



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2026 – 06:56 AM UTC

PDB ID : 4S20 / pdb_00004s20
Title : Structural basis for transcription reactivation by RapA
Authors : Liu, B.; Zuo, Y.; Steitz, T.A.
Deposited on : 2015-01-16
Resolution : 4.70 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
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

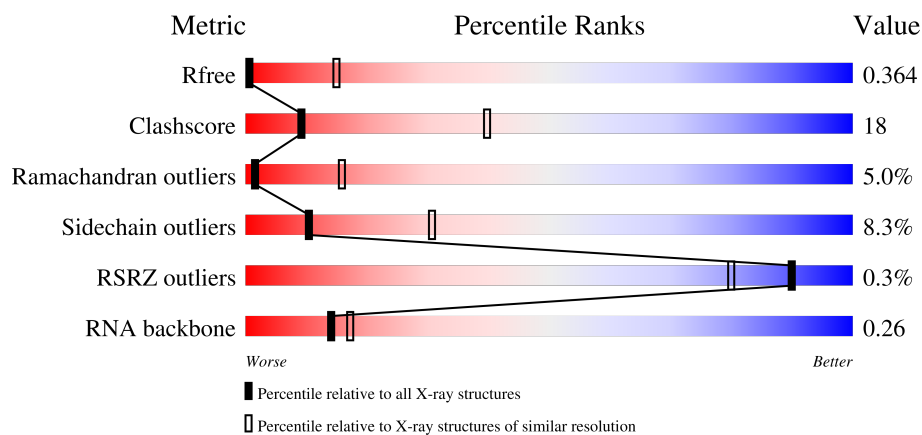
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.70 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.



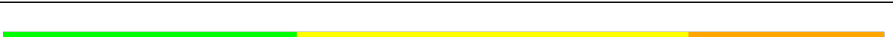
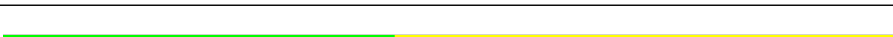
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	1020 (5.48-3.92)
Clashscore	190562	1005 (5.40-3.96)
Ramachandran outliers	187476	1095 (5.50-3.90)
Sidechain outliers	187428	1077 (5.50-3.90)
RSRZ outliers	180081	1015 (5.48-3.92)
RNA backbone	3983	1042 (6.22-3.10)

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	329	62% 7% 31%
1	B	329	62% 7% 31%
1	F	329	61% 7% 31%
1	G	329	61% 7% 31%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	C	1342	
2	H	1342	
3	D	1416	
3	I	1416	
4	E	90	
4	J	90	
5	K	974	
5	L	974	
6	M	15	
6	O	15	
7	N	9	
7	P	9	

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 61154 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 subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	227	Total	C	N	O	S	0	0	0
			1759	1096	311	346	6			
1	B	227	Total	C	N	O	S	0	0	0
			1759	1096	311	346	6			
1	F	227	Total	C	N	O	S	0	0	0
			1759	1096	311	346	6			
1	G	227	Total	C	N	O	S	0	0	0
			1759	1096	311	346	6			

- Molecule 2 is a protein called DNA-directed RNA polymerase subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1179	Total	C	N	O	S	0	0	0
			9285	5828	1620	1798	39			
2	H	1173	Total	C	N	O	S	0	0	0
			9235	5799	1609	1788	39			

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1149	Total	C	N	O	S	0	0	0
			8990	5649	1614	1681	46			
3	I	1160	Total	C	N	O	S	0	0	0
			9068	5701	1626	1695	46			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	1408	LEU	-	expression tag	UNP K0BCS5
D	1409	GLU	-	expression tag	UNP K0BCS5
D	1410	VAL	-	expression tag	UNP K0BCS5
D	1411	HIS	-	expression tag	UNP K0BCS5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	1412	HIS	-	expression tag	UNP K0BCS5
D	1413	HIS	-	expression tag	UNP K0BCS5
D	1414	HIS	-	expression tag	UNP K0BCS5
D	1415	HIS	-	expression tag	UNP K0BCS5
D	1416	HIS	-	expression tag	UNP K0BCS5
I	1408	LEU	-	expression tag	UNP K0BCS5
I	1409	GLU	-	expression tag	UNP K0BCS5
I	1410	VAL	-	expression tag	UNP K0BCS5
I	1411	HIS	-	expression tag	UNP K0BCS5
I	1412	HIS	-	expression tag	UNP K0BCS5
I	1413	HIS	-	expression tag	UNP K0BCS5
I	1414	HIS	-	expression tag	UNP K0BCS5
I	1415	HIS	-	expression tag	UNP K0BCS5
I	1416	HIS	-	expression tag	UNP K0BCS5

