



wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 03:03 am BST

PDB ID : 3KSB
Title : Detailed structural insight into the DNA cleavage complex of type IIA topoisomerases (re-sealed form)
Authors : Laponogov, I.; Pan, X.-S.; Veselkov, D.A.; McAuley, K.E.; Fisher, L.M.; Sanderson, M.R.
Deposited on : 2009-11-21
Resolution : 3.50 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

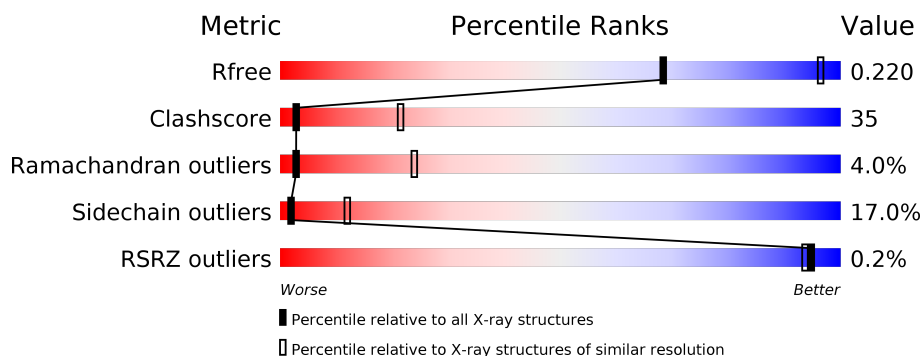
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.50 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	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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	496	
1	B	496	
2	C	268	
2	D	268	
3	E	34	
4	F	34	

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 10601 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 topoisomerase 4 subunit A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	471	Total	C	N	O	S	0	0	0
			3405	2169	591	631	14			
1	B	471	Total	C	N	O	S	0	0	0
			3395	2161	592	629	13			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	489	LEU	-	EXPRESSION TAG	UNP P72525
A	490	GLU	-	EXPRESSION TAG	UNP P72525
A	491	HIS	-	EXPRESSION TAG	UNP P72525
A	492	HIS	-	EXPRESSION TAG	UNP P72525
A	493	HIS	-	EXPRESSION TAG	UNP P72525
A	494	HIS	-	EXPRESSION TAG	UNP P72525
A	495	HIS	-	EXPRESSION TAG	UNP P72525
A	496	HIS	-	EXPRESSION TAG	UNP P72525
B	489	LEU	-	EXPRESSION TAG	UNP P72525
B	490	GLU	-	EXPRESSION TAG	UNP P72525
B	491	HIS	-	EXPRESSION TAG	UNP P72525
B	492	HIS	-	EXPRESSION TAG	UNP P72525
B	493	HIS	-	EXPRESSION TAG	UNP P72525
B	494	HIS	-	EXPRESSION TAG	UNP P72525
B	495	HIS	-	EXPRESSION TAG	UNP P72525
B	496	HIS	-	EXPRESSION TAG	UNP P72525

