



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 13, 2016 – 04:48 AM EST

PDB ID : 5TCT
Title : The Structure of SV40 Large T Hexameric Helicase in Complex with AT-rich Origin DNA
Authors : Gai, D.; Wang, D.; Li, S.X.; Chen, X.S.
Deposited on : 2016-09-15
Resolution : 2.90 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

<http://wwpdb.org/validation/2016/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
Mogul : 1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20028442
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20028442

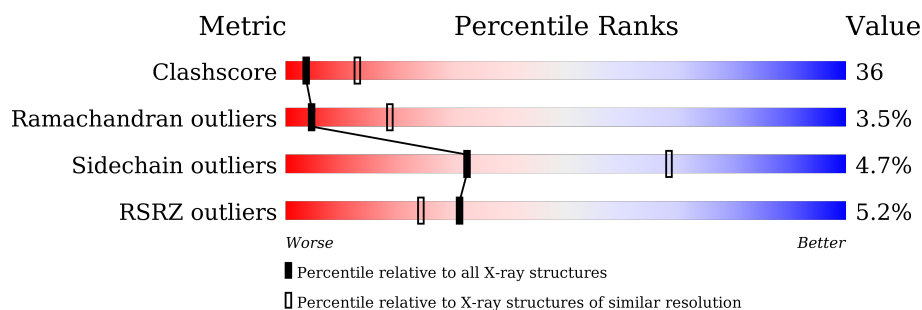
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)
RSRZ outliers	91569	1456 (2.90-2.90)

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 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 $\leq 5\%$. 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	362	
1	B	362	
1	C	362	
1	D	362	
1	E	362	
1	F	362	
2	W	33	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	K	33	<div> <div>30%</div> <div>6%76%18%</div> </div>

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 19123 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 Large T antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	362	Total	C	N	O	S	0	0	0
			2933	1888	493	531	21			
1	B	362	Total	C	N	O	S	0	0	0
			2933	1888	493	531	21			
1	E	362	Total	C	N	O	S	0	0	0
			2933	1888	493	531	21			
1	F	362	Total	C	N	O	S	0	0	0
			2933	1888	493	531	21			
1	C	362	Total	C	N	O	S	0	0	0
			2933	1888	493	531	21			
1	D	362	Total	C	N	O	S	0	0	0
			2933	1888	493	531	21			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	531	TYR	PHE	ENGINEERED MUTATION	UNP P03070
A	549	PRO	ALA	ENGINEERED MUTATION	UNP P03070
B	531	TYR	PHE	ENGINEERED MUTATION	UNP P03070
B	549	PRO	ALA	ENGINEERED MUTATION	UNP P03070
E	531	TYR	PHE	ENGINEERED MUTATION	UNP P03070
E	549	PRO	ALA	ENGINEERED MUTATION	UNP P03070
F	531	TYR	PHE	ENGINEERED MUTATION	UNP P03070
F	549	PRO	ALA	ENGINEERED MUTATION	UNP P03070
C	531	TYR	PHE	ENGINEERED MUTATION	UNP P03070
C	549	PRO	ALA	ENGINEERED MUTATION	UNP P03070
D	531	TYR	PHE	ENGINEERED MUTATION	UNP P03070
D	549	PRO	ALA	ENGINEERED MUTATION	UNP P03070

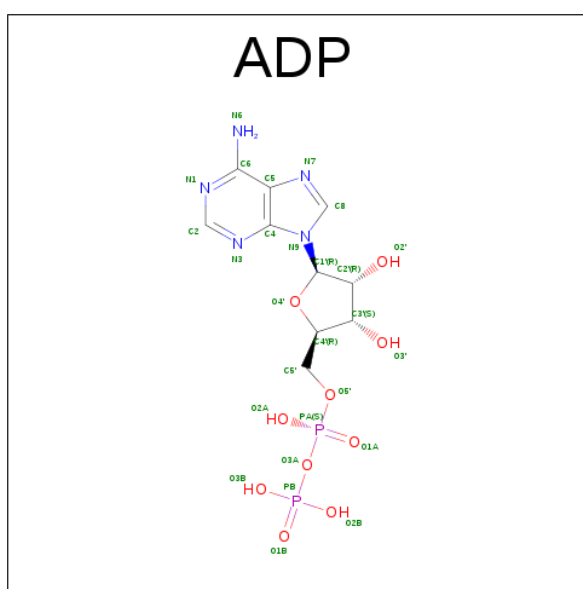
- Molecule 2 is a DNA chain called AT-rich Origin DNA (33-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	W	33	Total	C	N	O	P	0	0	0
			677	323	127	194	33			

