



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 08:46 am GMT

PDB ID : 4KN4
Title : X-ray crystal structure of the Escherichia coli RNA polymerase in complex with Benzoxazinorifamycin-2b
Authors : Murakami, K.S.
Deposited on : 2013-05-08
Resolution : 3.96 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
Mogul : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : trunk28620
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) : recalc28949

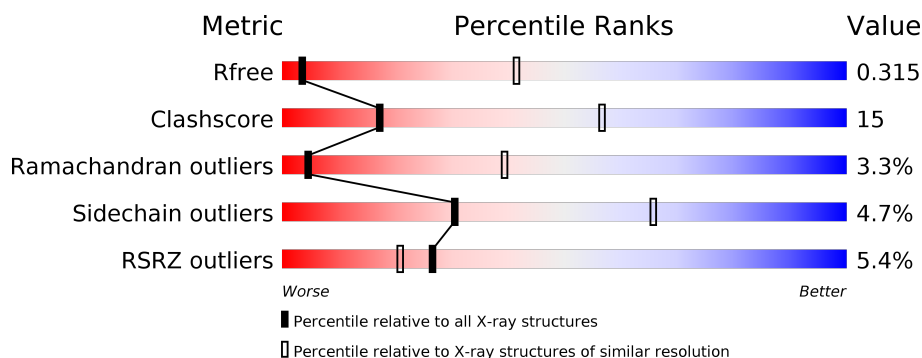
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.96 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	1057 (4.32-3.60)
Clashscore	112137	1154 (4.32-3.60)
Ramachandran outliers	110173	1111 (4.32-3.60)
Sidechain outliers	110143	1101 (4.32-3.60)
RSRZ outliers	101464	1069 (4.32-3.60)

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 $\leq 5\%$. 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	
1	B	329	
1	F	329	
1	G	329	
2	C	1342	
2	H	1342	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	1407	
3	I	1407	
4	E	91	
4	J	91	
5	X	613	
5	Y	613	

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
6	1RL	H	1401	-	-	-	X

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 56333 atoms, of which 122 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	323	Total	C	N	O	S	0	0	0
			2514	1571	443	492	8			
1	B	221	Total	C	N	O	S	0	0	0
			1706	1065	300	335	6			
1	F	229	Total	C	N	O	S	0	0	0
			1775	1106	313	350	6			
1	G	217	Total	C	N	O	S	0	0	0
			1671	1045	293	327	6			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1335	Total	C	N	O	S	0	0	0
			10523	6601	1836	2043	43			
2	H	1335	Total	C	N	O	S	0	0	0
			10523	6601	1836	2043	43			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1160	Total	C	N	O	S	0	0	0
			9060	5695	1621	1697	47			
3	I	1160	Total	C	N	O	S	0	0	0
			9060	5695	1621	1697	47			

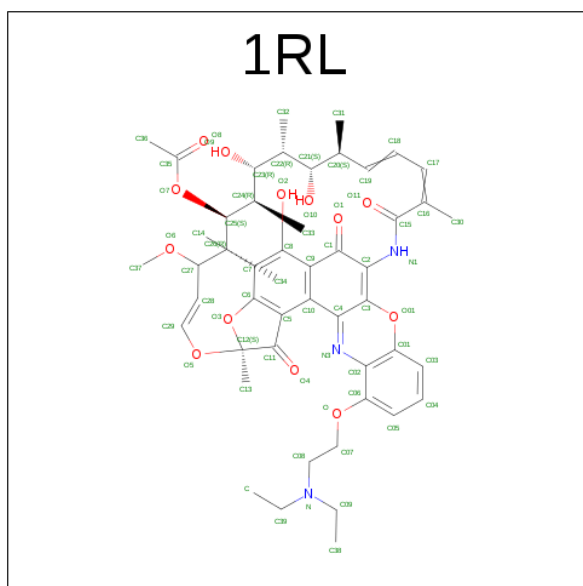
- Molecule 4 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	90	Total	C	N	O	S	0	0	0
			708	430	136	141	1			
4	J	76	Total	C	N	O	S	0	0	0
			605	368	115	121	1			

- Molecule 5 is a protein called RNA polymerase sigma factor RpoD.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	X	517	Total	C	N	O	S	0	0	0
			4198	2621	745	806	26			
5	Y	458	Total	C	N	O	S	0	0	0
			3732	2335	671	703	23			

- Molecule 6 is BENZOXAZINORIFAMYCIN-2B (three-letter code: 1RL) (formula: $C_{49}H_{61}N_3O_{13}$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	C	1	Total	C	H	N	O	0	0
			126	49	61	3	13		
6	H	1	Total	C	H	N	O	0	0
			126	49	61	3	13		

- Molecule 7 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

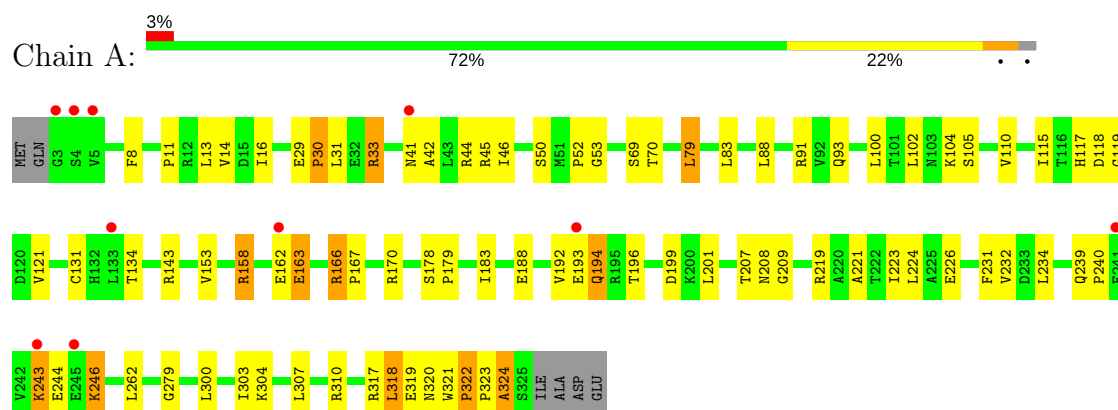
- Molecule 8 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

