) : NULL
REMARK 230  <I/SIGMA(I)> FOR SHELL         : 2.300
REMARK 230
REMARK 230 METHOD USED TO DETERMINE THE STRUCTURE: MOLECULAR REPLACEMENT
REMARK 230 SOFTWARE USED : CNS
REMARK 230 STARTING MODEL: PDB ENTRY 1UCR
REMARK 230
REMARK 230 REMARK: NULL
```

REMARK 240 (updated), Electron Crystallography Experiment Details

REMARK 240 is mandatory if data was obtained from an electron crystallography study. The format of the date in this remark is DD-MMM-YY. DD is the day of the month (a number 01 through 31), MMM is the English 3-letter abbreviation for the month, and YY is the year.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 240
REMARK 240 EXPERIMENTAL DETAILS
REMARK 240   RECONSTRUCTION METHOD          :
REMARK 240   SAMPLE TYPE                    :
REMARK 240   SPECIMEN TYPE                  :
REMARK 240 DATA ACQUISITION
REMARK 240   DATE OF DATA COLLECTION        :
REMARK 240   TEMPERATURE           (KELVIN) :
REMARK 240   PH                             :
REMARK 240   NUMBER OF CRYSTALS USED        :
REMARK 240   MICROSCOPE MODEL               :
REMARK 240   DETECTOR TYPE                  :
REMARK 240   ACCELERATION VOLTAGE (KV)      :
REMARK 240   NUMBER OF UNIQUE REFLECTIONS   :
REMARK 240   RESOLUTION RANGE HIGH      (A) :
REMARK 240   RESOLUTION RANGE LOW       (A) :
REMARK 240   DATA SCALING SOFTWARE          :
REMARK 240   COMPLETENESS FOR RANGE     (%) :
REMARK 240   DATA REDUNDANCY                :
REMARK 240   IN THE HIGHEST RESOLUTION SHELL
REMARK 240   HIGHEST RESOLUTION SHELL, RANGE HIGH (A) :
REMARK 240   HIGHEST RESOLUTION SHELL, RANGE LOW  (A) :
REMARK 240   COMPLETENESS FOR SHELL     (%) :
REMARK 240   DATA REDUNDANCY IN SHELL       :
REMARK 240   R MERGE FOR SHELL          (I) :
REMARK 240   METHOD USED TO DETERMINE THE STRUCTURE:
REMARK 240   SOFTWARE USED                  :
REMARK 240   STARTING MODEL                 :
```

Example

```
REMARK 240
REMARK 240 EXPERIMENTAL DETAILS
REMARK 240   RECONSTRUCTION METHOD          : CRYSTALLOGRAPHY
REMARK 240   SAMPLE TYPE                    : 2D CRYSTAL
REMARK 240   SPECIMEN TYPE                  : VITREOUS ICE (CRYO EM)
REMARK 240 DATA ACQUISITION
REMARK 240   DATE OF DATA COLLECTION        : 01-DEC-03
REMARK 240   TEMPERATURE           (KELVIN) : 300.0
REMARK 240   PH                             : 6.00
REMARK 240   NUMBER OF CRYSTALS USED        : 286
REMARK 240   MICROSCOPE MODEL               : JEM3000SFF
REMARK 240   DETECTOR TYPE                  : CCD
REMARK 240   ACCELERATION VOLTAGE (KV)      : 300
REMARK 240   NUMBER OF UNIQUE REFLECTIONS   : 22293
REMARK 240   RESOLUTION RANGE HIGH      (A) : 1.9
REMARK 240   RESOLUTION RANGE LOW       (A) : 20.000
REMARK 240   DATA SCALING SOFTWARE          : SOFTWARE
REMARK 240   COMPLETENESS FOR RANGE     (%) : 80.0
REMARK 240   DATA REDUNDANCY                : 5.700
```

```
REMARK 240   IN THE HIGHEST RESOLUTION SHELL.
REMARK 240   HIGHEST RESOLUTION SHELL, RANGE HIGH (A) : 1.90
REMARK 240   HIGHEST RESOLUTION SHELL, RANGE LOW  (A) : 2.0
REMARK 240   COMPLETENESS FOR SHELL     (%) : 82.0
REMARK 240   DATA REDUNDANCY IN SHELL       : 5.70
REMARK 240   R MERGE FOR SHELL          (I) : 0.166
REMARK 240   METHOD USED TO DETERMINE THE STRUCTURE: MOLECULAR
REMARK 240                                   REPLACEMENT
REMARK 240   SOFTWARE USED                  : CNS
REMARK 240   STARTING MODEL                 : PDB ENTRY 1SOR
```

REMARK 245 (updated), Electron Microscopy Experiment Details

REMARK 245 is mandatory if data was obtained from a EM study. The format of the date in this remark is DD-MMM-YY. DD is the day of the month (a number 01 through 31), MMM is the English 3-letter abbreviation for the month, and YY is the year.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 245
REMARK 245 EXPERIMENTAL DETAILS
REMARK 245   RECONSTRUCTION METHOD          :
REMARK 245   SPECIMEN TYPE                  :
REMARK 245
REMARK 245 ELECTRON MICROSCOPE SAMPLE
REMARK 245   SAMPLE TYPE                    :
REMARK 245   PARTICLE TYPE                  :
REMARK 245   NAME OF SAMPLE                 :
REMARK 245   SAMPLE CONCENTRATION (MG ML-1) :
REMARK 245   SAMPLE SUPPORT DETAILS         :
REMARK 245   SAMPLE VITRIFICATION DETAILS   :
REMARK 245   SAMPLE BUFFER                  :
REMARK 245   PH                             :
REMARK 245   SAMPLE DETAILS                 :
REMARK 245
REMARK 245 DATA ACQUISITION
REMARK 245   DATE OF EXPERIMENT                :
REMARK 245   NUMBER OF MICROGRAPHS-IMAGES      :
REMARK 245   TEMPERATURE (KELVIN)              :
REMARK 245   MICROSCOPE MODEL                  :
REMARK 245   DETECTOR TYPE                     :
REMARK 245   MINIMUM DEFOCUS (NM)              :
REMARK 245   MAXIMUM DEFOCUS (NM)              :
REMARK 245   MINIMUM TILT ANGLE (DEGREES)      :
REMARK 245   MAXIMUM TILT ANGLE (DEGREES)      :
REMARK 245   NOMINAL CS                        :
REMARK 245   IMAGING MODE                      :
REMARK 245   ELECTRON DOSE (ELECTRONS NM**-2)  :
REMARK 245   ILLUMINATION MODE                 :
REMARK 245   NOMINAL MAGNIFICATION             :
REMARK 245   CALIBRATED MAGNIFICATION          :
REMARK 245   SOURCE                            :
REMARK 245   ACCELERATION VOLTAGE (KV)         :
REMARK 245   IMAGING DETAILS                   :
```

Example

```
REMARK 245
REMARK 245 EXPERIMENTAL DETAILS
REMARK 245   RECONSTRUCTION METHOD          : SINGLE PARTICLE
REMARK 245   SPECIMEN TYPE                  : VITREOUS ICE (CRYO EM)
REMARK 245
REMARK 245 ELECTRON MICROSCOPE SAMPLE
REMARK 245   SAMPLE TYPE                    : PARTICLE
REMARK 245   PARTICLE TYPE                  : MIXED SYMMETRY
REMARK 245   NAME OF SAMPLE                 : BACTERIOPHAGE T4
REMARK 245   SAMPLE CONCENTRATION (MG ML-1) : 20.00
REMARK 245   SAMPLE SUPPORT DETAILS         : NULL
REMARK 245   SAMPLE VITRIFICATION DETAILS   : NULL
REMARK 245   SAMPLE BUFFER                  : H2O
REMARK 245   PH                             : 7.50
REMARK 245   SAMPLE DETAILS                 : PHAGE
REMARK 245
REMARK 245 DATA ACQUISITION
REMARK 245   DATE OF EXPERIMENT                : 06-JAN-02
REMARK 245   NUMBER OF MICROGRAPHS-IMAGES      : NULL
REMARK 245   TEMPERATURE (KELVIN)              : 100.00
REMARK 245   MICROSCOPE MODEL                  : FEI/PHILIPS CM300FEG/T
REMARK 245   DETECTOR TYPE                     : NULL
REMARK 245   MINIMUM DEFOCUS (NM)              : 500.00
REMARK 245   MAXIMUM DEFOCUS (NM)              : 3400.00
REMARK 245   MINIMUM TILT ANGLE (DEGREES)      : 0.00
REMARK 245   MAXIMUM TILT ANGLE (DEGREES)      : 0.00
REMARK 245   NOMINAL CS                        : 1.40
REMARK 245   IMAGING MODE                      : BRIGHT FIELD
REMARK 245   ELECTRON DOSE (ELECTRONS NM**-2)  : 20.00
REMARK 245   ILLUMINATION MODE                 : SPOT SCAN
REMARK 245   NOMINAL MAGNIFICATION             : 45000
REMARK 245   CALIBRATED MAGNIFICATION          : 47000
REMARK 245   SOURCE                            : FIELD EMISSION GUN
REMARK 245   ACCELERATION VOLTAGE (KV)         : 300
```

REMARK 247, Electron Microscopy details

REMARK 247 is mandatory if data was obtained from an EM study.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 247
REMARK 247 ELECTRON MICROSCOPY
REMARK 247   THE COORDINATES IN THIS ENTRY WERE GENERATED FROM
REMARK 247   ELECTRON MICROSCOPY DATA. PROTEIN DATA BANK CONVENTIONS
REMARK 247   REQUIRE THAT CRYST1 AND SCALE RECORDS BE INCLUDED,
REMARK 247   BUT THE VALUES ON THESE RECORDS ARE MEANINGLESS
REMARK 247   EXCEPT FOR THE CALCULATION OF THE STRUCTURE FACTORS
```

REMARK 250, Other Type of Experiment Details

REMARKs specific to other kinds of studies, not listed above.
REMARK 250 is mandatory if other than X-ray, NMR, neutron, or electron study.
The format of the date in this remark is DD-MMM-YY. DD is the day of the month (a number 01 through 31), MMM is the English 3-letter abbreviation for the month, and YY is the year.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 250
REMARK 250 EXPERIMENTAL DETAILS
REMARK 250  EXPERIMENT TYPE                : 
REMARK 250  DATE OF DATA COLLECTION        : 
REMARK 250
REMARK 250 REMARK:
```

REMARK 265, Solution Scattering Experiment Details

Examples

```
REMARK 265
REMARK 265 EXPERIMENTAL DETAILS
REMARK 265
REMARK 265 EXPERIMENT TYPE : SMALL ANGLE X-RAY SCATTERING
REMARK 265  DATA ACQUISITION
REMARK 265   RADIATION/NEUTRON SOURCE                  : SRS BEAMLINE 2.1
REMARK 265   SYNCHROTRON (Y/N)                         : Y
REMARK 265   BEAMLINE                                  : 2.1
REMARK 265   BEAMLINE INSTRUMENT                       : NULL
REMARK 265   DETECTOR TYPE                             : 500-CHANNEL QUADRANT
REMARK 265   DETECTOR MANUFACTURER DETAILS             : NULL
REMARK 265   TEMPERATURE (KELVIN)                      : 288
REMARK 265   PH                                        : NULL
REMARK 265   NUMBER OF TIME FRAMES USED                : 10
REMARK 265   PROTEIN CONCENTRATION RANGE (MG/ML)       : 0.7 - 14
REMARK 265   SAMPLE BUFFER                             : TRIS
REMARK 265   DATA REDUCTION SOFTWARE                   : OTOKO
REMARK 265   DATA ANALYSIS SOFTWARE                    : SCTPL5, GNOM
REMARK 265   GUINIER MEAN RADIUS OF GYRATION (NM)      : 11.1
REMARK 265   SIGMA MEAN RADIUS OF GYRATION             : 0.4
REMARK 265   R(XS-1) MEAN CROSS SECTIONAL RADII (NM)   : 4.4
REMARK 265   R(XS-1) SIGMA MEAN CROSS SECTIONAL RADII : 0.2
REMARK 265   R(XS-2) MEAN CROSS SECTIONAL RADII (NM)   : 1.7
REMARK 265   R(XS-2) SIGMA MEAN CROSS SECTIONAL RADII : 0.1
REMARK 265   P(R) PROTEIN LENGTH (NM)                  : 40
REMARK 265
REMARK 265 EXPERIMENT TYPE : SMALL ANGLE NEUTRON SCATTERING
REMARK 265  DATA ACQUISITION
REMARK 265   RADIATION/NEUTRON SOURCE                  : ILL
REMARK 265   SYNCHROTRON (Y/N)                         : N
REMARK 265   BEAMLINE TYPE                             : NULL
REMARK 265   BEAMLINE INSTRUMENT                       : D11, D22
REMARK 265   DETECTOR TYPE                             : AREA
REMARK 265   DETECTOR MANUFACTURER DETAILS             : NULL
REMARK 265   TEMPERATURE (KELVIN)                      : NULL
REMARK 265   PH                                        : NULL
REMARK 265   NUMBER OF TIME FRAMES USED                : NULL
REMARK 265   PROTEIN CONCENTRATION RANGE (MG/ML)       : 0.4 - 9.6
REMARK 265   SAMPLE BUFFER                             : PBS IN 99.9% D2O
REMARK 265   DATA REDUCTION SOFTWARE                   : DETEC, RNILS, SPOLLY
REMARK 265   DATA ANALYSIS SOFTWARE                    : SCTPL5, GNOM
REMARK 265   GUINIER MEAN RADIUS OF GYRATION (NM)      : 11.3
REMARK 265   SIGMA MEAN RADIUS OF GYRATION             : 0.4
REMARK 265   R(XS-1) MEAN CROSS SECTIONAL RADII (NM)   : 3.9
REMARK 265   R(XS-1) SIGMA MEAN CROSS SECTIONAL RADII : 0.2
REMARK 265   R(XS-2) MEAN CROSS SECTIONAL RADII (NM)   : 1.51
REMARK 265   R(XS-2) SIGMA MEAN CROSS SECTIONAL RADII : 0.06
REMARK 265   P(R) PROTEIN LENGTH (NM)                  : 37 - 39
REMARK 265
REMARK 265  DATA ACQUISITION
REMARK 265   RADIATION/NEUTRON SOURCE                  : ISIS
REMARK 265   SYNCHROTRON (Y/N)                         : N
REMARK 265   BEAMLINE TYPE                             : PULSED NEUTRON
REMARK 265   BEAMLINE INSTRUMENT                       : LOQ
REMARK 265   DETECTOR TYPE                             : AREA (TIME-OF-FLIGHT)
REMARK 265   TEMPERATURE (KELVIN)                      : NULL
REMARK 265   PH                                        : NULL
REMARK 265   NUMBER OF TIME FRAMES USED                : NULL
```

```
REMARK 265   PROTEIN CONCENTRATION RANGE (MG/ML)      : 3.7, 6.1
REMARK 265   SAMPLE BUFFER                            : PBS IN 99.9% D2O
REMARK 265   DATA REDUCTION SOFTWARE                  : COLLETTE
REMARK 265   DATA ANALYSIS SOFTWARE                   : SCTPL5, GNOM
REMARK 265   GUINIER MEAN RADIUS OF GYRATION (NM)     : 11.7
REMARK 265   SIGMA MEAN RADIUS OF GYRATION            : 0.5
REMARK 265   R(XS-1) MEAN CROSS SECTIONAL RADII (NM)  : NULL
REMARK 265   R(XS-1) SIGMA MEAN CROSS SECTIONAL RADII : NULL
REMARK 265   R(XS-2) MEAN CROSS SECTIONAL RADII (NM)  : NULL
REMARK 265   R(XS-2) SIGMA MEAN CROSS SECTIONAL RADII : NULL
REMARK 265   P(R) PROTEIN LENGTH (NM)                 : 40
REMARK 265
REMARK 265 DATA ANALYSIS AND MODEL FITTING:
REMARK 265  METHOD USED TO DETERMINE THE STRUCTURE: CONSTRAINED SCATTERING
REMARK 265                                          FITTING OF HOMOLOGY
REMARK 265                                          MODELS
REMARK 265  SOFTWARE USED    : INSIGHT II, HOMOLOGY, DISCOVERY,
REMARK 265                     BIOPOLYMER, DELPHI, SCTPL5, GNOM
REMARK 265  SOFTWARE AUTHORS : MSI
REMARK 265  STARTING MODEL   : PDB CODE 1HFI, 1HCC, 1HFH, 1VCC
REMARK 265
REMARK 265 CONFORMERS, NUMBER CALCULATED   : 2010
REMARK 265 CONFORMERS, NUMBER SUBMITTED    : 4
REMARK 265 CONFORMERS, SELECTION CRITERIA  :  THE MODELLED SCATTERING
REMARK 265   CURVES WERE ASSESSED BY CALCULATION OF THE
REMARK 265   RG, RSX-1 AND RXS-2 VALUES IN THE SAME Q RANGES
REMARK 265   USED IN THE EXPERIMENTAL GUINIER FITS. MODELS WERE
REMARK 265   THEN RANKED USING A GOODNESS-OF-FIT R-FACTOR
REMARK 265   DEFINED BY ANALOGY WITH PROTEIN CRYSTALLOGRAPHY
REMARK 265   AND BASED ON THE EXPERIMENTAL CURVES IN THE Q RANGE
REMARK 265   EXTENDING TO 1.4 NM-1.
REMARK 265
REMARK 265 REPRESENTATIVE CONFORMER IN THIS ENSEMBLE : 1
REMARK 265
REMARK 265  OTHER DETAILS: HOMOLOGY MODELS WERE BUILT FOR
REMARK 265   THE 17 SCR DOMAINS AND ENERGY MINIMISATIONS WERE
REMARK 265   PERFORMED TO IMPROVE THE CONNECTIVITY IN THE FH MODEL.
REMARK 265   TRIANTENNARY COMPLEX-TYPE CARBOHYDRATE STRUCTURES
REMARK 265   (MAN3GLCNAC6GAL3FUC3NEUNAC1) WERE ADDED TO EACH OF THE
REMARK 265   N-LINKED GLYCOSYLATION SITES. A LIBRARY OF LINKER PEPTIDE
REMARK 265   CONFORMATIONS WAS USED IN DOMAIN MODELLING CONSTRAINED
REMARK 265   BY THE SOLUTION SCATTERING FITS. MODELLING WITH THE
REMARK 265   SCATTERING DATA WAS ALSO CARRIED OUT BY ROTATIONAL
REMARK 265   SEARCH METHODS. THE X-RAY AND NEUTRON SCATTERING CURVE
REMARK 265   I(Q) WAS CALCULATED ASSUMING A UNIFORM SCATTERING DENSITY
REMARK 265   FOR THE SPHERES USING THE DEBYE EQUATION AS ADAPTED TO
REMARK 265   SPHERES. X-RAY CURVES WERE CALCULATED FROM THE HYDRATED
REMARK 265   SPHERE MODELS WITHOUT CORRECTIONS FOR WAVELENGTH SPREAD OR
REMARK 265   BEAM DIVERGENCE, WHILE THESE CORRECTIONS WERE APPLIED FOR
REMARK 265   THE NEUTRON CURVES BUT NOW USING UNHYDRATED MODELS.
```

REMARKs 280-290, Crystallographic Details

REMARK 280, Crystal

REMARK 280 presents information about the crystal. The solvent content and Matthews coefficient are provided for protein and polypeptide crystals. Crystallization conditions are in free text.

REMARK 280 is mandatory for single crystal studies.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 280
REMARK 280 CRYSTAL
REMARK 280 SOLVENT CONTENT, VS   (%):
REMARK 280 MATTHEWS COEFFICIENT, VM (ANGSTROMS**3/DA):
REMARK 280
REMARK 280 CRYSTALLIZATION CONDITIONS: FREE TEXT GOES HERE.
```

