



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 3, 2019 – 09:55 PM EST

PDB ID : 1Y1V
Title : Refined RNA Polymerase II-TFIIS complex
Authors : Kettenberger, H.; Armache, K.-J.; Cramer, P.
Deposited on : 2004-11-19
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	:	2.4
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.4

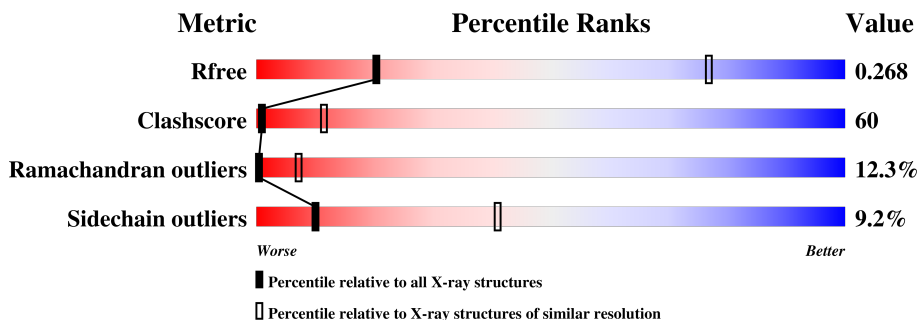
1 Overall quality at a glance ⓘ

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}	111664	1028 (4.02-3.58)
Clashscore	122126	1061 (4.00-3.60)
Ramachandran outliers	120053	1025 (4.00-3.60)
Sidechain outliers	120020	1019 (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. 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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	1733	21% 47% 13% • 18%
2	B	1224	24% 52% 13% • 9%
3	C	318	22% 48% 12% • 16%
4	D	221	33% 40% 6% 20%
5	E	215	34% 55% 11%
6	F	155	15% 30% 8% 46%
7	G	171	33% 59% 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	H	146	<div><div></div><div>30%53%8%9%</div></div>
9	I	122	<div><div></div><div>31%48%13%5%•</div></div>
10	J	70	<div><div></div><div>23%51%17%•7%</div></div>
11	K	120	<div><div></div><div>37%53%5%•5%</div></div>
12	L	70	<div><div></div><div>6%41%19%34%</div></div>
13	S	179	<div><div></div><div>62%28%•••</div></div>

2 Entry composition

There are 15 unique types of molecules in this entry. The entry contains 31803 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
1	A	1426	Total	C	N	O	S	0	0	0
			11214	7069	1959	2124	62			

- Molecule 2 is a protein called DNA-directed RNA polymerase II 140 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1112	Total	C	N	O	S	58	0	0
			8837	5594	1548	1640	55			

- Molecule 3 is a protein called DNA-directed RNA polymerase II 45 kDa 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 32 kDa polypeptide.

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

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide.

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

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide.

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

- Molecule 7 is a protein called DNA-directed RNA polymerase II 19 kDa polypeptide.

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

- 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
8	H	133	Total	C	N	O	S	0	0	0
			1068	673	180	211	4			

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit 9.

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

- Molecule 10 is a protein called DNA-directed RNA polymerases I/II/III subunit 10.

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

- Molecule 11 is a protein called DNA-directed RNA polymerase II 13.6 kDa polypeptide.

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

- 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
12	L	46	Total	C	N	O	S	0	0	0
			364	224	72	64	4			

- Molecule 13 is a protein called Transcription elongation factor S-II.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	S	174	Total 666	C 454	N 99	O 108	S 5	0	0	104