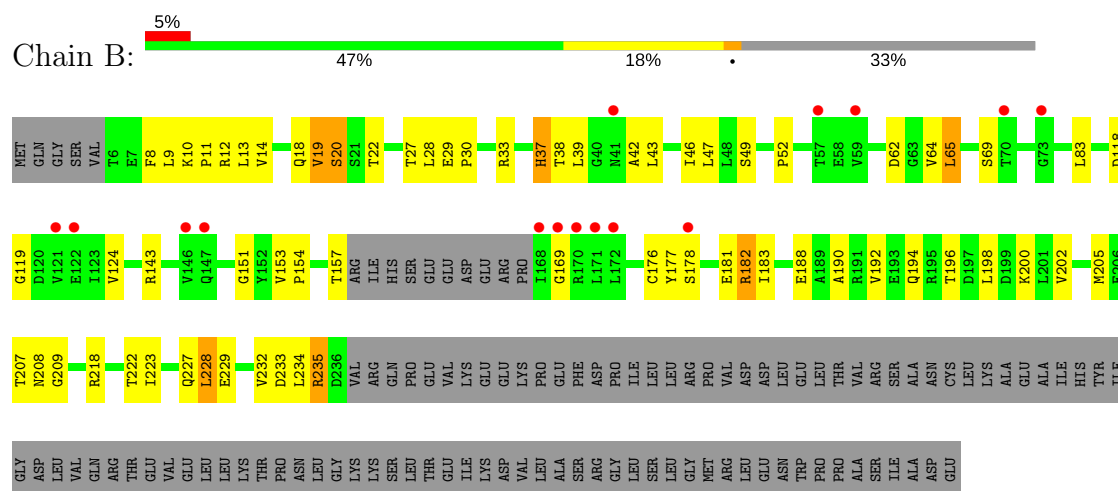
3 Residue-property plots

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

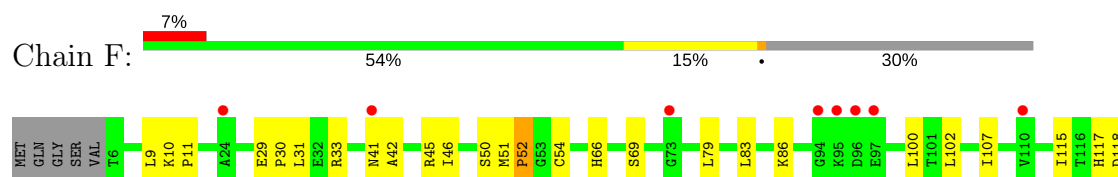
• Molecule 1: DNA-directed RNA polymerase subunit alpha

