



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 05:38 PM UTC

PDB ID : 4GZZ / pdb_00004gzz
Title : Crystal structures of bacterial RNA Polymerase paused elongation complexes
Authors : Weixlbaumer, A.; Leon, K.; Landick, R.; Darst, S.A.
Deposited on : 2012-09-06
Resolution : 4.29 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

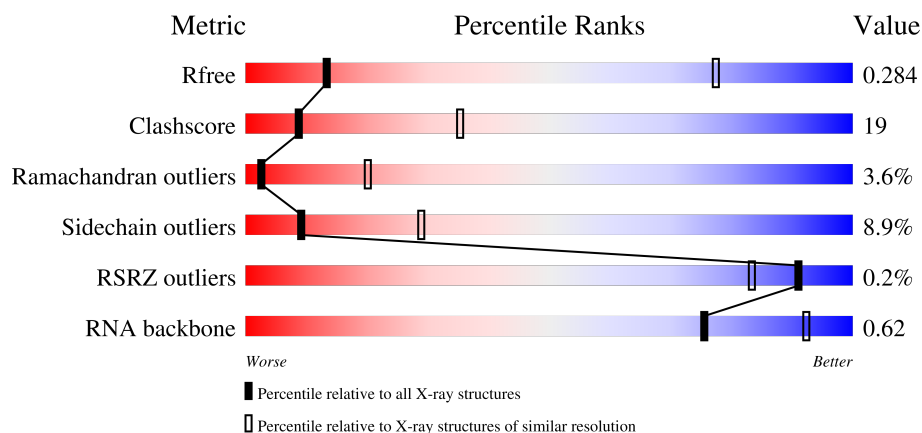
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.29 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	1052 (4.70-3.90)
Clashscore	190562	1097 (4.70-3.90)
Ramachandran outliers	187476	1001 (4.70-3.90)
Sidechain outliers	187428	1007 (4.72-3.88)
RSRZ outliers	180081	1049 (4.70-3.90)
RNA backbone	3983	1036 (5.50-3.10)

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	315	
1	B	315	
2	C	1119	
3	D	1534	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	E	99	41%41%10%6%
5	N	13	46%38%15%
6	R	16	6%44%6%44%
7	T	22	50%50%

2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 24400 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	223	Total	C	N	O	S	0	0	0
			1759	1123	306	328	2			
1	B	223	Total	C	N	O	S	0	0	0
			1759	1123	306	328	2			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1083	Total	C	N	O	S	0	0	0
			8548	5412	1524	1588	24			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1358	Total	C	N	O	S	0	0	0
			10714	6780	1900	2001	33			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	1525	HIS	-	expression tag	UNP Q8RQE8
D	1526	HIS	-	expression tag	UNP Q8RQE8
D	1527	HIS	-	expression tag	UNP Q8RQE8
D	1528	HIS	-	expression tag	UNP Q8RQE8
D	1529	HIS	-	expression tag	UNP Q8RQE8
D	1530	HIS	-	expression tag	UNP Q8RQE8
D	1531	HIS	-	expression tag	UNP Q8RQE8
D	1532	HIS	-	expression tag	UNP Q8RQE8
D	1533	HIS	-	expression tag	UNP Q8RQE8
D	1534	HIS	-	expression tag	UNP Q8RQE8

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	93	Total	C	N	O	S	0	0	0
			754	481	131	138	4			

- Molecule 5 is a DNA chain called non-template DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	N	11	Total	C	N	O	P	0	0	0
			225	107	43	64	11			

- Molecule 6 is a RNA chain called RNA transcript.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	R	9	Total	C	N	O	P	0	0	0
			191	85	31	66	9			

- Molecule 7 is a DNA chain called template DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	T	22	Total	C	N	O	P	0	0	0
			447	213	81	131	22			