- Molecule 4 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	75	Total	C	N	O	S	0	0	0
			600	365	114	120	1			
4	J	75	Total	C	N	O	S	0	0	0
			600	365	114	120	1			

- Molecule 5 is a protein called RNA polymerase-associated protein RapA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	K	961	Total	C	N	O	S	0	0	0
			7665	4797	1370	1468	30			
5	L	961	Total	C	N	O	S	0	0	0
			7665	4797	1370	1468	30			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	-5	HIS	-	expression tag	UNP P60240
K	-4	HIS	-	expression tag	UNP P60240
K	-3	HIS	-	expression tag	UNP P60240
K	-2	HIS	-	expression tag	UNP P60240
K	-1	HIS	-	expression tag	UNP P60240
K	0	HIS	-	expression tag	UNP P60240
K	350	CYS	ARG	engineered mutation	UNP P60240
L	-5	HIS	-	expression tag	UNP P60240

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
L	-4	HIS	-	expression tag	UNP P60240
L	-3	HIS	-	expression tag	UNP P60240
L	-2	HIS	-	expression tag	UNP P60240
L	-1	HIS	-	expression tag	UNP P60240
L	0	HIS	-	expression tag	UNP P60240
L	350	CYS	ARG	engineered mutation	UNP P60240

- Molecule 6 is a DNA chain called 5'-D(P*AP*CP*GP*AP*CP*TP*GP*AP*GP*CP*CP*GP*AP*TP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	M	15	Total	C	N	O	P	0	0	0
			311	146	61	89	15			
6	O	15	Total	C	N	O	P	0	0	0
			311	146	61	89	15			

- Molecule 7 is a RNA chain called 5'-R(P*AP*UP*CP*GP*GP*CP*UP*CP*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	N	9	Total	C	N	O	P	0	0	0
			191	85	33	64	9			
7	P	9	Total	C	N	O	P	0	0	0
			191	85	33	64	9			

- Molecule 8 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	D	2	Total	Zn	0	0
			2	2		
8	I	2	Total	Zn	0	0
			2	2		

- Molecule 9 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	D	1	Total	Mg	0	0
			1	1		
9	I	1	Total	Mg	0	0
			1	1		

Chain G:  61% 7% 31%

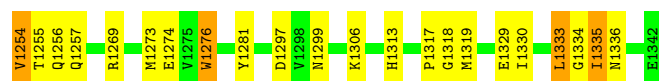
LEU	THR	GLU	GLU	LYS	ASP	VAL	LEU	ALA	SER	SER	GLY	SER	LEU	LEU	GLY	MET	ARG	LEU	GLU	TRP	PRO	PRO	ALA	SER	ILE	ALA	ASP	GLU																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

