



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 28, 2020 – 07:40 pm BST

PDB ID : 1I50  
Title : RNA POLYMERASE II CRYSTAL FORM II AT 2.8 Å RESOLUTION  
Authors : Cramer, P.; Bushnell, D.A.; Kornberg, R.D.  
Deposited on : 2001-02-23  
Resolution : 2.80 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](mailto: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) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

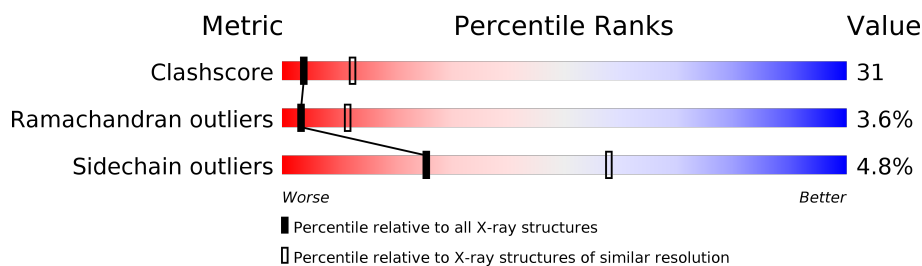
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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))
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)

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  $\geq 3$ , 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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	1733	44% 33% 5% 18%
2	B	1224	49% 37% • 11%
3	C	318	44% 36% • 16%
4	E	215	55% 42% •
5	F	155	27% 26% • 46%
6	H	146	32% 49% 8% • 9%
7	I	122	52% 43% 5%
8	J	70	50% 40% • 7%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	K	120	<div><div></div><div>52%</div><div>39%</div><div>• 5%</div></div>
10	L	70	<div><div></div><div>16%</div><div>37%</div><div>11%</div><div>•</div><div>34%</div></div>

## 2 Entry composition [i](#)

There are 13 unique types of molecules in this entry. The entry contains 28366 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 SUB-UNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1419	Total	C	N	O	S	0	0	0
			11154	7023	1952	2118	61			

- Molecule 2 is a protein called DNA-DIRECTED RNA POLYMERASE II 140KD POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1094	Total	C	N	O	S	0	0	0
			8711	5525	1519	1614	53			

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

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

- Molecule 4 is a protein called DNA-DIRECTED RNA POLYMERASE II 27KD POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	215	Total	C	N	O	S	0	0	0
			1760	1116	310	322	12			

- Molecule 5 is a protein called DNA-DIRECTED RNA POLYMERASE II 23KD POLYPEPTIDE.

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

- Molecule 6 is a protein called DNA-DIRECTED RNA POLYMERASE II 14.5KD POLYPEPTIDE.

TIDE.

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

- Molecule 7 is a protein called DNA-DIRECTED RNA POLYMERASE II 14.2KD POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	I	122	Total	C	N	O	S	0	0	0
			997	613	182	191	11			

- Molecule 8 is a protein called DNA-DIRECTED RNA POLYMERASE II 8.3KD POLYPEPTIDE.

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

- Molecule 9 is a protein called DNA-DIRECTED RNA POLYMERASE II 13.6KD POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	K	114	Total	C	N	O	S	0	0	0
			919	590	156	171	2			

- Molecule 10 is a protein called DNA-DIRECTED RNA POLYMERASE II 7.7KD POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	L	46	Total	C	N	O	S	0	0	0
			364	224	72	64	4			

- Molecule 11 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	J	1	Total	Zn	0	0
			1	1		
11	B	1	Total	Zn	0	0
			1	1		
11	I	2	Total	Zn	0	0
			2	2		
11	C	1	Total	Zn	0	0
			1	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	2	Total 2	Zn 2	0	0
11	L	1	Total 1	Zn 1	0	0

- Molecule 12 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	A	1	Total 1	Mn 1	0	0

- Molecule 13 is water.

