



wwPDB X-ray Structure Validation Summary Report ⓘ

May 28, 2020 – 08:50 pm BST

PDB ID : 1WCM
Title : Complete 12-Subunit RNA Polymerase II at 3.8 Angstrom
Authors : Armache, K.-J.; Mitterweger, S.; Meinhart, A.; Cramer, P.
Deposited on : 2004-11-17
Resolution : 3.80 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : 2.11
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.11

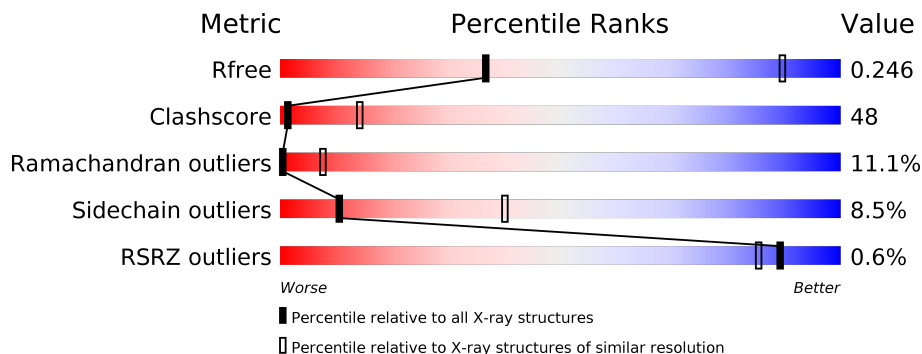
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.80 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}	130704	1212 (4.00-3.60)
Clashscore	141614	1288 (4.00-3.60)
Ramachandran outliers	138981	1243 (4.00-3.60)
Sidechain outliers	138945	1237 (4.00-3.60)
RSRZ outliers	127900	1121 (4.00-3.60)

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 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	1733	<p>27% 43% 10% 18%</p>
2	B	1224	<p>29% 48% 12% 10%</p>
3	C	318	<p>23% 48% 12% 16%</p>
4	D	177	<p>42% 46% 11%</p>
5	E	215	<p>40% 54% 6%</p>
6	F	155	<p>18% 30% 6% 46%</p>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	G	171	
8	H	146	 % % % %
9	I	122	 2% % % % % % % %
10	J	70	
11	K	120	
12	L	70	 % % % % % % % %

2 Entry composition

There are 14 unique types of molecules in this entry. The entry contains 30945 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 II LARGEST SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1416	11140	7021	1946	2111	62	0	0	0

- Molecule 2 is a protein called DNA-DIRECTED RNA POLYMERASE II SECOND LARGEST SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	1097	8720	5526	1523	1617	54	0	0	0

- Molecule 3 is a protein called DNA-DIRECTED RNA POLYMERASE II 45 KDA POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	266	2095	1317	348	417	13	0	0	0

- Molecule 4 is a protein called DNA-DIRECTED RNA POLYMERASE II 32 KDA POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	177	1356	840	241	273	2	0	0	0

- Molecule 5 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III 27 KDA POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	214	1752	1111	309	321	11	0	0	0

- Molecule 6 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III 23

KDA POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	F	84	679	434	115	127	3	0	0	0

- Molecule 7 is a protein called DNA-DIRECTED RNA POLYMERASE II 19 KD POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
7	G	171	1340	861	222	249	8	0	0	0

- Molecule 8 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III 14.5 KDA POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
8	H	133	1068	673	180	211	4	0	0	0

- Molecule 9 is a protein called DNA-DIRECTED RNA POLYMERASE II 14.2 KDA POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
9	I	119	971	596	179	186	10	0	0	0

- Molecule 10 is a protein called DNA-DIRECTED RNA POLYMERASES I, II AND III 8.3 KDA POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
10	J	65	532	339	93	94	6	0	0	0

- Molecule 11 is a protein called DNA-DIRECTED RNA POLYMERASE II 13.6 KDA POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
11	K	115	920	590	157	171	2	0	0	1

- Molecule 12 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III 7.7 KDA POLYPEPTIDE.