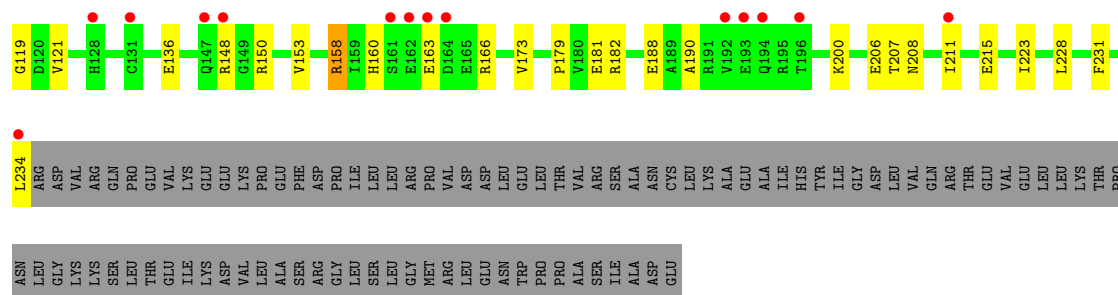


• Molecule 1: DNA-directed RNA polymerase subunit alpha

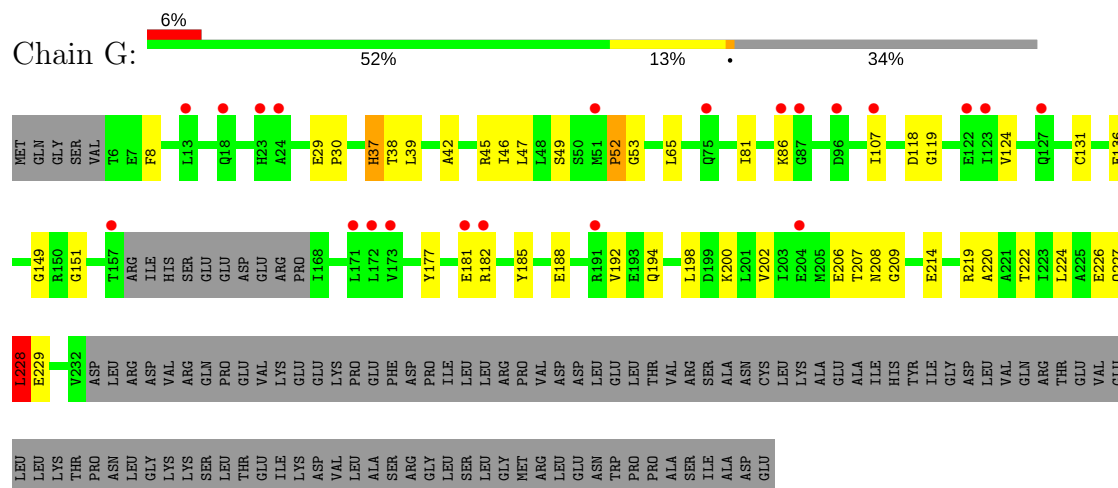


