



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 2, 2023 – 08:17 PM EDT

PDB ID : 1I6H  
Title : RNA POLYMERASE II ELONGATION COMPLEX  
Authors : Gnatt, A.L.; Cramer, P.; Kornberg, R.D.  
Deposited on : 2001-03-02  
Resolution : 3.30 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 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 : 4.02b-467  
Xtrriage (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) : 2.34

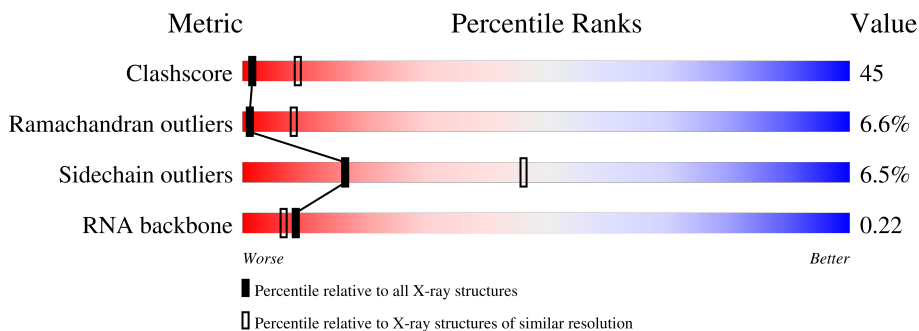
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.30 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	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RNA backbone	3102	1117 (3.70-2.90)

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  $\geq 3$ , 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	D	13	15% (green) 69% (yellow) 15% (orange)
2	R	9	11% (green) 56% (yellow) 11% (orange) 22% (red)
3	A	1733	31% (green) 41% (yellow) 7% (orange) 20% (grey)
4	B	1224	32% (green) 51% (yellow) 7% (orange) 10% (grey)
5	C	318	37% (green) 40% (yellow) 7% (orange) 16% (grey)
6	E	215	41% (green) 55% (yellow) 4% (orange) 0% (red) 0% (grey)

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
7	F	155	
8	H	146	
9	I	122	
10	J	70	
11	K	120	
12	L	70	

## 2 Entry composition [i](#)

There are 14 unique types of molecules in this entry. The entry contains 28430 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 DNA chain called 5'-D(P\*AP\*AP\*AP\*TP\*GP\*CP\*CP\*TP\*GP\*GP\*TP\*CP\*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	D	13	267	127	47	80	13	0	0	0

- Molecule 2 is a RNA chain called 5'-R(P\*GP\*AP\*CP\*CP\*AP\*GP\*GP\*CP\*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	R	9	196	87	39	61	9	0	0	0

- Molecule 3 is a protein called DNA-DIRECTED RNA POLYMERASE II LARGEST SUB-UNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	1381	10857	6851	1899	2046	61	0	0	0

- Molecule 4 is a protein called DNA-DIRECTED RNA POLYMERASE II 140KD POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	B	1097	8721	5526	1523	1618	54	0	0	0

- Molecule 5 is a protein called DNA-DIRECTED RNA POLYMERASE II 45KD POLYPEPTIDE.

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

- Molecule 6 is a protein called DNA-DIRECTED RNA POLYMERASE II 27KD POLYPEPTIDE.

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

- Molecule 7 is a protein called DNA-DIRECTED RNA POLYMERASE II 23KD POLYPEPTIDE.

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

- Molecule 8 is a protein called DNA-DIRECTED RNA POLYMERASE II 14.5KD 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 14.2KD 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 POLYMERASE II 8.3KD 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 13.6KD POLYPEPTIDE.

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

- Molecule 12 is a protein called DNA-DIRECTED RNA POLYMERASE II 7.7KD POLYPEPTIDE.

