W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 09:10 PM GMT

PDB ID	:	1NT9
Title	:	Complete 12-subunit RNA polymerase II
Authors	:	Armache, K.-J.; Kettenberger, H.; Cramer, P.
Deposited on	:	2003-01-29
Resolution	:	4.20 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

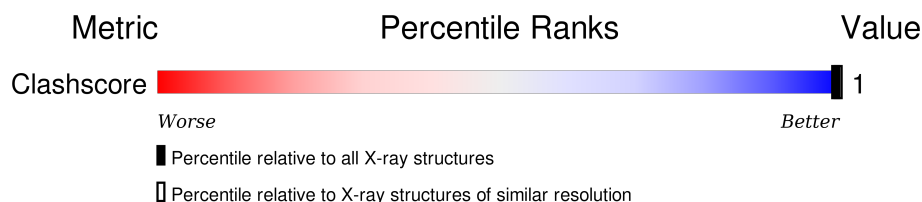|		
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 4.20 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

Metric	Percentile Ranks	Value
Clashscore		1

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	102246	1140 (4.80-3.60)

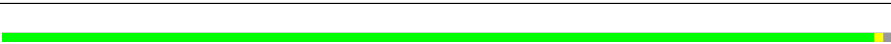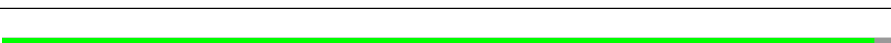The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	1733	80% 20%
2	B	1224	89% 11%
3	C	318	84% 16%
4	D	221	44% 56%
5	E	215	99%
6	F	155	54% 46%
7	G	171	98%
8	H	146	91% 9%
9	I	122	98%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
10	J	70	93% 7%
11	K	120	95% 5%
12	L	70	66% 34%

2 Entry composition (i)

There are 12 unique types of molecules in this entry. The entry contains 3779 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called DNA-DIRECTED RNA POLYMERASE II LARGEST SUB-UNIT.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
1	A	1383	Total 1383	C 1383	0	0	1383

- Molecule 2 is a protein called DNA-directed polymerase II SECOND LARGEST SUBUNIT.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
2	B	1089	Total 1089	C 1089	0	0	1089

- Molecule 3 is a protein called DNA-directed RNA polymerase II 45 kDa polypeptide.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
3	C	266	Total 266	C 266	0	0	266

- Molecule 4 is a protein called YEAST DNA-DIRECTED RNA POLYMERASE II 32 KD POLYPEPTIDE.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
4	D	97	Total 97	C 97	0	0	97

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
5	E	214	Total 214	C 214	0	0	214

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
6	F	84	Total 84	C 84	0	0	84

- Molecule 7 is a protein called DNA-DIRECTED RNA POLYMERASE II 19 KD POLYPEP-TIDE.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
7	G	169	Total 169	C 169	0	0	169

- Molecule 8 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III 14.5 KD POLYPEPTIDE.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
8	H	133	Total 133	C 133	0	0	133

- Molecule 9 is a protein called DNA-directed RNA polymerase II 14.2 kDa polypeptide.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
9	I	119	Total 119	C 119	0	0	119

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III 8.3 kDa polypep-tide.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
10	J	65	Total 65	C 65	0	0	65

- Molecule 11 is a protein called DNA-DIRECTED RNA POLYMERASE II 13.6 KD POLYPEP-TIDE.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
11	K	114	Total 114	C 114	0	0	114

- Molecule 12 is a protein called DNA-directed RNA polymerases I, II, and III 7.7 kDa polypep-tide.

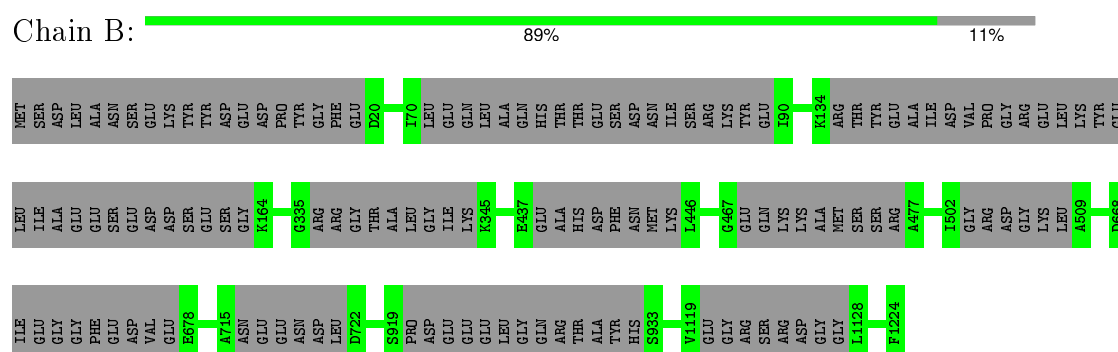| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace
12	L	46	Total 46	C 46	0	0	46

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

Note EDS was not executed.