• Molecule 1: DNA-directed RNA polymerase subunit alpha

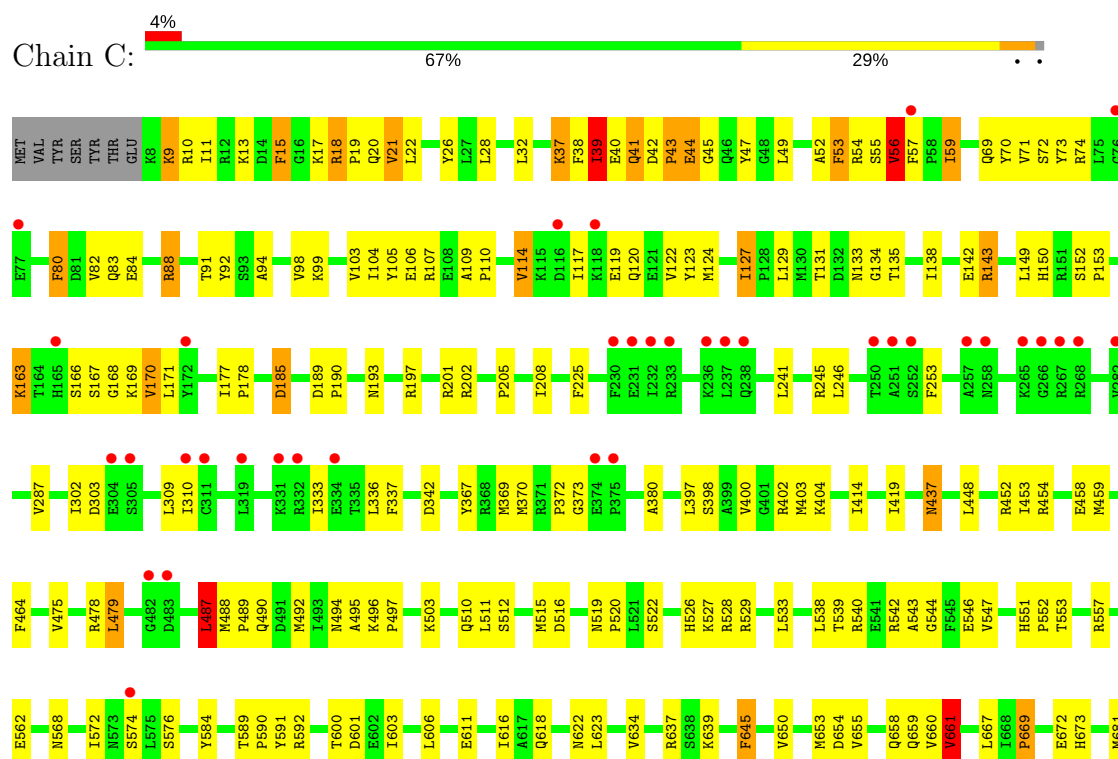


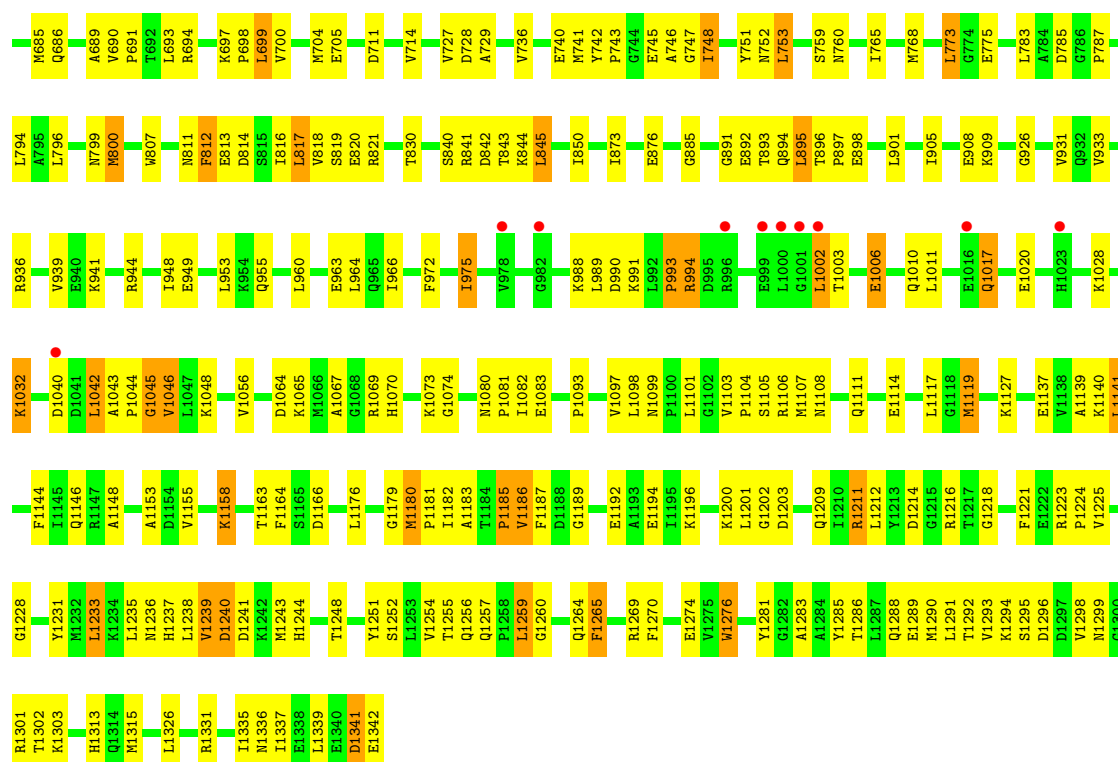


• Molecule 1: DNA-directed RNA polymerase subunit alpha