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

- Molecule 13 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	R	1	Total	Mg	0	0
			1	1		

- Molecule 14 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

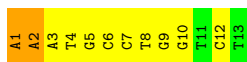
### 3 Residue-property plots [i](#)

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. 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: 5'-D(P\*AP\*AP\*AP\*TP\*GP\*CP\*CP\*TP\*GP\*GP\*TP\*CP\*T)-3'

Chain D: 



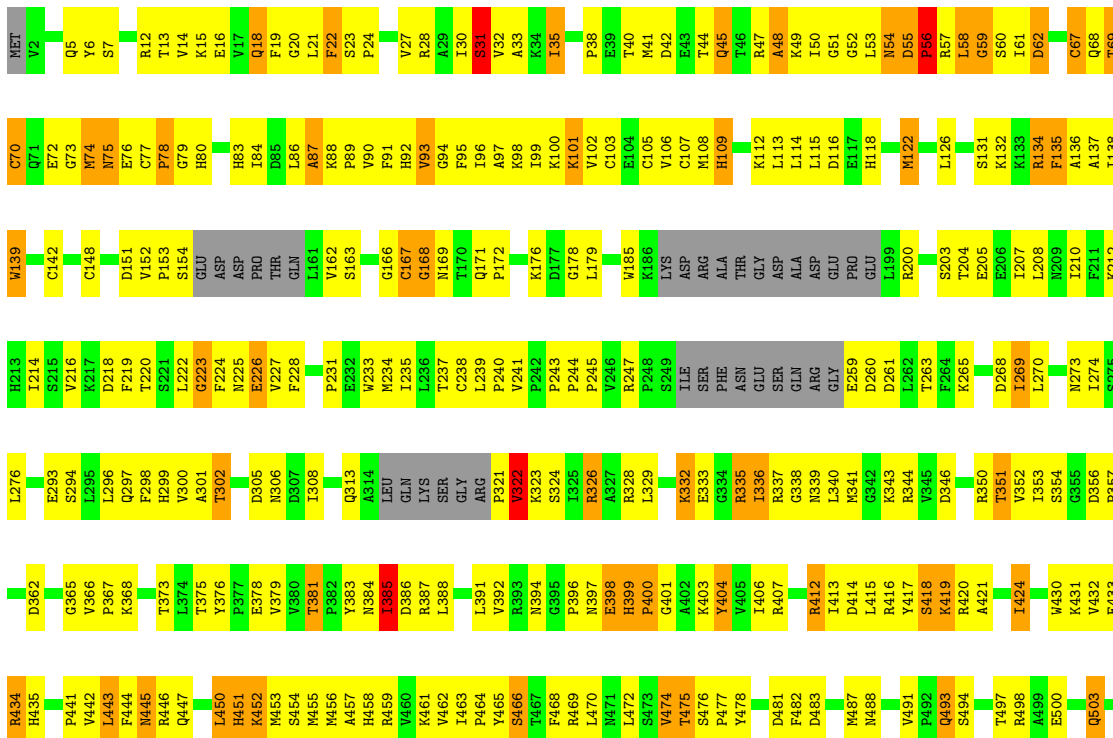
- Molecule 2: 5'-R(P\*GP\*AP\*CP\*CP\*AP\*GP\*GP\*CP\*A)-3'

Chain R: 