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

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

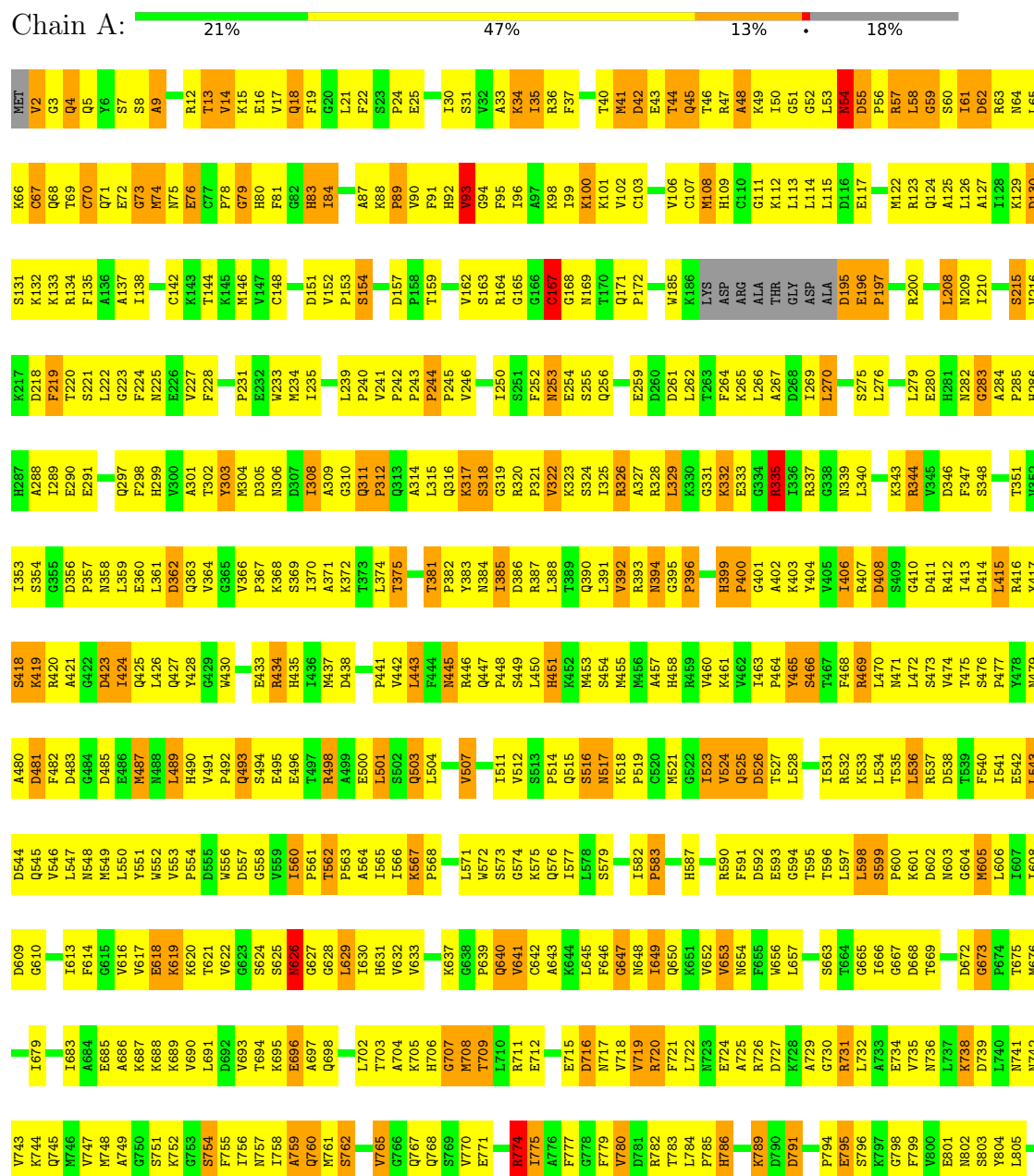
- Molecule 15 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	S	1	Total 1	Mg 1	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.