- Molecule 3 is a DNA chain called AT-rich Origin DNA (33-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	K	33	Total	C	N	O	P	0	0	0
			677	325	113	206	33			

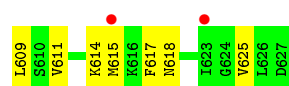
- Molecule 4 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula: $C_{10}H_{15}N_5O_{10}P_2$).



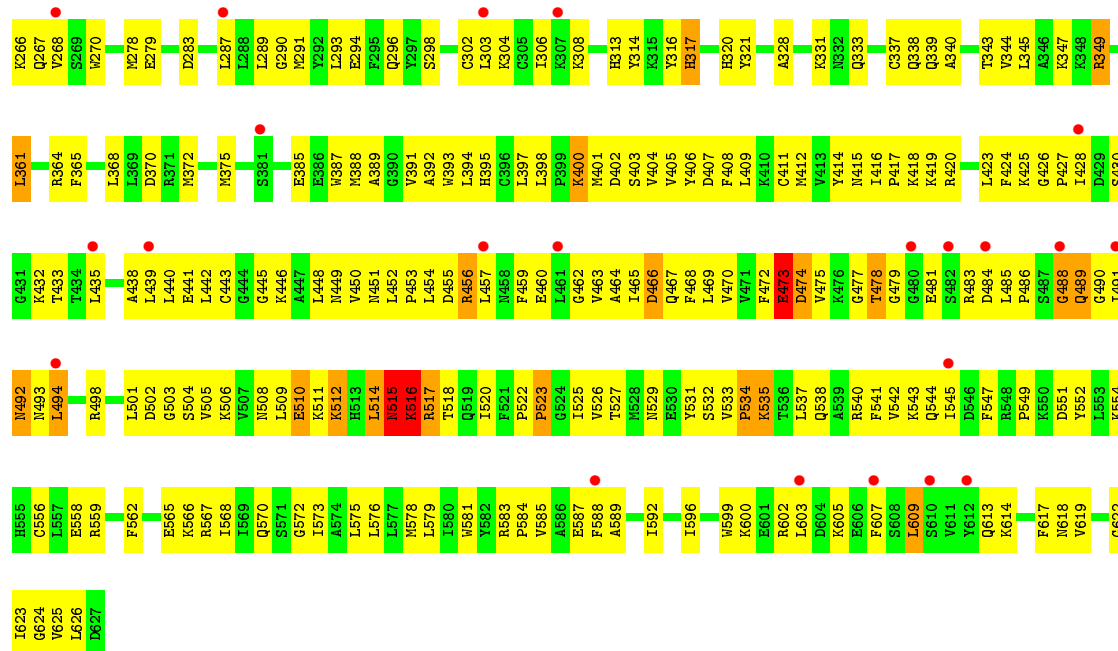
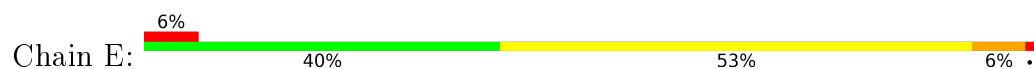
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	D	1	Total 1	Zn 1	0	0
5	E	1	Total 1	Zn 1	0	0
5	B	1	Total 1	Zn 1	0	0
5	C	1	Total 1	Zn 1	0	0
5	A	1	Total 1	Zn 1	0	0
5	F	1	Total 1	Zn 1	0	0