- Molecule 3: DNA-DIRECTED RNA POLYMERASE II LARGEST SUBUNIT

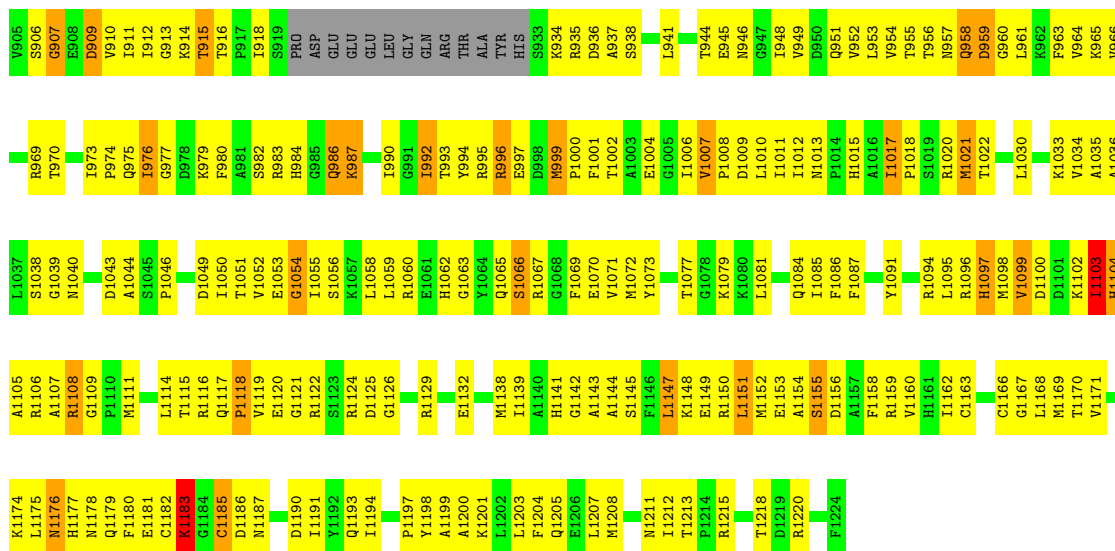
Chain A: 



