



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 05:35 PM UTC

PDB ID : 6BOG / pdb_00006bog
Title : Crystal structure of RapA, a Swi2/Snf2 protein that recycles RNA polymerase during transcription
Authors : Shaw, G.X.; Gan, J.; Zhou, Y.N.; Zhang, R.; Joachimiak, A.; Jin, D.J.; Ji, X.
Deposited on : 2017-11-20
Resolution : 3.21 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

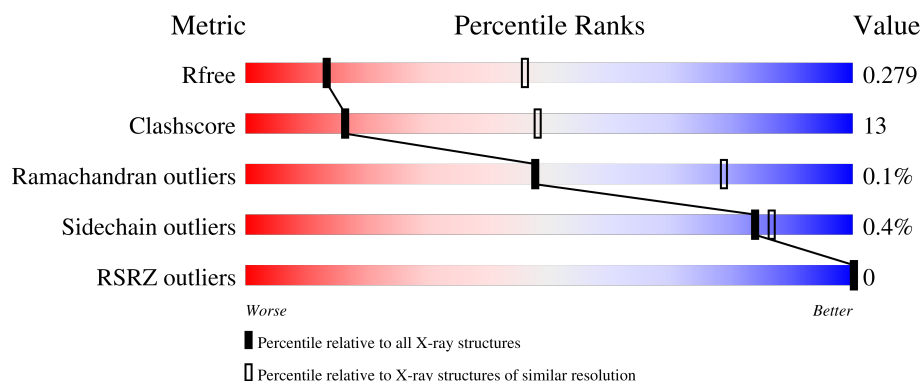
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.21 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}	180053	1466 (3.20-3.20)
Clashscore	190562	1573 (3.20-3.20)
Ramachandran outliers	187476	1548 (3.20-3.20)
Sidechain outliers	187428	1547 (3.20-3.20)
RSRZ outliers	180081	1466 (3.20-3.20)

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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. 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	968	70% 29%
1	B	968	73% 26%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	SO4	A	1003	-	-	X	-
2	SO4	B	1002	-	-	X	-
2	SO4	B	1003	-	-	X	-

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 15473 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 RNA polymerase-associated protein RapA.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	967	Total	C	N	O	S	Se	0	0	0
			7719	4831	1382	1477	6	23			
1	B	967	Total	C	N	O	S	Se	0	0	0
			7719	4831	1382	1477	6	23			

- Molecule 2 is SULFATE ION (CCD ID: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

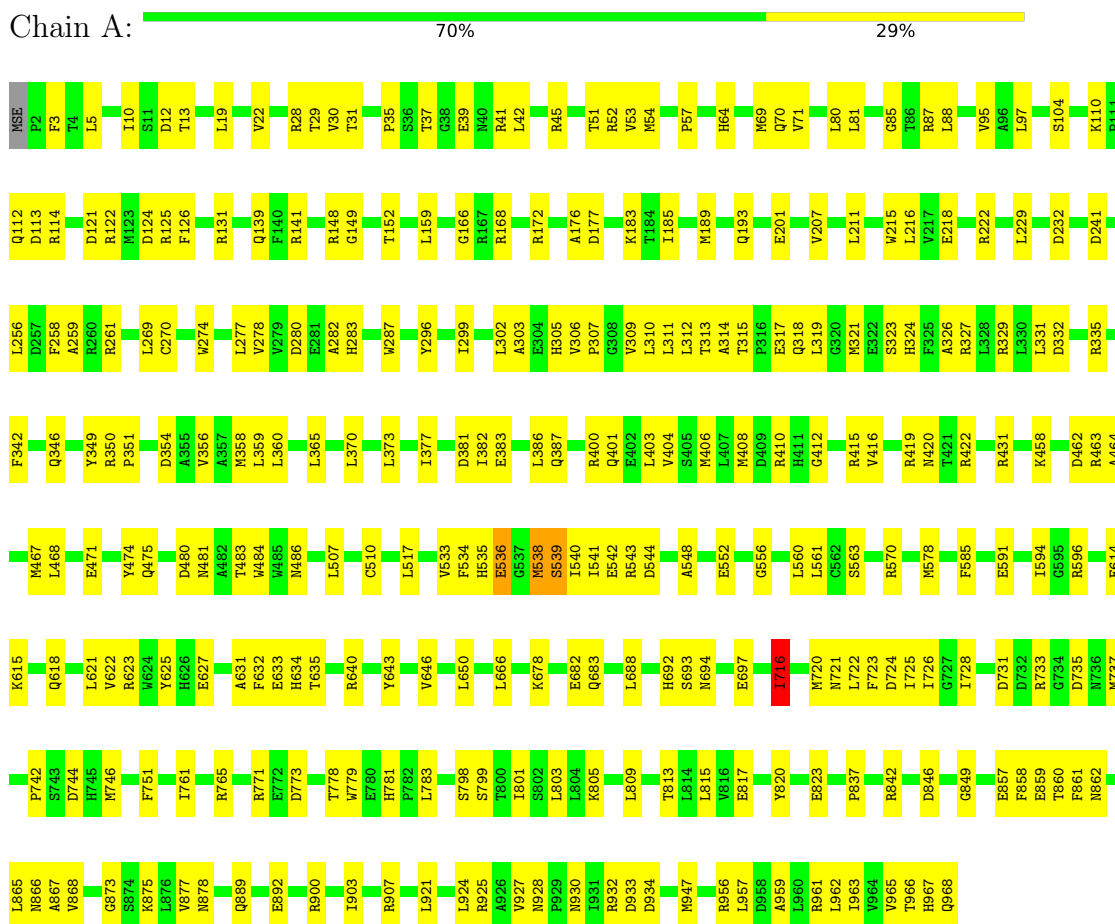
Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		