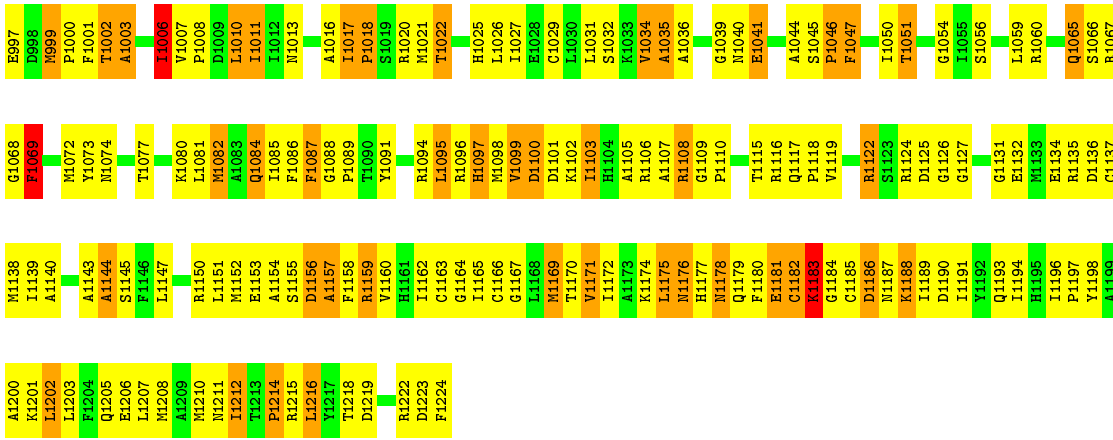
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
12	L	46	363	224	72	63	4	0	0	0

- Molecule 13 is ZINC ION (three-letter code: ZN) (formula: Zn).

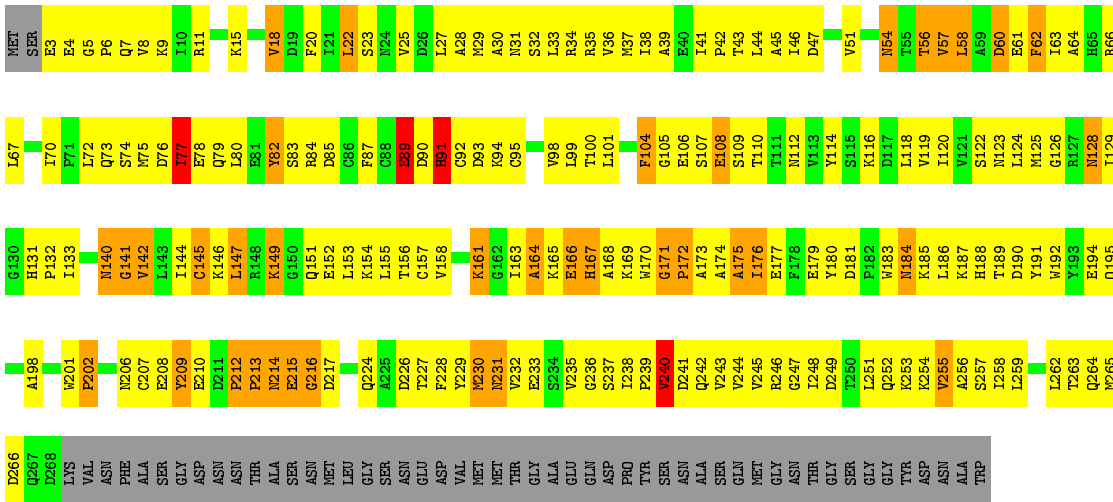
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	J	1	Total	Zn	0	0
			1	1		
13	B	1	Total	Zn	0	0
			1	1		
13	I	2	Total	Zn	0	0
			2	2		
13	C	1	Total	Zn	0	0
			1	1		
13	A	2	Total	Zn	0	0
			2	2		
13	L	1	Total	Zn	0	0
			1	1		

- Molecule 14 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

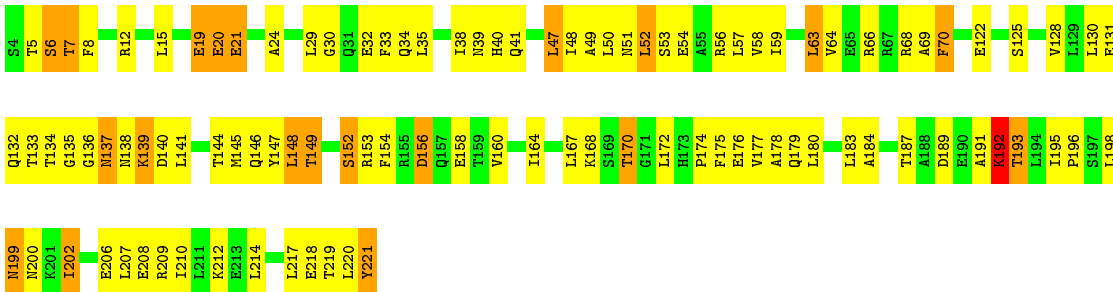
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	A	1	Total	Mg	0	0
			1	1		