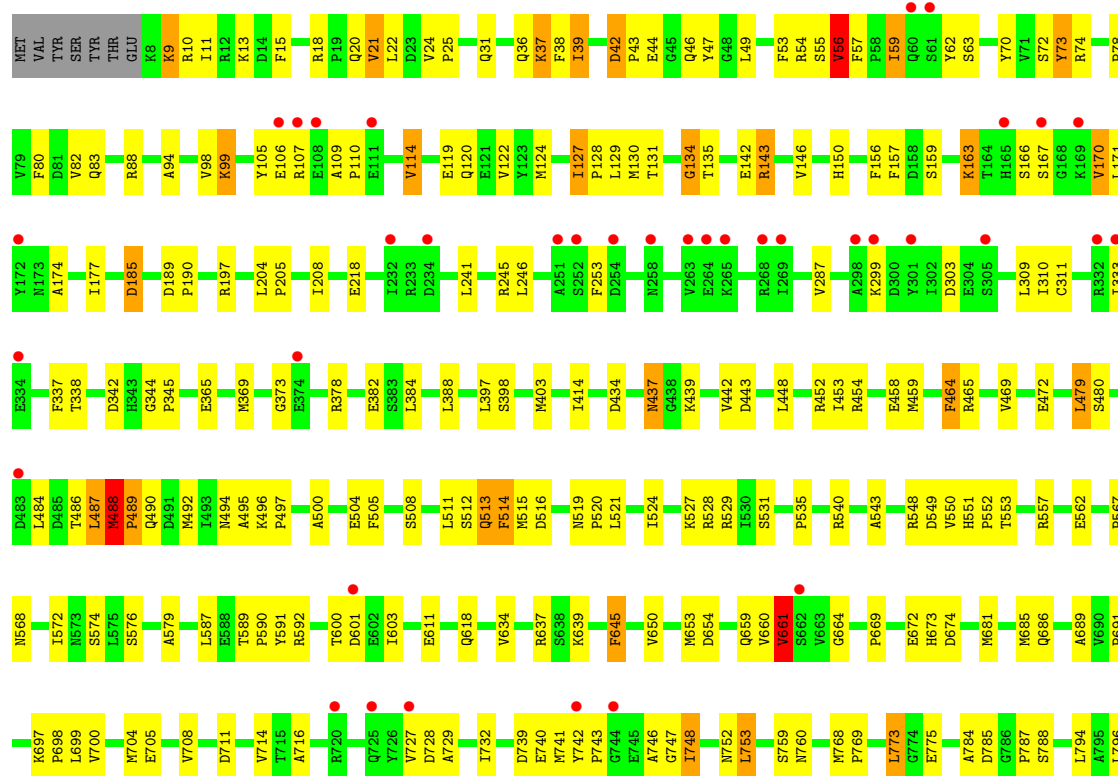


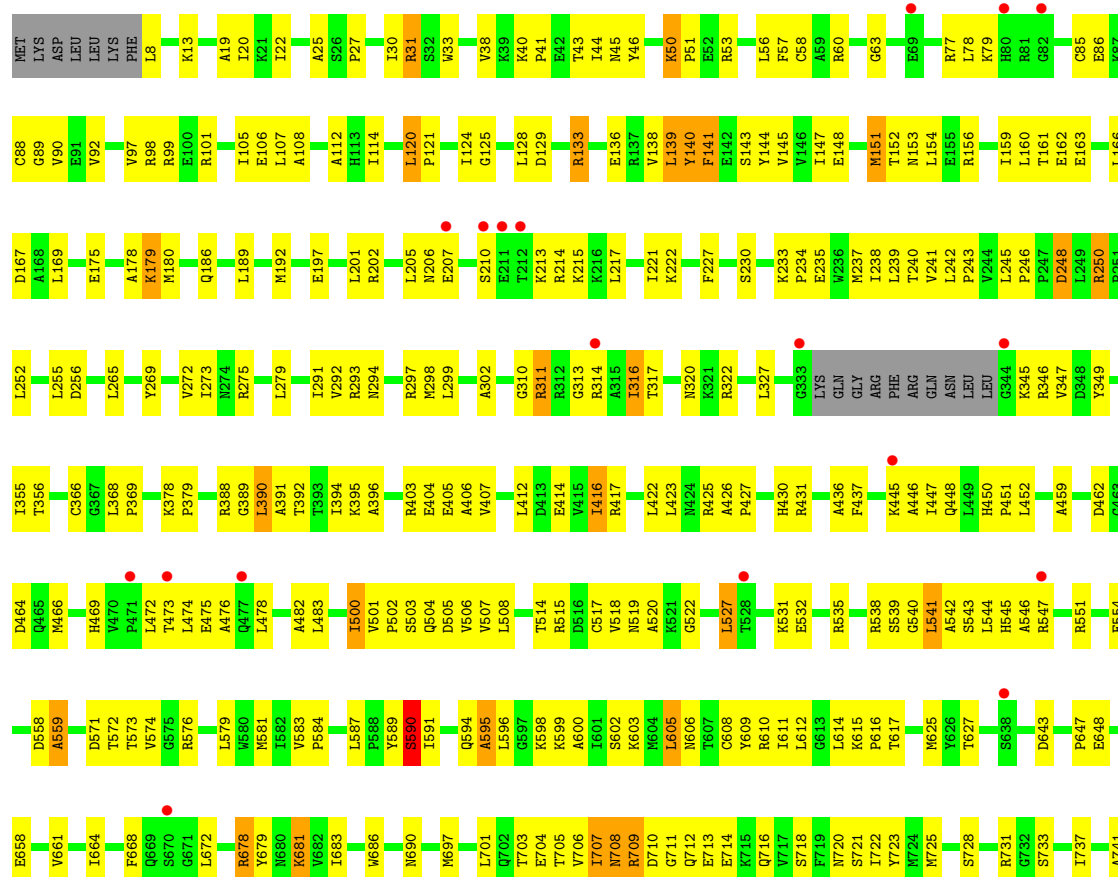
• Molecule 2: DNA-directed RNA polymerase subunit beta