Example

```
REMARK 280 CRYSTAL
REMARK 280 SOLVENT CONTENT, VS  (%): 36.85
REMARK 280 MATTHEWS COEFFICIENT, VM (ANGSTROMS**3/DA): 1.79
REMARK 280
REMARK 280 CRYSTALLIZATION CONDITIONS: 1.4M SODIUM ACETATE,
REMARK 280                              0.1M MES PH 6.5
```

REMARK 285, CRYST1

REMARK 285 presents information about the unit cell.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 285
REMARK 285 CRYST1
REMARK 285 FREE TEXT GOES HERE.
```

Example

```
REMARK 285
REMARK 285 CRYST1
REMARK 285 TEXT TO EXPLAIN UNUSUAL UNIT-CELL DATA: THE DATA WAS
REMARK 285 COLLECTED ON TWO-DIMENSIONAL CRYSTALS AND HENCE THE
REMARK 285 C-AXIS REPEAT DOES NOT CORRESPOND TO A REAL REPEAT, BUT
REMARK 285 INSTEAD REFERS TO THE SAMPLING THAT IS USED TO DESCRIBE
REMARK 285 THE CONTINUOUS TRANSFORM. THE C VALUE OF 100.9 IS
REMARK 285 THEREFORE THE VALUE WHICH SHOULD BE USED IN
REMARK 285 INTERPRETING THE MEANING OF THE L INDEX.
```

REMARK 290, Crystallographic Symmetry

REMARK 290 is mandatory for crystalline studies. The REMARK is automatically generated.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 290
REMARK 290 CRYSTALLOGRAPHIC SYMMETRY
REMARK 290 SYMMETRY OPERATORS FOR SPACE GROUP: P 21 21 21
REMARK 290
REMARK 290      SYMOP   SYMMETRY
REMARK 290     NNNMMM   OPERATOR
REMARK 290       1555   X,Y,Z
REMARK 290       2555   1/2-X,-Y,1/2+Z
REMARK 290       3555   -X,1/2+Y,1/2-Z
REMARK 290       4555   1/2+X,1/2-Y,-Z
REMARK 290
REMARK 290     WHERE NNN -> OPERATOR NUMBER
REMARK 290           MMM -> TRANSLATION VECTOR
REMARK 290
REMARK 290 CRYSTALLOGRAPHIC SYMMETRY TRANSFORMATIONS
REMARK 290 THE FOLLOWING TRANSFORMATIONS OPERATE ON THE ATOM/HETATM
REMARK 290 RECORDS IN THIS ENTRY TO PRODUCE CRYSTALLOGRAPHICALLY
REMARK 290 RELATED MOLECULES.
REMARK 290   SMTRY1   1  1.000000  0.000000  0.000000        0.00000
REMARK 290   SMTRY2   1  0.000000  1.000000  0.000000        0.00000
REMARK 290   SMTRY3   1  0.000000  0.000000  1.000000        0.00000
REMARK 290   SMTRY1   2 -1.000000  0.000000  0.000000       36.30027
REMARK 290   SMTRY2   2  0.000000 -1.000000  0.000000        0.00000
REMARK 290   SMTRY3   2  0.000000  0.000000  1.000000       59.50256
REMARK 290   SMTRY1   3 -1.000000  0.000000  0.000000        0.00000
REMARK 290   SMTRY2   3  0.000000  1.000000  0.000000       46.45545
REMARK 290   SMTRY3   3  0.000000  0.000000 -1.000000       59.50256
REMARK 290   SMTRY1   4  1.000000  0.000000  0.000000       36.30027
REMARK 290   SMTRY2   4  0.000000 -1.000000  0.000000       46.45545
REMARK 290   SMTRY3   4  0.000000  0.000000 -1.000000        0.00000
REMARK 290
```

REMARK 300 (updated), Biomolecule

Description of the biologically functional molecule (biomolecule) in free text.
Remark 300 is mandatory if REMARK 350 is provided.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 300
REMARK 300 BIOMOLECULE: 1
REMARK 300 SEE REMARK 350 FOR THE AUTHOR PROVIDED AND/OR PROGRAM
REMARK 300 GENERATED ASSEMBLY INFORMATION FOR THE STRUCTURE IN
REMARK 300 THIS ENTRY.  THE REMARK MAY ALSO PROVIDE INFORMATION ON
REMARK 300 BURIED SURFACE AREA.
REMARK 300 FREE TEXT GOES HERE.
```

Examples

```
REMARK 300
REMARK 300 BIOMOLECULE: 1, 2, 3, 4, 5, 6, 7, 8
REMARK 300 SEE REMARK 350 FOR THE AUTHOR PROVIDED AND/OR PROGRAM
REMARK 300 GENERATED ASSEMBLY INFORMATION FOR THE STRUCTURE IN
REMARK 300 THIS ENTRY.  THE REMARK MAY ALSO PROVIDE INFORMATION ON
REMARK 300 BURIED SURFACE AREA.

REMARK 300
REMARK 300 BIOMOLECULE: 1, 2, 3
REMARK 300 SEE REMARK 350 FOR THE AUTHOR PROVIDED AND/OR PROGRAM
REMARK 300 GENERATED ASSEMBLY INFORMATION FOR THE STRUCTURE IN
REMARK 300 THIS ENTRY.  THE REMARK MAY ALSO PROVIDE INFORMATION ON
REMARK 300 BURIED SURFACE AREA.
REMARK 300 DETAILS: THE CATALYTIC SUBUNIT OF LIVER ALCOHOL DEHYDROGENASE FROM
REMARK 300 EQUUS CABALLUS IS A HOMODIMER.
```

Example - Icosahedral virus

```
REMARK 300
REMARK 300 BIOMOLECULE: 1
REMARK 300 SEE REMARK 350 FOR THE AUTHOR PROVIDED AND/OR PROGRAM
REMARK 300 GENERATED ASSEMBLY INFORMATION FOR THE STRUCTURE IN
REMARK 300 THIS ENTRY.  THE REMARK MAY ALSO PROVIDE INFORMATION ON
REMARK 300 BURIED SURFACE AREA.
REMARK 300 DETAILS: THE ASSEMBLY REPRESENTED IN THIS ENTRY HAS REGULAR
REMARK 300 ICOSAHEDRAL POINT SYMMETRY (SCHOENFLIES SYMBOL = I).
```

Example - Helical viruses

```
REMARK 300
REMARK 300 BIOMOLECULE: 1
REMARK 300 SEE REMARK 350 FOR THE AUTHOR PROVIDED AND/OR PROGRAM
REMARK 300 GENERATED ASSEMBLY INFORMATION FOR THE STRUCTURE IN
REMARK 300 THIS ENTRY.  THE REMARK MAY ALSO PROVIDE INFORMATION ON
REMARK 300 BURIED SURFACE AREA.
REMARK 300 DETAILS: THE ASSEMBLY REPRESENTED IN THIS ENTRY HAS REGULAR
REMARK 300 HELICAL SYMMETRY WITH THE FOLLOWING PARAMETERS:
REMARK 300 ROTATION PER SUBUNIT (TWIST) = -33.23 DEGREES
REMARK 300 RISE PER SUBUNIT (HEIGHT) = 16.00 ANGSTROMS
REMARK 300 IN ADDITION, THERE IS 5-FOLD CIRCULAR
REMARK 300 SYMMETRY AROUND THE HELIX AXIS
```

Example - point symmetry crystal structure

```
REMARK 300
REMARK 300 BIOMOLECULE: 1
REMARK 300 SEE REMARK 350 FOR THE AUTHOR PROVIDED AND/OR PROGRAM
REMARK 300 GENERATED ASSEMBLY INFORMATION FOR THE STRUCTURE IN
REMARK 300 THIS ENTRY.  THE REMARK MAY ALSO PROVIDE INFORMATION ON
REMARK 300 BURIED SURFACE AREA.
REMARK 300 DETAILS: THE ASSEMBLY REPRESENTED IN THIS ENTRY HAS
REMARK 300 REGULAR DIHEDRAL POINT SYMMETRY (SCHOENFLIES SYMBOL = D17).
```

REMARK 350 (updated), Generating the Biomolecule

REMARK 350 presents all transformations, both crystallographic and non-crystallographic, needed to generate the biomolecule. These transformations operate on the coordinates in the entry. Both author and computational descriptions of assemblies are provided, if applicable.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 350 COORDINATES FOR A COMPLETE MULTIMER REPRESENTING THE KNOWN
REMARK 350 BIOLOGICALLY SIGNIFICANT OLIGOMERIZATION STATE OF THE
REMARK 350 MOLECULE CAN BE GENERATED BY APPLYING BIOMT TRANSFORMATIONS
REMARK 350 GIVEN BELOW.  BOTH NON-CRYSTALLOGRAPHIC AND
REMARK 350 CRYSTALLOGRAPHIC OPERATIONS ARE GIVEN.
REMARK 350
REMARK 350 BIOMOLECULE: 1
REMARK 350 AUTHOR DETERMINED BIOLOGICAL UNIT: DODECAMERIC
REMARK 350 SOFTWARE DETERMINED QUATERNARY STRUCTURE: DODECAMERIC
REMARK 350 SOFTWARE USED: PISA
REMARK 350 TOTAL BURIED SURFACE AREA: 2990 ANGSTROM**2
REMARK 350 SURFACE AREA OF THE COMPLEX: 9330 ANGSTROM**2
REMARK 350 CHANGE IN SOLVENT FREE ENERGY: -40.0 KCAL/MOL
REMARK 350 APPLY THE FOLLOWING TO CHAINS: A, B, C, D, E, F, G, H, I,
REMARK 350                    AND CHAINS: J, K, L
REMARK 350   BIOMT1   1  1.000000  0.000000  0.000000        0.00000
REMARK 350   BIOMT2   1  0.000000  1.000000  0.000000        0.00000
REMARK 350   BIOMT3   1  0.000000  0.000000  1.000000        0.00000
```

Note: If entry is part of a SPLIT record (larger multi-protein complex), REMARK 350 represents only the quaternary structure of that split entry.

```
REMARK 350 COORDINATES FOR A COMPLETE MULTIMER REPRESENTING THE KNOWN
REMARK 350 BIOLOGICALLY SIGNIFICANT OLIGOMERIZATION STATE OF THE
REMARK 350 MOLECULE CAN BE GENERATED BY APPLYING BIOMT TRANSFORMATIONS
REMARK 350 GIVEN BELOW.  BOTH NON-CRYSTALLOGRAPHIC AND
REMARK 350 CRYSTALLOGRAPHIC OPERATIONS ARE GIVEN.
REMARK 350
REMARK 350 BIOMOLECULE: 1
REMARK 350 QUATERNARY STRUCTURE FOR THIS ENTRY: 21MERIC
REMARK 350 APPLY THE FOLLOWING TO CHAINS: A, B, C, D, E, F, G, H, I,
REMARK 350                    AND CHAINS: J, K, L, M, N, O, P, Q, T,
REMARK 350                    AND CHAINS: S, T, U
REMARK 350   BIOMT1   1  1.000000  0.000000  0.000000        0.00000
REMARK 350   BIOMT2   1  0.000000  1.000000  0.000000        0.00000
REMARK 350   BIOMT3   1  0.000000  0.000000  1.000000        0.00000
```

Example – Author and computed assembly predictions agree

```
REMARK 350 COORDINATES FOR A COMPLETE MULTIMER REPRESENTING THE KNOWN
REMARK 350 BIOLOGICALLY SIGNIFICANT OLIGOMERIZATION STATE OF THE
REMARK 350 MOLECULE CAN BE GENERATED BY APPLYING BIOMT TRANSFORMATIONS
REMARK 350 GIVEN BELOW.  BOTH NON-CRYSTALLOGRAPHIC AND
REMARK 350 CRYSTALLOGRAPHIC OPERATIONS ARE GIVEN.
REMARK 350
REMARK 350 BIOMOLECULE: 1
REMARK 350 AUTHOR DETERMINED BIOLOGICAL UNIT: DODECAMERIC
REMARK 350 SOFTWARE DETERMINED QUATERNARY STRUCTURE: DODECAMERIC
REMARK 350 SOFTWARE USED: PISA
REMARK 350 TOTAL BURIED SURFACE AREA: 2990 ANGSTROM**2
REMARK 350 SURFACE AREA OF THE COMPLEX: 9330 ANGSTROM**2
REMARK 350 CHANGE IN SOLVENT FREE ENERGY: -40.0 KCAL/MOL
REMARK 350 APPLY THE FOLLOWING TO CHAINS: A, B, C, D, E, F, G, H, I,
REMARK 350                    AND CHAINS: J, K, L
REMARK 350   BIOMT1   1  1.000000  0.000000  0.000000        0.00000
REMARK 350   BIOMT2   1  0.000000  1.000000  0.000000        0.00000
REMARK 350   BIOMT3   1  0.000000  0.000000  1.000000        0.00000
```

Note: The value for the average buried surface area will be round to the nearest 10.

Example – Author and computed assembly predictions differ

```
REMARK 350 COORDINATES FOR A COMPLETE MULTIMER REPRESENTING THE KNOWN
REMARK 350 BIOLOGICALLY SIGNIFICANT OLIGOMERIZATION STATE OF THE
REMARK 350 MOLECULE CAN BE GENERATED BY APPLYING BIOMT TRANSFORMATIONS
REMARK 350 GIVEN BELOW. BOTH NON-CRYSTALLOGRAPHIC AND
REMARK 350 CRYSTALLOGRAPHIC OPERATIONS ARE GIVEN.
REMARK 350
REMARK 350 BIOMOLECULE: 1
REMARK 350 AUTHOR DETERMINED BIOLOGICAL UNIT: HEXAMERIC
REMARK 350 APPLY THE FOLLOWING TO CHAINS: A, B, C, D, E, F
REMARK 350   BIOMT1   1  1.000000  0.000000  0.000000        0.00000
REMARK 350   BIOMT2   1  0.000000  1.000000  0.000000        0.00000
REMARK 350   BIOMT3   1  0.000000  0.000000  1.000000        0.00000
REMARK 350
REMARK 350 BIOMOLECULE: 2
REMARK 350 AUTHOR DETERMINED BIOLOGICAL UNIT: HEXAMERIC
REMARK 350 APPLY THE FOLLOWING TO CHAINS: G, H, I, J, K, L
REMARK 350   BIOMT1   1  1.000000  0.000000  0.000000        0.00000
REMARK 350   BIOMT2   1  0.000000  1.000000  0.000000        0.00000
REMARK 350   BIOMT3   1  0.000000  0.000000  1.000000        0.00000
REMARK 350
REMARK 350 BIOMOLECULE: 3
REMARK 350 SOFTWARE DETERMINED QUATERNARY STRUCTURE: DODECAMERIC
REMARK 350 SOFTWARE USED: PISA
REMARK 350 TOTAL BURIED SURFACE AREA: 2990 ANGSTROM**2
REMARK 350 SURFACE AREA OF THE COMPLEX: 9330 ANGSTROM**2
REMARK 350 CHANGE IN SOLVENT FREE ENERGY: -40.0 KCAL/MOL
REMARK 350 APPLY THE FOLLOWING TO CHAINS: A, B, C, D, E, F, G, H, I,
REMARK 350                    AND CHAINS: J, K, L
REMARK 350   BIOMT1   1  1.000000  0.000000  0.000000        0.00000
REMARK 350   BIOMT2   1  0.000000  1.000000  0.000000        0.00000
REMARK 350   BIOMT3   1  0.000000  0.000000  1.000000        0.00000
```

Example – When there are no quaternary assemblies provided by either author or software

```
REMARK 350 COORDINATES FOR A COMPLETE MULTIMER REPRESENTING THE KNOWN
REMARK 350 BIOLOGICALLY SIGNIFICANT OLIGOMERIZATION STATE OF THE
```

```
REMARK 350 MOLECULE CAN BE GENERATED BY APPLYING BIOMT TRANSFORMATIONS
REMARK 350 GIVEN BELOW. BOTH NON-CRYSTALLOGRAPHIC AND
REMARK 350 CRYSTALLOGRAPHIC OPERATIONS ARE GIVEN.
REMARK 350
REMARK 350 BIOMOLECULE:  1
REMARK 350 SOFTWARE DETERMINED QUATERNARY STRUCTURE: MONOMERIC
REMARK 350 SOFTWARE USED: PISA
REMARK 350 APPLY THE FOLLOWING TO CHAINS: A
REMARK 350   BIOMT1   1  1.000000  0.000000  0.000000        0.00000
REMARK 350   BIOMT2   1  0.000000  1.000000  0.000000        0.00000
REMARK 350   BIOMT3   1  0.000000  0.000000  1.000000        0.00000
```

Note that the average buried surface area is not included in this example because the quaternary structure is a monomer.