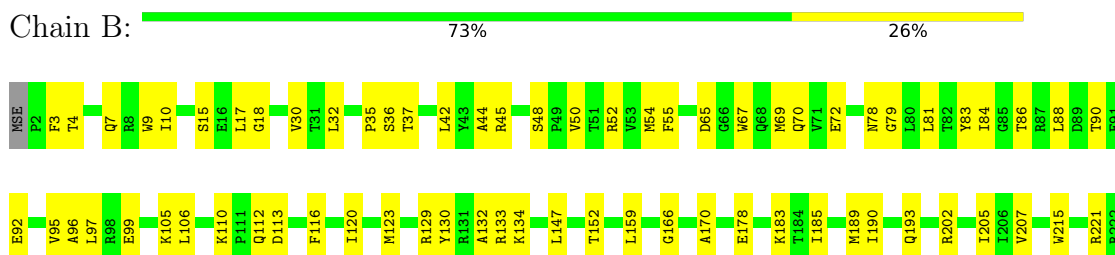
3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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.

- Molecule 1: RNA polymerase-associated protein RapA



- Molecule 1: RNA polymerase-associated protein RapA



F223	A357	K505	Y625	T813	I939
D232	M388	A511	H626	L814	V946
F246	L369	A514	E627	L815	M947
L256	L370	Q518	R640	I819	D951
D257	L386	V522	V646	E823	R966
R260	R400	R526	Y647	L830	L957
R261	V404	R543	L650	R834	D958
S262	L407	V533	R671	P838	A959
K263	M408	F534	G684	R842	L962
E267	D409	H535	R685	R842	I963
C270	R410	R538	D686	D846	V964
E271	R415	S539	R687	N850	V965
A272	N420	E542	L688	R851	T966
E273	K430	R543	L689	L852	H967
W274	H434	D544	E690	V856	Q968
D275	T435	F550	N694	E857	
Y296	L436	C562	I716	F858	
H305	L437	S563	M720	E859	
V309	L438	E564	N721	T860	
L310	P441	I565	I725	R863	
L311	Y444	E568	I726	N866	
L312	Q445	Q569	G727	A867	
T313	K449	R570	I728	Q873	
A314	K458	M578	Q730	V877	
E317	E461	D582	M737	R900	
Q318	D462	L583	I738	I903	
H324	R463	F584	T741	R907	
P326	E471	N586	D744	D911	
A326	R472	P587	H745	E912	
R327	Q475	D588	M746	K913	
L328	Q475	L589	R765	L918	
F336	D460	Q592	R771	L924	
V343	M481	R593	E772	R925	
K347	A482	L597	L768	A926	
N348	T483	D598	I801	V927	
Y349	W484	D605	K805	N928	
R350	W485	I608	N806	R932	
D354	N486	L613	K807	D933	
A355	P489	E614	A808	D934	
V356	W493	K615	L809	E935	
		T616			