- Molecule 2 is a protein called DNA topoisomerase 4 subunit B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	217	Total	C	N	O	S	0	0	0
			1544	987	271	280	6			
2	D	214	Total	C	N	O	S	0	0	0
			1521	974	267	274	6			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	380	MET	-	INITIATING METHIONINE	UNP Q59961
C	381	GLY	-	EXPRESSION TAG	UNP Q59961
C	382	HIS	-	EXPRESSION TAG	UNP Q59961
C	383	HIS	-	EXPRESSION TAG	UNP Q59961
C	384	HIS	-	EXPRESSION TAG	UNP Q59961
C	385	HIS	-	EXPRESSION TAG	UNP Q59961
C	386	HIS	-	EXPRESSION TAG	UNP Q59961
C	387	HIS	-	EXPRESSION TAG	UNP Q59961
C	388	HIS	-	EXPRESSION TAG	UNP Q59961
C	389	HIS	-	EXPRESSION TAG	UNP Q59961
C	390	HIS	-	EXPRESSION TAG	UNP Q59961
C	391	HIS	-	EXPRESSION TAG	UNP Q59961
C	392	SER	-	EXPRESSION TAG	UNP Q59961
C	393	SER	-	EXPRESSION TAG	UNP Q59961
C	394	GLY	-	EXPRESSION TAG	UNP Q59961
C	395	HIS	-	EXPRESSION TAG	UNP Q59961
C	396	ILE	-	EXPRESSION TAG	UNP Q59961
C	397	ASP	-	EXPRESSION TAG	UNP Q59961
C	398	ASP	-	EXPRESSION TAG	UNP Q59961
C	399	ASP	-	EXPRESSION TAG	UNP Q59961
C	400	ASP	-	EXPRESSION TAG	UNP Q59961
C	401	LYS	-	EXPRESSION TAG	UNP Q59961
C	402	HIS	-	EXPRESSION TAG	UNP Q59961
C	403	MET	-	EXPRESSION TAG	UNP Q59961
D	380	MET	-	INITIATING METHIONINE	UNP Q59961
D	381	GLY	-	EXPRESSION TAG	UNP Q59961
D	382	HIS	-	EXPRESSION TAG	UNP Q59961
D	383	HIS	-	EXPRESSION TAG	UNP Q59961
D	384	HIS	-	EXPRESSION TAG	UNP Q59961
D	385	HIS	-	EXPRESSION TAG	UNP Q59961
D	386	HIS	-	EXPRESSION TAG	UNP Q59961
D	387	HIS	-	EXPRESSION TAG	UNP Q59961
D	388	HIS	-	EXPRESSION TAG	UNP Q59961
D	389	HIS	-	EXPRESSION TAG	UNP Q59961
D	390	HIS	-	EXPRESSION TAG	UNP Q59961
D	391	HIS	-	EXPRESSION TAG	UNP Q59961
D	392	SER	-	EXPRESSION TAG	UNP Q59961
D	393	SER	-	EXPRESSION TAG	UNP Q59961
D	394	GLY	-	EXPRESSION TAG	UNP Q59961
D	395	HIS	-	EXPRESSION TAG	UNP Q59961
D	396	ILE	-	EXPRESSION TAG	UNP Q59961
D	397	ASP	-	EXPRESSION TAG	UNP Q59961

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	398	ASP	-	EXPRESSION TAG	UNP Q59961
D	399	ASP	-	EXPRESSION TAG	UNP Q59961
D	400	ASP	-	EXPRESSION TAG	UNP Q59961
D	401	LYS	-	EXPRESSION TAG	UNP Q59961
D	402	HIS	-	EXPRESSION TAG	UNP Q59961
D	403	MET	-	EXPRESSION TAG	UNP Q59961

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	18	Total	C	N	O	P	0	0	0
			366	176	70	103	17			

- Molecule 4 is a DNA chain called 5'-D(*CP*TP*GP*TP*TP*TP*TP*A*CP*GP*TP*GP*CP*AP*TP*AP*GP*TP*CP*AP*TP*TP*CP*AP*TP*GP*AP*CP*CP*TP*TP*GP*GP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	18	Total	C	N	O	P	0	0	0
			364	176	64	107	17			

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	D	1	Total	Mg	0	0
			1	1		
5	C	1	Total	Mg	0	0
			1	1		

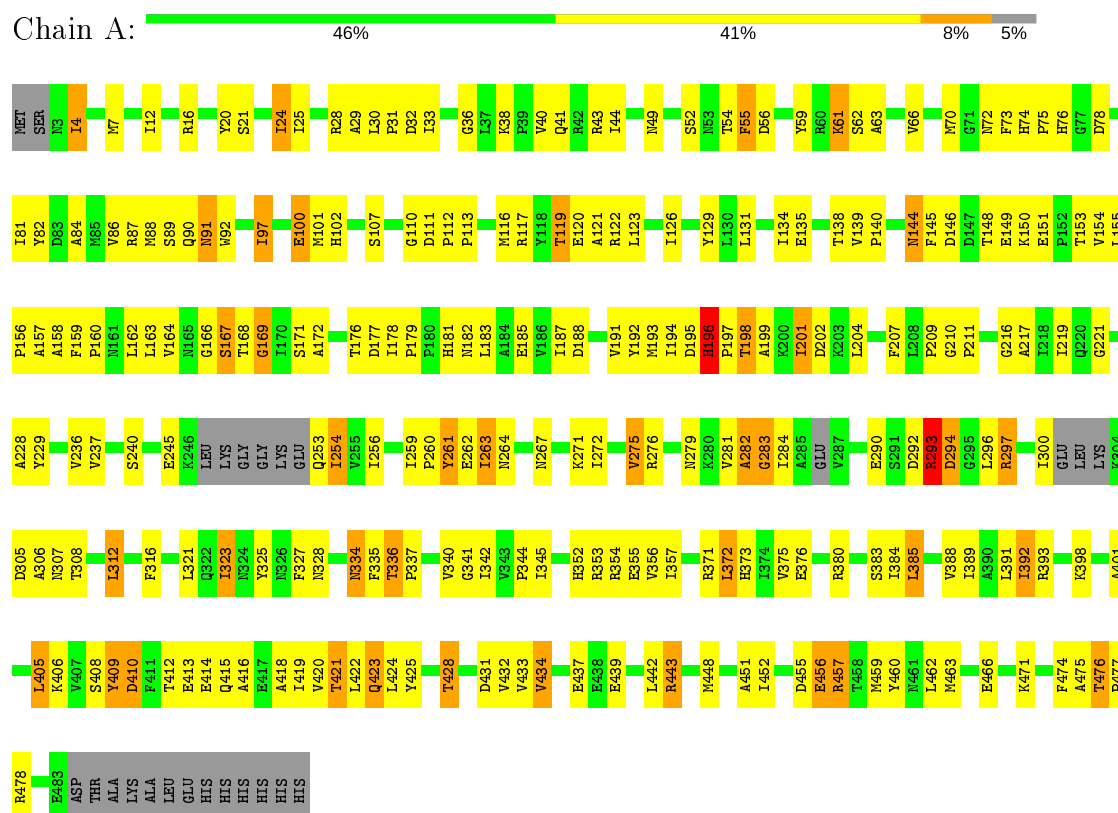
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	2	Total	O	0	0
			2	2		
6	B	2	Total	O	0	0
			2	2		