- Molecule 6 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

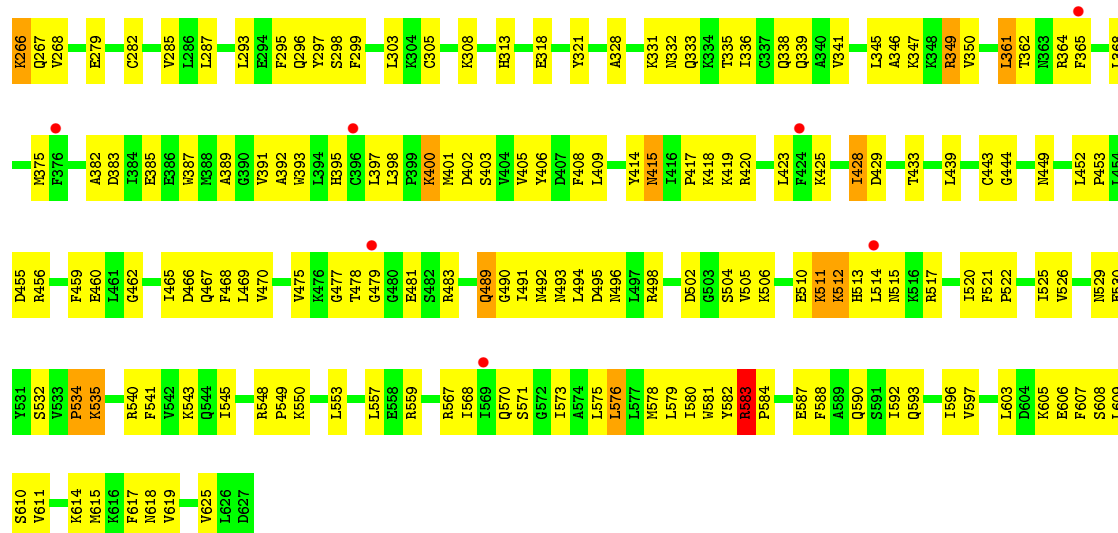
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	1	Total 1	Mn 1	0	0
6	D	1	Total 1	Mn 1	0	0
6	F	1	Total 1	Mn 1	0	0