4 Data and refinement statistics

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, α , β , γ	123.86Å 123.86Å 187.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.19 – 3.21 29.19 – 3.21	Depositor EDS
% Data completeness (in resolution range)	98.9 (29.19-3.21) 98.9 (29.19-3.21)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.38 (at 3.18Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.244 , 0.279 0.245 , 0.279	Depositor DCC
R_{free} test set	994 reflections (2.17%)	wwPDB-VP
Wilson B-factor (Å ²)	91.6	Xtriage
Anisotropy	0.222	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 72.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.449 for h,-k,-l	Xtriage
Reported twinning fraction	0.500 for h,-k,-l	Depositor
Outliers	0 of 45884 reflections	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	15473	wwPDB-VP
Average B, all atoms (Å ²)	105.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.02% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

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	0.31	1/7840 (0.0%)	0.57	1/10581 (0.0%)
1	B	0.26	1/7840 (0.0%)	0.56	0/10581
All	All	0.29	2/15680 (0.0%)	0.56	1/21162 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	538	MSE	CA-C	-5.69	1.45	1.53
1	A	536	GLU	CA-C	-5.05	1.45	1.52

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	716	ILE	CB-CA-C	-6.24	103.72	112.14

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	7719	0	7583	211	0
1	B	7719	0	7583	192	0
2	A	15	0	0	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	20	0	0	5	0
All	All	15473	0	15166	398	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 13.

The worst 5 of 398 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:B:533:VAL:O	1:B:538:MSE:HE1	1.35	1.22
1:B:737:MSE:HE1	1:B:771:ARG:NH1	1.90	0.85
1:A:540:ILE:HD13	1:A:543:ARG:NH2	1.95	0.81
1:B:70:GLN:HB3	1:B:86:THR:HB	1.62	0.80
1:A:540:ILE:CD1	1:A:543:ARG:NH2	2.47	0.77

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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	965/968 (100%)	930 (96%)	35 (4%)	0	100	100
1	B	965/968 (100%)	928 (96%)	36 (4%)	1 (0%)	48	79
All	All	1930/1936 (100%)	1858 (96%)	71 (4%)	1 (0%)	48	79

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	564	GLU

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	828/805 (103%)	824 (100%)	4 (0%)	81	85
1	B	828/805 (103%)	825 (100%)	3 (0%)	84	86
All	All	1656/1610 (103%)	1649 (100%)	7 (0%)	84	86

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	716	ILE
1	B	538	MSE
1	B	857	GLU
1	B	539	SER
1	A	694	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	504	GLN
1	B	851	ASN
1	B	848	ASN
1	A	694	ASN
1	B	475	GLN

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 [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

7 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
2	SO4	B	1001	-	4,4,4	0.23	0	6,6,6	0.15	0
2	SO4	B	1004	-	4,4,4	0.24	0	6,6,6	0.06	0
2	SO4	B	1002	-	4,4,4	0.24	0	6,6,6	0.14	0
2	SO4	B	1003	-	4,4,4	0.23	0	6,6,6	0.11	0
2	SO4	A	1001	-	4,4,4	0.24	0	6,6,6	0.18	0
2	SO4	A	1002	-	4,4,4	0.24	0	6,6,6	0.18	0
2	SO4	A	1003	-	4,4,4	0.24	0	6,6,6	0.16	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1001	SO4	1	0
2	B	1002	SO4	2	0
2	B	1003	SO4	2	0
2	A	1002	SO4	1	0
2	A	1003	SO4	2	0

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	944/968 (97%)	-1.59	0 100 100	58, 105, 147, 170	0
1	B	944/968 (97%)	-1.61	0 100 100	43, 99, 151, 199	0
All	All	1888/1936 (97%)	-1.60	0 100 100	43, 102, 149, 199	0

There are no RSRZ outliers to report.

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 oligosaccharides in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	SO4	A	1002	5/5	0.99	0.02	74,91,95,99	0
2	SO4	B	1002	5/5	0.99	0.03	74,75,82,103	0
2	SO4	B	1003	5/5	0.99	0.02	113,118,134,134	0
2	SO4	B	1004	5/5	0.99	0.03	84,98,114,122	0
2	SO4	A	1001	5/5	1.00	0.02	56,57,70,88	0
2	SO4	A	1003	5/5	1.00	0.02	95,102,125,133	0
2	SO4	B	1001	5/5	1.00	0.02	64,64,69,78	0

6.5 Other polymers [i](#)

There are no such residues in this entry.