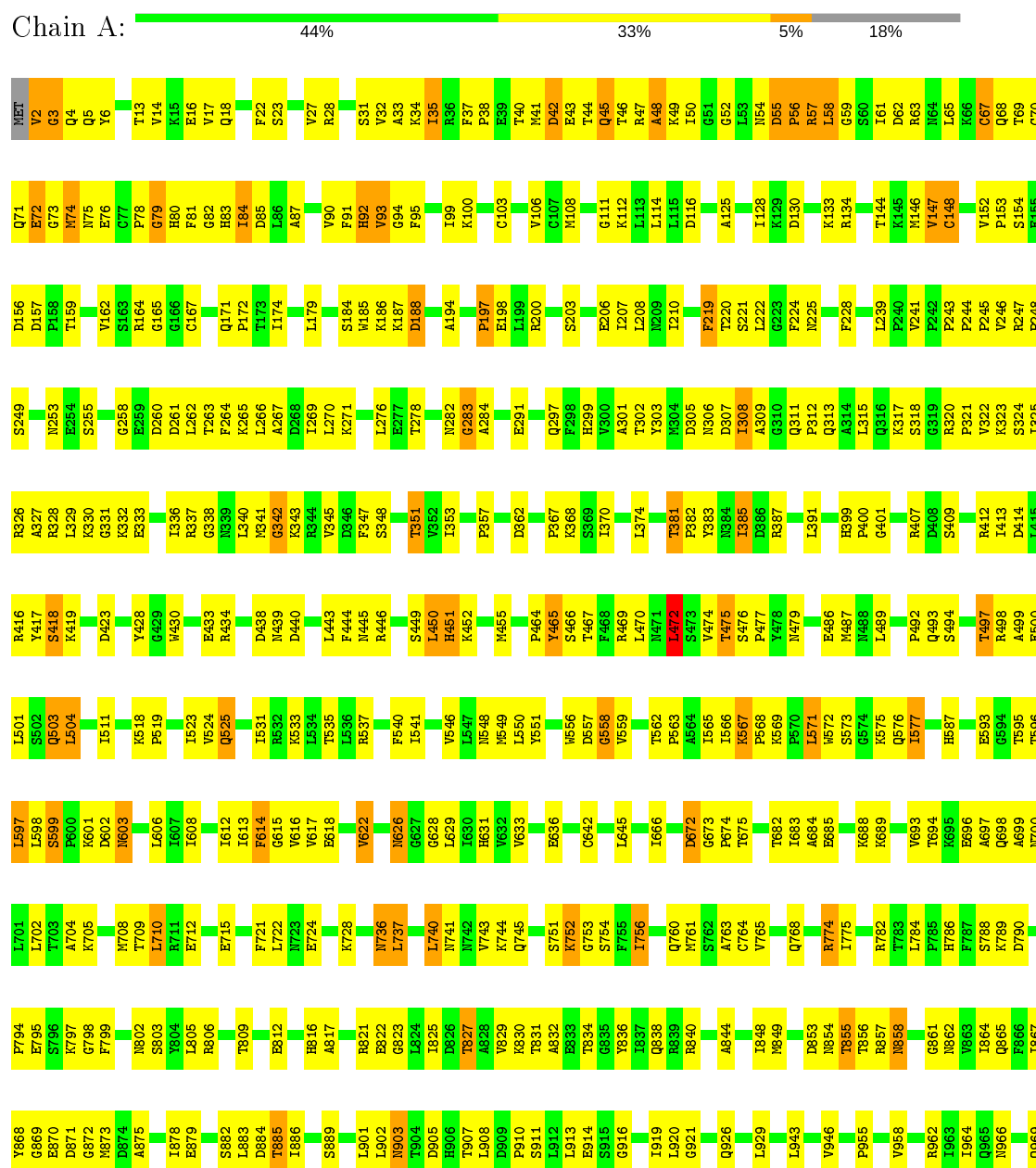
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	A	35	Total 35	O 35	0	0
13	B	23	Total 23	O 23	0	0
13	C	4	Total 4	O 4	0	0
13	E	7	Total 7	O 7	0	0
13	F	5	Total 5	O 5	0	0
13	I	2	Total 2	O 2	0	0
13	K	2	Total 2	O 2	0	0

### 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. 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. 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.

