



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 1, 2026 – 12:28 PM UTC

PDB ID : 6JCX / pdb_00006jcx
Title : Mycobacterium tuberculosis transcription initiation complex with ECF sigma factor sigma H and 6nt RNA
Authors : Li, L.; Zhang, Y.
Deposited on : 2019-01-30
Resolution : 2.90 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

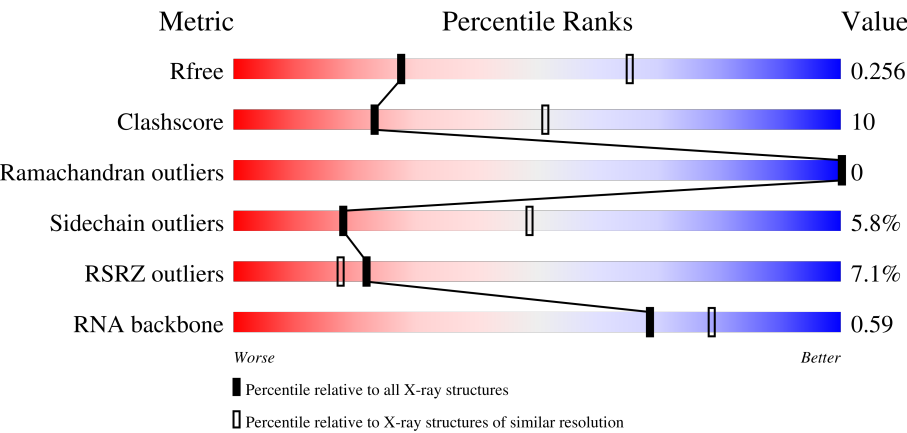
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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.



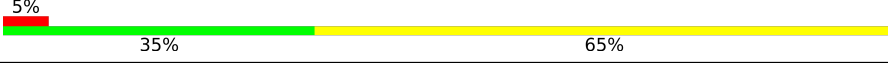
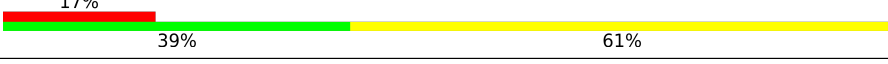
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R _{free}	180053	2481 (2.90-2.90)
Clashscore	190562	2690 (2.90-2.90)
Ramachandran outliers	187476	2623 (2.90-2.90)
Sidechain outliers	187428	2625 (2.90-2.90)
RSRZ outliers	180081	2481 (2.90-2.90)
RNA backbone	3983	1120 (3.10-2.70)

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	368	10% 42%16%.40%
1	B	368	10% 43%19%.37%
2	C	1174	8% 70%26%. .
3	D	1317	4% 70%24%. .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	E	110	
5	F	218	
6	H	20	
7	G	23	
8	I	6	