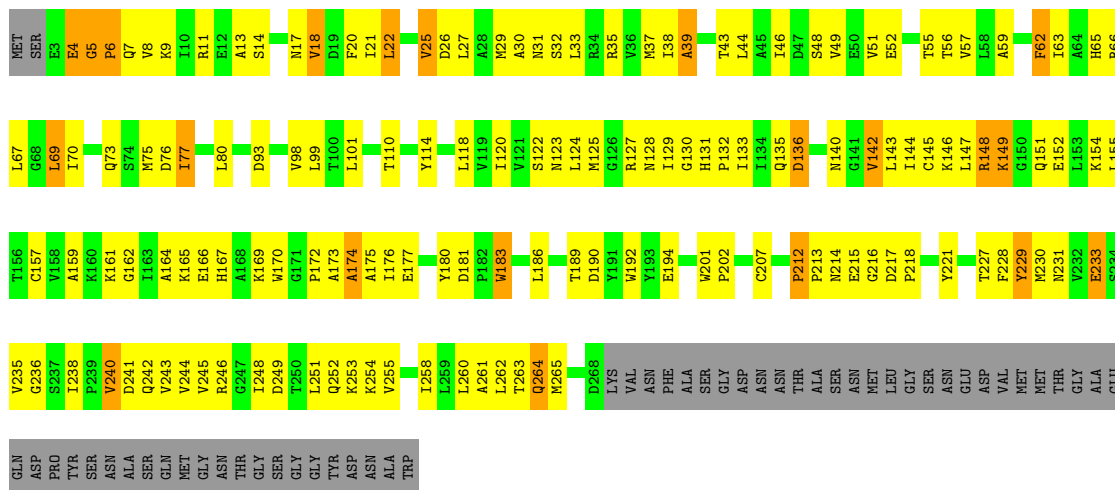




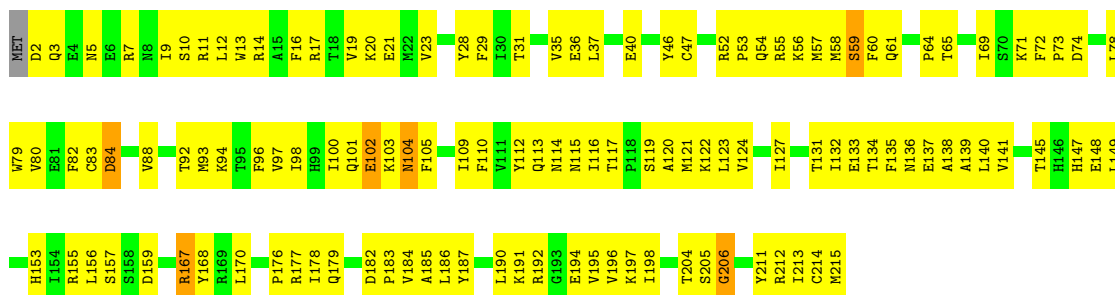





• Molecule 5: DNA-DIRECTED RNA POLYMERASE II 45KD POLYPEPTIDE

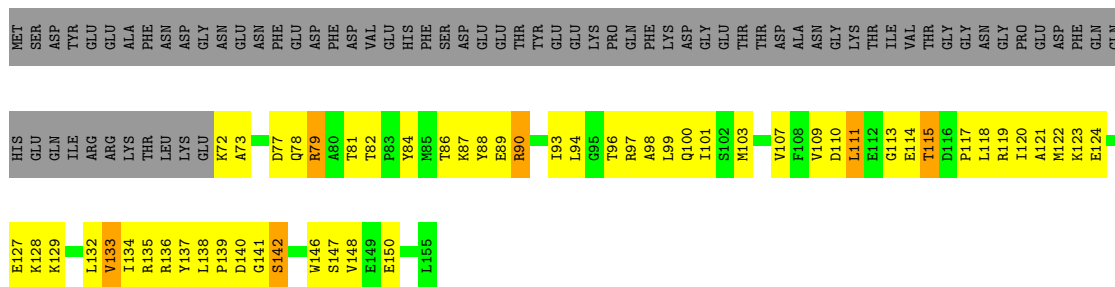


• Molecule 6: DNA-DIRECTED RNA POLYMERASE II 27KD POLYPEPTIDE



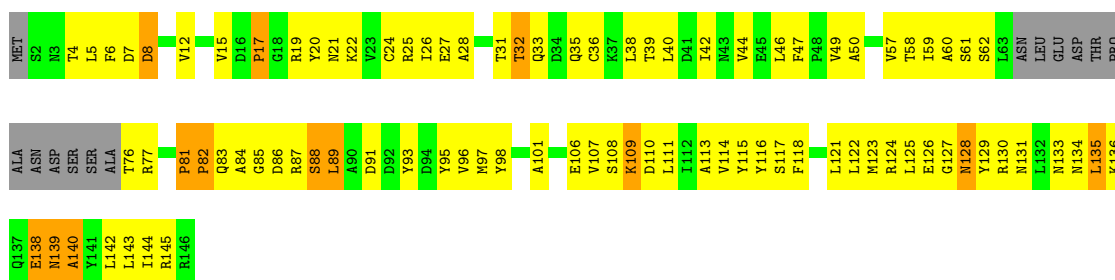
• Molecule 7: DNA-DIRECTED RNA POLYMERASE II 23KD POLYPEPTIDE

Chain F:  19% 32% 46%



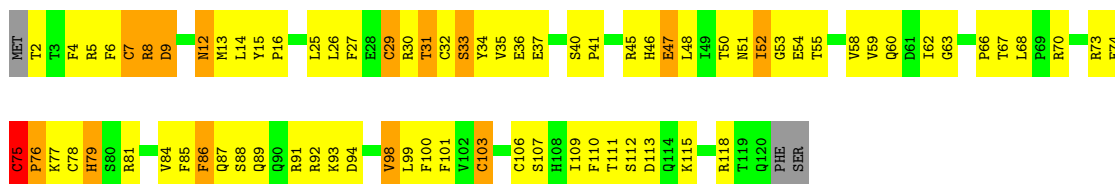
• Molecule 8: DNA-DIRECTED RNA POLYMERASE II 14.5KD POLYPEPTIDE

Chain H:  30% 52% 9% 9%




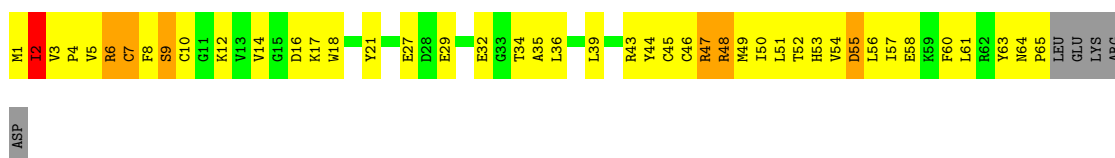
• Molecule 9: DNA-DIRECTED RNA POLYMERASE II 14.2KD POLYPEPTIDE

Chain I:  34% 51% 11% ..