Note EDS was not executed.

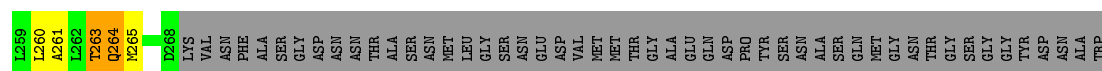
#### • Molecule 1: DNA-DIRECTED RNA POLYMERASE II LARGEST SUBUNIT





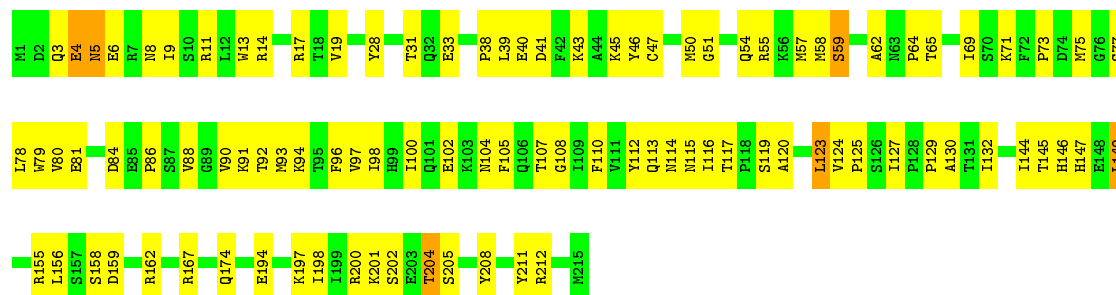






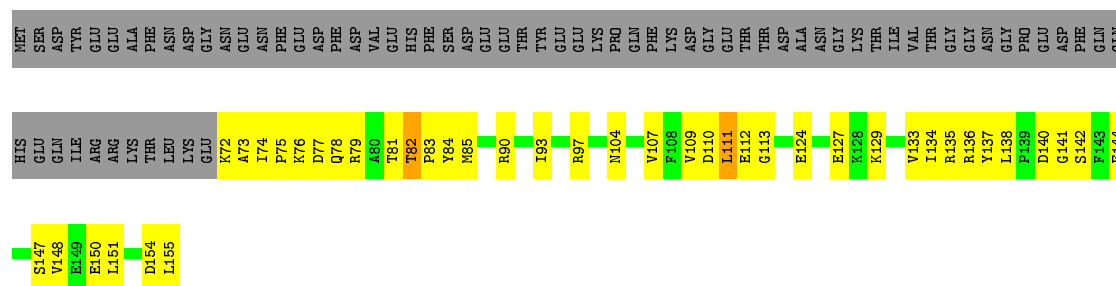
• Molecule 4: DNA-DIRECTED RNA POLYMERASE II 27KD POLYPEPTIDE

Chain E: 55% 42%



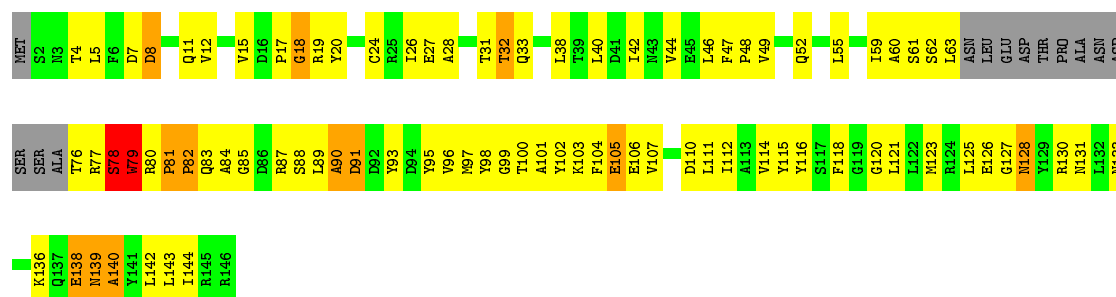
• Molecule 5: DNA-DIRECTED RNA POLYMERASE II 23KD POLYPEPTIDE