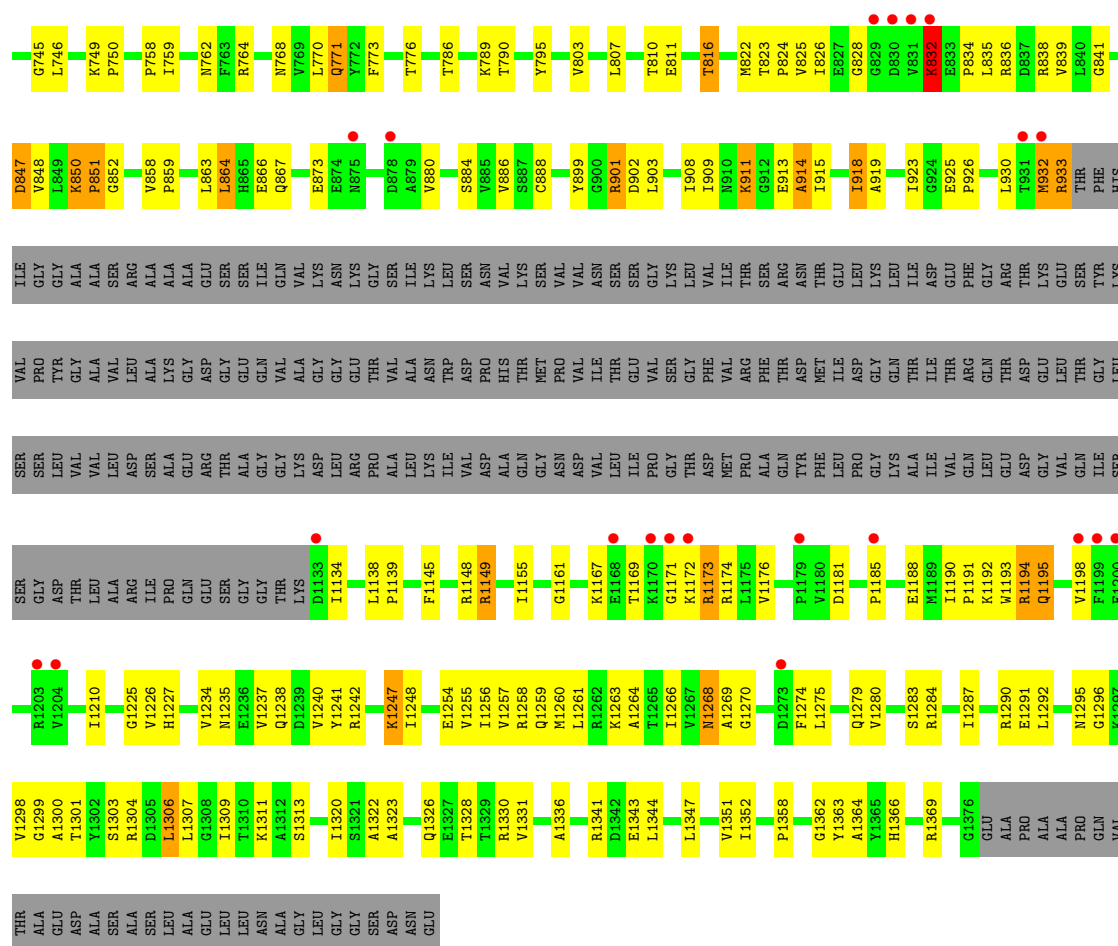




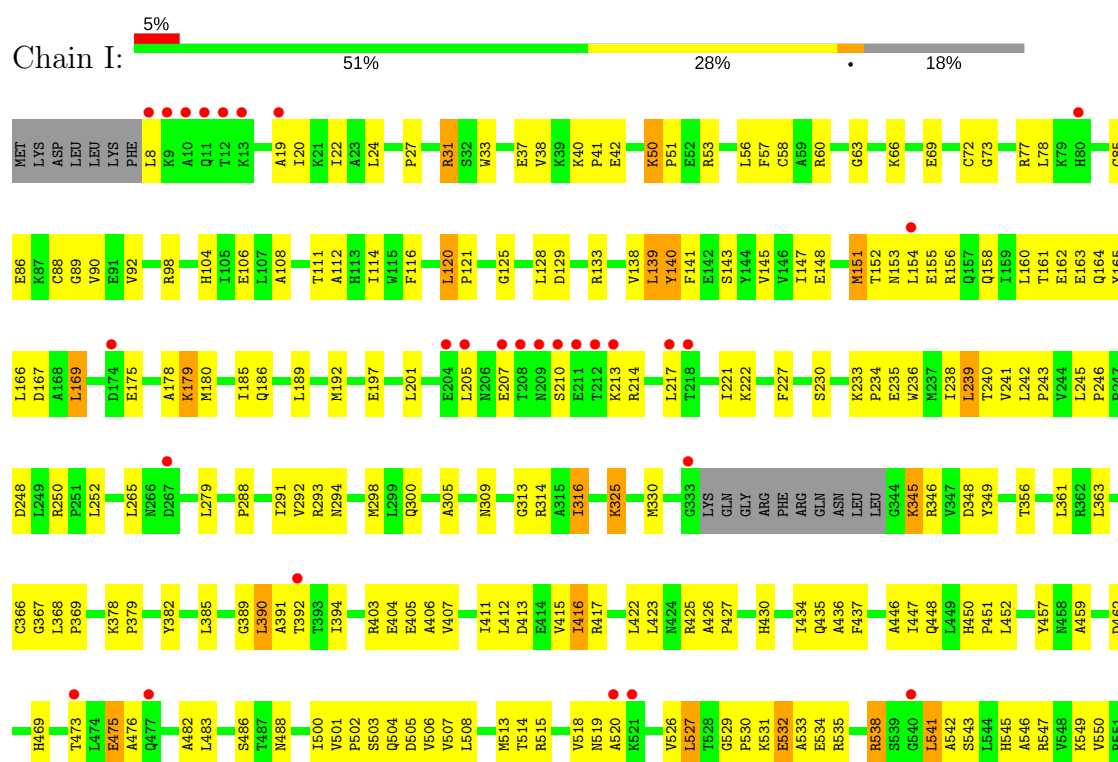
• Molecule 2: DNA-directed RNA polymerase subunit beta