- 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		

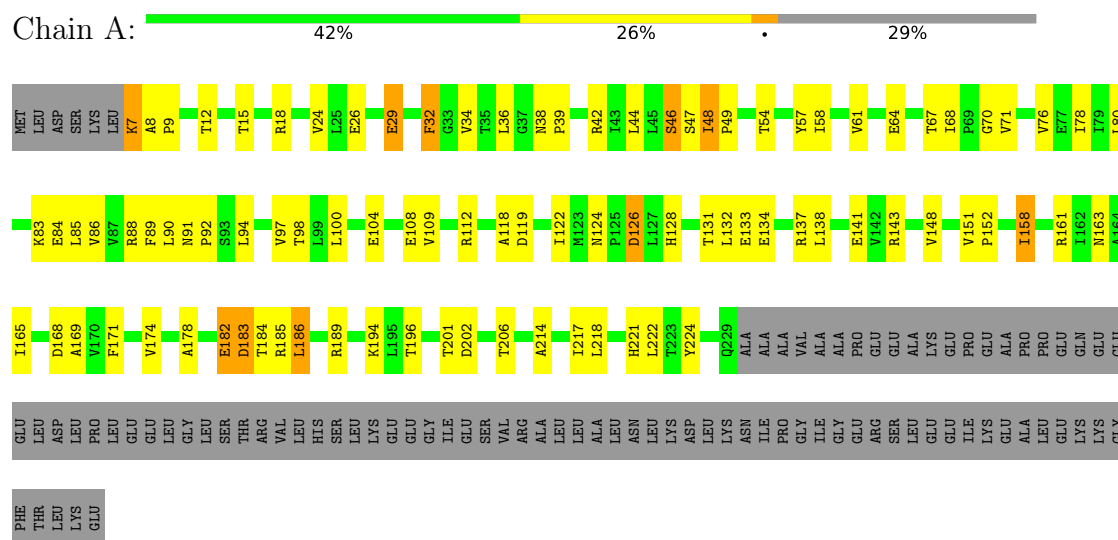
- Molecule 9 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

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

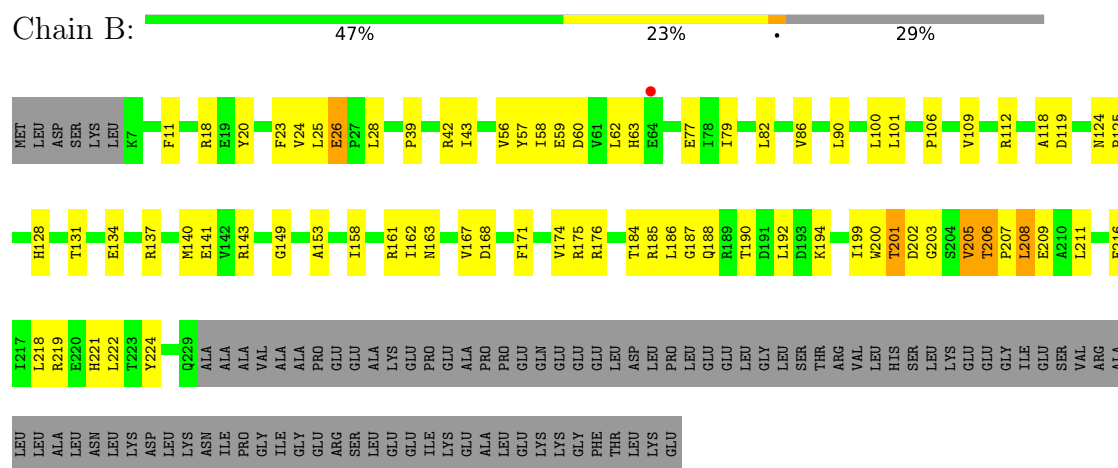
3 Residue-property plots

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: DNA-directed RNA polymerase subunit alpha