• Molecule 1: Large T antigen

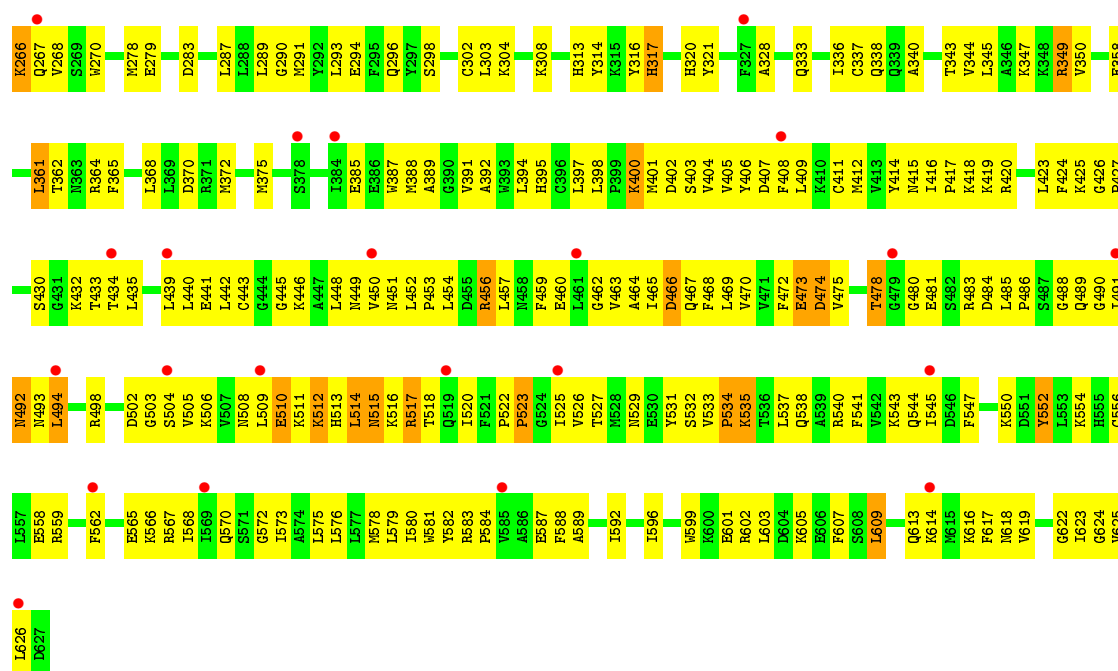


• Molecule 1: Large T antigen

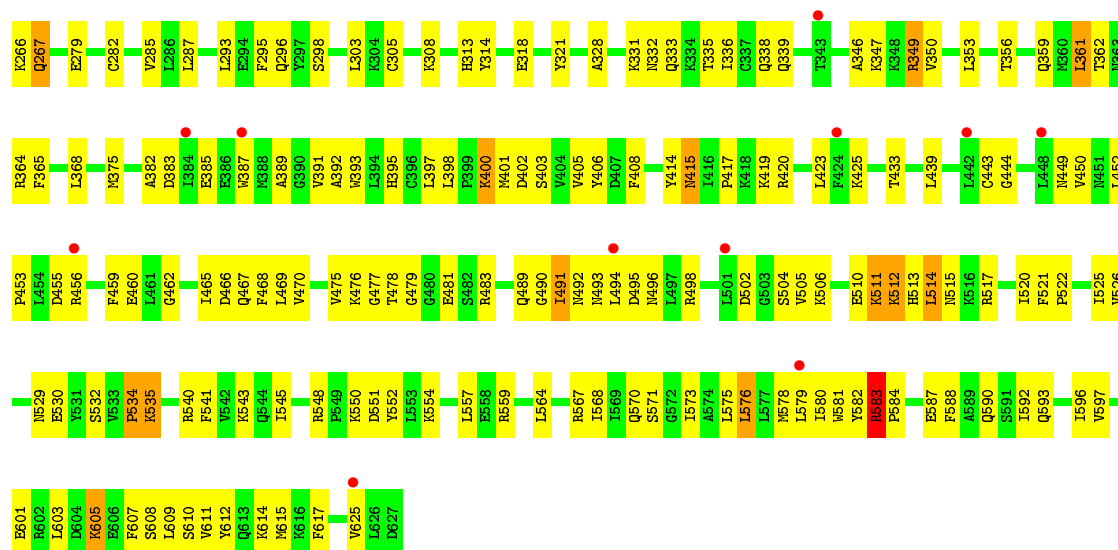


• Molecule 1: Large T antigen

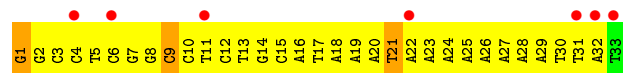




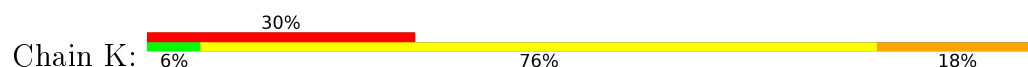
• Molecule 1: Large T antigen



• Molecule 2: AT-rich Origin DNA (33-MER)



• Molecule 3: AT-rich Origin DNA (33-MER)





4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	107.48Å 107.59Å 107.87Å 107.13° 106.95° 106.86°	Depositor
Resolution (Å)	50.00 – 2.90 49.87 – 2.91	Depositor EDS
% Data completeness (in resolution range)	95.8 (50.00-2.90) 75.8 (49.87-2.91)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.77 (at 2.91Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.278 , 0.304 0.271 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	88.1	Xtriage
Anisotropy	0.301	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 58.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.447 for k,l,h 0.447 for l,h,k 0.023 for -k,-h,-l 0.024 for -l,-k,-h 0.023 for -h,-l,-k	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	19123	wwPDB-VP
Average B, all atoms (Å ²)	136.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.08% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical 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, MN, ADP