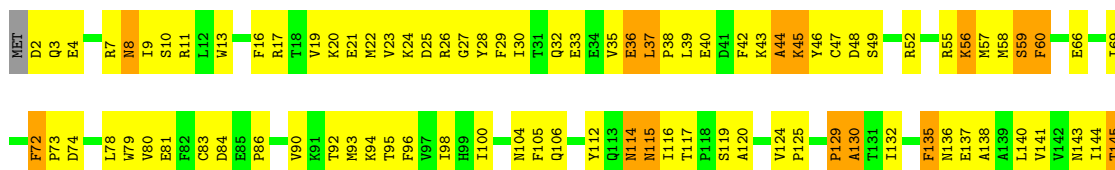
- Molecule 1: DNA-directed RNA polymerase II largest subunit



- Molecule 2: DNA-directed RNA polymerase II 140 kDa polypeptide



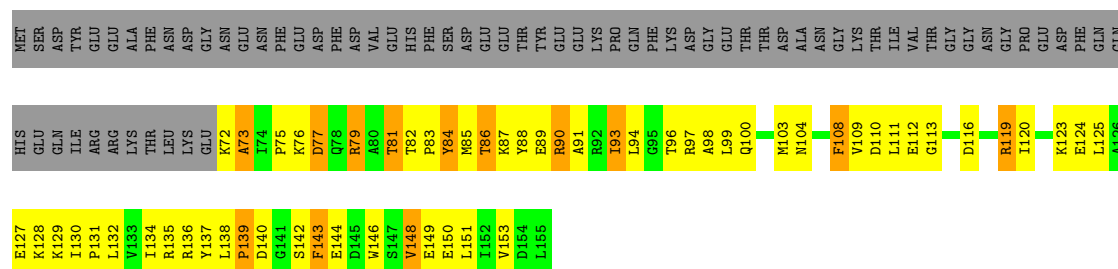






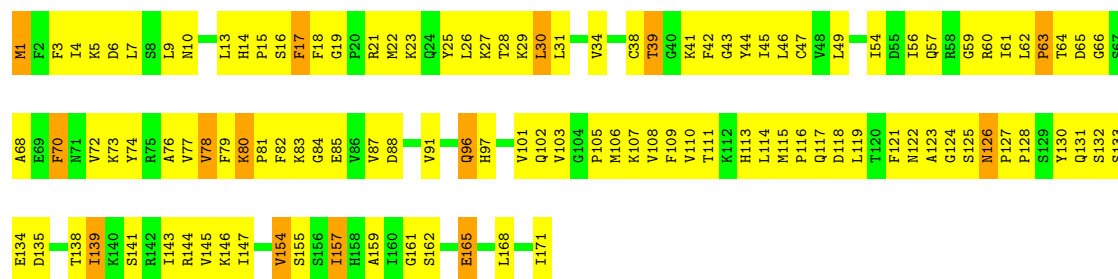
- Molecule 6: DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide

Chain F: 15% 30% 8% 46%



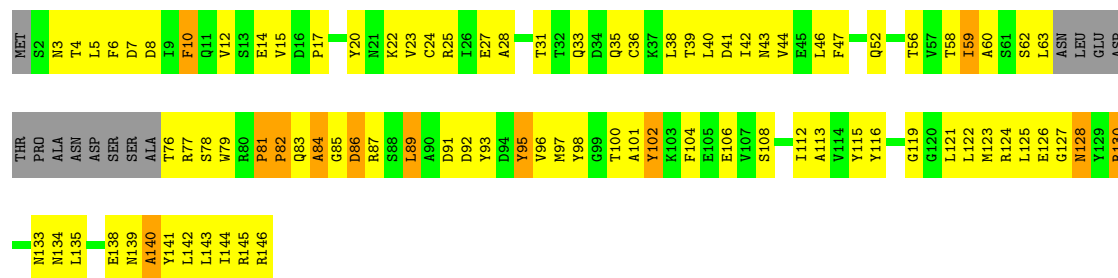
- Molecule 7: DNA-directed RNA polymerase II 19 kDa polypeptide