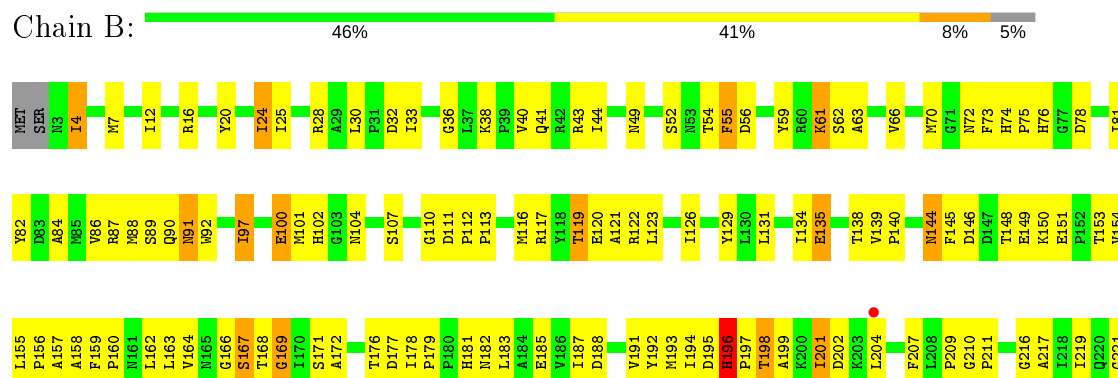
3 Residue-property plots

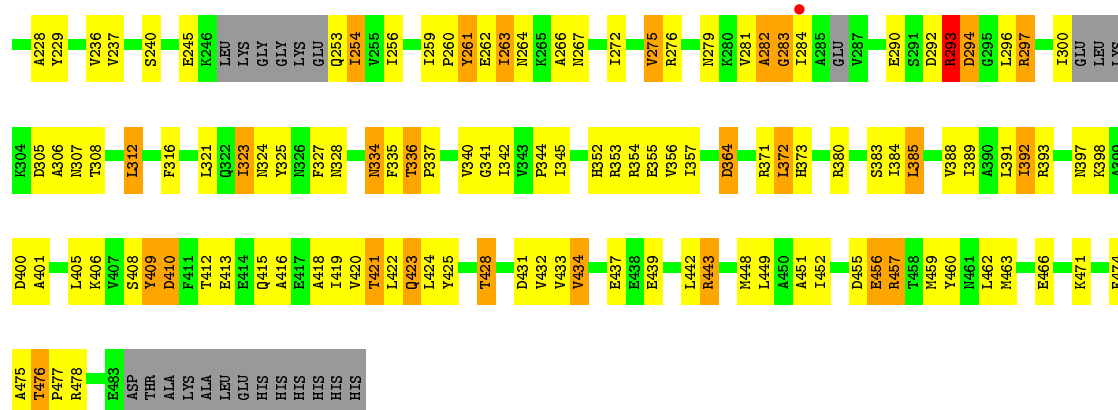
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). 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 topoisomerase 4 subunit A