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	0.35	0/2992	0.70	6/4030 (0.1%)
1	B	0.37	0/2992	0.57	0/4030
1	C	0.35	0/2992	0.58	2/4030 (0.0%)
1	D	0.35	0/2992	0.56	0/4030
1	E	0.33	0/2992	0.60	2/4030 (0.0%)
1	F	0.36	0/2992	0.57	0/4030
2	W	0.85	2/760 (0.3%)	0.78	0/1168
3	K	0.74	1/756 (0.1%)	0.79	0/1164
All	All	0.40	3/19468 (0.0%)	0.62	10/26512 (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	W	0	2
3	K	0	5
All	All	0	7

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	K	1	DT	OP3-P	-7.22	1.52	1.61
2	W	1	DG	OP3-P	-6.06	1.53	1.61
2	W	1	DG	P-O5'	5.00	1.64	1.59

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

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	338	GLN	N-CA-CB	15.53	138.55	110.60
1	A	483	ARG	N-CA-C	-15.15	70.09	111.00
1	E	516	LYS	N-CA-CB	14.36	136.46	110.60
1	A	483	ARG	CB-CA-C	13.24	136.88	110.40
1	C	480	GLY	N-CA-C	9.14	135.95	113.10

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	K	10	DT	Sidechain
3	K	5	DT	Sidechain
3	K	8	DT	Sidechain
2	W	21	DT	Sidechain
2	W	9	DC	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	2933	0	2982	235	1
1	B	2933	0	2981	164	0
1	C	2933	0	2982	226	0
1	D	2933	0	2981	176	0
1	E	2933	0	2981	234	0
1	F	2933	0	2981	168	0
2	W	677	0	372	129	0
3	K	677	0	378	118	1
4	A	27	0	12	0	0
4	B	27	0	12	2	0
4	C	27	0	12	0	0
4	D	27	0	12	0	0
4	E	27	0	12	0	0
4	F	27	0	12	3	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
5	C	1	0	0	0	0
5	D	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	E	1	0	0	0	0
5	F	1	0	0	0	0
6	B	1	0	0	0	0
6	D	1	0	0	0	0
6	F	1	0	0	0	0
All	All	19123	0	18710	1337	1

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 36.

The worst 5 of 1337 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:D:476:LYS:O	1:D:490:GLY:HA3	1.19	1.30
1:D:477:GLY:CA	1:D:490:GLY:O	1.88	1.22
1:D:476:LYS:O	1:D:490:GLY:CA	1.89	1.21
1:D:477:GLY:HA3	1:D:490:GLY:O	1.45	1.13
1:C:517:ARG:HH11	1:C:517:ARG:HB3	1.15	1.11

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:488:GLY:O	3:K:33:DT:OP1[1_444]	2.16	0.04

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	360/362 (99%)	285 (79%)	60 (17%)	15 (4%)	3	13
1	B	360/362 (99%)	302 (84%)	48 (13%)	10 (3%)	6	24

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	360/362 (99%)	288 (80%)	59 (16%)	13 (4%)	4	18
1	D	360/362 (99%)	302 (84%)	48 (13%)	10 (3%)	6	24
1	E	360/362 (99%)	286 (79%)	59 (16%)	15 (4%)	3	13
1	F	360/362 (99%)	305 (85%)	43 (12%)	12 (3%)	5	20
All	All	2160/2172 (99%)	1768 (82%)	317 (15%)	75 (4%)	4	18

5 of 75 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	338	GLN
1	A	474	ASP
1	A	478	THR
1	A	489	GLN
1	B	267	GLN