• Molecule 2: DNA-directed RNA polymerase subunit beta

Chain C:  62% 22% 12%

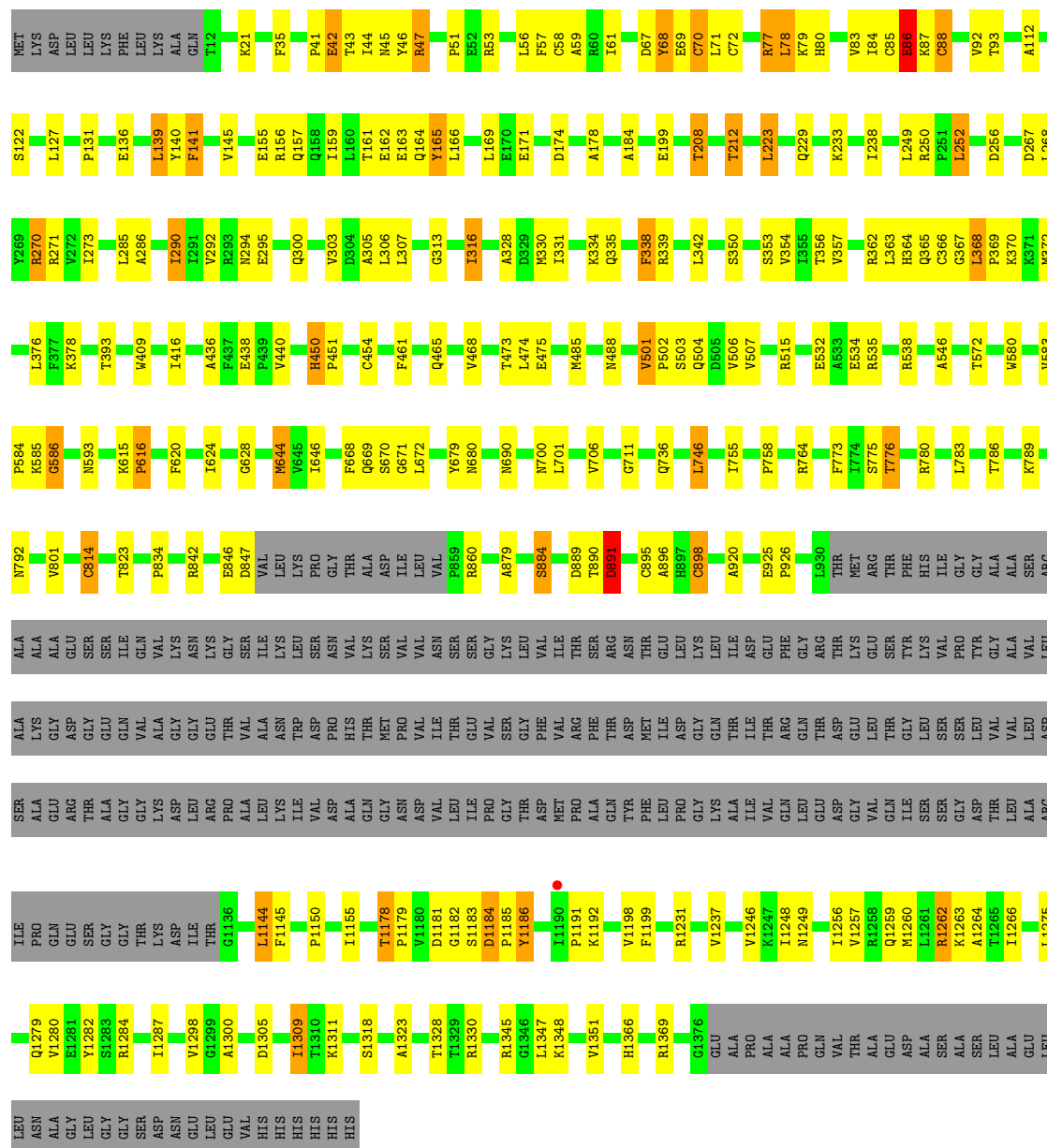
P1062	L971	W839	V733	E626	S509	F389	GLU	G140	MET
K1065	A977	S840	H737	G627	Q510	F390	TYR	VAL	Y3
M1066	VAL	D842	GLU	H628	F514	Y395	ASP	V144	T6
N1072	LEU	T843	ASP	E631	D516	L397	GLU	L149	R12
V1075	VAL	K844	GLU	V634	Q517	M403	THR	H150	K13
I1076	ALA	P847	TYR	R637	N518	N519	GLY	P153	D14
S1077	GLY	I850	PRO	S638	N519	P520	GLY	D158	F15
P1071	GLY	T850	GLY	K639	L521	L420	LEU		
F1081	VAL	K864	GLU		S522	V428	ILE	K161	R18
I1082	GLU	L865	ALA	S643	E523	M429	ALA	G162	V21
M1085	GLU	I748	G747	L644	I524	K431	PHE		
P1086	LYS	E867	D749	F645	K527	L433	ASP	H165	Y26
Y1087	LEU	T757	T757	T657	T539	I433	ILE	G168	L27
D1088	LYS	G879	N760	Q658	R540	I435	GLU	K169	I30
G1091	PRO	C885	Q761	Q659	E541	R436	ASN	V170	
I1096	ARG	R886	V782	V661	F545	M437	LEU	I176	Q36
V1097	ASP	V887		S662		E441	VAL	I177	
L1098	ARG	Q894	G786	V663			TYR	P173	E40
V1103	TRP	L895	F787	S666	H551	R454	ALA	G181	Q41
Q1111	LEU	T896	S788	L667	P552	S455	GLY	D185	D42
L1112	GLU	F906	T789	I668	H554	V456	LYS		P43
L1113	GLY	G907	D790	P669	R557	G457	ARG	F183	S61