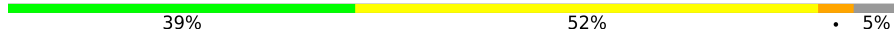


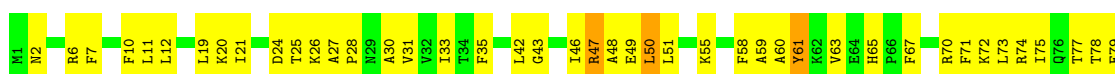
• Molecule 10: DNA-DIRECTED RNA POLYMERASE II 8.3KD POLYPEPTIDE

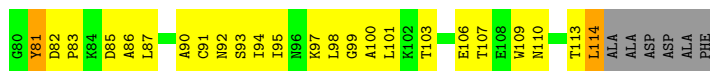
Chain J:  30% 53% 9% 7%



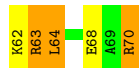
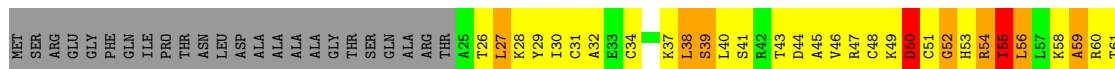
• Molecule 11: DNA-DIRECTED RNA POLYMERASE II 13.6KD POLYPEPTIDE

Chain K:  39% 52% 5% 4%





- Molecule 12: DNA-DIRECTED RNA POLYMERASE II 7.7KD POLYPEPTIDE



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	157.30Å 220.70Å 191.30Å 90.00° 97.50° 90.00°	Depositor
Resolution (Å)	40.00 – 3.30	Depositor
% Data completeness (in resolution range)	(Not available) (40.00-3.30)	Depositor
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS	Depositor
R, $R_{free}$	0.250 , 0.298	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	28430	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	63.0	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: 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	D	1.10	1/298 (0.3%)	1.03	0/456
2	R	1.70	4/219 (1.8%)	1.77	9/338 (2.7%)
3	A	0.41	0/11048	0.71	6/14936 (0.0%)
4	B	0.46	0/8891	0.71	2/11990 (0.0%)
5	C	0.48	0/2133	0.76	2/2891 (0.1%)
6	E	0.36	0/1788	0.65	0/2406
7	F	0.40	0/691	0.64	0/933
8	H	0.40	0/1086	0.73	0/1470
9	I	0.47	0/989	0.76	1/1331 (0.1%)
10	J	0.53	0/541	0.78	0/727
11	K	0.45	0/937	0.68	0/1265
12	L	0.48	0/366	0.78	0/485
All	All	0.47	5/28987 (0.0%)	0.73	20/39228 (0.1%)

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	D	0	1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	R	9	A	P-OP2	7.91	1.62	1.49
2	R	1	G	OP3-P	-7.91	1.51	1.61
1	D	1	DA	OP3-P	-6.84	1.52	1.61
2	R	9	A	P-OP1	6.39	1.59	1.49
2	R	9	A	C2'-O2'	5.17	1.48	1.41

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	R	9	A	N9-C1'-C2'	-14.41	95.27	114.00
2	R	3	C	O5'-P-OP2	12.38	125.55	110.70
3	A	1392	SER	N-CA-C	6.34	128.11	111.00
2	R	3	C	C5'-C4'-O4'	6.26	116.61	109.10
2	R	9	A	C1'-O4'-C4'	-6.17	104.97	109.90