• Molecule 3: DNA-DIRECTED RNA POLYMERASE II 45 KDA POLYPEPTIDE

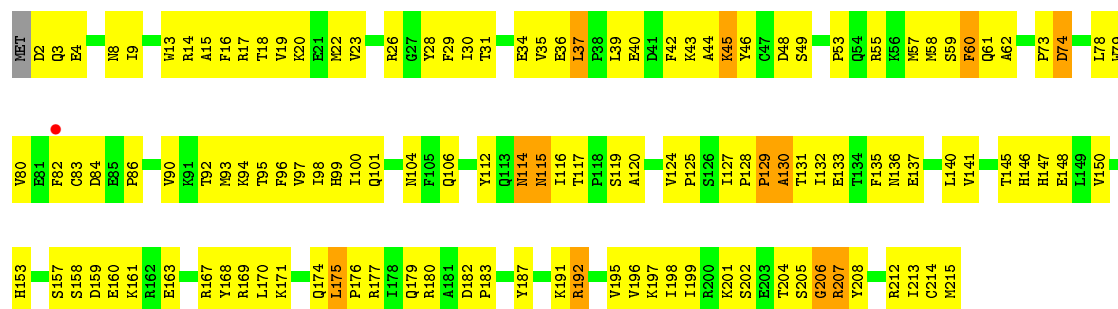


• Molecule 4: DNA-DIRECTED RNA POLYMERASE II 32 KDA POLYPEPTIDE

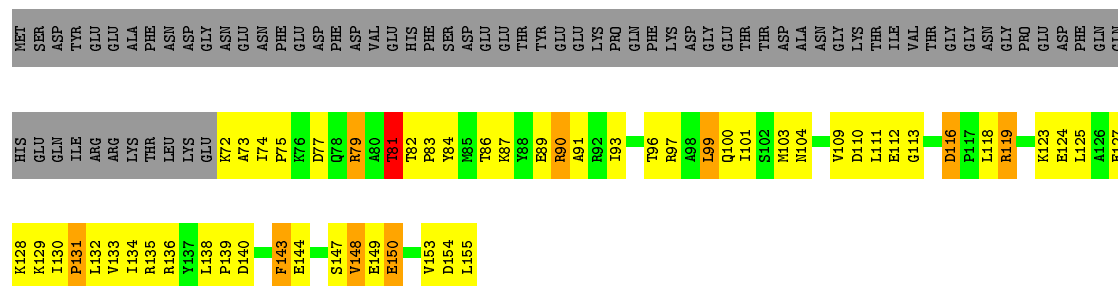


• Molecule 5: DNA-DIRECTED RNA POLYMERASES I, II, AND III 27 KDA POLYPEPTIDE

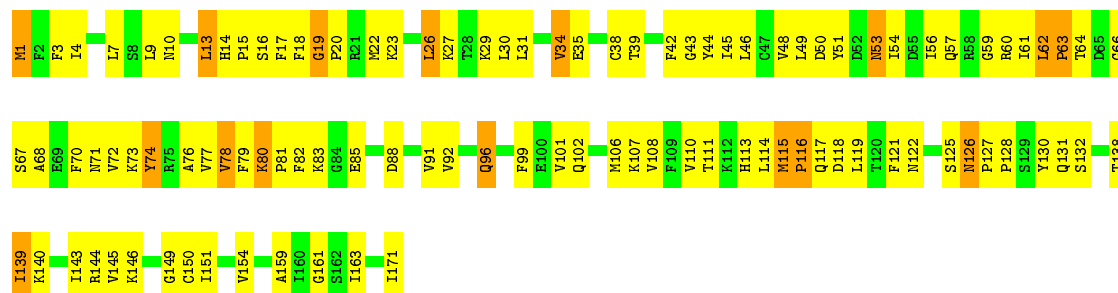




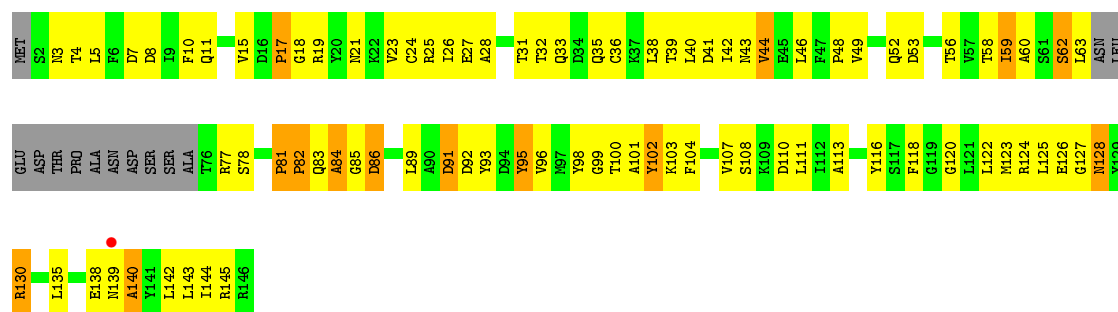
- Molecule 6: DNA-DIRECTED RNA POLYMERASES I, II, AND III 23 KDA POLYPEPTIDE



- Molecule 7: DNA-DIRECTED RNA POLYMERASE II 19 KD POLYPEPTIDE