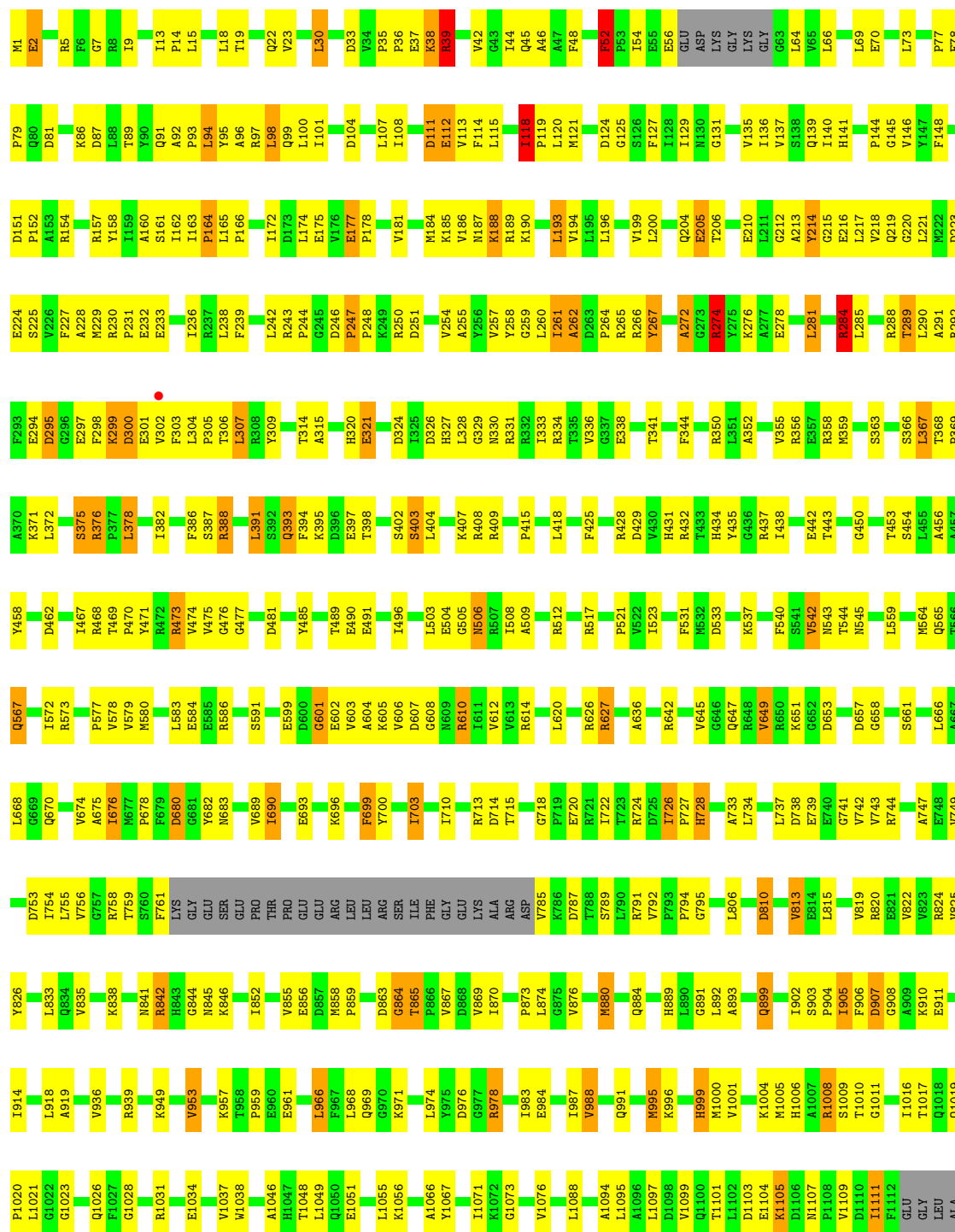


• Molecule 1: DNA-directed RNA polymerase subunit alpha



• Molecule 2: DNA-directed RNA polymerase subunit beta





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

Chain D: 49% 34% 5% 11%

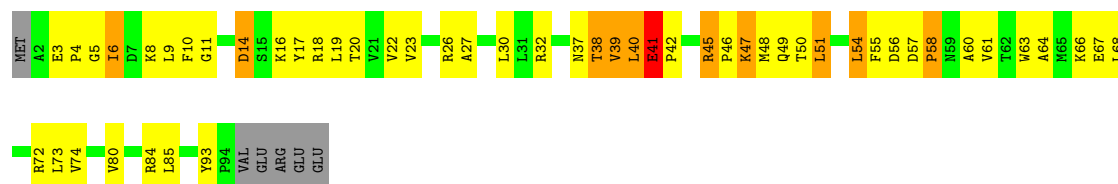


A1272	C1194	K1079	Q973	P877	L787	P706	S602	I513	G425	R346	PHE	LEU	E157	T81
V1273	Q1195	D1083	I974	I880	I792	T707	L603	P518	K426	V347	LEU	ARG	Y158	K82
I1274	T1196	L1086	Y978	L881	T793	H709	D605	P519	V427	N352	PRO	PRO	R159	S83
I1277	R1197	L1086	L983	L885	R796	R710	I606	L520	K428	V353	GLY	GLY	E160	I84
D1278	G1199	S1091	T984	E888	R797	G712	G612	P522	V431	V354	ALA	ALA	R162	R86
G1279	R1096	R1096	E987	E889	E798	I713	R613	D523	Y432	V355	THR	THR	Y163	R87
I1280	C1204	V1106	R988	V890	R800	G714	R614	L524	G430	P356	PRO	TRP	G164	Y86
E1281	Y1205	V1107	R989	E891	G801	A715	F614	R525	R434	G358	VAL	GLY	E167	H92
I1283	G1206	R1107	D990	D692	A802	W19	Q616	P526	V435	A359	VAL	VAL	E168	A100
E1284	R1108	R1108	Q881	G803	G803	V719	Q617	M527	V447	R361	HIS	GLY	Y169	H101
	E1109	E1109	I992	A803	A803	L720	V617	V528	V447	E362	GLY	ALA	P170	I102
	L1209	L1209	L993	L804	L804	V721	L618	T537	E448	E362	GLY	ALA	L171	W103
	M1211	C1112	E805	E805	E798	E722	K621	S538	S449	A363	VAL	PRO	P172	
	A1212	T1115	F806	F806	G723	G723	I452	D539	I452	G364	GLY	GLY	D176	D107
	R1213	S1119	A807	A807	Q724	Q724	S626	M541	A458	D365	ILE	ILE	A177	P109
			T808	T808	W725	W725	S627	D542	Q462	I367	GLY	GLY	L178	S110
			P809	P809	I726	I726	R628	L543		V368	GLN	PRO	A178	K111
			L813	L813	H729	H729	R629	L543		I371	PRO	ALA	L179	L115
			E817	E817	E735	E735	V630	L543		P373	GLY	GLY	K180	L116
			E820	E820	F736	F736	V633	L543		E467	ALA	ALA	D181	D117
			L823	L823	D739	D739	L637	R553	L473	V377	GLY	TYR	E183	L118