Chain F: 27% 26% 46%



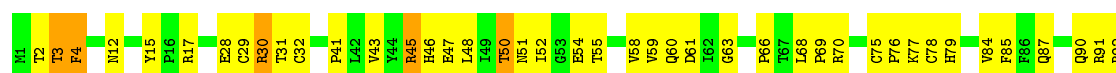
• Molecule 6: DNA-DIRECTED RNA POLYMERASE II 14.5KD POLYPEPTIDE

Chain H: 32% 49% 8% 9%



• Molecule 7: DNA-DIRECTED RNA POLYMERASE II 14.2KD POLYPEPTIDE

Chain I: 52% 43% 5%

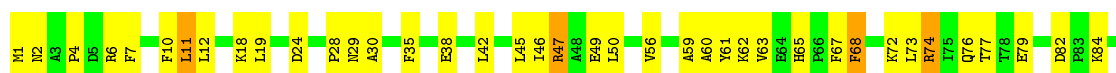




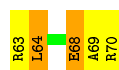
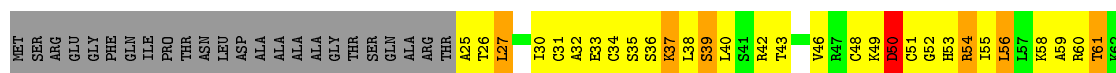
- Molecule 8: DNA-DIRECTED RNA POLYMERASE II 8.3KD POLYPEPTIDE



- Molecule 9: DNA-DIRECTED RNA POLYMERASE II 13.6KD POLYPEPTIDE



- Molecule 10: DNA-DIRECTED RNA POLYMERASE II 7.7KD POLYPEPTIDE



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	122.70 Å   223.00 Å   376.10 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	40.00 – 2.80	Depositor
% Data completeness (in resolution range)	(Not available) (40.00-2.80)	Depositor
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS	Depositor
R, $R_{free}$	0.229 , 0.282	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	28366	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	64.0	wwPDB-VP

## 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, MN

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	0.39	1/11352 (0.0%)	0.68	2/15352 (0.0%)
2	B	0.42	2/8882 (0.0%)	0.70	5/11976 (0.0%)
3	C	0.37	0/2133	0.65	0/2891
4	E	0.37	0/1796	0.64	0/2416
5	F	0.40	0/691	0.63	0/933
6	H	0.88	1/1086 (0.1%)	1.23	3/1470 (0.2%)
7	I	0.40	0/1016	0.63	0/1365
8	J	0.41	0/541	0.70	0/727
9	K	0.38	0/937	0.60	0/1265
10	L	0.42	0/366	0.72	0/485
All	All	0.43	4/28800 (0.0%)	0.71	10/38880 (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
2	B	0	2
6	H	0	3
All	All	0	5

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	H	79	TRP	C-N	-26.78	0.72	1.34
2	B	707	PRO	C-N	-11.27	1.08	1.34
2	B	710	LEU	C-N	-7.44	1.17	1.34
1	A	3	GLY	C-N	-5.20	1.22	1.34