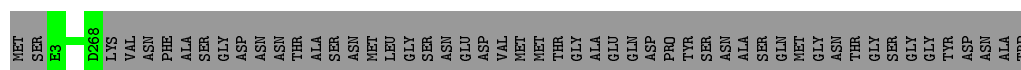• Molecule 1: DNA-DIRECTED RNA POLYMERASE II LARGEST SUBUNIT
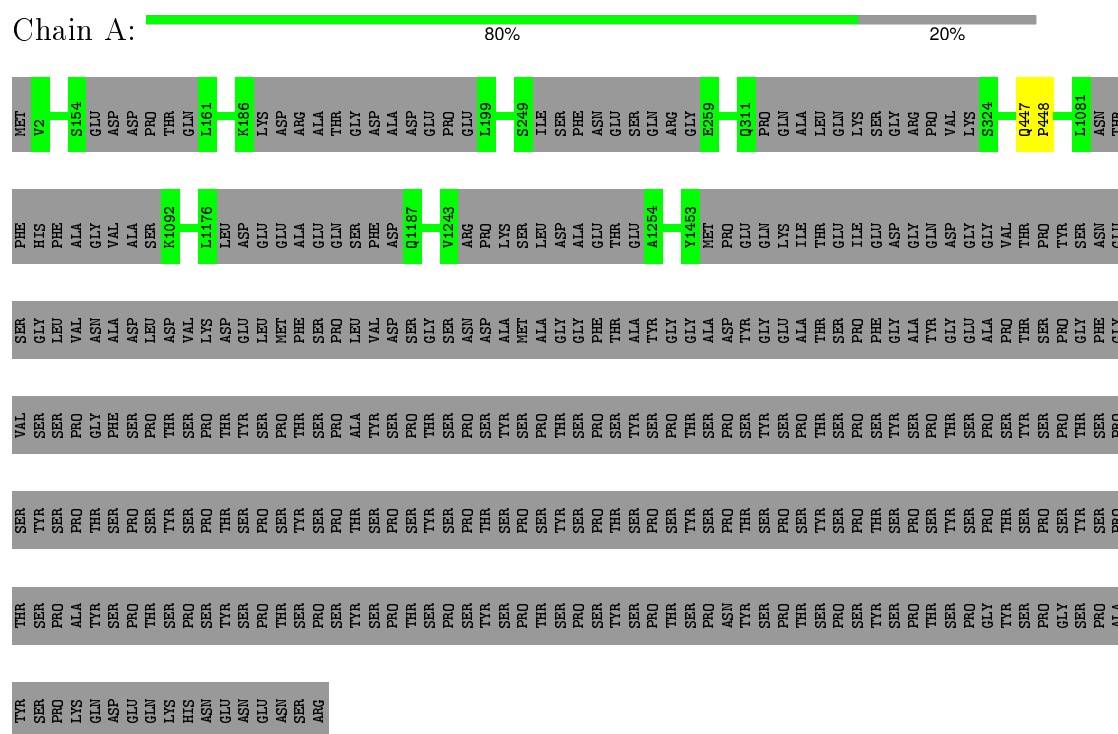
• Molecule 2: DNA-directed polymerase II SECOND LARGEST SUBUNIT
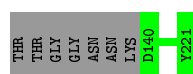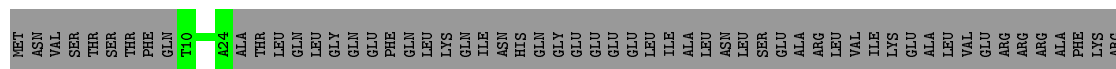

• Molecule 3: DNA-directed RNA polymerase II 45 kDa polypeptide

Chain C:

84%   16%

MET SER E3 D268 LYS VAL ASN PHE ALA SER GLY ASP ASN ASN ALA SER THR ALA SER MET LEU GLY SER ASN GLU ASP VAL MET MET THR GLY ALA GLU GLN ASP PRO TYR SER ASN SER GLN GLY MET GLY ASN THR GLY SER GLY GLY TYR ASP ASN ALA TRP

• Molecule 4: YEAST DNA-DIRECTED RNA POLYMERASE II 32 KD POLYPEPTIDE

Chain D:

44%   56%

MET ASN VAL SER SER THR SER PHE THR PHE GLN T10 A24 ALA THR LEU GLN LEU GLY GLN GLU GLU PHE GLN LEU LYS GLN ILE ASN HIS GLN GLY GLU GLU GLU GLU GLU LEU LEU ILE ALA LEU ASN LEU SER GLU ALA ARG LEU VAL ILE LYS GLU ALA LEU LEU VAL GLU ARG ARG ALA ALA PHE LYS ARG