Example – When software predicts multiple quaternary assemblies

For example, the author states the biological unit to be a dimer, but software predicts the quaternary structure to be either a dimer or a tetramer:

```
REMARK 300
REMARK 300
REMARK 300 BIOMOLECULE: 1, 2
REMARK 300 SEE REMARK 350 FOR THE AUTHOR PROVIDED AND/OR PROGRAM
REMARK 300 GENERATED ASSEMBLY INFORMATION FOR THE STRUCTURE IN
REMARK 300 THIS ENTRY.  THE REMARK MAY ALSO PROVIDE INFORMATION ON
REMARK 300 BURIED SURFACE AREA.
REMARK 300
REMARK 350 COORDINATES FOR A COMPLETE MULTIMER REPRESENTING THE KNOWN
REMARK 350 BIOLOGICALLY SIGNIFICANT OLIGOMERIZATION STATE OF THE
REMARK 350 MOLECULE CAN BE GENERATED BY APPLYING BIOMT TRANSFORMATIONS
REMARK 350 GIVEN BELOW. BOTH NON-CRYSTALLOGRAPHIC AND
REMARK 350 CRYSTALLOGRAPHIC OPERATIONS ARE GIVEN.
REMARK 350
REMARK 350 BIOMOLECULE:  1
REMARK 350 AUTHOR DETERMINED BIOLOGICAL UNIT: DIMERIC
REMARK 350 SOFTWARE DETERMINED QUATERNARY STRUCTURE: DIMERIC
REMARK 350 SOFTWARE USED: PISA
REMARK 350 TOTAL BURIED SURFACE AREA: 1460 ANGSTROM**2
REMARK 350 SURFACE AREA OF THE COMPLEX: 9330 ANGSTROM**2
REMARK 350 CHANGE IN SOLVENT FREE ENERGY: -40.0 KCAL/MOL
REMARK 350 APPLY THE FOLLOWING TO CHAINS: A, B
REMARK 350   BIOMT1   1  1.000000  0.000000  0.000000        0.00000
REMARK 350   BIOMT2   1  0.000000  1.000000  0.000000        0.00000
REMARK 350   BIOMT3   1  0.000000  0.000000  1.000000        0.00000
REMARK 350
REMARK 350 BIOMOLECULE:  2
REMARK 350 SOFTWARE DETERMINED QUATERNARY STRUCTURE: TETRAMERIC
REMARK 350 SOFTWARE USED: PISA
REMARK 350 TOTAL BURIED SURFACE AREA: 2860 ANGSTROM**2
REMARK 350 SURFACE AREA OF THE COMPLEX: 12330 ANGSTROM**2
REMARK 350 GAIN IN SOLVENT FREE ENERGY: -20.5 KCAL/MOL
REMARK 350 APPLY THE FOLLOWING TO CHAINS: A, B
REMARK 350   BIOMT1   1  1.000000  0.000000  0.000000        0.00000
REMARK 350   BIOMT2   1  0.000000  1.000000  0.000000        0.00000
REMARK 350   BIOMT3   1  0.000000  0.000000  1.000000        0.00000
REMARK 350   BIOMT1   2 -1.000000  0.000000  0.000000        0.00000
REMARK 350   BIOMT2   2  0.000000  1.000000  0.000000        0.00000
REMARK 350   BIOMT3   2  0.000000  0.000000 -1.000000        0.00000
```

REMARK 375 (updated), Special Position

REMARK 375 specifies atoms which lie within 0.15A of a symmetry-related atom and therefore, are considered to be on a special position, with cumulative occupancies of such atoms not exceeding 1.0.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 375
REMARK 375 SPECIAL POSITION
REMARK 375 FREE TEXT GOES HERE.
```

Example

```
REMARK 375
REMARK 375 SPECIAL POSITION
REMARK 375      HOH A 301  LIES ON A SPECIAL POSITION.
REMARK 375      HOH A  77  LIES ON A SPECIAL POSITION.
REMARK 375
REMARK 375 SPECIAL POSITION
REMARK 375      HOH A  13  LIES ON A SPECIAL POSITION.
REMARK 375      HOH A  28  LIES ON A SPECIAL POSITION.
REMARK 375      HOH A  36  LIES ON A SPECIAL POSITION.
```

REMARK 400, Compound

Further details about the macromolecular contents of the entry.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 400
REMARK 400 COMPOUND
REMARK 400 FREE TEXT GOES HERE.
```

Examples

```
REMARK 400 COMPOUND
REMARK 400 THE PRD1 SUS1 MUTANT LACKS THE PACKAGING PROTEIN P9
REMARK 400 AND PRODUCES ONLY EMPTY PARTICLES, WHICH REPRESENT
REMARK 400 AN ASSEMBLY INTERMEDIATE

REMARK 400
REMARK 400 COMPOUND
REMARK 400 COMPONENT OF NAPHTHALENE DIOXYGENASE (NDO)
REMARK 400 MULTICOMPONENT ENZYME SYSTEM WHICH CATALYZES THE INCORPORATION
REMARK 400 OF BOTH ATOMS OF MOLECULAR OXYGEN INTO NAPHTHALENE TO FORM
REMARK 400 CIS-NAPHTHALENE DIHYDRODIOL.
```

REMARK 450, Source

Further details about the biological source of the macromolecular contents of the entry.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 450
REMARK 450 SOURCE
REMARK 450 FREE TEXT GOES HERE.
```

REMARK 465 (updated), Missing residues

REMARK 465 lists the residues that are present in the SEQRES records but are completely absent from the coordinates section.

Template for non NMR entries

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 465
REMARK 465 MISSING RESIDUES
REMARK 465 THE FOLLOWING RESIDUES WERE NOT LOCATED IN THE
REMARK 465 EXPERIMENT. (M=MODEL NUMBER; RES=RESIDUE NAME; C=CHAIN
REMARK 465 IDENTIFIER; SSSEQ=SEQUENCE NUMBER; I=INSERTION CODE.)
REMARK 465
REMARK 465   M RES C SSSEQI
```

Example

```
REMARK 465
REMARK 465 MISSING RESIDUES
REMARK 465 THE FOLLOWING RESIDUES WERE NOT LOCATED IN THE
REMARK 465 EXPERIMENT. (M=MODEL NUMBER; RES=RESIDUE NAME; C=CHAIN
REMARK 465 IDENTIFIER; SSSEQ=SEQUENCE NUMBER; I=INSERTION CODE.)
REMARK 465
REMARK 465   M RES C SSSEQI
REMARK 465     ARG A    46
REMARK 465     GLY A    47
REMARK 465     ALA A    48
REMARK 465     ARG A    49
REMARK 465     MET A    50
```

Template for NMR entries (added)

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 465
REMARK 465 MISSING RESIDUES
REMARK 465 THE FOLLOWING RESIDUES WERE NOT LOCATED IN THE
REMARK 465 EXPERIMENT. (RES=RESIDUE NAME; C=CHAIN IDENTIFIER;
REMARK 465 SSSEQ=SEQUENCE NUMBER; I=INSERTION CODE.)
REMARK 465   MODELS X-YYY
REMARK 465     RES C SSSEQI
```

`The models is listed as a range, X-YYY.`

Example

```
REMARK 465
REMARK 465 MISSING RESIDUES
REMARK 465 THE FOLLOWING RESIDUES WERE NOT LOCATED IN THE
REMARK 465 EXPERIMENT. (RES=RESIDUE NAME; C=CHAIN IDENTIFIER;
REMARK 465 SSSEQ=SEQUENCE NUMBER; I=INSERTION CODE.)
REMARK 465   MODELS 1-20
REMARK 465     RES C SSSEQI
REMARK 465     MET A     1
REMARK 465     GLY A     2
```

REMARK 470 (updated), Missing Atom(s)

Non-hydrogen atoms of standard residues which are missing from the coordinates are listed. Missing HETATMs (atoms within hetetrogen groups) are not listed here.

Template for non NMR entries

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 470
REMARK 470 MISSING ATOM
REMARK 470 THE FOLLOWING RESIDUES HAVE MISSING ATOMS (M=MODEL NUMBER;
REMARK 470 RES=RESIDUE NAME; C=CHAIN IDENTIFIER; SSEQ=SEQUENCE NUMBER;
REMARK 470 I=INSERTION CODE):
REMARK 470   M RES CSSEQI  ATOMS
```

Example

```
REMARK 470
REMARK 470 MISSING ATOM
REMARK 470 THE FOLLOWING RESIDUES HAVE MISSING ATOMS (M=MODEL NUMBER;
REMARK 470 RES=RESIDUE NAME; C=CHAIN IDENTIFIER; SSEQ=SEQUENCE NUMBER;
REMARK 470 I=INSERTION CODE):
REMARK 470   M RES CSSEQI  ATOMS
REMARK 470     ARG A 412    CG   CD   NE   CZ   NH1  NH2
REMARK 470     ARG A 456    CG   CD   NE   CZ   NH1  NH2
REMARK 470     GLU A 486    CG   CD   OE1  OE2
REMARK 470     GLU A 547    CG   CD   OE1  OE2
REMARK 470     GLU A 548    CG   CD   OE1  OE2
REMARK 470     LYS A 606    CG   CD   CE   NZ
REMARK 470     ARG B 456    CG   CD   NE   CZ   NH1  NH2
REMARK 470     ASP B 484    CG   OD1  OD2
REMARK 470     GLN B 485    CG   CD   OE1  NE2
REMARK 470     GLU B 486    CG   CD   OE1  OE2
REMARK 470     ARG B 490    CG   CD   NE   CZ   NH1  NH2
REMARK 470     GLU B 522    CG   CD   OE1  OE2
REMARK 470     ARG B 576    CG   CD   NE   CZ   NH1  NH2
REMARK 470     ASP B 599    CG   OD1  OD2
```

Template for NMR entries (added)

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

```
REMARK 470 MISSING ATOM
REMARK 470 THE FOLLOWING RESIDUES HAVE MISSING ATOMS(RES=RESIDUE NAME;
REMARK 470 C=CHAIN IDENTIFIER; SSEQ=SEQUENCE NUMBER; I=INSERTION CODE):
REMARK 470   MODELS X-YYY
REMARK 470     RES CSSEQI  ATOMS
```

The models is listed as a range, X-YYY.

Example

```
REMARK 470 MISSING ATOM
REMARK 470 THE FOLLOWING RESIDUES HAVE MISSING ATOMS(RES=RESIDUE NAME;
REMARK 470 C=CHAIN IDENTIFIER; SSEQ=SEQUENCE NUMBER; I=INSERTION CODE):
REMARK 470   MODELS 1-25
REMARK 470     RES CSSEQI  ATOMS
REMARK 470     ILE A  20    CD1
REMARK 470     THR A  59    CG2
```

REMARK 475 (added), Residues modeled with zero occupancy

REMARK 475 enumerates residues modeled with zero occupancy.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 475
REMARK 475 ZERO OCCUPANCY RESIDUES
REMARK 475 THE FOLLOWING RESIDUES WERE MODELED WITH ZERO OCCUPANCY.
REMARK 475 THE LOCATION AND PROPERTIES OF THESE RESIDUES MAY NOT
REMARK 475 BE RELIABLE. (M=MODEL NUMBER; RES=RESIDUE NAME;
REMARK 475 C=CHAIN IDENTIFIER; SSEQ=SEQUENCE NUMBER; I=INSERTION CODE)
REMARK 475   M RES C SSEQI
```

Examples

```
REMARK 475
REMARK 475 ZERO OCCUPANCY RESIDUES
REMARK 475 THE FOLLOWING RESIDUES WERE MODELED WITH ZERO OCCUPANCY.
REMARK 475 THE LOCATION AND PROPERTIES OF THESE RESIDUES MAY NOT
REMARK 475 BE RELIABLE. (M=MODEL NUMBER; RES=RESIDUE NAME;
REMARK 475 C=CHAIN IDENTIFIER; SSEQ=SEQUENCE NUMBER; I=INSERTION CODE)
REMARK 475   M RES C SSEQI
REMARK 475      DG D    4

REMARK 475
REMARK 475 ZERO OCCUPANCY RESIDUES
REMARK 475 THE FOLLOWING RESIDUES WERE MODELED WITH ZERO OCCUPANCY.
REMARK 475 THE LOCATION AND PROPERTIES OF THESE RESIDUES MAY NOT
REMARK 475 BE RELIABLE. (M=MODEL NUMBER; RES=RESIDUE NAME;
REMARK 475 C=CHAIN IDENTIFIER; SSEQ=SEQUENCE NUMBER; I=INSERTION CODE)
REMARK 475   M RES C SSEQI
REMARK 475     GLY A   24
```

REMARK 480 (added), Polymer atoms modeled with zero occupancy

REMARK 480 enumerates non-hydrogen atoms in residues modeled with zero occupancy.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 480
REMARK 480 ZERO OCCUPANCY ATOM
REMARK 480 THE FOLLOWING RESIDUES HAVE ATOMS MODELED WITH ZERO
REMARK 480 OCCUPANCY. THE LOCATION AND PROPERTIES OF THESE ATOMS
REMARK 480 MAY NOT BE RELIABLE. (M=MODEL NUMBER; RES=RESIDUE NAME;
REMARK 480 C=CHAIN IDENTIFIER; SSEQ=SEQUENCE NUMBER; I=INSERTION CODE):
REMARK 480   M RES C SSEQI ATOMS
```

Examples

```
REMARK 480
REMARK 480 ZERO OCCUPANCY ATOM
REMARK 480 THE FOLLOWING RESIDUES HAVE ATOMS MODELED WITH ZERO
REMARK 480 OCCUPANCY. THE LOCATION AND PROPERTIES OF THESE ATOMS
REMARK 480 MAY NOT BE RELIABLE. (M=MODEL NUMBER; RES=RESIDUE NAME;
REMARK 480 C=CHAIN IDENTIFIER; SSEQ=SEQUENCE NUMBER; I=INSERTION CODE):
REMARK 480   M RES C SSEQI ATOMS
REMARK 480      DC D    3   C4'   O4'   C1'   C3'   O3'

REMARK 480
REMARK 480 ZERO OCCUPANCY ATOM
REMARK 480 THE FOLLOWING RESIDUES HAVE ATOMS MODELED WITH ZERO
REMARK 480 OCCUPANCY. THE LOCATION AND PROPERTIES OF THESE ATOMS
REMARK 480 MAY NOT BE RELIABLE. (M=MODEL NUMBER;
REMARK 480 RES=RESIDUE NAME; C=CHAIN IDENTIFIER; SSEQ=SEQUENCE NUMBER;
REMARK 480 I=INSERTION CODE):
REMARK 480   M RES C SSEQI  ATOMS
REMARK 480     HIS A   26    CG  ND1  CD2  CE1  NE2
REMARK 480     HIS B   26    CB   CG  ND1  CD2  CE1  NE2
REMARK 480     GLU B   52    CD  OE1  OE2
```

REMARK 500 (updated), Geometry and Stereochemistry

REMARK 500 provides further details about the stereochemistry of the structure. This REMARK is generated automatically and may incorporate comments provided by the author. It is currently divided into the subtopics:

- CLOSE CONTACTS IN SAME ASYMMETRIC UNIT,
- CLOSE CONTACTS,
- COVALENT BOND LENGTHS,
- COVALENT BOND ANGLES,
- TORSION ANGLES,
- NON-CIS & NON-TRANS,
- PLANAR GROUPS,
- MAIN CHAIN PLANARITY,
- CHIRAL CENTERS.

Additional subtopics may be added as needed. For close contacts, the cutoff limit is 2.2 Angstroms for non-hydrogen atoms and is 1.6 Angstroms for H and D atoms. These distances are listed in the REMARK 500 for close contacts symmetry.

All the calculations on RMSD deviations include all the atoms present in the coordinates including atoms with zero occupancy.

The calculation of bond and angle deviations for protein entries will be based on the updated Engh & Huber amino acid target values[1]. For nucleic acids, the Parkinson et al., statistics will be used for these calculations[2]. All bonds and angles that deviate more than 6 times from their standard target values will be flagged as a deviation. The PHI/PSI values are based on the Kleywegt-Jones calculations[3].

The improper CA-C-CB-N angles for chiral centers are calculated and are defined below with 10 degree allowed deviations.
+35 for L amino acids
-35 for D amino acids

- +25 to +45 degree range is defined as sp3, L.
  If D is expected, it gives "WRONG HAND" in the details. If the calculated value is positive and outside this range, it gives "OUTSIDE RANGE" in the details.
- -10 to +10 degree range is defined as sp2, planar.
  If it is expected to be sp2 and the value is outside this range, it gives "EXPECTING PLANAR" in the details. If it is expected to be sp3 and the value is within this range, it gives "EXPECTING SP3" in the details.
- -45 to -25 degree range is defined as sp3, D.

---

[1] Structure quality and target parameters. R. A. Engh and R. Huber. International Tables for Crystallography (2006). Vol. F, ch. 18.3, pp. 382-392

[2] "New Parameters for the Refinement of Nucleic Acid Containing Structures." G. Parkinson, J. Vojtechovsky, L. Clowney, A. Brunger*, and H. M. Berman. (1996) Acta Cryst. D 52, 57-64

[3] "PHI/PSI- Chology: Ramachandran revisited. " G.J. Kleywegt and T.A. Jones (1996) Structure 4, 1395-1400.