2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 25143 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	219	Total	C	N	O	S	0	0	0
			1654	1044	282	326	2			
1	B	232	Total	C	N	O	S	0	0	0
			1696	1073	287	334	2			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	MET	-	initiating methionine	UNP P9WGZ1
A	-19	GLY	-	expression tag	UNP P9WGZ1
A	-18	HIS	-	expression tag	UNP P9WGZ1
A	-17	HIS	-	expression tag	UNP P9WGZ1
A	-16	HIS	-	expression tag	UNP P9WGZ1
A	-15	HIS	-	expression tag	UNP P9WGZ1
A	-14	HIS	-	expression tag	UNP P9WGZ1
A	-13	HIS	-	expression tag	UNP P9WGZ1
A	-12	HIS	-	expression tag	UNP P9WGZ1
A	-11	HIS	-	expression tag	UNP P9WGZ1
A	-10	HIS	-	expression tag	UNP P9WGZ1
A	-9	HIS	-	expression tag	UNP P9WGZ1
A	-8	SER	-	expression tag	UNP P9WGZ1
A	-7	SER	-	expression tag	UNP P9WGZ1
A	-6	GLY	-	expression tag	UNP P9WGZ1
A	-5	HIS	-	expression tag	UNP P9WGZ1
A	-4	ILE	-	expression tag	UNP P9WGZ1
A	-3	GLU	-	expression tag	UNP P9WGZ1
A	-2	GLY	-	expression tag	UNP P9WGZ1
A	-1	ARG	-	expression tag	UNP P9WGZ1
A	0	HIS	-	expression tag	UNP P9WGZ1
B	-20	MET	-	initiating methionine	UNP P9WGZ1
B	-19	GLY	-	expression tag	UNP P9WGZ1
B	-18	HIS	-	expression tag	UNP P9WGZ1
B	-17	HIS	-	expression tag	UNP P9WGZ1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-16	HIS	-	expression tag	UNP P9WGZ1
B	-15	HIS	-	expression tag	UNP P9WGZ1
B	-14	HIS	-	expression tag	UNP P9WGZ1
B	-13	HIS	-	expression tag	UNP P9WGZ1
B	-12	HIS	-	expression tag	UNP P9WGZ1
B	-11	HIS	-	expression tag	UNP P9WGZ1
B	-10	HIS	-	expression tag	UNP P9WGZ1
B	-9	HIS	-	expression tag	UNP P9WGZ1
B	-8	SER	-	expression tag	UNP P9WGZ1
B	-7	SER	-	expression tag	UNP P9WGZ1
B	-6	GLY	-	expression tag	UNP P9WGZ1
B	-5	HIS	-	expression tag	UNP P9WGZ1
B	-4	ILE	-	expression tag	UNP P9WGZ1
B	-3	GLU	-	expression tag	UNP P9WGZ1
B	-2	GLY	-	expression tag	UNP P9WGZ1
B	-1	ARG	-	expression tag	UNP P9WGZ1
B	0	HIS	-	expression tag	UNP P9WGZ1

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1138	Total	C	N	O	S	0	0	0
			8731	5458	1528	1706	39			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	5	MET	-	initiating methionine	UNP P9WGY9
C	6	VAL	-	expression tag	UNP P9WGY9

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1258	Total	C	N	O	S	0	0	0
			9822	6153	1776	1852	41			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	MET	-	initiating methionine	UNP P9WGY7
D	1	VAL	-	expression tag	UNP P9WGY7

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	E	71	Total	C	N	O	0	0	0
			566	362	95	109			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	186	Total	C	N	O	S	0	0	0
			1477	926	260	286	5			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-1	GLY	-	expression tag	UNP P9WGH9
F	0	ALA	-	expression tag	UNP P9WGH9

- Molecule 6 is a DNA chain called DNA (5'-D(*TP*GP*CP*AP*TP*CP*CP*GP*TP*GP*AP*GP*TP*CP*GP*AP*GP*GP*GP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	H	20	Total	C	N	O	P	0	0	0
			412	196	77	120	19			

- Molecule 7 is a DNA chain called DNA (5'-D(*TP*TP*GP*TP*GP*GP*GP*AP*GP*CP*TP*GP*TP*CP*AP*CP*GP*GP*AP*TP*GP*CP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	23	Total	C	N	O	P	0	0	0
			475	226	89	138	22			

- Molecule 8 is a RNA chain called RNA (5'-R(*CP*CP*UP*CP*GP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	I	6	Total	C	N	O	P	0	0	0
			122	56	21	40	5			

- Molecule 9 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	D	2	Total	Zn	0	0
			2	2		

- Molecule 10 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

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

- Molecule 11 is water.