E1114	LEU	A910	L791	F670	V558	E453	ARG	Y62	S63
H1116	THR	S911	G792	L671	C559	E461	ILE	N193	
L1117	GLU	D912	E793	D674	P560	M462	THR	L194	E77
L1118	GLY	N913	L794	D675	I561	Q463	ALA		P78
G1118	LYS	K914	A795	A676	T561	F464	ARG	R197	V79
M1119	GLN	D915	Q798	N677	N568		ILE	A206	F80
A1120	ASN	S916	N799			L468	GLU		D81
I1121	GLN	G926	A803	L680	L571	E472	THR	T207	C85
K1122	L1011	T927	W807	A689	N573		GLN	I208	Q86
G1123	E1020	V928	N808	T692	V577	V475	LEU	I209	I87
D1126	F1025	I929	G809	L693	T581	L479	ASP	L213	R88
N1129	R1034	V933	Y810	P698	E588	L487	LEU	N214	G89
A1130	M1130	D937	F812	L699	E588	M488	ASP	Y215	Y90
L1132	E813	G938	T1037	V700	Y591	P489	VAL	T216	S93
K1133	Q1038	V939	Q1038	G701	R592	Q490	LYS		A94
Q1134	G1039		S815	T702	K593	D491	LEU	F225	
Q1135	L1042	R944	E820	M704	A608	M492	ILE	GLU	
V1138	L1047	E947	R821	V708		K496	GLY	LYS	
L1141	K1048	I948	W822	A709	N613	P497	VAL	VAL	V114
I1145	I1049	M951	E825			I498	GLY	PHE	M124
Y1149	V1056			S712	T616		TYR	GLU	
	K1057	L964	H832	V714			ILE	ILE	L129
	R1058		I833			A501	ALA	ARG	M130
			Q834	V714	S621	V502	GLY	ASP	T131
			E835			K503	VAL	LYS	D132
				S730	D624	E379	VAL	LEU	G134
					T625		ALA	GLN	T135