If L is expected, it gives "WRONG HAND" in the details. If the calculated value is negative and outside this range, it gives "OUTSIDE RANGE" in the details.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 500
REMARK 500 GEOMETRY AND STEREOCHEMISTRY
REMARK 500 SUBTOPIC:
REMARK 500
REMARK 500 FREE TEXT GOES HERE.
```

Example – Close Contacts in the same asymmetric unit

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 500
REMARK 500 GEOMETRY AND STEREOCHEMISTRY
REMARK 500 SUBTOPIC: CLOSE CONTACTS IN SAME ASYMMETRIC UNIT
REMARK 500
REMARK 500 THE FOLLOWING ATOMS ARE IN CLOSE CONTACT.
REMARK 500
REMARK 500  ATM1  RES C  SSEQI   ATM2  RES C  SSEQI           DISTANCE
REMARK 500   N    PHE 1     8     OD2  ASP 1    31              2.17
REMARK 500   OD2  ASP 1    31     N    PHE 1     8              2.17
REMARK 500
REMARK 500 THIS ENTRY HAS     104 CLOSE CONTACTS
REMARK 500
REMARK 500 REMARK: NULL
```

Example – Close Contacts

```
REMARK 500
REMARK 500 GEOMETRY AND STEREOCHEMISTRY
REMARK 500 SUBTOPIC: CLOSE CONTACTS
REMARK 500
REMARK 500 THE FOLLOWING ATOMS THAT ARE RELATED BY CRYSTALLOGRAPHIC
REMARK 500 SYMMETRY ARE IN CLOSE CONTACT.  AN ATOM LOCATED WITHIN 0.15
REMARK 500 ANGSTROMS OF A SYMMETRY RELATED ATOM IS ASSUMED TO BE ON A
REMARK 500 SPECIAL POSITION AND IS, THEREFORE, LISTED IN REMARK 375
REMARK 500 INSTEAD OF REMARK 500.  ATOMS WITH NON-BLANK ALTERNATE
REMARK 500 LOCATION INDICATORS ARE NOT INCLUDED IN THE CALCULATIONS.
REMARK 500
REMARK 500 DISTANCE CUTOFF:
REMARK 500 2.2 ANGSTROMS FOR CONTACTS NOT INVOLVING HYDROGEN ATOMS
REMARK 500 1.6 ANGSTROMS FOR CONTACTS INVOLVING HYDROGEN ATOMS
REMARK 500
REMARK 500  ATM1  RES C  SSEQI   ATM2  RES C  SSEQI  SSYMOP   DISTANCE
REMARK 500   O    ALA G   153     OD1  ASP H    46     2565     1.84
REMARK 500   CB   ALA G   153     OD1  ASP H    46     2565     2.18
REMARK 500
REMARK 500 THIS ENTRY HAS      64 SYMMETRY CONTACTS
REMARK 500
REMARK 500 REMARK: NULL
```

Example – Covalent bond lengths

```
         1         2         3         4         5         6         7
1234567890123456789012345678901234567890123456789012345678901234567890
REMARK 500
REMARK 500 GEOMETRY AND STEREOCHEMISTRY
REMARK 500 SUBTOPIC: COVALENT BOND LENGTHS
REMARK 500
REMARK 500 THE STEREOCHEMICAL PARAMETERS OF THE FOLLOWING RESIDUES
REMARK 500 HAVE VALUES WHICH DEVIATE FROM EXPECTED VALUES BY MORE
REMARK 500 THAN 6*RMSD (M=MODEL NUMBER; RES=RESIDUE NAME; C=CHAIN
REMARK 500 IDENTIFIER; SSEQ=SEQUENCE NUMBER; I=INSERTION CODE).
REMARK 500
REMARK 500 STANDARD TABLE:
REMARK 500 FORMAT: (10X,I3,1X,2(A3,1X,A1,I4,A1,1X,A4,3X),1X,F6.3)
REMARK 500
REMARK 500 EXPECTED VALUES PROTEIN: ENGH AND HUBER, 1999
REMARK 500 EXPECTED VALUES NUCLEIC ACID: CLOWNEY ET AL 1996
REMARK 500
REMARK 500  M RES CSSEQI ATM1   RES CSSEQI ATM2   DEVIATION
REMARK 500    ASN B 117   C     ASP B 117   O       0.129
REMARK 500    CYS J  29   CB    CYS J  29   SG     -0.111
REMARK 500
REMARK 500 REMARK: NULL
```

Example – Covalent bond angles

```
REMARK 500
REMARK 500 GEOMETRY AND STEREOCHEMISTRY
REMARK 500 SUBTOPIC: COVALENT BOND ANGLES
REMARK 500
REMARK 500 THE STEREOCHEMICAL PARAMETERS OF THE FOLLOWING RESIDUES
REMARK 500 HAVE VALUES WHICH DEVIATE FROM EXPECTED VALUES BY MORE
REMARK 500 THAN 6*RMSD (M=MODEL NUMBER; RES=RESIDUE NAME; C=CHAIN
REMARK 500 IDENTIFIER; SSEQ=SEQUENCE NUMBER; I=INSERTION CODE).
REMARK 500
REMARK 500 STANDARD TABLE:
REMARK 500 FORMAT: (10X,I3,1X,A3,1X,A1,I4,A1,3(1X,A4,2X),12X,F5.1)
REMARK 500
REMARK 500 EXPECTED VALUES: ENGH AND HUBER, 1999
REMARK 500 EXPECTED VALUES NUCLEIC ACID: CLOWNEY ET AL 1996
REMARK 500
REMARK 500  M RES CSSEQI ATM1   ATM2   ATM3
REMARK 500    VAL A 124   CB  -  CA  -  C   ANGL. DEV. = -12.0 DEGREES
REMARK 500    PRO B 109   CA  -  N   -  CD  ANGL. DEV. =  -3.7 DEGREES
REMARK 500
REMARK 500 REMARK: NULL
```

Example – Torsion angles

```
         1         2         3         4         5         6         7
1234567890123456789012345678901234567890123456789012345678901234567890
REMARK 500
REMARK 500 GEOMETRY AND STEREOCHEMISTRY
REMARK 500 SUBTOPIC: TORSION ANGLES
REMARK 500
REMARK 500 TORSION ANGLES OUTSIDE THE EXPECTED RAMACHANDRAN REGIONS:
REMARK 500 (M=MODEL NUMBER; RES=RESIDUE NAME; C=CHAIN IDENTIFIER;
REMARK 500 SSEQ=SEQUENCE NUMBER; I=INSERTION CODE).
REMARK 500
REMARK 500 STANDARD TABLE:
REMARK 500 FORMAT:(10X,I3,1X,A3,1X,A1,I4,A1,4X,F7.2,3X,F7.2)
REMARK 500
REMARK 500 EXPECTED VALUES: GJ KLEYWEGT AND TA JONES (1996). PHI/PSI-
REMARK 500 CHOLOGY: RAMACHANDRAN REVISITED. STRUCTURE 4, 1395 - 1400
REMARK 500
REMARK 500  M RES CSSEQI        PSI       PHI
REMARK 500    ASN A 100     -110.87   -163.72
REMARK 500    ILE A 166      -28.81    -31.64
REMARK 500
REMARK 500 THIS ENTRY HAS     108 RAMACHANDRAN OUTLIERS.
REMARK 500
REMARK 500 REMARK: NULL
```

Example – Cis/Trans geometry

```
REMARK 500
REMARK 500 GEOMETRY AND STEREOCHEMISTRY
REMARK 500 SUBTOPIC: NON-CIS, NON-TRANS
REMARK 500
REMARK 500 THE FOLLOWING PEPTIDE BONDS DEVIATE SIGNIFICANTLY FROM BOTH
REMARK 500 CIS AND TRANS CONFORMATION.  CIS BONDS, IF ANY, ARE LISTED
REMARK 500 ON CISPEP RECORDS.  TRANS IS DEFINED AS 180 +/- 30 AND
REMARK 500 CIS IS DEFINED AS 0 +/- 30 DEGREES.
REMARK 500                                 MODEL     OMEGA
REMARK 500 ARG A  413     ASP A  414          1       147.84
REMARK 500 ALA B  288     ASN B  289          2       -39.12
REMARK 500
REMARK 500 REMARK: NULL
```

Example – Planar groups

```
REMARK 500
REMARK 500 GEOMETRY AND STEREOCHEMISTRY
REMARK 500 SUBTOPIC: PLANAR GROUPS
REMARK 500
REMARK 500 PLANAR GROUPS IN THE FOLLOWING RESIDUES HAVE A TOTAL
REMARK 500 RMS DISTANCE OF ALL ATOMS FROM THE BEST-FIT PLANE
REMARK 500 BY MORE THAN AN EXPECTED VALUE OF 6*RMSD, WITH AN
REMARK 500 RMSD 0.02 ANGSTROMS, OR AT LEAST ONE ATOM HAS
REMARK 500 AN RMSD GREATER THAN THIS VALUE
REMARK 500 (M=MODEL NUMBER; RES=RESIDUE NAME; C=CHAIN IDENTIFIER;
REMARK 500 SSEQ=SEQUENCE NUMBER; I=INSERTION CODE).
REMARK 500
REMARK 500  M RES CSSEQI        RMS     TYPE
REMARK 500    TYR A  36         0.08    SIDE CHAIN
REMARK 500    TYR A 104         0.08    SIDE CHAIN
REMARK 500
REMARK 500 REMARK: NULL
```

Example – Main chain planarity

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 500
REMARK 500 GEOMETRY AND STEREOCHEMISTRY
REMARK 500 SUBTOPIC: MAIN CHAIN PLANARITY
REMARK 500
REMARK 500 THE FOLLOWING RESIDUES HAVE A PSEUDO PLANARITY
REMARK 500 TORSION ANGLE, C(I) - CA(I) - N(I+1) - O(I), GREATER
REMARK 500 10.0 DEGREES. (M=MODEL NUMBER; RES=RESIDUE NAME;
REMARK 500 C=CHAIN IDENTIFIER; SSEQ=SEQUENCE NUMBER;
REMARK 500 I=INSERTION CODE).
REMARK 500
REMARK 500  M RES CSSEQI        ANGLE
REMARK 500  1 GLY A 289        -10.28
REMARK 500
REMARK 500 REMARK: NULL
```

Example – Chiral centers

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 500
REMARK 500 GEOMETRY AND STEREOCHEMISTRY
REMARK 500 SUBTOPIC: CHIRAL CENTERS
REMARK 500
REMARK 500 UNEXPECTED CONFIGURATION OF THE FOLLOWING CHIRAL
REMARK 500 CENTER(S) USING IMPROPER C--N--CA--CB CHIRALITY
REMARK 500 FOR AMINO ACIDS AND C1'--O4'--N1(N9)--C2' FOR
REMARK 500 NUCLEIC ACIDS OR EQUIVALENT ANGLE
REMARK 500 M=MODEL NUMBER; RES=RESIDUE NAME; C=CHAIN
REMARK 500 IDENTIFIER; SSEQ=SEQUENCE NUMBER; I=INSERTION CODE
REMARK 500
REMARK 500 STANDARD TABLE:
REMARK 500 FORMAT: (10X,I3,1X,A3,1X,A1,I4,A1,6X,F5.1,6X,A1,10X,A1,3X,A16)
REMARK 500
REMARK 500  M RES CSSEQI    IMPROPER   EXPECTED   FOUND DETAILS
REMARK 500 16 LEU A  20         0.1      L          D   EXPECTING SP3
REMARK 500 16 VAL A  21       -96.0      L          D   OUTSIDE RANGE
REMARK 500 16 GLN A  22       -54.1      L          D   OUTSIDE RANGE
REMARK 500 16 THR A  24       -42.0      L          D   WRONG HAND
REMARK 500 16 LYS A  26       -96.9      L          D   OUTSIDE RANGE
REMARK 500 16 ARG A  29      -133.0      L          D   OUTSIDE RANGE
REMARK 500 16 LEU A  31        53.6      L          L   OUTSIDE RANGE
REMARK 500 16 LYS A  32       -45.4      L          D   OUTSIDE RANGE
REMARK 500 16 GLU A  33       -41.3      L          D   WRONG HAND
REMARK 500 16 ASP A  34       -43.2      L          D   WRONG HAND
REMARK 500
REMARK 500 REMARK: NULL
```

REMARK 525 (updated), Distant Solvent Atoms

REMARK 525 lists solvent atoms more than 5 Angstroms from any polymer chain.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 525
REMARK 525 SOLVENT
REMARK 525
REMARK 525  formatted text.
```

Example

```
REMARK 525
REMARK 525 SOLVENT
REMARK 525
REMARK 525 THE SOLVENT MOLECULES HAVE CHAIN IDENTIFIERS THAT
REMARK 525 INDICATE THE POLYMER CHAIN WITH WHICH THEY ARE MOST
REMARK 525 CLOSELY ASSOCIATED. THE REMARK LISTS ALL THE SOLVENT
REMARK 525 MOLECULES WHICH ARE MORE THAN 5A AWAY FROM THE
REMARK 525 NEAREST POLYMER CHAIN (M=MODEL NUMBER;
REMARK 525 RES=RESIDUE NAME; C=CHAIN IDENTIFIER; SSEQ=SEQUENCE
REMARK 525 NUMBER; I=INSERTION CODE):
REMARK 525
REMARK 525  M RES CSSEQI
REMARK 525    HOH B  89        DISTANCE =  6.29 ANGSTROMS
REMARK 525    HOH B  94        DISTANCE =  5.58 ANGSTROMS
```

REMARK 600, Heterogen

Further details on the heterogens in the entry.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 600
REMARK 600 HETEROGEN
REMARK 600
REMARK 600 FREE TEXT GOES HERE.
```

Example

```
REMARK 600
REMARK 600 HETEROGEN
REMARK 600
REMARK 600 CHAIN A ENDOTHIAPEPSIN:
REMARK 600 RESIDUES ASP 54 AND GLY 55 HAVE CYCLISED
REMARK 600  TO FORM A SUCCINIMIDE (RESIDUE SUI 54)
REMARK 600
REMARK 600 CHAIN B IN THIS PDB ENTRY IS THE
REMARK 600    GEM-DIOL INHIBITOR PD-135.040
```

REMARK 610 and REMARK 615 (added)

Ligands or hetgroups that are not part of any polymer (protein or nucleic acid) in the structure may also have missing atoms or atoms with zero occupancy. In such instances the name of the hetgroup or ligand, chain ID and model number (if applicable) will be listed in REMARK 610 (for missing atoms) or REMARK 615 (for atoms with 0.00 occupancy). As the list of specific atoms missing from a hetgroup may be really large, they will not listed in the remarks described above. The list of all missing atoms from the ligands may be easily derived by comparing the coordinates of the hetgroup to its definition in the ligand dictionary.

REMARK 610, Non-polymer residues with missing atoms

REMARK 610 enumerates non-polymer residues with missing atoms.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 610
REMARK 610 MISSING HETEROATOM
REMARK 610 THE FOLLOWING RESIDUES HAVE MISSING ATOMS (M=MODEL NUMBER;
REMARK 610 RES=RESIDUE NAME; C=CHAIN IDENTIFIER; SSEQ=SEQUENCE NUMBER;
REMARK 610 I=INSERTION CODE):
REMARK 610   M RES C SSEQI
REMARK 610     PPI    438
```

REMARK 615, Non-polymer residues containing atoms with zero occupancy

REMARK 615 enumerates non-polymer residues containing atoms modeled with zero occupancy.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 615
REMARK 615 ZERO OCCUPANCY ATOM
REMARK 615 THE FOLLOWING RESIDUES HAVE ATOMS MODELED WITH ZERO
REMARK 615 OCCUPANCY. THE LOCATION AND PROPERTIES OF THESE ATOMS
REMARK 615 MAY NOT BE RELIABLE. (M=MODEL NUMBER; RES=RESIDUE NAME;
REMARK 615 C=CHAIN IDENTIFIER; SSEQ=SEQUENCE NUMBER; I=INSERTION CODE):
REMARK 615   M RES C SSEQI
REMARK 615     PPI    438
```

REMARK 620 (added), Metal coordination

Details of metal coordination are provided in REMARK 620. By default, coordination angles for any metal coordination and surrounding residues (if present) will be provided in this REMARK.