			N824	N824	F740	F740	Q641	L554	E474	I378	LEU	PHE	E184	S119
			A825	A825	D741	D741	G642	L554		A379	LEU	ARG	V186	
			P826	P826	G742	G742	G643	L554		V385	ALA	ALA	E190	L123
			L827	L827	D743	D743	L644	L557	L478	H386	GLY	GLY	Q126	Q126
			K828	K828	Q744	Q744	P645	L566		L387	ARG	GLY	L126	L126
			W829	W829	V749	V749	K646	I567	H483	H388	GLN	GLY	Y128	Y128
			A830	A830	F750	F750	R647	R568	P484	P390	VAL	VAL	Y129	Y129
			C831	C831	L751	L751	M648	N569	S485	A391	ALA	ALA	S130	S130
			R832	R832	S752	S752	K659	M573	R486	S392	ALA	GLY	K131	K131
			E833	E833			K660	L574	A487	I393	GLN	LEU	Y132	Y132
			T834	T834	E758	E758	K664	M573	A487	L394	VAL	LEU	I133	I133
			S835	S835	A759	A759		M573	R488	V395	VAL	GLY	V134	V134
			V836	V836	R760	R760	K664	L574	R489	V396	ALA	GLY	L135	L135
			G837	G837	I761	I761	E886	Q575	R493	K397	GLY	GLY	D136	D136
			R838	R838	Q762	Q762	V687	L574	K494	R206	GLY	GLY	P137	P137
			Y841	Y841	R763	R763	M688	Q575	R495	V400	GLY	GLY	G139	G139
			F843	F843	S765	S765	D889	Q575	L496	D406	GLY	ALA	L142	L142
			A844	A844	A766	A766	A890	Q575	V499	V407	THR	THR	V210	V210
			A852	A852	H767	H767	L691	L581	R500	E408	VAL	VAL	V211	V211
			V853	V853	N768	N768	E692	L582	A501	V409	THR	THR	R212	R212
			L857	L857	L769	L769	E694	D583	F502	S410	LEU	ARG	V213	V213
			L860	L860	L770	L770	H696	N584	S505	R414	THR	THR	E214	E214
			E965	E965	P777	P777	H696	P590	G506	V415	LEU	GLY	Y215	Y215
			E966	E966	K780	K780	V700	N593	G506	A416	LEU	GLY	R150	R150
			A967	A967			L701	P594	N507	P417	LEU	PRO	Q151	Q151
			D968	D968			R783	P594	E508	G418	VAL	VAL	L152	L152
			R969	R969			R704	L600	P509	D419	ALA	ALA	T154	T154
			K970	K970			A705	L601	M511	V420	THR	THR	T155	T155
			L971	L971					M512	L421			E156	E156
			L972	L972									GLY	GLY



