



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 22, 2018 – 06:11 PM EST

PDB ID : 6C2J
Title : Bacteriophage N4 RNA polymerase II and DNA complex
Authors : Molodtsov, V.; Murakami, K.S.
Deposited on : 2018-01-08
Resolution : 2.50 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 ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030736
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20030736

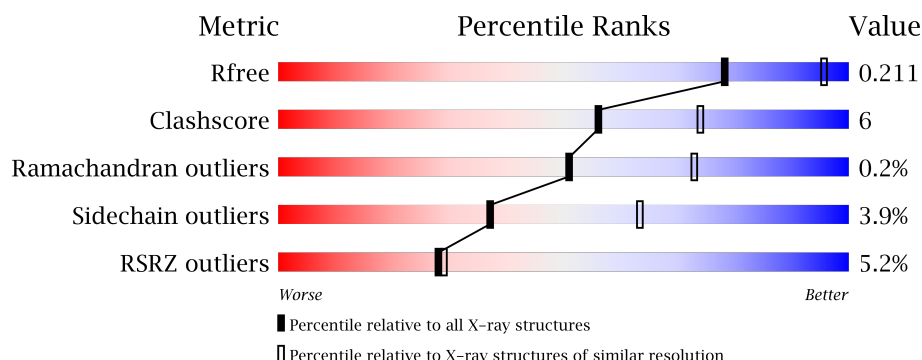
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	3846 (2.50-2.50)
Clashscore	112137	4554 (2.50-2.50)
Ramachandran outliers	110173	4463 (2.50-2.50)
Sidechain outliers	110143	4465 (2.50-2.50)
RSRZ outliers	101464	3876 (2.50-2.50)

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 $\leq 5\%$. 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.

Mol	Chain	Length	Quality of chain
1	A	269	<div> <div>6%</div> <div> <div></div> <div>85%</div> <div>13%</div> <div>..</div> </div> </div>
2	B	404	<div> <div>3%</div> <div> <div></div> <div>86%</div> <div>11%</div> <div>.</div> </div> </div>
3	C	48	<div> <div>10%</div> <div> <div></div> <div>13%</div> <div>81%</div> <div>.</div> </div> </div>

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 5640 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 RNAP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	265	Total	C	N	O	S	0	0	0
			2169	1379	378	399	13			

- Molecule 2 is a protein called RNAP2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	404	Total	C	N	O	S	0	0	0
			3256	2073	545	617	21			

- Molecule 3 is a DNA chain called DNA (5'-D(P*CP*CP*AP*CP*CP*TP*GP*CP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	9	Total	C	N	O	P	0	0	0
			179	85	32	53	9			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	21	Total	O	0	0
			21	21		
4	B	14	Total	O	0	0
			14	14		
4	C	1	Total	O	0	0
			1	1		

4 Data and refinement statistics

Property	Value	Source
Space group	I 4	Depositor
Cell constants a, b, c, α , β , γ	173.32Å 173.32Å 76.32Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.33 – 2.50 43.33 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.9 (43.33-2.50) 99.9 (43.33-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.69 (at 2.51Å)	Xtriage
Refinement program	PHENIX (1.11.1_2575: ???)	Depositor
R, R_{free}	0.189 , 0.214 0.186 , 0.211	Depositor DCC
R_{free} test set	1961 reflections (4.99%)	DCC
Wilson B-factor (Å ²)	65.7	Xtriage
Anisotropy	0.027	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 48.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.020 for -k,-h,-l	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	5640	wwPDB-VP
Average B, all atoms (Å ²)	79.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.20% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z > 5$	RMSZ	$\# Z > 5$
1	A	0.49	0/2221	0.73	3/2997 (0.1%)
2	B	0.50	0/3323	0.71	1/4496 (0.0%)
3	C	0.87	0/199	1.08	1/303 (0.3%)
All	All	0.51	0/5743	0.73	5/7796 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	B	0	1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	177	ASP	CB-CG-OD2	-9.15	110.07	118.30
1	A	177	ASP	CB-CG-OD1	6.74	124.36	118.30
2	B	238	MET	CB-CG-SD	6.22	131.07	112.40
3	C	1	DC	O4'-C4'-C3'	-5.14	102.44	104.50
1	A	177	ASP	CB-CA-C	5.08	120.56	110.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	129	GLY	Mainchain

5.2 Too-close contacts ⓘ

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	2169	0	2099	23	0
2	B	3256	0	3214	41	0
3	C	179	0	101	9	0
4	A	21	0	0	2	0
4	B	14	0	0	0	0
4	C	1	0	0	0	0
All	All	5640	0	5414	64	0

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 6.

The worst 5 of 64 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:4:ILE:H	1:A:4:ILE:HD12	1.54	0.73
2:B:366:ALA:O	2:B:387:LYS:HE2	1.92	0.69
2:B:299:THR:HG22	2:B:301:PHE:H	1.58	0.69
2:B:290:MET:HG2	2:B:317:LEU:HD22	1.75	0.67
1:A:258:ALA:O	4:A:301:HOH:O	2.12	0.67

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	261/269 (97%)	250 (96%)	11 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
2	B	402/404 (100%)	391 (97%)	10 (2%)	1 (0%)	51 73
All	All	663/673 (98%)	641 (97%)	21 (3%)	1 (0%)	51 73

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	130	ALA

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	235/243 (97%)	227 (97%)	8 (3%)	42 69
2	B	351/351 (100%)	336 (96%)	15 (4%)	33 58
All	All	586/594 (99%)	563 (96%)	23 (4%)	37 63

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	126	LEU
2	B	142	MET
2	B	299	THR
2	B	132	ILE
2	B	181	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

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](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2		OWAB(Å ²)	Q<0.9
1	A	265/269 (98%)	0.26	17 (6%)	20 21	43, 74, 115, 146	0
2	B	404/404 (100%)	0.14	13 (3%)	48 51	43, 72, 108, 136	0
3	C	9/48 (18%)	2.27	5 (55%)	0 0	156, 177, 209, 214	0
All	All	678/721 (94%)	0.22	35 (5%)	28 29	43, 73, 118, 214	0

The worst 5 of 35 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	192	PRO	7.3
1	A	197	THR	6.5
1	A	193	LYS	6.2
3	C	4	DC	6.1
1	A	199	GLN	5.6

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

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.