• Molecule 3: DNA-directed RNA polymerase subunit beta'

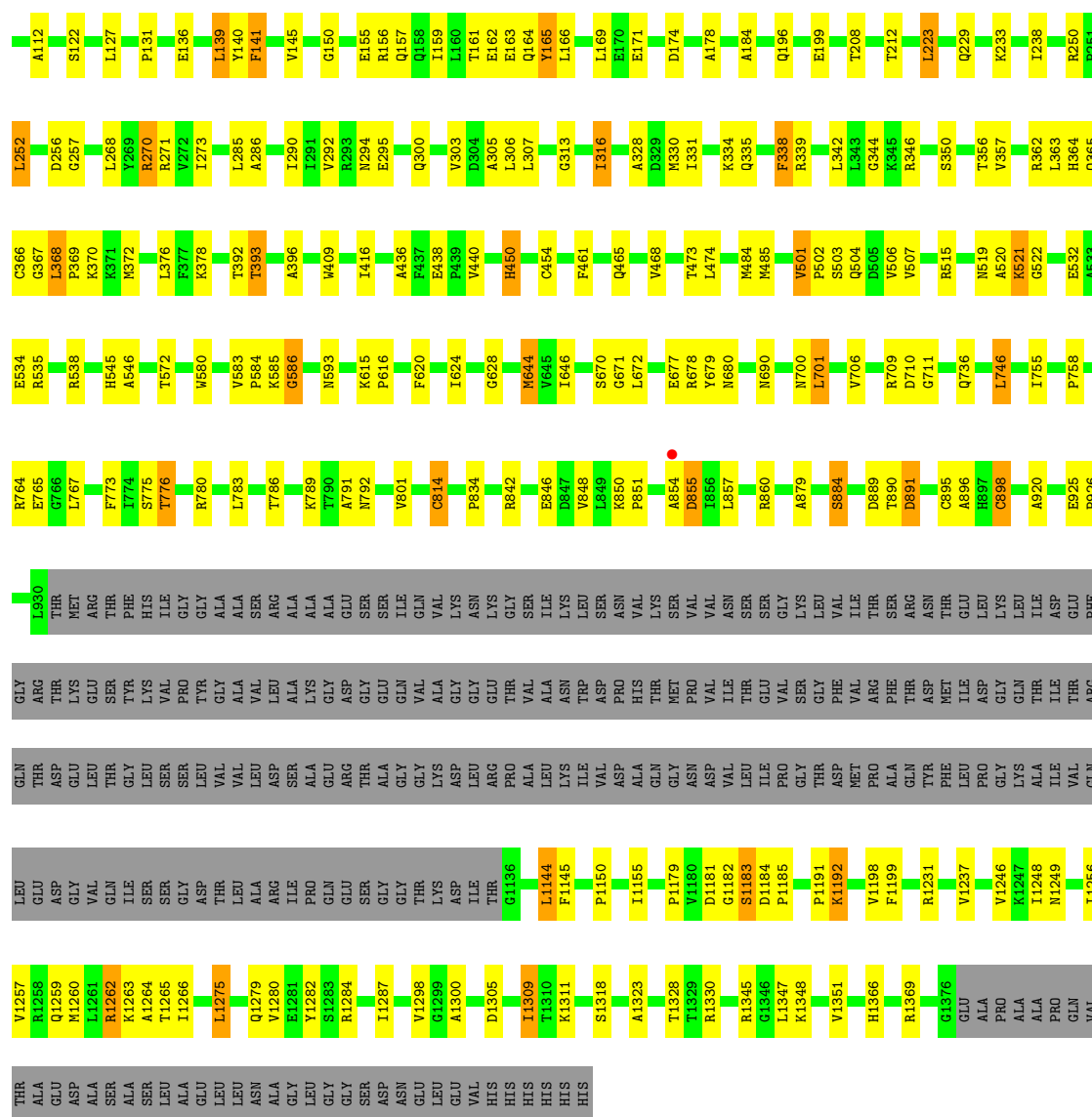
Chain D: 64% 15% 19%

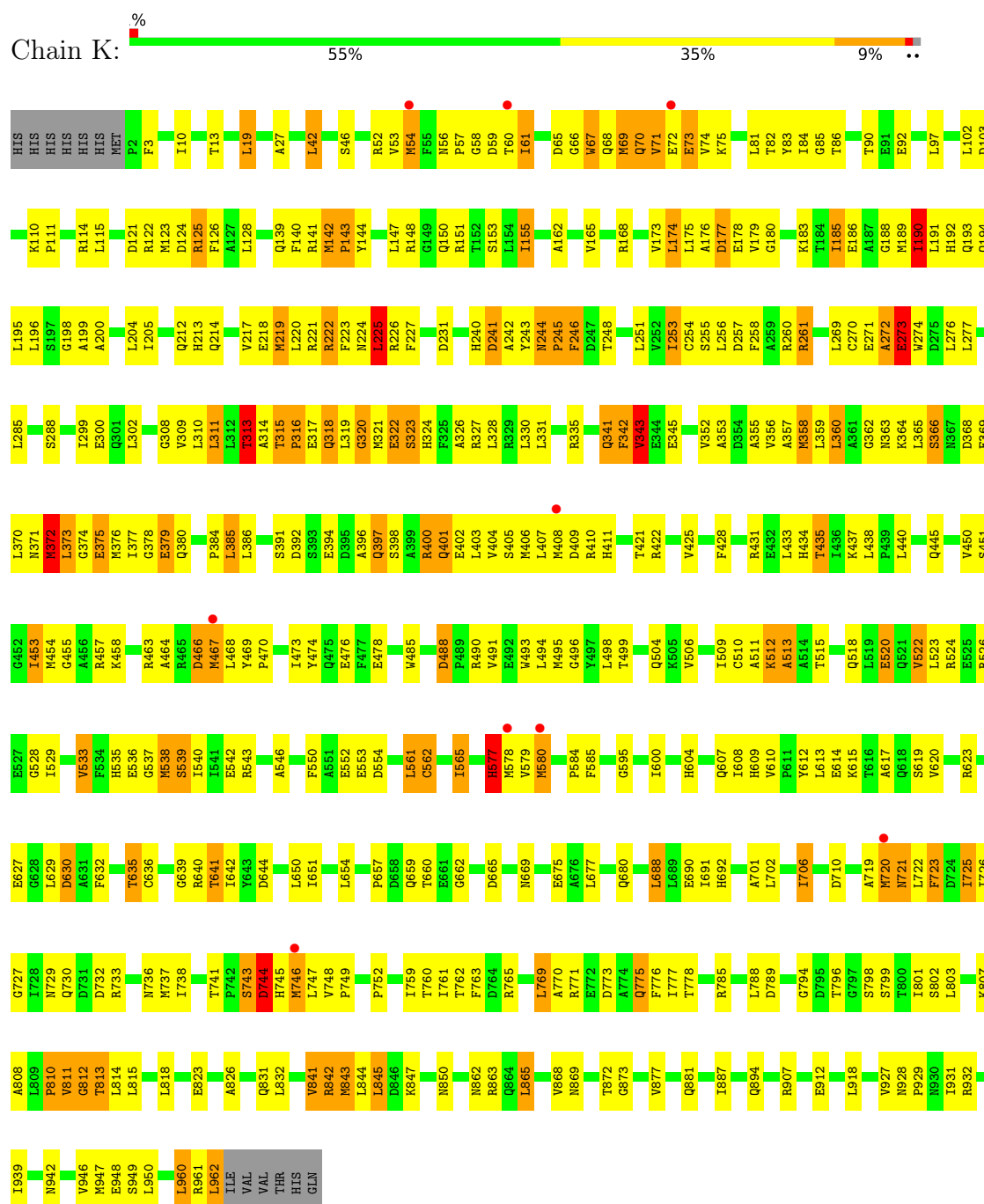


• Molecule 3: DNA-directed RNA polymerase subunit beta'

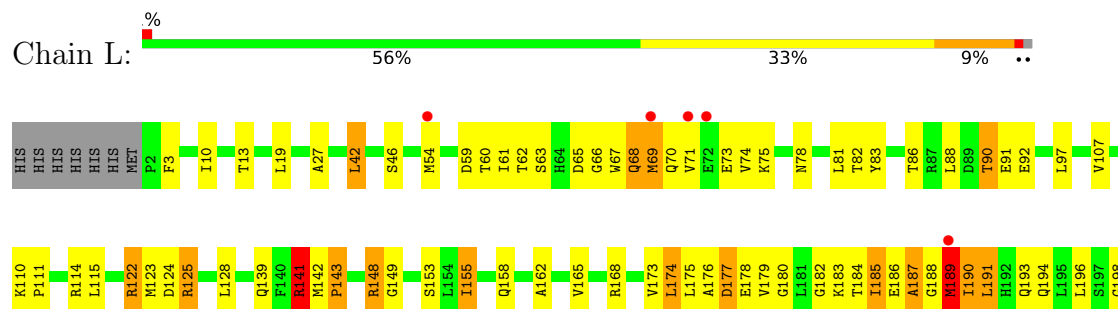
Chain I: 63% 16% 18%





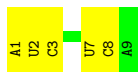


• Molecule 5: RNA polymerase-associated protein RapA



- Molecule 7: 5'-R(P*AP*UP*CP*GP*GP*CP*UP*CP*A)-3'

Chain P:  44% 56%



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	336.09Å 158.93Å 255.01Å 90.00° 101.28° 90.00°	Depositor
Resolution (Å)	20.00 – 4.70 20.00 – 4.70	Depositor EDS
% Data completeness (in resolution range)	93.2 (20.00-4.70) 93.2 (20.00-4.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.21 (at 4.74Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.273 , 0.369 0.288 , 0.364	Depositor DCC
R_{free} test set	3240 reflections (4.77%)	wwPDB-VP
Wilson B-factor (Å ²)	196.2	Xtriage
Anisotropy	0.118	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.20 , 251.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	61154	wwPDB-VP
Average B, all atoms (Å ²)	292.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.40% 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