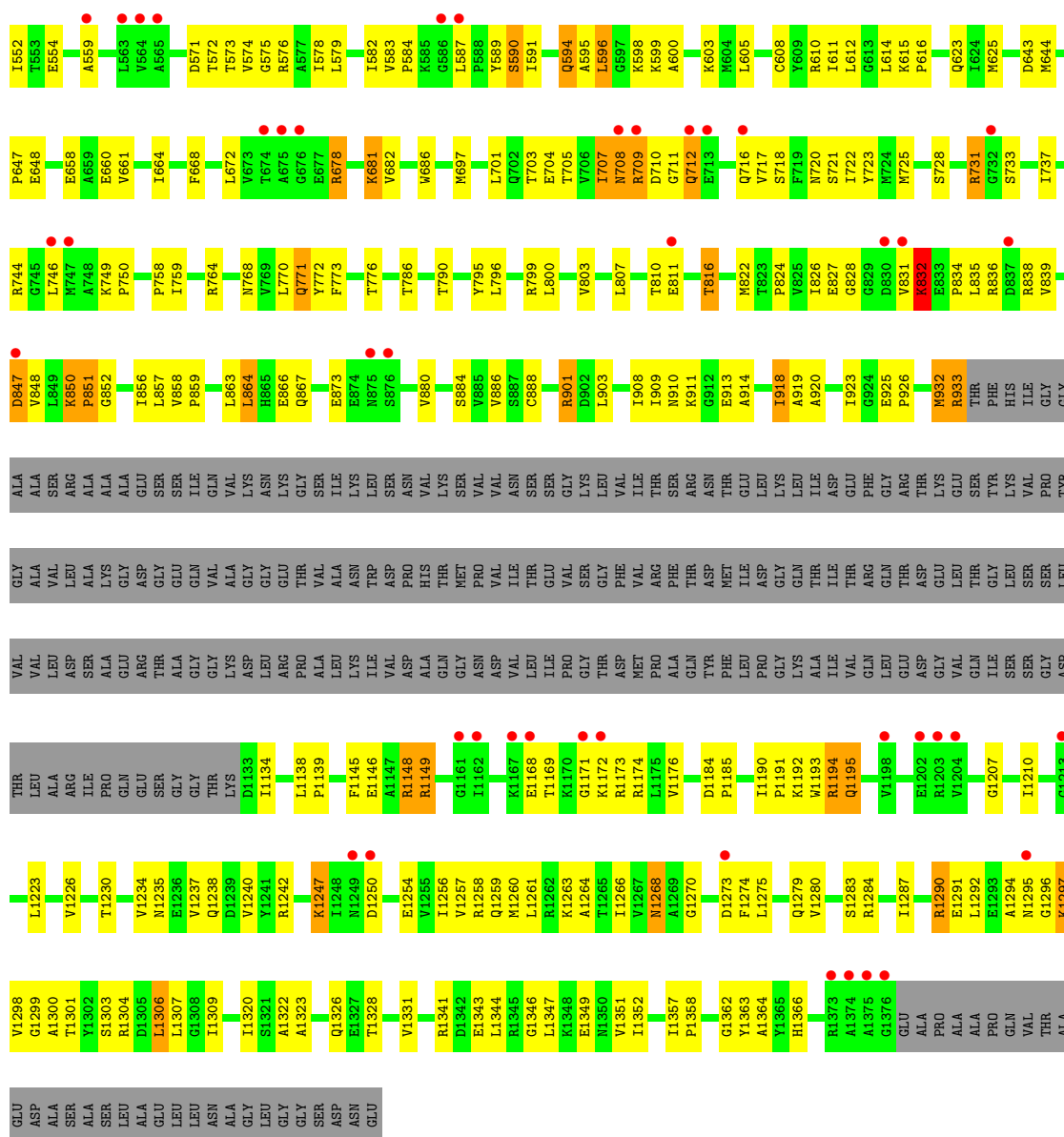




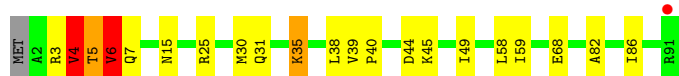
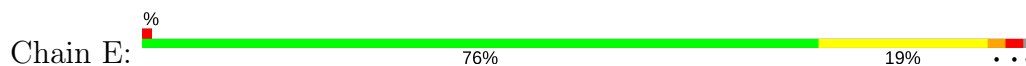


• Molecule 3: DNA-directed RNA polymerase subunit beta'

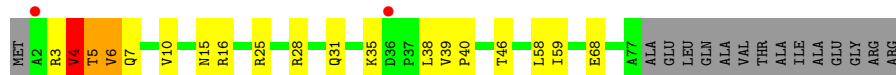




- Molecule 4: DNA-directed RNA polymerase subunit omega



- Molecule 4: DNA-directed RNA polymerase subunit omega



- Molecule 5: RNA polymerase sigma factor RpoD

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	184.68Å 203.97Å 307.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.77 – 3.96 29.77 – 3.96	Depositor EDS
% Data completeness (in resolution range)	94.5 (29.77-3.96) 90.0 (29.77-3.96)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.06 (at 3.98Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.251 , 0.315 0.248 , 0.315	Depositor DCC
R_{free} test set	4764 reflections (4.99%)	DCC
Wilson B-factor (Å ²)	148.7	Xtriage
Anisotropy	0.103	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 54.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.22$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	56333	wwPDB-VP
Average B, all atoms (Å ²)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.42% 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: ZN, 1RL, MG

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	0.19	0/2548	0.37	0/3454
1	B	0.19	0/1725	0.40	0/2337
1	F	0.20	0/1797	0.40	0/2436
1	G	0.19	0/1690	0.40	1/2290 (0.0%)
2	C	0.20	0/10690	0.39	0/14423
2	H	0.20	0/10690	0.39	0/14423
3	D	0.20	0/9198	0.40	0/12413
3	I	0.20	0/9198	0.40	0/12413
4	E	0.19	0/710	0.36	0/956
4	J	0.19	0/607	0.36	0/817
5	X	0.20	0/4253	0.37	0/5719
5	Y	0.20	0/3783	0.36	0/5083
All	All	0.20	0/56889	0.39	1/76764 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	228	LEU	CA-CB-CG	5.14	127.13	115.30

There are no chirality outliers.

There are no planarity outliers.

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	2514	0	2566	66	0
1	B	1706	0	1738	62	0
1	F	1775	0	1800	40	0
1	G	1671	0	1706	45	0
2	C	10523	0	10546	373	0
2	H	10523	0	10546	344	0
3	D	9060	0	9256	382	0
3	I	9060	0	9255	346	0
4	E	708	0	719	20	0
4	J	605	0	612	16	0
5	X	4198	0	4250	109	0
5	Y	3732	0	3809	74	0
6	C	65	61	61	6	0
6	H	65	61	61	10	0
7	D	2	0	0	0	0
7	I	2	0	0	0	0
8	D	1	0	0	0	0
8	I	1	0	0	0	0
All	All	56211	122	56925	1751	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 15.

The worst 5 of 1751 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)
3:I:1173:ARG:HA	3:I:1174:ARG:HB2	1.32	1.10
2:H:1119:MET:HG2	2:H:1228:GLY:HA2	1.34	1.09
3:D:1173:ARG:HA	3:D:1174:ARG:HB2	1.30	1.06
3:D:310:GLY:HA3	3:D:311:ARG:HB2	1.34	1.04
2:C:42:ASP:HB3	2:C:43:PRO:HD2	1.35	1.03

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	321/329 (98%)	263 (82%)	42 (13%)	16 (5%)	2	29
1	B	217/329 (66%)	187 (86%)	25 (12%)	5 (2%)	7	46
1	F	227/329 (69%)	195 (86%)	27 (12%)	5 (2%)	8	47
1	G	213/329 (65%)	188 (88%)	22 (10%)	3 (1%)	13	55
2	C	1333/1342 (99%)	1087 (82%)	204 (15%)	42 (3%)	5	40
2	H	1333/1342 (99%)	1091 (82%)	197 (15%)	45 (3%)	4	39
3	D	1154/1407 (82%)	933 (81%)	180 (16%)	41 (4%)	4	37
3	I	1154/1407 (82%)	934 (81%)	180 (16%)	40 (4%)	4	38
4	E	88/91 (97%)	77 (88%)	6 (7%)	5 (6%)	2	26
4	J	74/91 (81%)	64 (86%)	5 (7%)	5 (7%)	1	22
5	X	511/613 (83%)	453 (89%)	44 (9%)	14 (3%)	6	43
5	Y	454/613 (74%)	414 (91%)	30 (7%)	10 (2%)	8	47
All	All	7079/8222 (86%)	5886 (83%)	962 (14%)	231 (3%)	4	39