Template:

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 620
REMARK 620 METAL COORDINATION
REMARK 620  (M=MODEL NUMBER; RES=RESIDUE NAME; C=CHAIN IDENTIFIER;
REMARK 620   SSEQ=SEQUENCE NUMBER; I=INSERTION CODE):
REMARK 620
REMARK 620 COORDINATION ANGLES FOR:  M RES CSSEQI METAL
REMARK 620 N RES CSSEQI ATOM
```

Example

```
REMARK 620
REMARK 620 METAL COORDINATION
REMARK 620  (M=MODEL NUMBER; RES=RESIDUE NAME; C=CHAIN IDENTIFIER;
REMARK 620   SSEQ=SEQUENCE NUMBER; I=INSERTION CODE):
REMARK 620
REMARK 620 COORDINATION ANGLES FOR:  M RES CSSEQI METAL
REMARK 620                             F3S A 107  FE1
REMARK 620 N RES CSSEQI ATOM
REMARK 620 1 CYS A  39   SG
REMARK 620 2 F3S A 107  FE3  142.2
REMARK 620 3 F3S A 107  FE4  154.3  59.7
REMARK 620 4 F3S A 107   S1  120.2  53.8  55.7
REMARK 620 5 F3S A 107   S2  113.0 103.5  54.3 106.6
REMARK 620 6 F3S A 107   S3  103.8  53.0 101.7 103.2 109.2
REMARK 620 N                    1     2     3     4     5
REMARK 620
REMARK 620 COORDINATION ANGLES FOR:  M RES CSSEQI METAL
REMARK 620                             F3S A 107  FE3
REMARK 620 N RES CSSEQI ATOM
REMARK 620 1 F3S A 107  FE1
REMARK 620 2 F3S A 107  FE4   59.0
REMARK 620 3 F3S A 107   S1   52.7  55.1
REMARK 620 4 F3S A 107   S3   52.9 101.0 102.1
REMARK 620 5 CYS A  45   SG  146.5 146.2 115.6 112.8
REMARK 620 6 F3S A 107   S4  103.5  54.5 106.3 109.6 110.0
REMARK 620 N                    1     2     3     4     5
REMARK 620
REMARK 620 COORDINATION ANGLES FOR:  M RES CSSEQI METAL
REMARK 620                             F3S A 107  FE4
REMARK 620 N RES CSSEQI ATOM
REMARK 620 1 F3S A 107  FE1
REMARK 620 2 F3S A 107  FE3   61.3
REMARK 620 3 F3S A 107   S1   53.4  53.9
REMARK 620 4 F3S A 107   S2   54.4 105.0 104.5
REMARK 620 5 CYS A  20   SG  142.7 140.2 109.0 114.5
REMARK 620 6 F3S A 107   S4  105.1  54.1 104.8 111.7 111.6
REMARK 620 N                    1     2     3     4     5
REMARK 620
REMARK 620 COORDINATION ANGLES FOR:  M RES CSSEQI METAL
REMARK 620                             F3S A 108  FE1
REMARK 620 N RES CSSEQI ATOM
REMARK 620 1 F3S A 108   S3
```

```
REMARK 620 2 CYS A  16   SG  120.1
REMARK 620 3 F3S A 108  FE3   51.4 145.9
REMARK 620 4 F3S A 108  FE4   54.3 148.5  59.9
REMARK 620 5 F3S A 108   S1   98.3 110.0  50.6 101.5
REMARK 620 6 F3S A 108   S2  104.2 109.4 104.5  53.3 114.7
REMARK 620 N                    1     2     3     4     5
REMARK 620
```

REMARK 630 (added), Inhibitor Description

Details of inhibitor/peptide inhibitor which is presented as a chemical component (het group) are provided in REMARK 630. By default, molecule type and inhibitor's name will be provided in this REMARK.

Template:

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 630 MOLECULE TYPE:
REMARK 630 MOLECULE NAME:
REMARK 630 (M=MODEL NUMBER; RES=RESIDUE NAME; C=CHAIN IDENTIFIER;
REMARK 630  SSSEQ=SEQUENCE NUMBER; I=INSERTION CODE.)
REMARK 630
REMARK 630   M RES C SSSEQI
REMARK 630 SOURCE:
REMARK 630 SUBCOMP:
REMARK 630 STRUCTURE DETAILS:
REMARK 630 OTHER DETAILS:
```

REMARK 650, Helix

Further details on the helical portions of the entry.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 650
REMARK 650 HELIX
REMARK 650 FREE TEXT GOES HERE.
```

Examples

```
REMARK 650
REMARK 650 HELIX
REMARK 650 DETERMINATION METHOD: KDSSP
REMARK 650 THE MAJOR DOMAINS ARE: "N" FOR N-TERMINAL DOMAIN, "B" FOR
REMARK 650 BETA-BARREL DOMAIN, AND "C" FOR C-TERMINAL DOMAIN. "F"
REMARK 650 REFERS TO THE ACTIVE SITE FLAP.  ALPHA HELICES ARE NAMED
REMARK 650 WITH TWO CHARACTERS, THE FIRST REFERRING TO THE DOMAIN
REMARK 650 IN WHICH THEY OCCUR.

REMARK 650
REMARK 650 HELIX
REMARK 650 DETERMINATION METHOD: AUTHOR PROVIDED.
```

REMARK 700, Sheet

Further details on the sheet content of the structure. Several standard templates are shown.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 700
REMARK 700 SHEET
REMARK 700 FREE TEXT GOES HERE.
```

Examples

```
REMARK 700
REMARK 700 SHEET
REMARK 700 DETERMINATION METHOD:
REMARK 700 THE SHEET STRUCTURE OF THIS MOLECULE IS BIFURCATED.  IN
REMARK 700 ORDER TO REPRESENT THIS FEATURE IN THE SHEET RECORDS BELOW,
REMARK 700 TWO SHEETS ARE DEFINED.  STRANDS N1, N2, N3 AND N4 OF SHEET
REMARK 700 XXX AND XXX ARE IDENTICAL.

REMARK 700
REMARK 700 SHEET
REMARK 700 DETERMINATION METHOD:
REMARK 700 THE SHEET PRESENTED AS XXX ON SHEET RECORDS BELOW IS
REMARK 700 ACTUALLY AN N-STRANDED BETA-BARREL.  THIS IS
REMARK 700 REPRESENTED BY A N+1-STRANDED SHEET IN WHICH THE FIRST AND
REMARK 700 LAST STRANDS ARE IDENTICAL.

REMARK 700
REMARK 700 SHEET
REMARK 700 DETERMINATION METHOD:
REMARK 700 THERE ARE SEVERAL BIFURCATED SHEETS IN THIS STRUCTURE.
REMARK 700 EACH IS REPRESENTED BY TWO SHEETS WHICH HAVE ONE OR MORE
REMARK 700 IDENTICAL STRANDS.
REMARK 700 SHEETS XXX AND XXX REPRESENT ONE BIFURCATED SHEET.
REMARK 700 SHEETS XXX AND XXX REPRESENT ONE BIFURCATED SHEET.
```

N1, N2, N3 and N4 represent strand numbers, and XXX represents sheet identifiers.

When the remark for several bifurcated sheets is used, its last line is repeated for the appropriate number of bifurcated sheets, as shown in the last template above.

Examples

```
REMARK 700
REMARK 700 SHEET
REMARK 700 THE SHEET STRUCTURE OF THIS MOLECULE IS BIFURCATED.  IN
REMARK 700 ORDER TO REPRESENT THIS FEATURE IN THE SHEET RECORDS BELOW,
REMARK 700 TWO SHEETS are defined.  STRANDS 3, 4, AND 5
REMARK 700 OF SHEET *B2A* AND *B2B* ARE IDENTICAL.  STRANDS 3, 4, AND
REMARK 700 5 OF SHEET *B2C* AND *B2D* ARE IDENTICAL.
REMARK 700
```

```
REMARK 700 SHEET
REMARK 700 STRANDS 1 TO 4 OF THE BETA-SHEET HAVE GREEK-KEY TOPOLOGY.
REMARK 700 THE SHEET FORMS A FIVE-STRANDED BETA-BARREL WITH BULGES IN
REMARK 700 STRANDS 3 AND 5.  IN ORDER TO REPRESENT THIS FEATURE IN THE
REMARK 700 SHEET RECORDS BELOW, TWO SHEETS ARE DEFINED.

REMARK 700
REMARK 700 SHEET
REMARK 700 THE SHEET PRESENTED AS S5 ON SHEET RECORDS BELOW IS
REMARK 700 ACTUALLY A 6-STRANDED BETA-BARREL.  THIS IS
REMARK 700 REPRESENTED BY A 7-STRANDED SHEET IN WHICH THE FIRST AND
REMARK 700 LAST STRANDS ARE IDENTICAL.


REMARK 700
REMARK 700 SHEET
REMARK 700 DETERMINATION METHOD: AUTHOR PROVIDED.
```

REMARK 800 (updated), Important Sites

Further details on important sites of the entry. REMARK 800 is mandatory if SITE records exist.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 800
REMARK 800 SITE
REMARK 800 SITE_IDENTIFIER: FREE TEXT GOES HERE.
REMARK 800 EVIDENCE_CODE: (AUTHOR or SOFTWARE or UNKNOWN)
REMARK 800 SITE_DESCRIPTION: FREE TEXT GOES HERE.
```

* Site identifiers are 3-letter codes in a character range of AC1-ZZ9 if it is software determined.

Examples

```
REMARK 800
REMARK 800 SITE
REMARK 800 SITE_IDENTIFIER: RCA
REMARK 800 EVIDENCE_CODE: AUTHOR
REMARK 800 SITE_DESCRIPTION: DESIGNATED RECOGNITION REGION IN PRIMARY
REMARK 800 REFERENCE.  PROPOSED TO AFFECT SUBSTRATE SPECIFICITY.
REMARK 800
REMARK 800 SITE_IDENTIFIER: RCB
REMARK 800 EVIDENCE_CODE: AUTHOR
REMARK 800 SITE_DESCRIPTION: DESIGNATED RECOGNITION REGION IN PRIMARY
REMARK 800  REFERENCE.  PROPOSED TO AFFECT SUBSTRATE SPECIFICITY.

REMARK 800
REMARK 800 SITE
REMARK 800 SITE_IDENTIFIER: AC1
REMARK 800 EVIDENCE_CODE: SOFTWARE
REMARK 800 SITE_DESCRIPTION: BINDING SITE FOR RESIDUE BAT A 19
REMARK 800
REMARK 800 SITE_IDENTIFIER: AC2
REMARK 800 EVIDENCE_CODE: SOFTWARE
REMARK 800 SITE_DESCRIPTION: BINDING SITE FOR RESIDUE CA A 1
REMARK 800
REMARK 800 SITE_IDENTIFIER: AC3
REMARK 800 EVIDENCE_CODE: SOFTWARE
REMARK 800 SITE_DESCRIPTION: BINDING SITE FOR RESIDUE BIL A 20
```

Relationship to other records:

Remark 800 is mandatory if site records exist.

REMARK 900, Related Entries

This REMARK provides information about other PDB entries related to the entry. These may include coordinate entries deposited as a related set, an EMDB identifier for the related EM map, a BMRB identifier for the related NMR chemical shifts, or a structural genomics target identifier.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 900
REMARK 900 RELATED ENTRIES
REMARK 900 FREE TEXT GOES HERE.
```

Examples

```
REMARK 900
REMARK 900 RELATED ENTRIES
REMARK 900 RELATED ID: 2GB8   RELATED DB: PDB
REMARK 900 SOLUTION STRUCTURE OF WT CC-CCP COMPLEX
REMARK 900 RELATED ID: 2PCC   RELATED DB: PDB
REMARK 900 CRYSTAL STRUCTURE OF WT CC-CCP COMPLEX
REMARK 900 RELATED ID: 1YCC   RELATED DB: PDB
REMARK 900 CRYSTAL STRUCTURE OF YEAST ISO-1-CYTOCHROME C
REMARK 900 RELATED ID: 1ZBY   RELATED DB: PDB
REMARK 900 HIGH-RESOLUTION CRYSTAL STRUCTURE OF YEAST CYTOCHROME C
REMARK 900 PEROXIDASE

REMARK 900
REMARK 900 RELATED ENTRIES
REMARK 900 RELATED ID: STR82   RELATED DB: TARGETDB
REMARK 900 RELATED ID: 15386   RELATED DB: BMRB
```

REMARK 999, Sequence

This remark is a free text remark which describes anything unusual about a particular polymer sequence in SEQRES records.

For examples,

1. If the exact sequence of the sample is not known, due to, for example, proteolysis, the sequence should match the coordinates and a REMARK 999 can be added.
2. The information about a sequence region of a chimeric protein which does not match the UNP entry, such as a linker region, can be added to REMARK 999.
3. Sequence conflicts which are listed in the UNP reference can also be described in REMARK 999. A full explanation of the microheterogeneity for all residues at a particular residue number can be elaborated in REMARK 999.
4. If the coordinates alignment with the sequence is unknown and the residue numbering is arbitrary. The sequence would be poly UNK. The sequence, if it is known, would be listed in the REMARK 999

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
REMARK 999
REMARK 999 SEQUENCE
REMARK 999 FREE TEXT GOES HERE.
```

Example

```
REMARK 999
REMARK 999 SEQUENCE
REMARK 999 THE N-TERMINAL 19 RESIDUES 'GSHMVPGQKQHYVQPTAAN'
REMARK 999 CORRESPOND TO A PHAGE-DISPLAY DERIVED PEPTIDE,
REMARK 999 WHICH IS FUSED TO THE SECRETION CHAPERONE PROTEIN

REMARK 999
REMARK 999 SEQUENCE
REMARK 999 THE SEQUENCE USED IS THAT PROVIDED BY THE CDNA, WHICH
REMARK 999 CORRECTS SEVERAL ASP/ASN AND GLU/GLN MISASSIGNMENTS.

REMARK 999
REMARK 999 SEQUENCE
REMARK 999 THR AT POSITION 74 WAS FOUND BY WOLMAN ET AL., JOURNAL OF
REMARK 999 BIOCHEMISTRY 263, 15506 (1988).

REMARK 999 SEQUENCE
REMARK 999 THE INSERTED RESIDUES AT THE N-TERMINUS OF THE PROTEIN
REMARK 999 CORRESPOND TO A 32-RESIDUE DSE3 LANTHIDE-BINDING TAG
REMARK 999 THE RESIDUES NUMBERED 66 TO 100 IN THIS ENTRY CORRESPOND
REMARK 999 TO RESIDUES -4 TO 13 AND -1' TO 15' IN THE PRIMARY CITATION.
```

3. Primary Structure Section

The primary structure section of a PDB formatted file contains the sequence of residues in each chain of the macromolecule(s). Embedded in these records are chain identifiers and sequence numbers that allow other records to link into the sequence.

DBREF (standard format)

The DBREF record provides cross-reference links between PDB sequences (what appears in SEQRES record) and a corresponding database sequence.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"DBREF "	
8 - 11	IDcode	idCode	ID code of this entry.
13	Character	chainID	Chain identifier.
15 - 18	Integer	seqBegin	Initial sequence number of the PDB sequence segment.
19	AChar	insertBegin	Initial insertion code of the PDB sequence segment.
21 - 24	Integer	seqEnd	Ending sequence number of the PDB sequence segment.
25	AChar	insertEnd	Ending insertion code of the PDB sequence segment.
27 - 32	LString	database	Sequence database name.
34 - 41	LString	dbAccession	Sequence database accession code.
43 - 54	LString	dbIdCode	Sequence database identification code.
56 - 60	Integer	dbseqBegin	Initial sequence number of the database seqment.
61	AChar	idbnsBeg	Insertion code of initial residue of the segment, if PDB is the reference.
63 - 67	Integer	dbseqEnd	Ending sequence number of the database segment.
68	AChar	dbinsEnd	Insertion code of the ending residue of the segment, if PDB is the reference.

Note: By default this format is used as long as the information entered into these fields fits. For sequence databases that use longer accession code or long sequence numbering, the new DBREF1/DBREF2 format can be used.

Details

* PDB entries contain multi-chain molecules with sequences that may be wild type, variant, or synthetic. Sequences may also have been modified through site-directed mutagenesis experiments (engineered). A number of PDB entries report structures of individual domains cleaved from larger molecules.

The DBREF records present sequence correlations between PDB SEQRES records and corresponding GenBank (for nucleic acids) or UNIPROT/Norine (for proteins) entries. PDB entries containing heteropolymers are linked to different sequence database entries.

* Database names and their abbreviations as used on DBREF records.

Database name	Database abbreviations (columns 27 - 32)
GenBank	GB
Protein Data Bank	PDB
UNIPROT	UNP
Norine	NORINE

* wwPDB does not guarantee that all possible references to the listed databases will be provided. In most cases, only one reference to a sequence database will be provided.

* If no reference is found in the sequence databases, then the PDB entry itself is given as the reference.

* Selection of the appropriate sequence database entry or entries to be linked to a PDB entry is done on the basis of the sequence and its biological source. Questions on entry assignment that may arise are resolved by consultation with the database.

Verification/Validation/Value Authority Control

The sequence database entry found during PDB's search is compared to that provided by the depositor and any differences are resolved or annotated.

All polymers in the entry will be assigned a DBREF record.

Relationships to Other Record Types

DBREF represents the sequence as found in SEQRES records.

DBREF1/DBREF2 replaces DBREF when the accession codes or sequence numbering does not fit the DBREF format.

Examples

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
DBREF  2JHQ A    1   226  UNP    Q9KPK8   UNG_VIBCH        1    226
```

```
DBREF  3AKY A    1   219  UNP    P07170   KAD1_YEAST       3    221

DBREF  1HAN A    2   298  UNP    P47228   BPHC_BURCE       1    297

DBREF  3D3I A    0   760  UNP    P42592   YGJK_ECOLI      23    783
DBREF  3D3I B    0   760  UNP    P42592   YGJK_ECOLI      23    783

DBREF  3C2J A    1     8  PDB    3C2J     3C2J             1      8
DBREF  3C2J B  101   108  PDB    3C2J     3C2J           101    108

DBREF  1FFK 0    2  2923  GB     3377779  AF034620      2597   5518
DBREF  1FFK 9    1   122  GB     3377779  AF034620      5658   5779

DBREF  1UNJ X    6    11  NORINE NOR00228 NOR00228         6     11
```

DBREF1 / DBREF2 (added)

Details

This updated two-line format is used when the accession code or sequence numbering does not fit the space allotted in the standard DBREF format. This includes some GenBank sequence numbering (greater than 5 characters) and UNIMES accession numbers (greater than 12 characters).

Record Format

DBREF1

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"DBREF1"	
8 - 11	IDcode	idCode	ID code of this entry.
13	Character	chainID	Chain identifier.
15 - 18	Integer	seqBegin	Initial sequence number of the PDB sequence segment, right justified.
19	AChar	insertBegin	Initial insertion code of the PDB sequence segment.
21 - 24	Integer	seqEnd	Ending sequence number of the PDB sequence segment, right justified.
25	AChar	insertEnd	Ending insertion code of the PDB sequence segment.
27 - 32	LString	database	Sequence database name.
48 - 67	LString	dbIdCode	Sequence database identification code, left justified.