SER GLN LYS LYS HIS LYS LYS HIS LEU LYS HIS GLU ASN ALA ASN ASP GLU THR ALA VAL GLU ASP ASP ASP GLU ASP VAL ASN ALA ASP ASP ASP PHE MET HIS SER GLU THR ARG GLU LYS GLU LEU GLU SER ILE ASP VAL LEU LEU GLU GLN

THR THR GLY GLY ASN ASN LYS D140 Y221

• Molecule 5: DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide

Chain E:

99%   .

MET D2 P128 P129 M215

• Molecule 6: DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide

Chain F:

54%   46%

MET SER ASP TYR GLU GLU ALA PHE ASN ASP GLY ASN ASP GLN ASN PHE GLU ASP PHE ASP VAL GLU HIS PHE SER ASP GLU GLU THR GLU GLU LYS TYR GLU GLU LYS PRO GLN PHE LYS ASP GLY GLU THR ASP ALA ASN GLY LYS THR ILE VAL THR GLY GLY ASN GLY PRO GLU ASP PHE GLN GLN

HIS GLU GLN ILE ARG ARG LYS THR LEU LYS GLU K72 L155
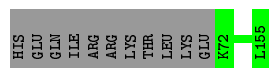
• Molecule 7: DNA-DIRECTED RNA POLYMERASE II 19 KD POLYPEPTIDE

Chain G:

98%   ..

M1 H97 G169 ALA ILE

• Molecule 8: DNA-DIRECTED RNA POLYMERASES I, II, AND III 14.5 KD POLYPEP-TIDE

Chain H:

91%   9%

MET S2 L63 ASN LEU GLU ASP THR PRO ALA ASN ASP SER SER ALA T76 R146
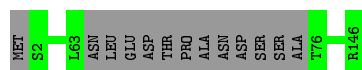
• Molecule 9: DNA-directed RNA polymerase II 14.2 kDa polypeptide

Chain I:

98%   .

MET T2 Q120 PHE SER
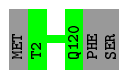
• Molecule 10: DNA-directed RNA polymerases I, II, and III 8.3 kDa polypeptide

Chain J:

93%   7%

M1 P65 LEU GLU LYS ARG ASP

• Molecule 11: DNA-DIRECTED RNA POLYMERASE II 13.6 KD POLYPEPTIDE
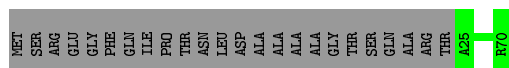
Chain K:

95%   5%

M1 L114 ALA ALA ASP ASP ALA PHE

• Molecule 12: DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide

Chain L:

66%   34%

MET SER ARG GLU GLY PHE GLN ILE PRO THR ASN LEU ASP ALA ALA ALA GLY THR SER GLU ALA ARG THR A25 R70

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	220.40Å 391.40Å 282.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 4.20	Depositor
% Data completeness (in resolution range)	93.8 (50.00-4.20)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
Refinement program	O, CNS	Depositor
R, R_{free}	(Not available) , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3779	wwPDB-VP
Average B, all atoms (Å²)	20.0	wwPDB-VP

5 Model quality (i)

5.1 Standard geometry (i)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

There are no protein, RNA or DNA chains available to summarize Z scores of covalent bonds and angles.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1383	0	0	1	0
2	B	1089	0	0	0	0
3	C	266	0	0	0	0
4	D	97	0	0	0	0
5	E	214	0	0	1	0
6	F	84	0	0	0	0
7	G	169	0	0	0	1
8	H	133	0	0	0	0
9	I	119	0	0	0	0
10	J	65	0	0	0	0
11	K	114	0	0	0	0
12	L	46	0	0	0	0
All	All	3779	0	0	2	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 1.

All (2) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:447:GLN:CA	1:A:448:PRO:CA	2.90	0.49
5:E:128:PRO:CA	5:E:129:PRO:CA	2.91	0.48

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:G:97:HIS:CA	7:G:97:HIS:CA[3_654]	1.77	0.43

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

There are no protein backbone outliers to report in this entry.

5.3.2 Protein sidechains ⓘ

There are no protein residues with a non-rotameric sidechain to report in this entry.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section will therefore be empty.

6.2 Non-standard residues in protein, DNA, RNA chains (i)

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates (i)

EDS was not executed - this section will therefore be empty.

6.4 Ligands (i)

EDS was not executed - this section will therefore be empty.

6.5 Other polymers (i)

EDS was not executed - this section will therefore be empty.