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	2	DA	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	D	267	0	148	30	0
2	R	196	0	100	26	0
3	A	10857	0	10959	1050	0
4	B	8721	0	8746	897	0
5	C	2095	0	2052	167	0
6	E	1752	0	1776	130	0
7	F	679	0	701	62	0
8	H	1068	0	1040	136	0
9	I	971	0	933	101	0
10	J	532	0	544	77	0
11	K	919	0	929	87	0
12	L	364	0	388	55	0
13	R	1	0	0	0	0
14	A	2	0	0	0	0
14	B	1	0	0	0	0
14	C	1	0	0	0	0
14	I	2	0	0	1	0
14	J	1	0	0	1	0
14	L	1	0	0	0	0
All	All	28430	0	28316	2555	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 45.

The worst 5 of 2555 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:A:567:LYS:HB2	3:A:568:PRO:HD2	1.18	1.17
12:L:60:ARG:HG3	12:L:61:THR:H	1.05	1.15
9:I:111:THR:HG22	9:I:113:ASP:H	1.05	1.13
4:B:345:LYS:HA	4:B:348:ARG:HE	1.11	1.13
3:A:1364:ASN:ND2	3:A:1366:ARG:HG2	1.65	1.11

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	
3	A	1365/1733 (79%)	1021 (75%)	253 (18%)	91 (7%)	1	8
4	B	1077/1224 (88%)	839 (78%)	169 (16%)	69 (6%)	1	9
5	C	264/318 (83%)	208 (79%)	39 (15%)	17 (6%)	1	9
6	E	212/215 (99%)	170 (80%)	33 (16%)	9 (4%)	3	17
7	F	82/155 (53%)	64 (78%)	15 (18%)	3 (4%)	3	20
8	H	129/146 (88%)	93 (72%)	21 (16%)	15 (12%)	0	2
9	I	117/122 (96%)	93 (80%)	15 (13%)	9 (8%)	1	6
10	J	63/70 (90%)	48 (76%)	11 (18%)	4 (6%)	1	9
11	K	112/120 (93%)	96 (86%)	15 (13%)	1 (1%)	17	48
12	L	44/70 (63%)	22 (50%)	12 (27%)	10 (23%)	0	0
All	All	3465/4173 (83%)	2654 (77%)	583 (17%)	228 (7%)	1	8

5 of 228 Ramachandran outliers are listed below:



Mol	Chain	Res	Type
3	A	31	SER
3	A	48	ALA
3	A	55	ASP
3	A	56	PRO
3	A	74	MET

### 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	
3	A	1206/1520 (79%)	1129 (94%)	77 (6%)	17	46
4	B	952/1061 (90%)	886 (93%)	66 (7%)	15	44
5	C	234/274 (85%)	222 (95%)	12 (5%)	24	54
6	E	196/197 (100%)	189 (96%)	7 (4%)	35	63
7	F	74/137 (54%)	68 (92%)	6 (8%)	11	36
8	H	117/128 (91%)	112 (96%)	5 (4%)	29	59
9	I	113/116 (97%)	104 (92%)	9 (8%)	12	37
10	J	60/65 (92%)	56 (93%)	4 (7%)	16	45
11	K	99/102 (97%)	90 (91%)	9 (9%)	9	31
12	L	40/57 (70%)	35 (88%)	5 (12%)	4	19
All	All	3091/3657 (84%)	2891 (94%)	200 (6%)	17	46

5 of 200 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	B	762	ASN
5	C	22	LEU
12	L	70	ARG
4	B	835	GLN
4	B	999	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 75 such sidechains are listed below:

Mol	Chain	Res	Type
5	C	112	ASN
11	K	52	ASN
5	C	167	HIS
6	E	104	ASN
3	A	994	GLN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	R	8/9 (88%)	2 (25%)	0

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	R	3	C
2	R	9	A

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 monosaccharides 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](#)

There are no chain breaks in this entry.

## 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.