DBREF2

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"DBREF2"	
8 - 11	IDcode	idCode	ID code of this entry.
13	Character	chainID	Chain identifier.
19 - 40	LString	dbAccession	Sequence database accession code, left justified.
46 - 55	Integer	seqBegin	Initial sequence number of the Database segment, right justified.
58 - 67	Integer	seqEnd	Ending sequence number of the Database segment, right justified.

Details

* The DBREF1/DBREF2 record presents sequence correlations between PDB SEQRES records and corresponding GenBank (for nucleic acids) or UNIMES (for proteins) entries. Several cases are easily represented by means of pointers between the databases using DBREF.

* Database names and their abbreviations as used as in DBREF records.

Database name	Database abbreviations (columns 27 - 32)
GenBank	GB
UNIMES	UNIMES

* wwPDB does not guarantee that all possible references to the listed databases will be provided. In most cases, only one reference to a sequence database will be provided.

Verification/Validation/Value Authority Control

The sequence database entry found by wwPDB staff is compared to answers provided by the depositor; any differences are resolved or annotated appropriately.

Relationships to Other Record Types

DBREF1/DBREF2 represents the sequence as found in SEQRES records.

Template

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
DBREF1 2J83 A    61  322  XXXXXX               YYYYYYYYYYYYYYYYYYYY
DBREF2 2J83 A     ZZZZZZZZZZZZZZZZZZZZZZ     nnnnnnnnnn  mmmmmmmmmm
```

Examples

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
DBREF1 2J83 A    61  322  UNIMES               UPI000148A153
DBREF2 2J83 A     MES00005880000                     61         322
```

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
DBREF1 2J83 A    61  322  GB                   AE017221
DBREF2 2J83 A     46197919                      1534489     1537377
```

SEQADV

Overview

The SEQADV record identifies differences between sequence information in the SEQRES records of the PDB entry and the sequence database entry given in DBREF. Please note that these records were designed to identify differences and not errors. No assumption is made as to which database contains the correct data. A comment explaining any engineered differences in the sequence between the PDB and the sequence database may also be included here.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"SEQADV"	
8 - 11	IDcode	idCode	ID code of this entry.
13 - 15	Residue name	resName	Name of the PDB residue in conflict.
17	Character	chainID	PDB chain identifier.
19 - 22	Integer	seqNum	PDB sequence number.
23	AChar	iCode	PDB insertion code.
25 - 28	LString	database	
30 - 38	LString	dbIdCode	Sequence database accession number.
40 - 42	Residue name	dbRes	Sequence database residue name.
44 - 48	Integer	dbSeq	Sequence database sequence number.
50 - 70	LString	conflict	Conflict comment.

Details

* In a number of cases, conflicts between the sequences found in PDB entries and in sequence database reference entries have been noted. There are several possible reasons for these conflicts, including natural variants or engineered sequences (mutants), polymorphic sequences, or ambiguous or conflicting experimental results. These discrepancies are reported in SEQADV. Additional details may be included in remark 999.

* When conflicts arise which are not classifiable by these terms, a reference to either a published paper, a PDB entry, or a REMARK within the entry is given.

* The comment "SEE REMARK 999" is included when the explanation for the conflict is too long to fit the SEQADV record.
* Some of the possible conflict comments:

- Cloning artifact
- Expression tag
- Conflict
- Engineered
- Variant
- Insertion
- Deletion
- Microheterogeneity
- Chromophore

.

* Microheterogeneity is to be represented as a variant with one of the possible residues in the site being selected (arbitrarily) as the primary residue. The residues which do not match to the UNP reference will be listed in SEQADV records with the explanation of “microheterogeneity”.

Verification/Validation/Value Authority Control

SEQADV records are automatically generated.

Relationships to Other Record Types

SEQADV refers to the sequence as found in the SEQRES records, and to the sequence database reference found on DBREF.

REMARK 999 contains text that explains discrepancies when the explanation is too lengthy to fit in SEQADV.

Examples

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

SEQADV 3ABC MET A   -1  UNP  P10725              EXPRESSION TAG

SEQADV 3ABC GLY A   50  UNP  P10725    VAL    50 ENGINEERED

SEQADV 2QLE CRO A   66  UNP  P42212    SER    65 CHROMOPHORE

SEQADV 2OKW LEU A   64  UNP  P42212    PHE    64 SEE REMARK 999
```

SEQRES (updated)

Overview

SEQRES records contain a listing of the consecutive chemical components covalently linked in a linear fashion to form a polymer. The chemical components included in this listing may be standard or modified amino acid and nucleic acid residues. It may also include other residues that are linked to the standard backbone in the polymer. Chemical components or groups covalently linked to side-chains (in peptides) or sugars and/or bases (in nucleic acid polymers) will not be listed here.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"SEQRES"	
8 - 10	Integer	serNum	Serial number of the SEQRES record for the current chain. Starts at 1 and increments by one each line. Reset to 1 for each chain.
12	Character	chainID	Chain identifier. This may be any single legal character, including a blank which is is used if there is only one chain.
14 - 17	Integer	numRes	Number of residues in the chain. This value is repeated on every record.
20 - 22	Residue name	resName	Residue name.
24 - 26	Residue name	resName	Residue name.
28 - 30	Residue name	resName	Residue name.
32 - 34	Residue name	resName	Residue name.
36 - 38	Residue name	resName	Residue name.
40 - 42	Residue name	resName	Residue name.
44 - 46	Residue name	resName	Residue name.
48 - 50	Residue name	resName	Residue name.
52 - 54	Residue name	resName	Residue name.
56 - 58	Residue name	resName	Residue name.
60 - 62	Residue name	resName	Residue name.
64 - 66	Residue name	resName	Residue name.
68 - 70	Residue name	resName	Residue name.

Verification/Validation/Value Authority Control

The residues presented in the ATOM records must agree with those on the SEQRES records.

The SEQRES records are checked using sequence databases and information provided by the depositor.

SEQRES is compared to the ATOM records during processing, and both are checked against the sequence databases. All discrepancies are either resolved or annotated appropriately in the entry.

The ribo- and deoxyribonucleotides in the SEQRES records are distinguished. The ribo- forms of these residues are identified with the residue names A, C, G, U and I. The deoxy- forms of these residues are identified with the residue names DA, DC, DG, DT and DI. Modified nucleotides in the sequence are identified by separate 3-letter residue codes. The *plus* character prefix to label modified nucleotides (e.g. +A, +C, +T) is no longer used.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
SEQRES   1 A   21  GLY ILE VAL GLU GLN CYS CYS THR SER ILE CYS SER LEU
SEQRES   2 A   21  TYR GLN LEU GLU ASN TYR CYS ASN
SEQRES   1 B   30  PHE VAL ASN GLN HIS LEU CYS GLY SER HIS LEU VAL GLU
SEQRES   2 B   30  ALA LEU TYR LEU VAL CYS GLY GLU ARG GLY PHE PHE TYR
SEQRES   3 B   30  THR PRO LYS ALA
SEQRES   1 C   21  GLY ILE VAL GLU GLN CYS CYS THR SER ILE CYS SER LEU
SEQRES   2 C   21  TYR GLN LEU GLU ASN TYR CYS ASN
SEQRES   1 D   30  PHE VAL ASN GLN HIS LEU CYS GLY SER HIS LEU VAL GLU
SEQRES   2 D   30  ALA LEU TYR LEU VAL CYS GLY GLU ARG GLY PHE PHE TYR
SEQRES   3 D   30  THR PRO LYS ALA

SEQRES   1 A    8   DA  DA  DC  DC  DG  DG  DT  DT
SEQRES   1 B    8   DA  DA  DC  DC  DG  DG  DT  DT

SEQRES   1 X   39    U   C   C   C   C   C   G   U   G   C   C   C   A
SEQRES   2 X   39    U   A   G   C   G   G   C   G   U   G   G   A   A
SEQRES   3 X   39    C   C   A   C   C   C   G   U   U   C   C   C   A
```

Known Problems

Polysaccharides do not lend themselves to being represented in SEQRES.

There is no mechanism provided to describe the sequence order if their starting position is unknown.

For cyclic peptides, a residue is arbitrarily assigned as the N-terminus.

MODRES (updated)

Overview

The MODRES record provides descriptions of modifications (e.g., chemical or post-translational) to protein and nucleic acid residues. Included are correlations between residue names given in a PDB entry and standard residues.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"MODRES"	
8 - 11	IDcode	idCode	ID code of this entry.
13 - 15	Residue name	resName	Residue name used in this entry.
17	Character	chainID	Chain identifier.
19 - 22	Integer	seqNum	Sequence number.
23	AChar	iCode	Insertion code.
25 - 27	Residue name	stdRes	Standard residue name.
30 - 70	String	comment	Description of the residue modification.

Details

* Residues modified post-translationally, enzymatically, or by design are described in MODRES records. In those cases where the wwPDB has opted to use a non-standard residue name for the residue, MODRES also correlates the new name to the precursor standard residue name.

* Modified nucleotides in the sequence are now identified by separate 3-letter residue codes. The *plus* character prefix to label modified nucleotides (e.g. +A, +C, +T) is no longer used.

* MODRES is mandatory when modified standard residues exist in the entry. Examples of some modification descriptions:

- Glycosylation site
- Post-translational modification
- Designed chemical modification
- Phosphorylation site
- D-configuration

* A MODRES record is not required if coordinate records are not provided for the modified residue.

* D-amino acids are given their own residue name (resName), i.e., DAL for D-alanine. This resName appears in the SEQRES records, and has the associated MODRES, HET, and FORMUL records. The coordinates are given as HETATMs within the ATOM records and occur in the correct order within the chain. This ordering is an exception to the stated Order of Records.

* When a standard residue name is used to describe a modified site, resName (columns 13-15) and stdRES (columns 25-27) contain the same value.

Verification/Validation/Value Authority Control

MODRES is generated by the wwPDB.

Relationships to Other Record Types

MODRES maps ATOM and HETATM records to the standard residue names. HET, and FORMUL may also appear.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
MODRES 2R0L ASN A   74  ASN  GLYCOSYLATION SITE

MODRES 1IL2 1MG D 1937    G  1N-METHYLGUANOSINE-5'-MONOPHOSPHATE

MODRES 4ABC MSE B   32  MET  SELENOMETHIONINE
```

4. Heterogen Section (updated)

The heterogen section of a PDB formatted file contains the complete description of non-standard residues in the entry. Detailed chemical definitions of non-polymer chemical components are described in the Chemical Component Dictionary (ftp://ftp.wwpdb.org/pub/pdb/data/monomers)

HET

HET records are used to describe non-standard residues, such as prosthetic groups, inhibitors, solvent molecules, and ions for which coordinates are supplied. Groups are considered HET if they are not part of a biological polymer described in SEQRES and considered to be a molecule bound to the polymer, or they are a chemical species that constitute part of a biological polymer and is not one of the following:

- standard amino acids, or
- standard nucleic acids (C, G, A, U, I, DC, DG, DA, DU, DT and DI), or
- unknown amino acid (UNK) or nucleic acid (N) where UNK and N are used to indicate the unknown residue name.

HET records also describe chemical components for which the chemical identity is unknown, in which case the group is assigned the hetID UNL (Unknown Ligand).

The heterogen section of a PDB formatted file contains the complete description of non-standard residues in the entry.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"HET "	
8 - 10	LString(3)	hetID	Het identifier, right-justified.
13	Character	ChainID	Chain identifier.
14 - 17	Integer	seqNum	Sequence number.
18	AChar	iCode	Insertion code.
21 - 25	Integer	numHetAtoms	Number of HETATM records for the group present in the entry.
31 - 70	String	text	Text describing Het group.

Details

* Each HET group is assigned a hetID of not more than three (3) alphanumeric characters. The sequence number, chain identifier, insertion code, and number of coordinate records are given for each occurrence of the HET group in the entry. The chemical name of the HET group is given in the HETNAM record and synonyms for the chemical name are given in the HETSYN records, see ftp://ftp.wwpdb.org/pub/pdb/data/monomers .

* There is a separate HET record for each occurrence of the HET group in an entry.

* A particular HET group is represented in the PDB archive with a unique hetID.

* PDB entries do not have HET records for water molecules, deuterated water, or methanol (when used as solvent).

* Unknown atoms or ions will be represented as UNX with the chemical formula X1. Unknown ligands are UNL; unknown amino acids are UNK.

Verification/Validation/Value Authority Control

For each het group that appears in the entry, the wwPDB checks that the corresponding HET, HETNAM, HETSYN, FORMUL, HETATM, and CONECT records appear, if applicable. The HET record is generated automatically using the Chemical Component Dictionary and information from the HETATM records.

Each unique hetID represents a unique molecule.

Relationships to Other Record Types

For each het group that appears in the entry, there must be corresponding HET, HETNAM, HETSYN, FORMUL,HETATM, and CONECT records. LINK records may also be created.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
HET    TRS    975       8

HET    UDP  A1457      25
HET    B3P  A1458      19

HET    NAG  Y   3      15
HET    FUC  Y   4      10
HET    NON  Y   5      12

HET    UNK  A 161       1
```

HETNAM

Overview

This record gives the chemical name of the compound with the given hetID.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"HETNAM"	
9 - 10	Continuation	continuation	Allows concatenation of multiple records.
12 - 14	LString(3)	hetID	Het identifier, right-justified.
16 - 70	String	text	Chemical name.

Details

* Each hetID is assigned a unique chemical name for the HETNAM record, see ftp://ftp.wwpdb.org/pub/pdb/data/monomers .

* Other names for the group are given on HETSYN records.

* PDB entries follow IUPAC/IUB naming conventions to describe groups systematically.

* The special character “~” is used to indicate superscript in a heterogen name. For example: N^6 will be listed in the HETNAM section as N~6~, with the ~ character indicating both the start and end of the superscript in the name, e.g.,

```
N-(BENZYLSULFONYL)SERYL-N~1~-{4-[AMINO(IMINO)METHYL]BENZYL}GLYCINAMIDE
```

* Continuation of chemical names onto subsequent records is allowed.

* Only one HETNAM record is included for a given hetID, even if the same hetID appears on more than one HET record.

Verification/Validation/Value Authority Control

For each het group that appears in the entry, the corresponding HET, HETNAM, FORMUL, HETATM, and CONECT records must appear. The HETNAM record is generated automatically using the Chemical Component Dictionary and information from HETATM records.

Relationships to Other Record Types

For each het group that appears in the entry, there must be corresponding HET, HETNAM, FORMUL, HETATM, and CONECT records. HETSYN and LINK records may also be created.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
HETNAM     NAG N-ACETYL-D-GLUCOSAMINE
HETNAM     SAD BETA-METHYLENE SELENAZOLE-4-CARBOXAMIDE ADENINE
HETNAM  2  SAD DINUCLEOTIDE

HETNAM     UDP URIDINE-5'-DIPHOSPHATE

HETNAM     UNX UNKNOWN ATOM OR ION
HETNAM     UNL UNKNOWN LIGAND

HETNAM     B3P 2-[3-(2-HYDROXY-1,1-DIHYDROXYMETHYL-ETHYLAMINO)-
HETNAM   2 B3P  PROPYLAMINO]-2-HYDROXYMETHYL-PROPANE-1,3-DIOL
```

HETSYN

Overview

This record provides synonyms, if any, for the compound in the corresponding (i.e., same hetID) HETNAM record. This is to allow greater flexibility in searching for HET groups.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"HETSYN"	
9 - 10	Continuation	continuation	Allows concatenation of multiple records.
12 - 14	LString(3)	hetID	Het identifier, right-justified.
16 - 70	SList	hetSynonyms	List of synonyms.

Details

* The wwPDB does not guarantee a complete list of possible synonyms. New synonyms may be added. The list can be continued onto additional HETSYN records. Even if the same hetID appears on more than one HET record, only one set of HETSYN records is included for the hetID.

Verification/Validation/Value Authority Control

For each HETSYN record in the entry, the corresponding HET, HETNAM, FORMUL, HETATM, and CONECT records must appear.

Relationships to Other Record Types

If there is a HETSYN record there must be corresponding HET, HETNAM, FORMUL, HETATM, and CONECT records. LINK records may also be created.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
HETSYN     HV5 3-METHYL-L-VALINE
HETSYN     AB1 ABT-378; LOPINAVIR
HETSYN     CMP CYCLIC AMP; CAMP
HETSYN     TRS TRIS BUFFER;
```

FORMUL

Overview

The FORMUL record presents the chemical formula and charge of a non-standard group.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"FORMUL"	
9 - 10	Integer	compNum	Component number.
13 - 15	LString(3)	hetID	Het identifier.
17 - 18	Integer	continuation	Continuation number.
19	Character	asterisk	"*" for water.
20 - 70	String	text	Chemical formula.

Details

* The elements of the chemical formula are given in the order C, H, N, and O, with other elements following in alphabetical order, each separated by a single space.

* The number of each atom type present immediately follows its chemical symbol without an intervening blank space. There will be no number indicated if there is only one atom for a particular atom type.

* Each set of SEQRES records and each HET group is assigned a component number in an entry. These numbers are assigned serially, beginning with 1 for the first set of SEQRES records. In addition:

- If a HET group is presented on a SEQRES record its FORMUL is assigned the component number of the chain in which it appears.

- If the HET group occurs more than once and is not presented on SEQRES records, the component number of its first occurrence is used.

* All occurrences of the HET group within a chain are grouped together with a multiplier. The remaining occurrences are also grouped with a multiplier. The sum of the multipliers is the number equaling the number of times that that HET group appears in the entry.

* A continuation field is provided in the event that more space is needed for the formula.
Columns 17 - 18 are used in order to maintain continuity with the existing format.

Verification/Validation/Value Authority Control

For each het group that appears in the entry, the corresponding HET, HETNAM, FORMUL, HETATM, and CONECT records must appear. The FORMUL record is generated automatically by PDB processing programs using the het group template file and information from HETATM records. UNL, UNK and UNX will not be listed in FORMUL even though these het groups present in the coordinate section.

Relationships to Other Record Types

For each het group that appears in the entry, the corresponding HET, HETNAM, FORMUL, HETATM, and CONECT records must appear.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
FORMUL   2  SO4     2(O4 S 2-)
FORMUL   3  GLC     C6 H12 O6
FORMUL   3  FOL     2(C19 H17 N7 O6 2-)
FORMUL   4   CL     2(CL 1-)
FORMUL   5   CA     2(CA 2+)

FORMUL   1  ACE     C2 H4 O
FORMUL   2  UDP     C9 H14 N2 O12 P2
FORMUL   3  B3P     C11 H26 N2 O6

FORMUL   8  HOH    *463(H2 O)
```

Known Problems

Partially deuterated centers are not well represented in this record.