- Molecule 4: DNA-directed RNA polymerase subunit omega

Chain E: 41% 41% 10% 6%



- Molecule 5: non-template DNA

Chain N: 46% 38% 15%



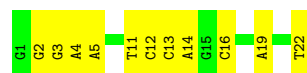
- Molecule 6: RNA transcript

Chain R: 6% 44% 6% 44%



- Molecule 7: template DNA

Chain T: 50% 50%



4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α , β , γ	286.55Å 286.55Å 199.41Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	38.78 – 4.29 38.78 – 4.29	Depositor EDS
% Data completeness (in resolution range)	99.0 (38.78-4.29) 87.7 (38.78-4.29)	Depositor EDS
R_{merge}	0.44	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.08 (at 4.28Å)	Xtriage
Refinement program	PHENIX 1.8_1069, CNS	Depositor
R, R_{free}	0.234 , 0.285 0.236 , 0.284	Depositor DCC
R_{free} test set	2074 reflections (3.34%)	wwPDB-VP
Wilson B-factor (Å ²)	164.3	Xtriage
Anisotropy	0.112	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 80.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.047 for h,-h-k,-l	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	24400	wwPDB-VP
Average B, all atoms (Å ²)	121.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.23% 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: 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	0.37	0/1791	0.82	1/2436 (0.0%)
1	B	0.37	0/1791	0.82	2/2436 (0.1%)
2	C	0.38	0/8711	0.86	20/11784 (0.2%)
3	D	0.38	0/10897	0.83	16/14726 (0.1%)
4	E	0.37	0/768	0.91	4/1035 (0.4%)
5	N	0.10	0/252	0.51	0/386
6	R	0.10	0/212	0.22	0/328
7	T	0.13	0/500	0.54	0/768
All	All	0.37	0/24922	0.83	43/33899 (0.1%)

There are no bond length outliers.

The worst 5 of 43 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	705	ALA	CA-C-N	-10.62	106.56	119.84
3	D	705	ALA	C-N-CA	-10.62	106.56	119.84
2	C	726	ILE	CA-C-N	7.64	129.40	119.84
2	C	726	ILE	C-N-CA	7.64	129.40	119.84
2	C	247	PRO	N-CA-C	7.46	119.81	110.70

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	1759	0	1805	66	0
1	B	1759	0	1805	51	1
2	C	8548	0	8650	381	0
3	D	10714	0	10936	427	1
4	E	754	0	769	41	0
5	N	225	0	124	3	0
6	R	191	0	95	10	0
7	T	447	0	248	10	0
8	D	2	0	0	0	0
9	D	1	0	0	0	0
All	All	24400	0	24432	912	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 19.