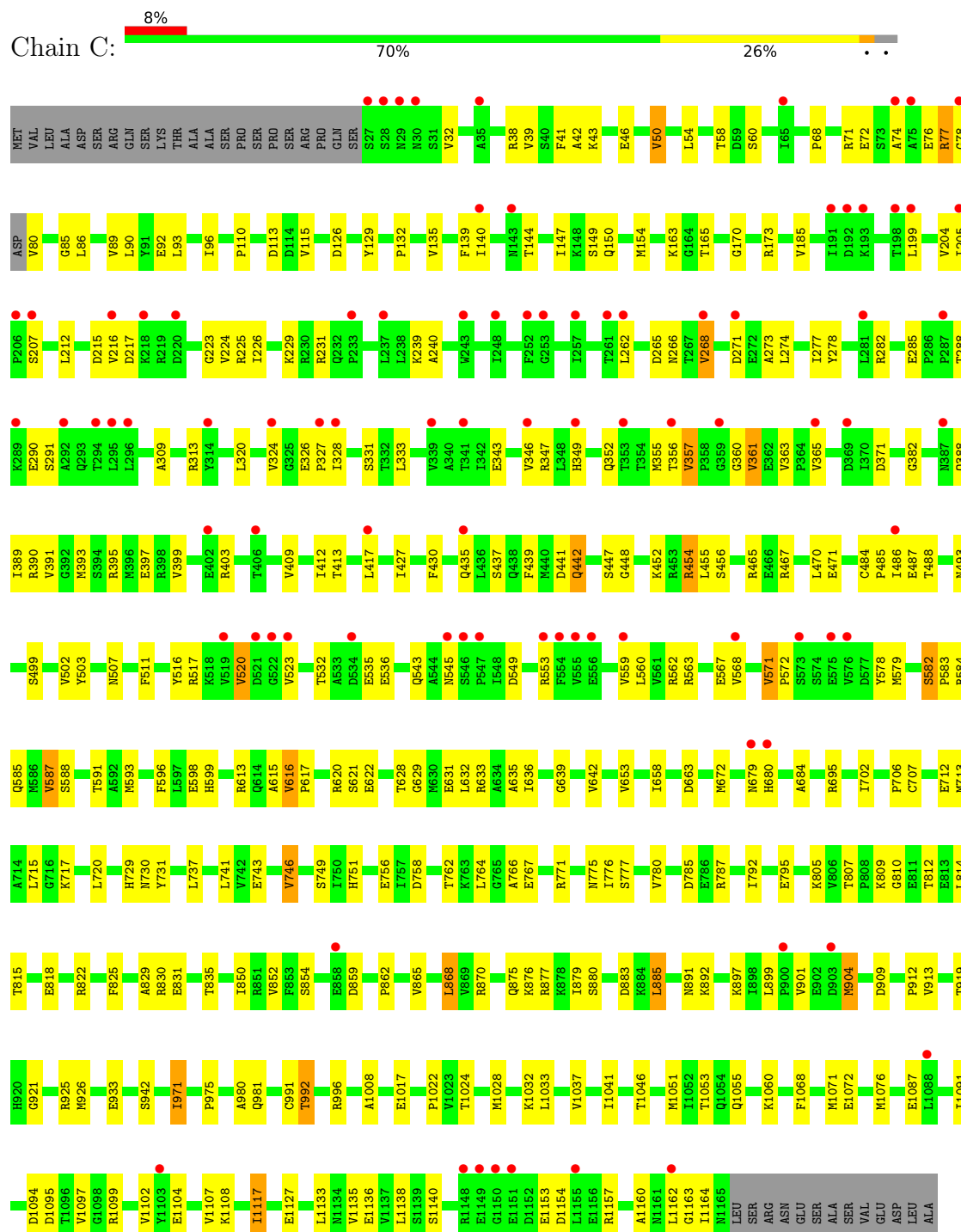
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	9	Total 9	O 9	0	0
11	B	9	Total 9	O 9	0	0
11	C	66	Total 66	O 66	0	0
11	D	71	Total 71	O 71	0	0
11	E	3	Total 3	O 3	0	0
11	F	11	Total 11	O 11	0	0
11	H	7	Total 7	O 7	0	0
11	G	4	Total 4	O 4	0	0
11	I	5	Total 5	O 5	0	0

i

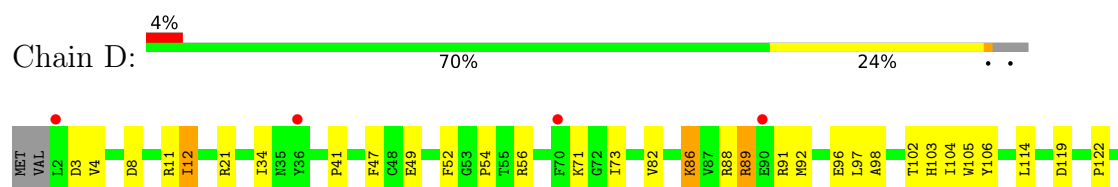
- Molecule 1: DNA-directed RNA polymerase subunit alpha