5. Secondary Structure Section

The secondary structure section of a PDB formatted file describes helices, sheets, and turns 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
```

6. Connectivity Annotation Section

The connectivity annotation section allows the depositors to specify the existence and location of disulfide bonds and other linkages.

SSBOND (updated)

The SSBOND record identifies each disulfide bond in protein and polypeptide structures by identifying the two residues involved in the bond.

The disulfide bond distance is included after the symmetry operations at the end of the SSBOND record.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"SSBOND"	
8 - 10	Integer	serNum	Serial number.
12 - 14	LString(3)	"CYS"	Residue name.
16	Character	chainID1	Chain identifier.
18 - 21	Integer	seqNum1	Residue sequence number.
22	AChar	icode1	Insertion code.
26 - 28	LString(3)	"CYS"	Residue name.
30	Character	chainID2	Chain identifier.
32 - 35	Integer	seqNum2	Residue sequence number.
36	AChar	icode2	Insertion code.
60 - 65	SymOP	sym1	Symmetry operator for residue 1.
67 - 72	SymOP	sym2	Symmetry operator for residue 2.
74 – 78	Real(5.2)	Length	Disulfide bond distance

Details

* Bond distances between the sulfur atoms must be close to expected value.
* sym1 and sym2 are right justified and are always given even when identity operator (no cell translation) is to be applied to the residue.

Verification/Validation/Value Authority Control

wwPDB processing programs generate these records automatically

Relationships to Other Record Types

CONECT records are generated for the disulfide bonds when SG atoms of both cysteines are present in the coordinate records.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
SSBOND   1 CYS A    6    CYS A  127                            1555   1555  2.03
SSBOND   2 CYS A   30    CYS A  115                            1555   1555  2.07
SSBOND   3 CYS A   64    CYS A   80                            1555   1555  2.06
SSBOND   4 CYS A   76    CYS A   94                            1555   1555  2.04
```

Known Problems

If SG of cysteine is disordered then there are possible alternate linkages. wwPDB practice is to put together all possible SSBOND records. This is problematic because the alternate location identifier is not specified in the SSBOND record.

LINK (updated)

Overview

The LINK records specify connectivity between residues that is not implied by the primary structure. Connectivity is expressed in terms of the atom names. They also include the distance associated with the each linkage following the symmetry operations at the end of each record.

This record supplements information given in CONECT records and is provided here for convenience in searching.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"LINK "	
13 - 16	Atom	name1	Atom name.
17	Character	altLoc1	Alternate location indicator.
18 - 20	Residue name	resName1	Residue name.
22	Character	chainID1	Chain identifier.
23 - 26	Integer	resSeq1	Residue sequence number.
27	AChar	iCode1	Insertion code.
43 - 46	Atom	name2	Atom name.
47	Character	altLoc2	Alternate location indicator.
48 - 50	Residue name	resName2	Residue name.
52	Character	chainID2	Chain identifier.
53 - 56	Integer	resSeq2	Residue sequence number.
57	AChar	iCode2	Insertion code.
60 - 65	SymOP	sym1	Symmetry operator atom 1.
67 - 72	SymOP	sym2	Symmetry operator atom 2.
74 – 78	Real(5.2)	Length	Link distance

Details

* The atoms involved in bonds between HET groups or between a HET group and standard residue are listed.

* Inter-residue linkages not implied by the primary structure are listed (e.g., reduced peptide bond).

* Non-standard linkages between residues, e.g., side-chain to side-chain, are listed.

* Each LINK record specifies one linkage.

* These records do not specify connectivity within a HET group (see CONECT) or disulfide bridges (see SSBOND).

* sym1 and sym2 are right justified and are given as blank when the identity operator (and no cell translation) is to be applied to the atom.

- For NMR entries, only one set (or model) of LINK records will be supplied.
- Coordinate bonds are also listed as LINKs.

Verification/Validation/Value Authority Control

The distance between the pair of atoms listed must be consistent with the bonding.

Relationships to Other Record Types

CONECT records are generated from LINKs when both atoms are present in the entry. If symmetry operators are given to generate one of the residues involved in the bond, REMARK 290 defines the symmetry transformation.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
LINK         O   GLY A  49                NA    NA A6001     1555   1555 2.98
LINK         OG1 THR A  51                NA    NA A6001     1555   1555 2.72
LINK         OD2 ASP A  66                NA    NA A6001     1555   1555 2.72
LINK         NE  ARG A  68                NA    NA A6001     1555   1555 2.93
LINK         C21 2EG A   7                 C22 2EG B  19     1555   1555  1.56
```

CISPEP

Overview

CISPEP records specify the prolines and other peptides found to be in the *cis* conformation. Each cis peptide is listed on a separate line, with a consecutive numbering sequence.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"CISPEP"	
8 - 10	Integer	serNum	Record serial number.
12 - 14	LString(3)	pep1	Residue name.
16	Character	chainID1	Chain identifier.
18 - 21	Integer	seqNum1	Residue sequence number.
22	AChar	icode1	Insertion code.
26 - 28	LString(3)	pep2	Residue name.
30	Character	chainID2	Chain identifier.
32 - 35	Integer	seqNum2	Residue sequence number.
36	AChar	icode2	Insertion code.
44 - 46	Integer	modNum	Identifies the specific model.
54 - 59	Real(6.2)	measure	Angle measurement in degrees.

Details

* Cis peptides are those with omega angles of 0°±30°. Deviations larger than 30° are listed in REMARK 500.

Verification/Validation/Value Authority Control

These records are generated automatically.

Relationships to Other Record Types

Peptide bonds which deviate significantly from either the *cis* or *trans* conformation are annotated in REMARK 500.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
CISPEP   1 SER A   58    GLY A   59          0        20.91
CISPEP   1 GLY A  116    GLY A  117          0        18.50
```

```
CISPEP   1 MET A    1    SER A    2          0        -3.69
```

7. Miscellaneous Features Section

The miscellaneous features section may describe properties in the molecule such as environments surrounding a non-standard residue or the assembly of an active site. Other features may be described in the remarks section but are not given a specific record type so far.

SITE

Overview

* Site records specify residues comprising catalytic, co-factor, anti-codon, regulatory or other essential sites or environments surrounding ligands present in the structure.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"SITE "	
8 - 10	Integer	seqNum	Sequence number.
12 - 14	LString(3)	siteID	Site name.
16 - 17	Integer	numRes	Number of residues that compose the site.
19 - 21	Residue name	resName1	Residue name for first residue that creates the site.
23	Character	chainID1	Chain identifier for first residue of site.
24 - 27	Integer	seq1	Residue sequence number for first residue of the site.
28	AChar	iCode1	Insertion code for first residue of the site.
30 - 32	Residue name	resName2	Residue name for second residue that creates the site.
34	Character	chainID2	Chain identifier for second residue of the site.
35 - 38	Integer	seq2	Residue sequence number for second residue of the site.
39	AChar	iCode2	Insertion code for second residue of the site.
41 - 43	Residue name	resName3	Residue name for third residue that creates the site.
45	Character	chainID3	Chain identifier for third residue of the site.
46 - 49	Integer	seq3	Residue sequence number for third residue of the site.

50	AChar	iCode3	Insertion code for third residue of the site.
52 - 54	Residue name	resName4	Residue name for fourth residue that creates the site.
56	Character	chainID4	Chain identifier for fourth residue of the site.
57 - 60	Integer	seq4	Residue sequence number for fourth residue of the site.
61	AChar	iCode4	Insertion code for fourth residue of the site.

Details

* The sequence number (columns 8 - 10) is reset to 1 for each new site.

* SITE identifiers (columns 12 - 14) should be fully explained in remark 800.

* If a site is composed of more than four residues, these may be specified on additional records bearing the same site identifier.

* SITE records can include HET groups.

Verification/Validation/Value Authority Control

Every SITE must have a corresponding description in remark 800. The numbering of sequential SITE records and format of each one is verified, as well as the existence of each residue in the ATOM records.

Relationships to Other Record Types

Each listed SITE needs a corresponding REMARK 800 that details its significance.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
SITE     1 AC1  3 HIS A  94  HIS A  96  HIS A 119
SITE     1 AC2  5 ASN A  62  GLY A  63  HIS A  64  HOH A 328
SITE     2 AC2  5 HOH A 634
SITE     1 AC3  5 GLN A 136  GLN A 137  PRO A 138  GLU A 205
SITE     2 AC3  5 CYS A 206
SITE     1 AC4 11 HIS A  64  HIS A  94  HIS A  96  HIS A 119
SITE     2 AC4 11 LEU A 198  THR A 199  THR A 200  TRP A 209
SITE     3 AC4 11 HOH A 572  HOH A 582  HOH A 635
```

8. Crystallographic and Coordinate Transformation Section

This section describes the geometry of the crystallographic experiment and the coordinate system transformations.

CRYST1

Overview

The CRYST1 record presents the unit cell parameters, space group, and Z value. If the structure was not determined by crystallographic means, CRYST1 simply provides the unitary values, with an appropriate REMARK.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"CRYST1"	
7 - 15	Real(9.3)	a	a (Angstroms).
16 - 24	Real(9.3)	b	b (Angstroms).
25 - 33	Real(9.3)	c	c (Angstroms).
34 - 40	Real(7.2)	alpha	alpha (degrees).
41 - 47	Real(7.2)	beta	beta (degrees).
48 - 54	Real(7.2)	gamma	gamma (degrees).
56 - 66	LString	sGroup	Space group.
67 - 70	Integer	z	Z value.

Details

* If the entry describes a structure determined by a technique other than X-ray crystallography, CRYST1 contains a = b = c = 1.0, alpha = beta = gamma = 90 degrees, space group = P 1, and Z = 1.

* The Hermann-Mauguin space group symbol is given without parenthesis, e.g., P 43 21 2. Please note that the screw axis is described as a two digit number.

* The full International Table’s Hermann-Mauguin symbol is used, e.g., P 1 21 1 instead of P 21.

* For a rhombohedral space group in the hexagonal setting, the lattice type symbol used is H.

* The Z value is the number of polymeric chains in a unit cell. In the case of heteropolymers, Z is the number of occurrences of the most populous chain.

As an example, given two chains A and B, each with a different sequence, and the space group P 2 that has two equipoints in the standard unit cell, the following table gives the correct Z value.

Asymmetric Unit Content	Z value
A	2
AA	4
AB	2
AAB	4
AABB	4

* In the case of a polycrystalline fiber diffraction study, CRYST1 and SCALE contain the normal unit cell data.

Verification/Validation/Value Authority Control

The given space group and Z values are checked during processing for correctness and internal consistency. The calculated SCALE factor is compared to that supplied by the depositor. Packing is also computed, and close contacts of symmetry-related molecules are diagnosed.

Relationships to Other Record Types

The unit cell parameters are used to calculate SCALE. If the EXPDTA record is NMR, Electron microscopy, or Fiber Diffraction, the CRYST1 record is predefined as in the last example (see below). In these cases, an explanatory REMARK must also appear in the entry. Some fiber diffraction structures will be done this way, while others will have a CRYST1 record containing measured values.

Examples

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
CRYST1   52.000   58.600   61.900  90.00  90.00  90.00 P 21 21 21    8

CRYST1   42.544   69.085   50.950  90.00  95.55  90.00 P 1 21 1      2
```

Example of experimental method other than X-ray crystallography or fiber diffraction

```
CRYST1    1.000    1.000    1.000  90.00  90.00  90.00 P 1           1
```

Known Problems

No standard deviations are given.

ORIGXn

Overview

The ORIGXn (n = 1, 2, or 3) records present the transformation from the orthogonal coordinates contained in the entry to the submitted coordinates.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"ORIGXn"	n=1, 2, or 3
11 - 20	Real(10.6)	o[n][1]	On1
21 - 30	Real(10.6)	o[n][2]	On2
31 - 40	Real(10.6)	o[n][3]	On3
46 - 55	Real(10.5)	t[n]	Tn

Details

* This information is included in the file even if the transformation is an identity transformation (unitary matrix, null translation vector). See the SCALE section of this document for a definition of the default orthogonal Angstroms system.

* If the original submitted coordinates are Xsub, Ysub, Zsub and the orthogonal Angstroms coordinates contained in the data entry are X, Y, Z, then:

```
Xsub = O11X + O12Y + O13Z + T1

Ysub = O21X + O22Y + O23Z + T2

Zsub = O31X + O32Y + O33Z + T3
```

Verification/Validation/Value Authority Control

If the coordinates are submitted in the same orthogonal Angstrom coordinate frame as they appear in the entry (the usual case), then ORIGX is an identity matrix with a null translation vector.
If the transformation is not an identity matrix with a null translation vector, then applying this transformation to the coordinates in the entry yields the coordinates of the original deposited file.

Relationships to Other Record Types

ORIGX relates the coordinates in the ATOM and HETATM records to the coordinates in the file.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
ORIGX1      0.963457  0.136613  0.230424       16.61000
```

```
ORIGX2      -0.158977  0.983924  0.081383        13.72000
ORIGX3      -0.215598 -0.115048  0.969683        37.65000
```

SCALEn

Overview

The SCALEn (n = 1, 2, or 3) records present the transformation from the orthogonal coordinates as contained in the entry to fractional crystallographic coordinates. Non-standard coordinate systems should be explained in the remarks.

Record Format

```
COLUMNS         DATA TYPE      FIELD             DEFINITION
------------------------------------------------------------------
 1 -  6         Record name    "SCALEn"          n=1, 2, or 3

11 - 20         Real(10.6)     s[n][1]           Sn1

21 - 30         Real(10.6)     s[n][2]           Sn2

31 - 40         Real(10.6)     s[n][3]           Sn3

46 - 55         Real(10.5)     u[n]              Un
```

Details

* The standard orthogonal Angstroms coordinate system used is related to the axial system of the unit cell supplied (CRYST1 record) by the following definition:

> If vector a, vector b, vector c describe the crystallographic cell edges, and vector A, vector B, vector C are unit cell vectors in the default orthogonal Angstroms system, then vector A, vector B, vector C and vector a, vector b, vector c have the same origin; vector A is parallel to vector a, vector B is parallel to vector C times vector A, and vector C is parallel to vector a times vector b (i.e., vector c*). * If the orthogonal Angstroms coordinates are X, Y, Z, and the fractional cell coordinates are xfrac, yfrac, zfrac, then:

```
xfrac = S11X + S12Y + S13Z + U1

yfrac = S21X + S22Y + S23Z + U2

zfrac = S31X + S32Y + S33Z + U3
```

* For NMR, fiber diffraction, and EM entries, SCALE is given as an identity matrix with no translation.

Verification/Validation/Value Authority Control

The inverse of the determinant of the SCALE matrix equals the volume of the cell. This volume is calculated and compared to the SCALE matrix supplied by the depositor.

Relationships to Other Record Types

The SCALE transformation is related to the CRYST1 record, as the inverse of the determinant of the SCALE matrix equals the cell volume.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
SCALE1      0.019231  0.000000  0.000000        0.00000
SCALE2      0.000000  0.017065  0.000000        0.00000
SCALE3      0.000000  0.000000  0.016155        0.00000
```

MTRIXn

Overview

The MTRIXn (n = 1, 2, or 3) records present transformations expressing non-crystallographic symmetry. MTRIXn will appear only when such transformations are required to generate an entire asymmetric unit, such as a large viral structure.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"MTRIXn"	n=1, 2, or 3
8 - 10	Integer	serial	Serial number.
11 - 20	Real(10.6)	m[n][1]	Mn1
21 - 30	Real(10.6)	m[n][2]	Mn2
31 - 40	Real(10.6)	m[n][3]	Mn3
46 - 55	Real(10.5)	v[n]	Vn
60	Integer	iGiven	1 if coordinates for the representations which are approximately related by the transformations of the molecule are contained in the entry. Otherwise, blank.

Details

* The MTRIX transformations operate on the coordinates in the entry to yield equivalent representations of the molecule in the same coordinate frame. One trio of MTRIX records with a constant serial number is given for each non-crystallographic symmetry operation defined. If coordinates for the representations which are approximately related by the given transformation are present in the file, the last “iGiven” field is set to 1. Otherwise, this field is blank.

Verification/Validation/Value Authority Control

All MTRIX records are verified using records from the author and review.