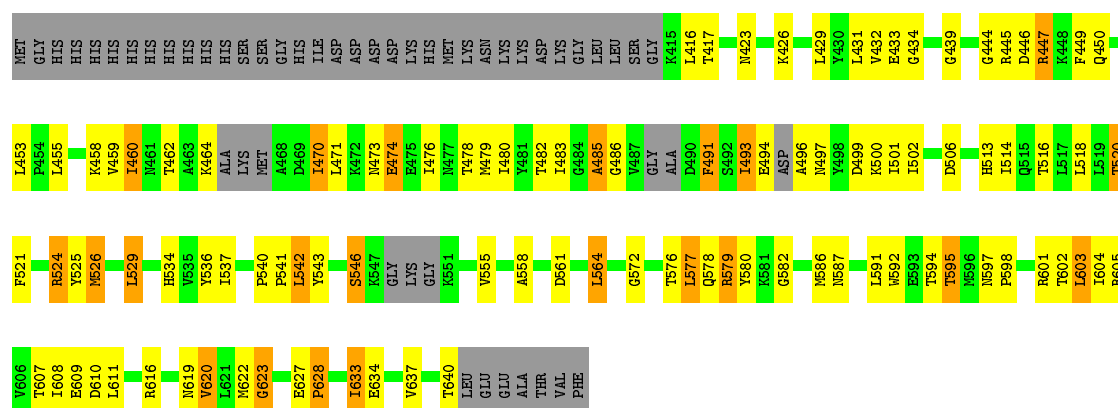
• Molecule 1: DNA topoisomerase 4 subunit A





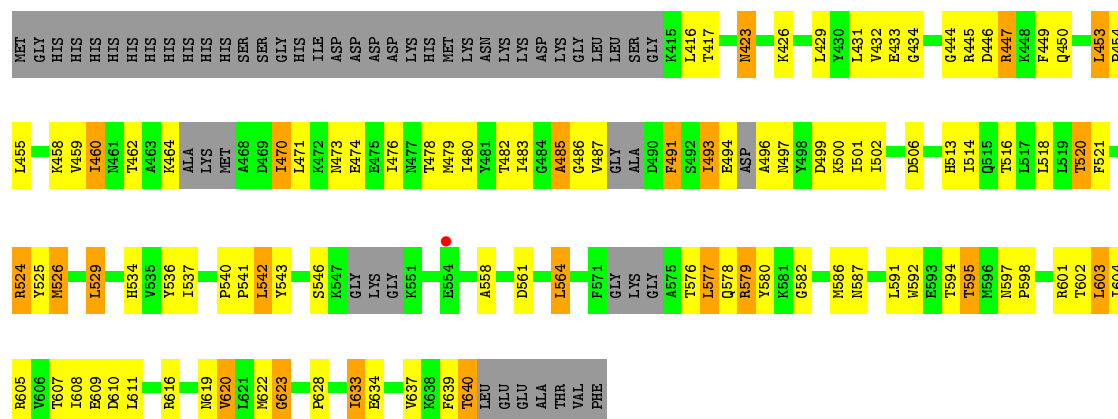
• Molecule 2: DNA topoisomerase 4 subunit B

Chain C: 43% 30% 8% 19%




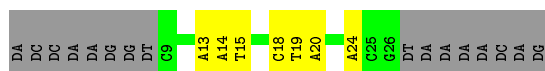
• Molecule 2: DNA topoisomerase 4 subunit B

Chain D: 42% 30% 8% 20%

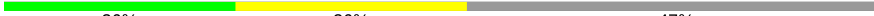


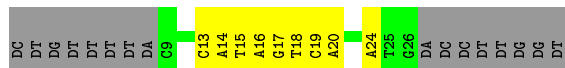
• Molecule 3: 5'-D(*AP*CP*CP*AP*AP*GP*GP*T*CP*AP*TP*GP*AP*AP*TP*GP*AP*CP*TP*AP*TP*GP*CP*AP*CP*GP*TP*AP*AP*AP*AP*CP*AP*G)-3'

Chain E:  32% 21% 47%



- Molecule 4: 5'-D(*CP*TP*GP*TP*TP*TP*TP*A*CP*GP*TP*GP*CP*AP*TP*AP*GP*TP*CP*AP*TP*TP*CP*AP*TP*GP*AP*CP*CP*TP*TP*GP*GP*T)-3'