Chain G: 33% 59% 8%



- Molecule 8: DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide

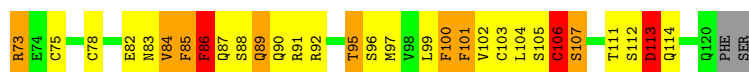
Chain H: 30% 53% 8% 9%



- Molecule 9: DNA-directed RNA polymerase II subunit 9

Chain I: 31% 48% 13% 5%





- Molecule 10: DNA-directed RNA polymerases I/II/III subunit 10

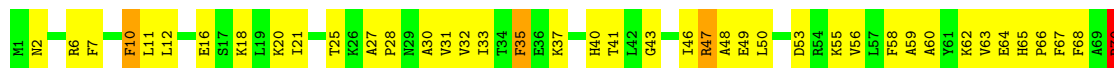
Chain J: 23% 51% 17% 7%



GLU
LYS
ARG
ASP

- Molecule 11: DNA-directed RNA polymerase II 13.6 kDa polypeptide

Chain K: 37% 53% 5% 5%



- Molecule 12: DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide

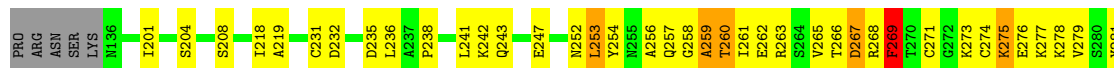
Chain L: 6% 41% 19% 34%



T61, K62, R63, L64, V65, Q66, F67, E68, A69, R70

- Molecule 13: Transcription elongation factor S-II

Chain S: 62% 28% 10% 2%



Y282, Q283, L284, Q285, T286, R287, S288, A289, D290, E291, P292, L293, T294, T295, F296, C297, T298, C299, E300, A301, C302, G303, N304, R305, K306, K307, F308, S309

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	218.90Å 395.30Å 281.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.80 49.41 – 3.74	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-3.80) 86.6 (49.41-3.74)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.52 (at 3.77Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.282 , 0.294 0.257 , 0.268	Depositor DCC
R_{free} test set	2439 reflections (1.99%)	wwPDB-VP
Wilson B-factor (Å ²)	68.0	Xtriage
Anisotropy	0.336	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 42.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.30$, $\langle L^2 \rangle = 0.13$	Xtriage
Estimated twinning fraction	0.199 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.206 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o, F_c correlation	0.81	EDS
Total number of atoms	31803	wwPDB-VP
Average B, all atoms (Å ²)	72.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.74% 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	0.50	2/11417 (0.0%)	0.78	4/15442 (0.0%)
2	B	0.51	4/9009 (0.0%)	0.76	8/12146 (0.1%)
3	C	0.48	0/2133	0.77	1/2891 (0.0%)
4	D	0.41	0/1365	0.64	0/1837
5	E	0.43	0/1788	0.66	0/2406
6	F	0.52	0/691	0.77	0/933
7	G	0.49	0/1368	0.72	0/1844
8	H	0.38	0/1086	0.65	1/1470 (0.1%)
9	I	0.46	0/989	0.77	1/1331 (0.1%)
10	J	0.48	0/541	0.75	0/727
11	K	0.45	0/937	0.67	0/1265
12	L	0.54	0/366	0.79	0/485
13	S	1.31	4/571 (0.7%)	1.64	7/765 (0.9%)
All	All	0.51	10/32261 (0.0%)	0.77	22/43542 (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
2	B	0	3
13	S	0	2
All	All	0	5

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
13	S	269	PHE	C-N	-16.91	0.95	1.34
2	B	467	GLY	C-O	-11.91	1.04	1.23
13	S	260	THR	CA-CB	10.48	1.80	1.53

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	468	GLU	CB-CG	8.39	1.68	1.52
13	S	268	ARG	CG-CD	6.05	1.67	1.51

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
13	S	269	PHE	O-C-N	-19.02	92.27	122.70
13	S	269	PHE	C-N-CA	16.73	163.51	121.70
13	S	269	PHE	CA-C-N	16.08	152.57	117.20
1	A	195	ASP	N-CA-C	9.35	136.25	111.00
2	B	510	LYS	CB-CA-C	-7.63	95.14	110.40

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	510	LYS	Mainchain
2	B	785	TYR	Sidechain
2	B	833	TYR	Sidechain
13	S	269	PHE	Sidechain,Peptide

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	11214	0	11281	1514	0
2	B	8837	0	8871	1206	0
3	C	2095	0	2052	260	0
4	D	1356	0	1319	101	0
5	E	1752	0	1776	200	0
6	F	679	0	701	82	0
7	G	1340	0	1357	159	0
8	H	1068	0	1040	115	0
9	I	971	0	929	110	0
10	J	532	0	542	103	0
11	K	919	0	929	96	0
12	L	364	0	387	68	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
13	S	666	0	553	105	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	0	0
14	J	1	0	0	0	0
14	L	1	0	0	0	0
14	S	1	0	0	0	0
15	S	1	0	0	0	0
All	All	31803	0	31737	3774	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 60.