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

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	0.51	0/1781	0.76	1/2414 (0.0%)
1	B	0.51	0/1781	0.77	1/2414 (0.0%)
1	F	0.52	0/1781	0.76	1/2414 (0.0%)
1	G	0.52	0/1781	0.77	1/2414 (0.0%)
2	C	0.54	0/9435	0.84	8/12729 (0.1%)
2	H	0.54	0/9385	0.83	9/12662 (0.1%)
3	D	0.55	0/9128	0.82	7/12313 (0.1%)
3	I	0.55	0/9208	0.82	2/12426 (0.0%)
4	E	0.59	0/602	0.74	0/810
4	J	0.57	0/602	0.73	0/810
5	K	0.58	3/7808 (0.0%)	0.94	30/10576 (0.3%)
5	L	0.57	2/7808 (0.0%)	0.97	35/10576 (0.3%)
6	M	0.28	0/349	0.56	0/535
6	O	0.28	0/349	0.59	0/535
7	N	0.41	0/212	0.68	0/326
7	P	0.40	0/212	0.72	0/326
All	All	0.55	5/62222 (0.0%)	0.85	95/84280 (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
5	L	0	1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	L	68	GLN	CA-C	5.98	1.59	1.52
5	K	69	MET	CA-C	5.81	1.59	1.52

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	L	947	MET	C-N	-5.77	1.25	1.33
5	K	70	GLN	CA-C	5.35	1.59	1.52
5	K	68	GLN	CA-C	5.20	1.59	1.52

The worst 5 of 95 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	L	461	GLU	N-CA-C	11.43	123.74	111.28
5	L	947	MET	N-CA-C	-11.41	100.15	112.93
5	L	465	ARG	N-CA-C	11.15	123.43	111.28
5	L	141	ARG	N-CA-C	-10.90	99.29	112.59
5	K	70	GLN	CA-C-N	-10.65	109.26	122.90

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	L	148	ARG	Sidechain

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	1759	0	1785	10	0
1	B	1759	0	1785	11	0
1	F	1759	0	1785	11	0
1	G	1759	0	1785	11	0
2	C	9285	0	9291	166	0
2	H	9235	0	9242	176	0
3	D	8990	0	9174	160	0
3	I	9068	0	9265	159	0
4	E	600	0	607	10	0
4	J	600	0	607	6	0
5	K	7665	0	7524	748	0
5	L	7665	0	7524	824	0
6	M	311	0	168	19	0
6	O	311	0	168	5	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	N	191	0	98	12	0
7	P	191	0	98	7	0
8	D	2	0	0	1	0
8	I	2	0	0	0	0
9	D	1	0	0	0	0
9	I	1	0	0	0	0
All	All	61154	0	60906	2212	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 18.

The worst 5 of 2212 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)
5:K:69:MET:CG	5:K:97:LEU:HD22	1.28	1.61
5:L:184:THR:HG22	5:L:312:LEU:CD2	1.17	1.61
5:L:219:MET:HG2	5:L:227:PHE:CE1	1.37	1.60
5:L:468:LEU:HA	5:L:617:ALA:CB	1.24	1.59
5:L:720:MET:CE	5:L:723:PHE:HE2	1.17	1.57

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	225/329 (68%)	202 (90%)	19 (8%)	4 (2%)	6	33
1	B	225/329 (68%)	205 (91%)	20 (9%)	0	100	100
1	F	225/329 (68%)	202 (90%)	19 (8%)	4 (2%)	6	33
1	G	225/329 (68%)	205 (91%)	19 (8%)	1 (0%)	30	67
2	C	1171/1342 (87%)	894 (76%)	211 (18%)	66 (6%)	1	14

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	H	1165/1342 (87%)	889 (76%)	206 (18%)	70 (6%)	1	13
3	D	1143/1416 (81%)	867 (76%)	235 (21%)	41 (4%)	2	20
3	I	1156/1416 (82%)	869 (75%)	240 (21%)	47 (4%)	2	18
4	E	73/90 (81%)	59 (81%)	8 (11%)	6 (8%)	0	9
4	J	73/90 (81%)	59 (81%)	8 (11%)	6 (8%)	0	9
5	K	959/974 (98%)	689 (72%)	201 (21%)	69 (7%)	1	11
5	L	959/974 (98%)	694 (72%)	200 (21%)	65 (7%)	1	12
All	All	7599/8960 (85%)	5834 (77%)	1386 (18%)	379 (5%)	1	16

5 of 379 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	51	MET
2	C	43	PRO
2	C	63	SER
2	C	115	LYS
2	C	216	THR