Chain F:  26% 26% 47%



4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α , β , γ	116.76Å 116.76Å 182.81Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.32 – 3.50 29.32 – 3.50	Depositor EDS
% Data completeness (in resolution range)	97.5 (29.32-3.50) 97.5 (29.32-3.50)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.07	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.51 (at 3.47Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, R_{free}	0.181 , 0.226 0.173 , 0.220	Depositor DCC
R_{free} test set	3524 reflections (10.30%)	wwPDB-VP
Wilson B-factor (Å ²)	134.4	Xtriage
Anisotropy	0.108	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 140.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.477 for -h,-k,l 0.031 for h,-h-k,-l 0.030 for -k,-h,-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	10601	wwPDB-VP
Average B, all atoms (Å ²)	147.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.76% 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: 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	0.59	0/3462	0.78	0/4721
1	B	0.58	0/3452	0.78	0/4711
2	C	0.53	0/1568	0.76	1/2140 (0.0%)
2	D	0.51	0/1544	0.76	1/2107 (0.0%)
3	E	0.44	0/411	1.09	0/633
4	F	0.45	0/407	1.10	1/627 (0.2%)
All	All	0.56	0/10844	0.80	3/14939 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	633	ILE	CB-CA-C	-5.63	100.33	111.60
2	C	633	ILE	CB-CA-C	-5.29	101.03	111.60
4	F	16	DA	O4'-C4'-C3'	-5.14	102.44	104.50

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

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	3405	0	3148	245	0
1	B	3395	0	3124	251	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	1544	0	1417	119	0
2	D	1521	0	1391	112	0
3	E	366	0	201	10	0
4	F	364	0	203	11	0
5	C	1	0	0	0	0
5	D	1	0	0	0	0
6	A	2	0	0	1	0
6	B	2	0	0	1	0
All	All	10601	0	9484	704	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 35.

The worst 5 of 704 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:B:457:ARG:HG2	1:B:457:ARG:HH11	1.10	1.14
1:A:457:ARG:HG2	1:A:457:ARG:HH11	1.11	1.10
1:A:55:PHE:HA	1:A:122:ARG:HD2	1.30	1.09
2:C:524:ARG:HH11	2:C:524:ARG:HG3	1.13	1.08
1:B:55:PHE:HA	1:B:122:ARG:HD2	1.32	1.08

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	463/496 (93%)	396 (86%)	47 (10%)	20 (4%)	2	22
1	B	463/496 (93%)	396 (86%)	47 (10%)	20 (4%)	2	22
2	C	207/268 (77%)	170 (82%)	30 (14%)	7 (3%)	3	28

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	202/268 (75%)	167 (83%)	29 (14%)	6 (3%)	4	30
All	All	1335/1528 (87%)	1129 (85%)	153 (12%)	53 (4%)	3	24

5 of 53 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	282	ALA
1	A	293	ARG
1	A	294	ASP
1	A	456	GLU
1	B	282	ALA

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	310/431 (72%)	262 (84%)	48 (16%)	2	16
1	B	307/431 (71%)	258 (84%)	49 (16%)	2	14
2	C	133/224 (59%)	106 (80%)	27 (20%)	1	6
2	D	130/224 (58%)	104 (80%)	26 (20%)	1	7
All	All	880/1310 (67%)	730 (83%)	150 (17%)	2	12

5 of 150 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	198	THR
1	B	357	ILE
2	D	564	LEU
1	B	254	ILE
1	B	297	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	68	ASN
1	B	102	HIS
2	C	578	GLN
1	B	76	HIS
1	B	91	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 2 ligands modelled in this entry, 2 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	471/496 (94%)	-0.35	0 100 100	94, 143, 193, 247	0
1	B	471/496 (94%)	-0.38	2 (0%) 92 90	95, 143, 193, 248	0
2	C	217/268 (80%)	-0.32	0 100 100	105, 150, 212, 247	0
2	D	214/268 (79%)	-0.29	1 (0%) 91 88	104, 150, 204, 247	0
3	E	18/34 (52%)	-0.50	0 100 100	117, 137, 180, 184	0
4	F	18/34 (52%)	-0.47	0 100 100	119, 137, 186, 189	0
All	All	1409/1596 (88%)	-0.35	3 (0%) 95 93	94, 146, 197, 248	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	204	LEU	2.6
2	D	554	GLU	2.3
1	B	284	ILE	2.2

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(\AA^2)	Q<0.9
5	MG	C	742	1/1	0.71	0.10	168,168,168,168	0
5	MG	D	742	1/1	0.87	0.17	174,174,174,174	0

6.5 Other polymers [i](#)

There are no such residues in this entry.