The worst 5 of 3774 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)
13:S:260:THR:CA	13:S:260:THR:CB	1.80	1.53
13:S:269:PHE:CZ	13:S:297:CYS:SG	2.04	1.50
13:S:269:PHE:CE2	13:S:297:CYS:SG	2.14	1.39
1:A:1230:GLU:OE2	13:S:201:ILE:CA	1.75	1.32
1:A:1283:VAL:CG1	13:S:256:ALA:O	1.78	1.31

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	1418/1733 (82%)	914 (64%)	316 (22%)	188 (13%)	0 5
2	B	1096/1224 (90%)	726 (66%)	223 (20%)	147 (13%)	0 5

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	C	264/318 (83%)	169 (64%)	62 (24%)	33 (12%)	0	7
4	D	173/221 (78%)	129 (75%)	27 (16%)	17 (10%)	1	12
5	E	212/215 (99%)	141 (66%)	50 (24%)	21 (10%)	1	12
6	F	82/155 (53%)	60 (73%)	15 (18%)	7 (8%)	1	15
7	G	169/171 (99%)	123 (73%)	34 (20%)	12 (7%)	1	20
8	H	129/146 (88%)	93 (72%)	26 (20%)	10 (8%)	1	18
9	I	117/122 (96%)	80 (68%)	22 (19%)	15 (13%)	0	6
10	J	63/70 (90%)	36 (57%)	14 (22%)	13 (21%)	0	2
11	K	112/120 (93%)	82 (73%)	25 (22%)	5 (4%)	3	29
12	L	44/70 (63%)	18 (41%)	14 (32%)	12 (27%)	0	0
13	S	68/179 (38%)	51 (75%)	10 (15%)	7 (10%)	0	11
All	All	3947/4744 (83%)	2622 (66%)	838 (21%)	487 (12%)	0	7

5 of 487 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	4	GLN
1	A	48	ALA
1	A	55	ASP
1	A	58	LEU
1	A	62	ASP

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	1246/1520 (82%)	1133 (91%)	113 (9%)	10	42
2	B	964/1061 (91%)	880 (91%)	84 (9%)	11	44
3	C	234/274 (85%)	205 (88%)	29 (12%)	5	28
4	D	140/200 (70%)	126 (90%)	14 (10%)	8	37
5	E	196/197 (100%)	184 (94%)	12 (6%)	20	57

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
6	F	74/137 (54%)	63 (85%)	11 (15%)	3	22
7	G	152/152 (100%)	143 (94%)	9 (6%)	21	58
8	H	117/128 (91%)	110 (94%)	7 (6%)	21	58
9	I	113/116 (97%)	97 (86%)	16 (14%)	3	25
10	J	60/65 (92%)	55 (92%)	5 (8%)	12	46
11	K	99/102 (97%)	91 (92%)	8 (8%)	13	47
12	L	40/57 (70%)	33 (82%)	7 (18%)	2	15
13	S	62/156 (40%)	55 (89%)	7 (11%)	6	32
All	All	3497/4165 (84%)	3175 (91%)	322 (9%)	10	41

5 of 322 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	629	ASP
2	B	1047	PHE
10	J	46	CYS
2	B	658	ILE
2	B	856	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 91 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	465	ASN
2	B	957	ASN
9	I	89	GLN
2	B	515	HIS
2	B	657	HIS

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 [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 10 ligands modelled in this entry, 10 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
13	S	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	S	269:PHE	C	270:THR	N	0.95

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.