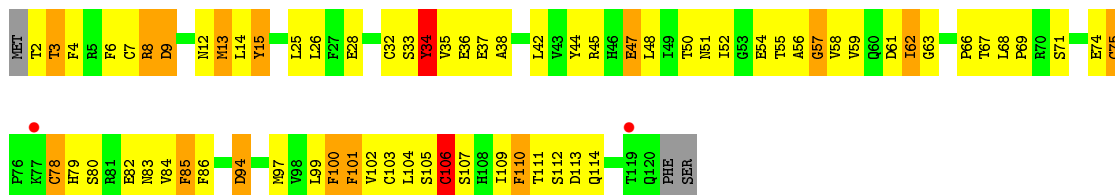


- Molecule 8: DNA-DIRECTED RNA POLYMERASES I, II, AND III 14.5 KDA POLYPEPTIDE




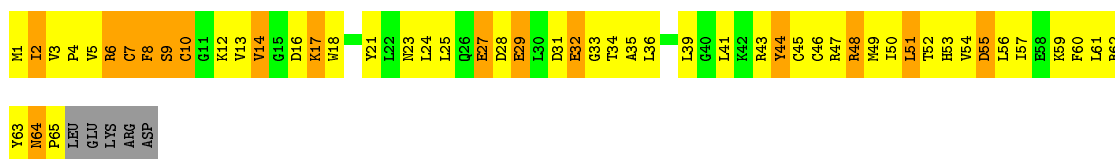
- Molecule 9: DNA-DIRECTED RNA POLYMERASE II 14.2 KDA POLYPEPTIDE

Chain I: 



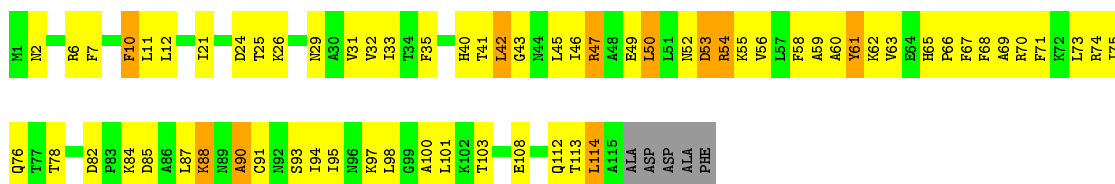
- Molecule 10: DNA-DIRECTED RNA POLYMERASES I, II AND III 8.3 KDA POLYPEPTIDE

Chain J: 




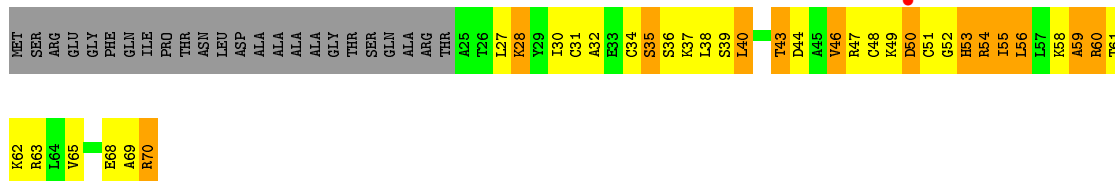
- Molecule 11: DNA-DIRECTED RNA POLYMERASE II 13.6 KDA POLYPEPTIDE