The worst 5 of 912 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)
2:C:428:ARG:HH12	3:D:1086:LEU:HD21	1.24	0.98
2:C:610:ARG:HG3	2:C:610:ARG:HH11	1.27	0.96
2:C:846:LYS:NZ	6:R:29:U:OP1	1.99	0.96
1:B:188:GLN:O	3:D:646:LYS:NZ	2.00	0.94
1:A:80:LEU:HD21	2:C:573:ARG:HH11	1.34	0.92

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)
1:B:143:ARG:NH1	3:D:1297:GLU:OE2[3_455]	2.18	0.02

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	221/315 (70%)	186 (84%)	29 (13%)	6 (3%)	4	25
1	B	221/315 (70%)	193 (87%)	25 (11%)	3 (1%)	9	39
2	C	1077/1119 (96%)	897 (83%)	144 (13%)	36 (3%)	3	21
3	D	1352/1534 (88%)	1097 (81%)	198 (15%)	57 (4%)	2	17
4	E	91/99 (92%)	73 (80%)	13 (14%)	5 (6%)	1	15
All	All	2962/3382 (88%)	2446 (83%)	409 (14%)	107 (4%)	2	20

5 of 107 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	118	ALA
1	A	184	THR
2	C	2	GLU
2	C	111	ASP
2	C	164	PRO

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	196/273 (72%)	187 (95%)	9 (5%)	24	46
1	B	196/273 (72%)	186 (95%)	10 (5%)	21	44
2	C	912/941 (97%)	831 (91%)	81 (9%)	9	29
3	D	1147/1289 (89%)	1034 (90%)	113 (10%)	7	24
4	E	82/88 (93%)	69 (84%)	13 (16%)	2	13
All	All	2533/2864 (88%)	2307 (91%)	226 (9%)	9	29

5 of 226 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	142	LEU
4	E	49	GLN
3	D	686	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	E	41	GLU
3	D	1327	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	125	GLN
3	D	1465	ASN
3	D	575	GLN
3	D	1227	GLN
3	D	388	HIS

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
6	R	8/16 (50%)	1 (12%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
6	R	22	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 3 ligands modelled in this entry, 3 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

There are no such residues in this entry.

5.8 Polymer linkage issues

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	223/315 (70%)	-0.74	0 100 100	52, 116, 169, 205	0
1	B	223/315 (70%)	-0.66	1 (0%) 88 77	53, 107, 163, 198	0
2	C	1083/1119 (96%)	-0.60	1 (0%) 92 88	34, 108, 177, 264	0
3	D	1358/1534 (88%)	-0.62	4 (0%) 90 80	31, 110, 196, 250	0
4	E	93/99 (93%)	-0.66	0 100 100	65, 121, 179, 228	0
5	N	11/13 (84%)	-0.26	0 100 100	340, 362, 386, 405	0
6	R	9/16 (56%)	-0.66	0 100 100	201, 206, 235, 243	0
7	T	22/22 (100%)	-0.15	0 100 100	219, 336, 388, 406	0
All	All	3022/3433 (88%)	-0.62	6 (0%) 91 83	31, 111, 192, 406	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	802	ALA	3.6
2	C	302	VAL	2.6
3	D	777	PRO	2.5
3	D	393	ILE	2.2
3	D	949	ILE	2.1

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	1601	1/1	0.92	0.11	70,70,70,70	0
9	MG	D	1603	1/1	0.95	0.04	70,70,70,70	0
8	ZN	D	1602	1/1	1.00	0.02	70,70,70,70	0

6.5 Other polymers [i](#)

There are no such residues in this entry.