Relationships to Other Record Types

None.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
MTRIX1   1 -1.000000  0.000000  0.000000        0.00000    1
MTRIX2   1  0.000000  1.000000  0.000000        0.00000    1
MTRIX3   1  0.000000  0.000000 -1.000000        0.00000    1
```

9. Coordinate Section

The Coordinate Section contains the collection of atomic coordinates as well as the MODEL and ENDMDL records.

MODEL

Overview

The MODEL record specifies the model serial number when multiple models of the same structure are presented in a single coordinate entry, as is often the case with structures determined by NMR.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"MODEL "	
11 - 14	Integer	serial	Model serial number.

Details

*This record is used only when more than one model appears in an entry. Generally, it is employed mainly for NMR structures. The chemical connectivity should be the same for each model. ATOM, HETATM, ANISOU, and TER records for each model structure and are interspersed as needed between MODEL and ENDMDL records.

*The numbering of models is sequential, beginning with 1.

* All models in a deposition should be superimposed in an appropriate author determined manner and only one superposition method should be used. Structures from different experiments, or different domains of a structure should not be superimposed and deposited as models of a deposition.

* All models in an NMR ensemble should be homogeneous – each model should have the exact same atoms (hydrogen and heavy atoms), sequence and chemistry.

*Deposition of minimized average structure (if available) must be accompanied with ensemble and must be homogeneous with ensemble.

*If a collection contains more than 99,999 total atoms, then more than one entry must be made.

Verification/Validation/Value Authority Control

Entries with multiple models in the NUMMDL record are checked for corresponding pairs of MODEL/

ENDMDL records, and for consecutively numbered models.

Relationships to Other Record Types

Each MODEL must have a corresponding ENDMDL record.

Examples

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
MODEL        1
ATOM      1  N   ALA A   1      11.104   6.134  -6.504  1.00  0.00           N
ATOM      2  CA  ALA A   1      11.639   6.071  -5.147  1.00  0.00           C
…
…
…
ATOM    293 1HG  GLU A  18     -14.861  -4.847   0.361  1.00  0.00           H
ATOM    294 2HG  GLU A  18     -13.518  -3.769   0.084  1.00  0.00           H
TER     295      GLU A  18
ENDMDL
MODEL        2
ATOM    296  N   ALA A   1      10.883   6.779  -6.464  1.00  0.00           N
ATOM    297  CA  ALA A   1      11.451   6.531  -5.142  1.00  0.00           C
…
…
ATOM    588 1HG  GLU A  18     -13.363  -4.163  -2.372  1.00  0.00           H
ATOM    589 2HG  GLU A  18     -12.634  -3.023  -3.475  1.00  0.00           H
TER     590      GLU A  18
ENDMDL
```

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
MODEL        1
ATOM      1  N  AALA A   1      72.883  57.697  56.410  0.50 83.80           N
ATOM      2  CA AALA A   1      73.796  56.531  56.644  0.50 84.78           C
ATOM      3  C  AALA A   1      74.549  56.551  57.997  0.50 85.05           C
ATOM      4  O  AALA A   1      73.951  56.413  59.075  0.50 84.77           O
…
…
…
HETATM37900  O  AHOH   490     -24.915 147.513  36.413  0.50 41.86           O
HETATM37901  O  AHOH   491     -28.699 130.471  22.248  0.50 36.06           O
HETATM37902  O  AHOH   492     -33.309 184.488  26.176  0.50 15.00           O
ENDMDL
MODEL        2
ATOM      1  N  BALA A   1      72.883  57.697  56.410  0.50 83.80           N
ATOM      2  CA BALA A   1      73.796  56.531  56.644  0.50 84.78           C
ATOM      3  C  BALA A   1      74.549  56.551  57.997  0.50 85.05           C
ATOM      4  O  BALA A   1      73.951  56.413  59.075  0.50 84.77           O
ATOM      5  CB BALA A   1      74.804  56.369  55.453  0.50 84.29           C
ATOM      6  N  BASP A   2      75.872  56.703  57.905  0.50 85.59           N
ATOM      7  CA BASP A   2      76.801  56.651  59.048  0.50 85.67           C
ATOM      8  C  BASP A   2      76.283  57.361  60.309  0.50 84.80           C
…
```

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.

* No ordering is specified for polysaccharides.

* 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.

* If an atom is provided in more than one position, then a non-blank alternate location indicator must be used as the alternate location indicator (Ares, Bres, where res=3-letter code for amino acid) 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.

* 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.

* The insertion code is commonly used in sequence numbering.

* 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.

Example

```
         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
-----------------------------------------------------------------
 1 - 6        Record name   "ANISOU"
```

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	Insertion code.
29 - 35	Integer	u[0][0]	U(1,1)
36 - 42	Integer	u[1][1]	U(2,2)
43 - 49	Integer	u[2][2]	U(3,3)
50 - 56	Integer	u[0][1]	U(1,2)
57 - 63	Integer	u[0][2]	U(1,3)
64 - 70	Integer	u[1][2]	U(2,3)
77 - 78	LString(2)	element	Element symbol, right-justified.
79 - 80	LString(2)	charge	Charge on the atom.

Details

* Columns 7 - 27 and 73 - 80 are identical to the corresponding ATOM/HETATM record.

* The anisotropic temperature factors (columns 29 - 70) are scaled by a factor of 10**4 (Angstroms**2) and are presented as integers.

* The anisotropic temperature factors are stored in the same coordinate frame as the atomic coordinate records.

* ANISOU values are listed only if they have been provided by the depositor.

Verification/Validation/Value Authority Control

The depositor provides ANISOU records, and the wwPDB verifies their format.

Relationships to Other Record Types

The anisotropic temperature factors are related to the corresponding ATOM/HETATM isotropic temperature factors as ,B(eq), as described in the ATOM and HETATM sections.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
ATOM    107  N   GLY A  13      12.681  37.302 -25.211 1.000 15.56           N
ANISOU  107  N   GLY A  13     2406   1892   1614    198    519   -328       N
ATOM    108  CA  GLY A  13      11.982  37.996 -26.241 1.000 16.92           C
ANISOU  108  CA  GLY A  13     2748   2004   1679    -21    155   -419       C
ATOM    109  C   GLY A  13      11.678  39.447 -26.008 1.000 15.73           C
ANISOU  109  C   GLY A  13     2555   1955   1468     87    357   -109       C
ATOM    110  O   GLY A  13      11.444  40.201 -26.971 1.000 20.93           O
ANISOU  110  O   GLY A  13     3837   2505   1611    164   -121    189       O
ATOM    111  N   ASN A  14      11.608  39.863 -24.755 1.000 13.68           N
ANISOU  111  N   ASN A  14     2059   1674   1462     27    244    -96       N
```

Relationships to Other Record Types

The standard deviations for the anisotropic temperature factors are related to the corresponding ATOM/ HETATM ANISOU temperature factors.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
ATOM    107  N   GLY A  13      12.681  37.302 -25.211 1.000 15.56           N
ANISOU  107  N   GLY A  13     2406   1892   1614    198    519   -328       N
SIGUIJ  107  N   GLY A  13       10     10     10     10    10      10       N
ATOM    108  CA  GLY A  13      11.982  37.996 -26.241 1.000 16.92           C
ANISOU  108  CA  GLY A  13     2748   2004   1679    -21    155   -419       C
SIGUIJ  108  CA  GLY A  13       10     10     10     10    10      10       C
ATOM    109  C   GLY A  13      11.678  39.447 -26.008 1.000 15.73           C
ANISOU  109  C   GLY A  13     2555   1955   1468     87    357   -109       C
SIGUIJ  109  C   GLY A  13       10     10     10     10    10      10       C
ATOM    110  O   GLY A  13      11.444  40.201 -26.971 1.000 20.93           O
ANISOU  110  O   GLY A  13     3837   2505   1611    164   -121    189       O
SIGUIJ  110  O   GLY A  13       10     10     10     10    10      10       O
ATOM    111  N   ASN A  14      11.608  39.863 -24.755 1.000 13.68           N
ANISOU  111  N   ASN A  14     2059   1674   1462     27    244    -96       N
SIGUIJ  111  N   ASN A  14       10     10     10     10    10      10       N
```

TER

Overview

The TER record indicates the end of a list of ATOM/HETATM records for a chain.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"TER "	
7 - 11	Integer	serial	Serial number.
18 - 20	Residue name	resName	Residue name.
22	Character	chainID	Chain identifier.
23 - 26	Integer	resSeq	Residue sequence number.
27	AChar	iCode	Insertion code.

Details

* Every chain of ATOM/HETATM records presented on SEQRES records is terminated with a TER record.

* The TER records occur in the coordinate section of the entry, and indicate the last residue presented for each polypeptide and/or nucleic acid chain for which there are determined coordinates. For proteins, the residue defined on the TER record is the carboxy-terminal residue; for nucleic acids it is the 3'-terminal residue.

* For a cyclic molecule, the choice of termini is arbitrary.

* Terminal oxygen atoms are presented as OXT for proteins, and as O5’ or OP3 for nucleic acids. These atoms are present only if the last residue in the polymer is truly the last residue in the SEQRES.

* The TER record has the same residue name, chain identifier, sequence number and insertion code as the terminal residue. The serial number of the TER record is one number greater than the serial number of the ATOM/HETATM preceding the TER.

Verification/Validation/Value Authority Control

TER must appear at the terminal carboxyl end or 3' end of a chain. For proteins, there is usually a terminal oxygen, labeled OXT. The validation program checks for the occurrence of TER and OXT records.

Relationships to Other Record Types

The residue name appearing on the TER record must be the same as the residue name of the immediately preceding ATOM or non-water HETATM record.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
ATOM    601  N   LEU A  75     -17.070 -16.002   2.409  1.00 55.63           N
ATOM    602  CA  LEU A  75     -16.343 -16.746   3.444  1.00 55.50           C
ATOM    603  C   LEU A  75     -16.499 -18.263   3.300  1.00 55.55           C
ATOM    604  O   LEU A  75     -16.645 -18.789   2.195  1.00 55.50           O
ATOM    605  CB  LEU A  75     -16.776 -16.283   4.844  1.00 55.51           C
TER     606      LEU A  75
…
ATOM   1185  O   LEU B  75      26.292  -4.310  16.940  1.00 55.45           O
ATOM   1186  CB  LEU B  75      23.881  -1.551  16.797  1.00 55.32           C
TER    1187      LEU B  75
HETATM 1188  H2  SRT A1076     -17.263  11.260  28.634  1.00 59.62           H
HETATM 1189  HA  SRT A1076     -19.347  11.519  28.341  1.00 59.42           H
HETATM 1190  H3  SRT A1076     -17.157  14.303  28.677  1.00 58.00           H
HETATM 1191  HB  SRT A1076     -15.110  13.610  28.816  1.00 57.77           H
HETATM 1192  O1  SRT A1076     -17.028  11.281  31.131  1.00 62.63           O

ATOM    295  HB2 ALA A  18       4.601  -9.393   7.275  1.00  0.00           H
ATOM    296  HB3 ALA A  18       3.340  -9.147   6.043  1.00  0.00           H
TER     297      ALA A  18
ENDMDL
```

HETATM

Overview

Non-polymer or other “non-standard” chemical coordinates, such as water molecules or atoms presented in HET groups use the HETATM record type. They also present the occupancy and temperature factor for each atom. The ATOM records present the atomic coordinates for standard residues. The element symbol is always present on each HETATM record; charge is optional.

Changes in ATOM/HETATM records will require standardization in 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	"HETATM"	
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.
39 - 46	Real(8.3)	y	Orthogonal coordinates for Y.
47 - 54	Real(8.3)	z	Orthogonal coordinates for Z.
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

* The x, y, z coordinates are in Angstrom units.

* No ordering is specified for polysaccharides.

* See the HET section of this document regarding naming of heterogens. See the Chemical Component Dictionary for residue names, formulas, and topology of the HET groups that have appeared so far in the PDB (see ftp://ftp.wwpdb.org/pub/pdb/data/monomers).

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

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

* Insertion codes and element naming are fully described in the ATOM section of this document.

Verification/Validation/Value Authority Control

Processing programs check ATOM/HETATM records for PDB file format, sequence information, and packing.

Relationships to Other Record Types

HETATM records must have corresponding HET, HETNAM, FORMUL and CONECT records, except for waters.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
HETATM 8237 MG    MG A1001      13.872  -2.555 -29.045  1.00 27.36          MG

HETATM 3835 FE   HEM A   1      17.140   3.115  15.066  1.00 14.14          FE
HETATM 8238  S   SO4 A2001      10.885 -15.746 -14.404  1.00 47.84           S
HETATM 8239  O1  SO4 A2001      11.191 -14.833 -15.531  1.00 50.12           O
HETATM 8240  O2  SO4 A2001       9.576 -16.338 -14.706  1.00 48.55           O
HETATM 8241  O3  SO4 A2001      11.995 -16.703 -14.431  1.00 49.88           O
HETATM 8242  O4  SO4 A2001      10.932 -15.073 -13.100  1.00 49.91           O
```

ENDMDL

Overview

The ENDMDL records are paired with MODEL records to group individual structures found in a coordinate entry.

Record Format

```
COLUMNS        DATA TYPE     FIELD        DEFINITION
------------------------------------------------------------------
 1 - 6         Record name   "ENDMDL"
```

Details

* MODEL/ENDMDL records are used only when more than one structure is presented in the entry, as is often the case with NMR entries.

* All the models in a multi-model entry must represent the same structure.

* Every MODEL record has an associated ENDMDL record.

Verification/Validation/Value Authority Control

Entries with multiple structures in the NUMMDL record are checked for corresponding pairs of MODEL/ ENDMDL records, and for consecutively numbered models.

Relationships to Other Record Types

There must be a corresponding MODEL record.

In the case of an NMR entry, the NUMMDL record states the number of model structures that are present in the individual entry.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
...
...
ATOM  14550 1HG  GLU   122     -14.364  14.787 -14.258  1.00  0.00           H
ATOM  14551 2HG  GLU   122     -13.794  13.738 -12.961  1.00  0.00           H
TER   14552      GLU   122
ENDMDL
MODEL        9
ATOM  14553  N   SER     1     -28.280   1.567  12.004  1.00  0.00           N
ATOM  14554  CA  SER     1     -27.749   0.392  11.256  1.00  0.00           C
...
...
ATOM  16369 1HG  GLU   122      -3.757  18.546  -8.439  1.00  0.00           H
ATOM  16370 2HG  GLU   122      -3.066  17.166  -7.584  1.00  0.00           H
```

```
TER   16371      GLU   122
ENDMDL
```

10. Connectivity Section

This section provides information on atomic connectivity. LINK, SSBOND, and CISPEP are found in the Connectivity Annotation section.

CONECT

Overview

The CONECT records specify connectivity between atoms for which coordinates are supplied. The connectivity is described using the atom serial number as shown in the entry. CONECT records are mandatory for HET groups (excluding water) and for other bonds not specified in the standard residue connectivity table. These records are generated automatically.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"CONECT"	
7 - 11	Integer	serial	Atom serial number
12 - 16	Integer	serial	Serial number of bonded atom
17 - 21	Integer	serial	Serial number of bonded atom
22 - 26	Integer	serial	Serial number of bonded atom
27 - 31	Integer	serial	Serial number of bonded atom

Details

* CONECT records are present for:

- Intra-residue connectivity within non-standard (HET) residues (excluding water).
- Inter-residue connectivity of HET groups to standard groups (including water) or to other HET groups.
- Disulfide bridges specified in the SSBOND records have corresponding records.

* No differentiation is made between atoms with delocalized charges (excess negative or positive charge).

* Atoms specified in the CONECT records have the same numbers as given in the coordinate section.

* All atoms connected to the atom with serial number in columns 7 - 11 are listed in the remaining

fields of the record.

* If more than four fields are required for non-hydrogen and non-salt bridges, a second CONECT record with the same atom serial number in columns 7 - 11 will be used.

* These CONECT records occur in increasing order of the atom serial numbers they carry in columns 7 - 11. The target-atom serial numbers carried on these records also occur in increasing order.

* The connectivity list given here is redundant in that each bond indicated is given twice, once with each of the two atoms involved specified in columns 7 - 11.

* For hydrogen bonds, when the hydrogen atom is present in the coordinates, a CONECT record between the hydrogen atom and its acceptor atom is generated.

* For NMR entries, CONECT records for one model are generated describing heterogen connectivity and others for LINK records assuming that all models are homogeneous models.

Verification/Validation/Value Authority Control

Connectivity is checked for unusual bond lengths.

Relationships to Other Record Types

CONECT records must be present in an entry that contains either non-standard groups or disulfide bonds.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
CONECT 1179  746 1184 1195 1203
CONECT 1179 1211 1222
CONECT 1021  544 1017 1020 1022 1211 1222      1311
```

Known Problems

CONECT records involving atoms for which the coordinates are not present in the entry (e.g., symmetry-generated) are not given.

CONECT records involving atoms for which the coordinates are missing due to disorder, are also not provided.

11. Bookkeeping Section

The Bookkeeping Section provides some final information about the file itself.

MASTER

Overview

The MASTER record is a control record for bookkeeping. It lists the number of lines in the coordinate entry or file for selected record types. MASTER records only the first model when there are multiple models in the coordinates.

Record Format

COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"MASTER"	
11 - 15	Integer	numRemark	Number of REMARK records
16 - 20	Integer	"0"	
21 - 25	Integer	numHet	Number of HET records
26 - 30	Integer	numHelix	Number of HELIX records
31 - 35	Integer	numSheet	Number of SHEET records
36 - 40	Integer	numTurn	deprecated
41 - 45	Integer	numSite	Number of SITE records
46 - 50	Integer	numXform	Number of coordinate transformation records (ORIGX+SCALE+MTRIX)
51 - 55	Integer	numCoord	Number of atomic coordinate records records (ATOM+HETATM)
56 - 60	Integer	numTer	Number of TER records
61 - 65	Integer	numConect	Number of CONECT records
66 - 70	Integer	numSeq	Number of SEQRES records

Details

* MASTER gives checksums of the number of records in the entry, for selected record types.
* MASTER records only the first model when there are multiple models in the coordinates.

Verification/Validation/Value Authority Control

The MASTER line is automatically generated.

Relationships to Other Record Types

MASTER presents a checksum of the lines present for each of the record types listed above.

Example

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
MASTER       40    0    0    0    0    0    0    6 2930    2    0   29
```

END

Overview

The END record marks the end of the PDB file.

Record Format

```
COLUMNS       DATA TYPE     FIELD         DEFINITION
-------------------------------------------------------
 1 -  6       Record name   "END   "
```

Details

* END is the final record of a coordinate entry.

Verification/Validation/Value Authority Control

END must appear in every coordinate entry.

Relationships to Other Record Types

This is the final record in the entry.

Example

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