Chain K: 



- Molecule 12: DNA-DIRECTED RNA POLYMERASES I, II, AND III 7.7 KDA POLYPEPTIDE

Chain L: 



4 Data and refinement statistics i

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	222.72Å 395.13Å 284.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.80 47.39 – 3.80	Depositor EDS
% Data completeness (in resolution range)	99.1 (50.00-3.80) 99.2 (47.39-3.80)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.21 (at 3.77Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.257 , 0.285 0.208 , 0.246	Depositor DCC
R_{free} test set	2439 reflections (2.00%)	wwPDB-VP
Wilson B-factor (Å ²)	116.1	Xtriage
Anisotropy	0.510	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 70.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.015 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.021 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	30945	wwPDB-VP
Average B, all atoms (Å ²)	98.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.81% 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: 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	0.48	0/11339	0.73	4/15334 (0.0%)
2	B	0.47	0/8890	0.70	1/11990 (0.0%)
3	C	0.52	0/2133	0.76	0/2891
4	D	0.45	0/1365	0.71	1/1837 (0.1%)
5	E	0.43	0/1788	0.64	0/2406
6	F	0.53	0/691	0.78	0/933
7	G	0.53	0/1368	0.74	0/1844
8	H	0.40	0/1086	0.66	0/1470
9	I	0.48	0/989	0.77	0/1331
10	J	0.54	0/541	0.89	1/727 (0.1%)
11	K	0.50	0/938	0.68	0/1267
12	L	0.55	0/365	0.79	0/485
All	All	0.48	0/31493	0.72	7/42515 (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
1	A	0	1
3	C	0	1
All	All	0	2

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
10	J	10	CYS	CA-CB-SG	8.66	129.59	114.00
1	A	1403	GLU	N-CA-C	5.38	125.53	111.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	567	LYS	C-N-CD	5.34	139.62	128.40
2	B	1185	CYS	N-CA-C	-5.30	96.69	111.00
1	A	452	LYS	N-CA-C	-5.21	96.94	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	303	TYR	Sidechain
3	C	82	TYR	Sidechain

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	11140	0	11217	1180	0
2	B	8720	0	8745	919	0
3	C	2095	0	2051	244	0
4	D	1356	0	1319	114	0
5	E	1752	0	1776	154	0
6	F	679	0	701	84	0
7	G	1340	0	1357	150	0
8	H	1068	0	1040	104	0
9	I	971	0	927	94	0
10	J	532	0	542	93	0
11	K	920	0	929	83	0
12	L	363	0	386	45	0
13	A	2	0	0	0	0
13	B	1	0	0	0	0
13	C	1	0	0	0	0
13	I	2	0	0	0	0
13	J	1	0	0	0	0
13	L	1	0	0	0	0
14	A	1	0	0	0	0
All	All	30945	0	30990	2984	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 48.

The worst 5 of 2984 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:A:77:CYS:O	1:A:78:PRO:O	1.65	1.14
10:J:5:VAL:HG12	10:J:6:ARG:HG3	1.30	1.11
7:G:138:THR:HG22	7:G:139:ILE:H	1.12	1.09
1:A:53:LEU:HD23	1:A:54:ASN:N	1.69	1.06
4:D:40:HIS:HB3	7:G:73:LYS:NZ	1.69	1.06

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	1406/1733 (81%)	949 (68%)	293 (21%)	164 (12%)	0	6
2	B	1077/1224 (88%)	735 (68%)	221 (20%)	121 (11%)	0	7
3	C	264/318 (83%)	159 (60%)	66 (25%)	39 (15%)	0	4
4	D	173/177 (98%)	122 (70%)	34 (20%)	17 (10%)	0	10
5	E	212/215 (99%)	148 (70%)	49 (23%)	15 (7%)	1	17
6	F	82/155 (53%)	64 (78%)	14 (17%)	4 (5%)	2	24
7	G	169/171 (99%)	131 (78%)	26 (15%)	12 (7%)	1	17
8	H	129/146 (88%)	84 (65%)	29 (22%)	16 (12%)	0	6
9	I	117/122 (96%)	80 (68%)	29 (25%)	8 (7%)	1	18
10	J	63/70 (90%)	37 (59%)	10 (16%)	16 (25%)	0	1
11	K	113/120 (94%)	89 (79%)	18 (16%)	6 (5%)	2	23
12	L	44/70 (63%)	19 (43%)	14 (32%)	11 (25%)	0	1
All	All	3849/4521 (85%)	2617 (68%)	803 (21%)	429 (11%)	0	7