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	195/286 (68%)	191 (98%)	4 (2%)	47	65
1	B	195/286 (68%)	187 (96%)	8 (4%)	27	49
1	F	195/286 (68%)	191 (98%)	4 (2%)	47	65
1	G	195/286 (68%)	187 (96%)	8 (4%)	27	49
2	C	1016/1157 (88%)	937 (92%)	79 (8%)	11	32
2	H	1011/1157 (87%)	931 (92%)	80 (8%)	11	32
3	D	962/1177 (82%)	884 (92%)	78 (8%)	11	31
3	I	971/1177 (82%)	895 (92%)	76 (8%)	11	32
4	E	65/74 (88%)	62 (95%)	3 (5%)	24	46

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	J	65/74 (88%)	62 (95%)	3 (5%)	24	46
5	K	822/835 (98%)	720 (88%)	102 (12%)	4	18
5	L	822/835 (98%)	727 (88%)	95 (12%)	5	19
All	All	6514/7630 (85%)	5974 (92%)	540 (8%)	10	31

5 of 540 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	L	254	CYS
5	L	360	LEU
5	L	253	ILE
5	L	769	LEU
2	H	403	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 108 such sidechains are listed below:

Mol	Chain	Res	Type
5	K	70	GLN
5	K	592	GLN
5	L	669	ASN
5	K	212	GLN
5	K	348	ASN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
7	N	8/9 (88%)	3 (37%)	0
7	P	8/9 (88%)	1 (12%)	0
All	All	16/18 (88%)	4 (25%)	0

All (4) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
7	N	2	U
7	N	6	C
7	N	7	U
7	P	7	U

There are no RNA pucker outliers to report.

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 6 are monoatomic - leaving 0 for Mogul analysis.

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.

No monomer is involved in short contacts.

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

6.1 Protein, DNA and RNA chains

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	227/329 (68%)	-0.51	0 100 100	178, 268, 367, 400	0
1	B	227/329 (68%)	-0.45	0 100 100	237, 380, 491, 539	0
1	F	227/329 (68%)	-0.62	0 100 100	192, 268, 357, 401	0
1	G	227/329 (68%)	-0.47	0 100 100	256, 379, 484, 537	0
2	C	1179/1342 (87%)	-0.58	3 (0%) 90 80	126, 241, 378, 510	0
2	H	1173/1342 (87%)	-0.60	2 (0%) 91 82	143, 251, 375, 532	0
3	D	1149/1416 (81%)	-0.53	1 (0%) 92 86	181, 301, 441, 641	0
3	I	1160/1416 (81%)	-0.59	1 (0%) 92 86	175, 282, 433, 662	0
4	E	75/90 (83%)	-0.40	0 100 100	285, 343, 385, 396	0
4	J	75/90 (83%)	-0.50	0 100 100	329, 349, 365, 376	0
5	K	961/974 (98%)	-0.46	9 (0%) 81 66	178, 278, 461, 664	0
5	L	961/974 (98%)	-0.45	9 (0%) 81 66	169, 269, 469, 638	0
6	M	15/15 (100%)	-0.21	0 100 100	321, 353, 461, 465	0
6	O	15/15 (100%)	-0.13	0 100 100	331, 351, 543, 550	0
7	N	9/9 (100%)	-0.69	0 100 100	258, 268, 381, 391	0
7	P	9/9 (100%)	-0.73	0 100 100	236, 245, 375, 395	0
All	All	7689/9008 (85%)	-0.53	25 (0%) 90 80	126, 279, 441, 664	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	K	578	MET	5.8
5	K	580	MET	4.2
5	L	72	GLU	3.9
5	K	408	MET	3.3
5	K	54	MET	3.1

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($\text{\AA}^2$)	Q<0.9
8	ZN	D	1501	1/1	0.79	0.17	106,106,106,106	0
8	ZN	I	1501	1/1	0.79	0.13	131,131,131,131	0
9	MG	I	1503	1/1	0.93	0.12	52,52,52,52	0
8	ZN	I	1502	1/1	0.97	0.12	117,117,117,117	0
8	ZN	D	1502	1/1	0.98	0.20	84,84,84,84	0
9	MG	D	1503	1/1	0.99	0.17	10,10,10,10	0

6.5 Other polymers [i](#)

There are no such residues in this entry.