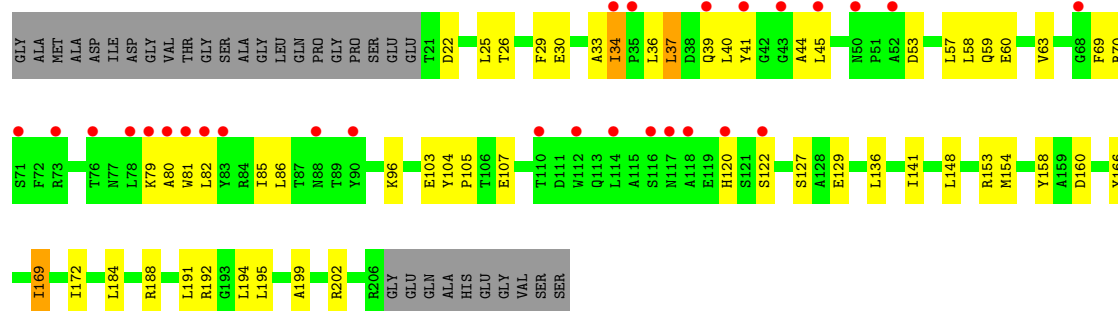


• Molecule 2: DNA-directed RNA polymerase subunit beta



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

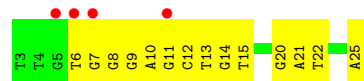




- Molecule 6: DNA (5'-D(*TP*GP*CP*AP*TP*CP*CP*GP*TP*GP*AP*GP*TP*CP*GP*AP*GP*GP*GP*T)-3')



- Molecule 7: DNA (5'-D(*TP*TP*GP*TP*GP*GP*GP*AP*GP*CP*TP*GP*TP*CP*AP*CP*GP*GP*AP*TP*GP*CP*A)-3')



- Molecule 8: RNA (5'-R(*CP*CP*UP*CP*GP*A)-3')



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	125.14Å 161.00Å 128.85Å 90.00° 117.06° 90.00°	Depositor
Resolution (Å)	36.09 – 2.90 36.09 – 2.90	Depositor EDS
% Data completeness (in resolution range)	97.5 (36.09-2.90) 97.6 (36.09-2.90)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.13 (at 2.90Å)	Xtriage
Refinement program	PHENIX 1.14_3260	Depositor
R, R_{free}	0.210 , 0.256 (Not available) , 0.256	Depositor DCC
R_{free} test set	2446 reflections (2.49%)	wwPDB-VP
Wilson B-factor (Å ²)	67.4	Xtriage
Anisotropy	0.566	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 60.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.015 for l,-k,h	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	25143	wwPDB-VP
Average B, all atoms (Å ²)	77.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.60% 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, 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	0.22	0/1679	0.44	0/2285
1	B	0.17	0/1722	0.40	0/2351
2	C	0.22	0/8889	0.43	0/12067
3	D	0.22	0/9984	0.44	2/13500 (0.0%)
4	E	0.21	0/577	0.43	0/786
5	F	0.21	0/1506	0.41	0/2041
6	H	0.29	0/462	0.50	0/713
7	G	0.27	0/533	0.66	2/823 (0.2%)
8	I	0.24	0/135	0.52	0/208
All	All	0.22	0/25487	0.44	4/34774 (0.0%)

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
3	D	0	1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	G	8	DG	OP1-P-O3'	-8.63	82.12	108.00
7	G	8	DG	OP2-P-O3'	-8.32	83.03	108.00
3	D	578	ARG	CA-C-N	5.61	132.25	121.54
3	D	578	ARG	C-N-CA	5.61	132.25	121.54

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	D	579	LEU	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	1654	0	1687	38	0
1	B	1696	0	1683	41	0
2	C	8731	0	8564	205	0
3	D	9822	0	9862	209	0
4	E	566	0	567	11	0
5	F	1477	0	1422	39	0
6	H	412	0	227	11	0
7	G	475	0	261	18	0
8	I	122	0	66	4	0
9	D	2	0	0	0	0
10	D	1	0	0	0	0
11	A	9	0	0	0	0
11	B	9	0	0	1	0
11	C	66	0	0	4	0
11	D	71	0	0	2	0
11	E	3	0	0	0	0
11	F	11	0	0	0	0
11	G	4	0	0	0	0
11	H	7	0	0	0	0
11	I	5	0	0	0	0
All	All	25143	0	24339	507	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 10.