5 of 429 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	5	GLN
1	A	48	ALA
1	A	54	ASN
1	A	55	ASP
1	A	57	ARG

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	1239/1520 (82%)	1135 (92%)	104 (8%)	11	40
2	B	952/1061 (90%)	866 (91%)	86 (9%)	9	37
3	C	234/274 (85%)	212 (91%)	22 (9%)	8	35
4	D	140/159 (88%)	124 (89%)	16 (11%)	5	28
5	E	196/197 (100%)	187 (95%)	9 (5%)	27	56
6	F	74/137 (54%)	65 (88%)	9 (12%)	5	25
7	G	152/152 (100%)	142 (93%)	10 (7%)	16	48
8	H	117/128 (91%)	111 (95%)	6 (5%)	24	54
9	I	113/116 (97%)	99 (88%)	14 (12%)	4	24
10	J	60/65 (92%)	54 (90%)	6 (10%)	7	32
11	K	99/102 (97%)	92 (93%)	7 (7%)	14	45
12	L	40/57 (70%)	37 (92%)	3 (8%)	13	44
All	All	3416/3968 (86%)	3124 (92%)	292 (8%)	10	40

5 of 292 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	485	ARG
2	B	956	THR
9	I	85	PHE
2	B	516	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	701	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 78 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	236	HIS
2	B	734	HIS
9	I	12	ASN
2	B	363	HIS
2	B	484	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 9 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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
4	D	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	76:LYS	C	118:THR	N	35.50

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	1416/1733 (81%)	-0.37	10 (0%) 87 83	19, 87, 162, 200	0
2	B	1097/1224 (89%)	-0.33	8 (0%) 87 83	23, 97, 166, 194	0
3	C	266/318 (83%)	-0.42	0 100 100	37, 81, 139, 160	0
4	D	177/177 (100%)	-0.36	0 100 100	52, 108, 147, 165	0
5	E	214/215 (99%)	-0.34	1 (0%) 91 87	57, 142, 187, 193	0
6	F	84/155 (54%)	-0.65	0 100 100	25, 59, 105, 124	0
7	G	171/171 (100%)	-0.31	0 100 100	57, 84, 114, 138	0
8	H	133/146 (91%)	0.04	1 (0%) 86 81	101, 139, 175, 184	0
9	I	119/122 (97%)	-0.11	2 (1%) 70 62	74, 130, 159, 200	0
10	J	65/70 (92%)	-0.60	0 100 100	42, 79, 120, 127	0
11	K	115/120 (95%)	-0.34	0 100 100	42, 83, 114, 123	0
12	L	46/70 (65%)	-0.24	1 (2%) 62 54	76, 137, 168, 177	0
All	All	3903/4521 (86%)	-0.34	23 (0%) 89 85	19, 94, 166, 200	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1176	LEU	7.7
2	B	882	THR	4.2
9	I	119	THR	3.3
2	B	92	PHE	3.2
2	B	919	SER	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 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. 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
13	ZN	I	1122	1/1	0.91	0.06	156,156,156,156	0
13	ZN	A	2456	1/1	0.97	0.07	86,86,86,86	0
14	MG	A	2458	1/1	0.98	0.21	56,56,56,56	0
13	ZN	I	1121	1/1	0.99	0.14	90,90,90,90	0
13	ZN	L	1071	1/1	0.99	0.12	115,115,115,115	0
13	ZN	B	2225	1/1	0.99	0.16	44,44,44,44	0
13	ZN	J	1066	1/1	0.99	0.16	65,65,65,65	0
13	ZN	C	1269	1/1	1.00	0.08	39,39,39,39	0
13	ZN	A	2457	1/1	1.00	0.10	44,44,44,44	0

6.5 Other polymers [i](#)

There are no such residues in this entry.