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	H	78	SER	O-C-N	-31.77	71.87	122.70
6	H	78	SER	CA-C-N	20.38	162.03	117.20
2	B	707	PRO	O-C-N	-13.53	101.05	122.70
6	H	78	SER	C-N-CA	12.14	152.06	121.70
1	A	472	LEU	CA-CB-CG	-6.84	99.57	115.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	707	PRO	Mainchain
2	B	710	LEU	Mainchain
6	H	78	SER	Mainchain,Peptide
6	H	79	TRP	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	11154	0	11222	754	0
2	B	8711	0	8734	539	0
3	C	2095	0	2051	141	0
4	E	1760	0	1788	87	0
5	F	679	0	701	50	0
6	H	1068	0	1038	121	0
7	I	997	0	956	64	0
8	J	532	0	542	47	0
9	K	919	0	929	65	0
10	L	364	0	387	68	0
11	A	2	0	0	0	0
11	B	1	0	0	0	0
11	C	1	0	0	0	0
11	I	2	0	0	0	0
11	J	1	0	0	0	0
11	L	1	0	0	0	0
12	A	1	0	0	0	0
13	A	35	0	0	2	0
13	B	23	0	0	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
13	C	4	0	0	0	0
13	E	7	0	0	1	0
13	F	5	0	0	0	0
13	I	2	0	0	0	0
13	K	2	0	0	0	0
All	All	28366	0	28348	1743	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 31.

The worst 5 of 1743 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)
6:H:79:TRP:C	6:H:80:ARG:CA	1.90	1.41
2:B:708:GLU:O	2:B:711:GLU:HG3	1.20	1.34
6:H:79:TRP:CA	6:H:80:ARG:N	1.94	1.29
6:H:79:TRP:O	6:H:80:ARG:N	1.67	1.26
3:C:56:THR:HG23	3:C:58:LEU:H	1.14	1.13

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	1411/1733 (81%)	1203 (85%)	149 (11%)	59 (4%)	3	9
2	B	1074/1224 (88%)	953 (89%)	100 (9%)	21 (2%)	7	24
3	C	264/318 (83%)	228 (86%)	30 (11%)	6 (2%)	6	21
4	E	213/215 (99%)	184 (86%)	24 (11%)	5 (2%)	6	21
5	F	82/155 (53%)	74 (90%)	6 (7%)	2 (2%)	6	20

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
6	H	129/146 (88%)	83 (64%)	28 (22%)	18 (14%)	0	0
7	I	120/122 (98%)	103 (86%)	14 (12%)	3 (2%)	5	19
8	J	63/70 (90%)	58 (92%)	4 (6%)	1 (2%)	9	31
9	K	112/120 (93%)	101 (90%)	11 (10%)	0	100	100
10	L	44/70 (63%)	21 (48%)	13 (30%)	10 (23%)	0	0
All	All	3512/4173 (84%)	3008 (86%)	379 (11%)	125 (4%)	3	11

5 of 125 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	4	GLN
1	A	46	THR
1	A	48	ALA
1	A	55	ASP
1	A	56	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	1239/1520 (82%)	1172 (95%)	67 (5%)	22	53
2	B	950/1061 (90%)	908 (96%)	42 (4%)	28	61
3	C	234/274 (85%)	222 (95%)	12 (5%)	24	55
4	E	197/197 (100%)	191 (97%)	6 (3%)	41	75
5	F	74/137 (54%)	72 (97%)	2 (3%)	44	78
6	H	117/128 (91%)	116 (99%)	1 (1%)	78	94
7	I	116/116 (100%)	109 (94%)	7 (6%)	19	48
8	J	60/65 (92%)	56 (93%)	4 (7%)	16	43
9	K	99/102 (97%)	93 (94%)	6 (6%)	18	48
10	L	40/57 (70%)	36 (90%)	4 (10%)	7	22
All	All	3126/3657 (86%)	2975 (95%)	151 (5%)	25	58



5 of 151 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	20	ASP
2	B	485	ARG
8	J	28	ASP
2	B	120	ARG
2	B	217	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 85 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	300	HIS
2	B	648	HIS
7	I	12	ASN
2	B	366	GLN
2	B	516	ASN

### 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 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
2	B	2
6	H	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	710:LEU	C	711:GLU	N	1.16
1	B	707:PRO	C	708:GLU	N	1.08
1	H	79:TRP	C	80:ARG	N	0.72

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.