The worst 5 of 507 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:485:PRO:HB2	3:D:853:THR:HG21	1.52	0.89
2:C:1051:MET:O	3:D:89:ARG:NH2	2.08	0.87
2:C:1053:THR:HG23	2:C:1055:GLN:H	1.40	0.86

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:1024:THR:H	3:D:730:THR:HG21	1.43	0.84
3:D:505:HIS:HD2	3:D:1005:GLU:HG3	1.49	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	215/368 (58%)	210 (98%)	5 (2%)	0	100	100
1	B	228/368 (62%)	213 (93%)	15 (7%)	0	100	100
2	C	1134/1174 (97%)	1084 (96%)	50 (4%)	0	100	100
3	D	1250/1317 (95%)	1205 (96%)	45 (4%)	0	100	100
4	E	67/110 (61%)	62 (92%)	5 (8%)	0	100	100
5	F	184/218 (84%)	180 (98%)	4 (2%)	0	100	100
All	All	3078/3555 (87%)	2954 (96%)	124 (4%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains [i](#)

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	186/315 (59%)	171 (92%)	15 (8%)	11	33

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	181/315 (58%)	164 (91%)	17 (9%)	8	26
2	C	940/995 (94%)	885 (94%)	55 (6%)	18	48
3	D	1040/1096 (95%)	987 (95%)	53 (5%)	21	53
4	E	62/90 (69%)	61 (98%)	1 (2%)	55	83
5	F	149/175 (85%)	141 (95%)	8 (5%)	20	51
All	All	2558/2986 (86%)	2409 (94%)	149 (6%)	18	49

5 of 149 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	820	MET
5	F	45	LEU
3	D	863	THR
3	D	1084	GLN
2	C	447	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	1109	GLN
4	E	105	HIS
3	D	1125	GLN
5	F	117	ASN
4	E	64	ASN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
8	I	5/6 (83%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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](#)

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 [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	219/368 (59%)	0.32	5 (2%) 61 52	46, 62, 95, 132	0
1	B	232/368 (63%)	1.14	36 (15%) 5 4	64, 100, 128, 144	0
2	C	1138/1174 (96%)	0.51	89 (7%) 19 16	36, 64, 143, 173	0
3	D	1258/1317 (95%)	0.43	48 (3%) 44 36	39, 67, 111, 144	0
4	E	71/110 (64%)	0.90	12 (16%) 4 3	65, 84, 112, 120	0
5	F	186/218 (85%)	1.19	28 (15%) 5 4	24, 92, 131, 141	12 (6%)
6	H	20/20 (100%)	0.15	1 (5%) 34 27	47, 70, 137, 142	0
7	G	23/23 (100%)	1.03	4 (17%) 4 3	80, 102, 159, 169	3 (13%)
8	I	6/6 (100%)	-0.42	0 100 100	47, 52, 64, 69	0
All	All	3153/3604 (87%)	0.56	223 (7%) 22 17	24, 70, 129, 173	15 (0%)

The worst 5 of 223 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	F	114	LEU	6.7
5	F	68	GLY	5.8
2	C	75	ALA	5.0
1	B	60	LEU	5.0
2	C	1151	GLU	4.8

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
10	MG	D	2003	1/1	0.89	0.17	46,46,46,46	0
9	ZN	D	2002	1/1	0.99	0.02	91,91,91,91	0
9	ZN	D	2001	1/1	0.99	0.02	71,71,71,71	0

6.5 Other polymers [i](#)

There are no such residues in this entry.