5 of 231 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	52	PRO
1	A	319	GLU
1	B	20	SER
2	C	21	VAL
2	C	39	ILE

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	281/286 (98%)	272 (97%)	9 (3%)	44	74
1	B	189/286 (66%)	183 (97%)	6 (3%)	44	74

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	F	197/286 (69%)	193 (98%)	4 (2%)	60	83
1	G	185/286 (65%)	184 (100%)	1 (0%)	91	96
2	C	1150/1157 (99%)	1090 (95%)	60 (5%)	27	63
2	H	1150/1157 (99%)	1090 (95%)	60 (5%)	27	63
3	D	971/1168 (83%)	916 (94%)	55 (6%)	24	61
3	I	971/1168 (83%)	916 (94%)	55 (6%)	24	61
4	E	74/75 (99%)	72 (97%)	2 (3%)	50	77
4	J	65/75 (87%)	64 (98%)	1 (2%)	70	87
5	X	460/540 (85%)	443 (96%)	17 (4%)	39	71
5	Y	407/540 (75%)	392 (96%)	15 (4%)	39	71
All	All	6100/7024 (87%)	5815 (95%)	285 (5%)	30	66

5 of 285 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	1341	ARG
2	H	46	GLN
3	I	1149	ARG
5	X	21	TYR
5	X	476	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 44 such sidechains are listed below:

Mol	Chain	Res	Type
4	E	31	GLN
5	X	446	GLN
3	I	274	ASN
5	X	8	GLN
5	X	54	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 ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 8 ligands modelled in this entry, 6 are monoatomic - leaving 2 for Mogul analysis.

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	$\# Z > 2$
6	1RL	C	1401	-	65,70,70	3.03	19 (29%)	89,104,104	2.00	18 (20%)
6	1RL	H	1401	-	65,70,70	3.02	19 (29%)	89,104,104	2.09	21 (23%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	1RL	C	1401	-	-	0/65/84/84	0/1/6/6
6	1RL	H	1401	-	-	0/65/84/84	0/1/6/6

The worst 5 of 38 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	H	1401	1RL	C12-C11	-4.22	1.33	1.54
6	C	1401	1RL	C12-C11	-4.21	1.33	1.54
6	C	1401	1RL	C3-C4	-3.08	1.41	1.47
6	H	1401	1RL	C3-C4	-3.04	1.41	1.47
6	H	1401	1RL	C34-C26	-2.68	1.47	1.53

The worst 5 of 39 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	H	1401	1RL	C20-C19-C18	-7.33	108.88	126.06
6	H	1401	1RL	C12-O3-C6	-6.03	94.83	107.57
6	C	1401	1RL	C20-C19-C18	-6.01	111.97	126.06
6	C	1401	1RL	C12-O3-C6	-5.87	95.17	107.57
6	H	1401	1RL	O4-C11-C5	-4.27	122.06	131.84

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	C	1401	1RL	6	0
6	H	1401	1RL	10	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 ⓘ

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	323/329 (98%)	0.01	10 (3%)	49	39	0, 68, 172, 270	0
1	B	221/329 (67%)	0.33	15 (6%)	18	13	6, 90, 201, 277	0
1	F	229/329 (69%)	0.41	22 (9%)	9	7	9, 123, 209, 249	0
1	G	217/329 (65%)	0.40	21 (9%)	8	7	34, 112, 176, 230	0
2	C	1335/1342 (99%)	-0.05	47 (3%)	44	35	0, 43, 165, 367	0
2	H	1335/1342 (99%)	0.11	80 (5%)	23	16	0, 81, 205, 388	0
3	D	1160/1407 (82%)	-0.03	39 (3%)	46	36	0, 34, 158, 313	0
3	I	1160/1407 (82%)	0.11	72 (6%)	21	16	0, 50, 190, 338	0
4	E	90/91 (98%)	-0.35	1 (1%)	80	72	0, 34, 108, 170	0
4	J	76/91 (83%)	0.16	2 (2%)	56	46	3, 75, 160, 201	0
5	X	517/613 (84%)	0.27	41 (7%)	13	11	1, 99, 236, 425	0
5	Y	458/613 (74%)	0.19	35 (7%)	15	11	1, 102, 234, 374	0
All	All	7121/8222 (86%)	0.09	385 (5%)	26	21	0, 66, 199, 425	0

The worst 5 of 385 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	X	319	ALA	12.8
5	X	36	VAL	11.2
5	Y	239	GLY	9.7
5	X	35	ILE	9.1
3	I	11	GLN	8.8

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

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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
6	1RL	H	1401	65/65	0.79	0.42	1.57	20,20,20,20	0
6	1RL	C	1401	65/65	0.82	0.36	0.63	20,20,20,20	0
7	ZN	D	1502	1/1	0.96	0.19	-0.84	8,8,8,8	0
7	ZN	I	1502	1/1	0.98	0.11	-1.43	49,49,49,49	0
7	ZN	I	1501	1/1	0.95	0.03	-1.48	60,60,60,60	0
7	ZN	D	1501	1/1	0.94	0.05	-1.77	54,54,54,54	0
8	MG	D	1503	1/1	0.96	0.32	-	24,24,24,24	0
8	MG	I	1503	1/1	0.92	0.66	-	20,20,20,20	0

6.5 Other polymers [i](#)

There are no such residues in this entry.