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	323/323 (100%)	308 (95%)	15 (5%)	33	69
1	B	323/323 (100%)	308 (95%)	15 (5%)	33	69
1	C	323/323 (100%)	307 (95%)	16 (5%)	30	65
1	D	323/323 (100%)	308 (95%)	15 (5%)	33	69
1	E	323/323 (100%)	307 (95%)	16 (5%)	30	65
1	F	323/323 (100%)	309 (96%)	14 (4%)	35	71
All	All	1938/1938 (100%)	1847 (95%)	91 (5%)	32	68

5 of 91 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	512	LYS
1	F	349	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	459	PHE
1	E	515	ASN
1	F	266	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 63 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	492	ASN
1	F	363	ASN
1	D	489	GLN
1	E	529	ASN
1	F	267	GLN

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 15 ligands modelled in this entry, 9 are monoatomic - leaving 6 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
4	ADP	A	701	-	24,29,29	1.40	4 (16%)	23,45,45	2.17	1 (4%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	ADP	B	701	6	24,29,29	1.30	3 (12%)	23,45,45	2.15	1 (4%)
4	ADP	C	701	-	24,29,29	1.53	5 (20%)	23,45,45	2.18	1 (4%)
4	ADP	D	701	6	24,29,29	1.45	4 (16%)	23,45,45	2.14	2 (8%)
4	ADP	E	701	-	24,29,29	1.39	4 (16%)	23,45,45	2.19	1 (4%)
4	ADP	F	702	6	24,29,29	1.28	3 (12%)	23,45,45	2.20	2 (8%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	ADP	A	701	-	-	0/12/32/32	0/3/3/3
4	ADP	B	701	6	-	0/12/32/32	0/3/3/3
4	ADP	C	701	-	-	0/12/32/32	0/3/3/3
4	ADP	D	701	6	-	0/12/32/32	0/3/3/3
4	ADP	E	701	-	-	0/12/32/32	0/3/3/3
4	ADP	F	702	6	-	0/12/32/32	0/3/3/3

The worst 5 of 23 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	701	ADP	C8-N7	-2.39	1.30	1.34
4	D	701	ADP	C8-N7	-2.25	1.30	1.34
4	C	701	ADP	C8-N7	-2.23	1.30	1.34
4	E	701	ADP	PB-O2B	-2.19	1.47	1.54
4	B	701	ADP	C8-N7	-2.18	1.30	1.34

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	E	701	ADP	N3-C2-N1	-9.58	121.35	128.87
4	C	701	ADP	N3-C2-N1	-9.54	121.38	128.87
4	F	702	ADP	N3-C2-N1	-9.47	121.43	128.87
4	A	701	ADP	N3-C2-N1	-9.45	121.44	128.87
4	B	701	ADP	N3-C2-N1	-9.30	121.56	128.87

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	701	ADP	2	0
4	F	702	ADP	3	0

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 ⓘ

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	362/362 (100%)	0.26	26 (7%)	18	12	88, 148, 197, 205	0
1	B	362/362 (100%)	0.09	11 (3%)	54	47	82, 122, 157, 176	0
1	C	362/362 (100%)	0.21	22 (6%)	25	18	87, 148, 197, 205	0
1	D	362/362 (100%)	0.09	11 (3%)	54	47	83, 122, 157, 176	0
1	E	362/362 (100%)	0.23	22 (6%)	25	18	88, 148, 197, 205	0
1	F	362/362 (100%)	0.05	7 (1%)	70	66	84, 122, 157, 176	0
2	W	33/33 (100%)	0.79	7 (21%)	1	1	98, 152, 202, 215	32 (96%)
3	K	33/33 (100%)	1.84	10 (30%)	1	0	88, 160, 237, 250	32 (96%)
All	All	2238/2238 (100%)	0.19	116 (5%)	31	24	82, 133, 185, 250	64 (2%)

The worst 5 of 116 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	K	20	DA	14.4
1	C	494	LEU	8.4
1	A	312	SER	8.2
3	K	21	DG	8.2
3	K	22	DA	7.1

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å ²)	Q<0.9
5	ZN	E	702	1/1	0.56	0.21	1.59	160,160,160,160	0
5	ZN	D	702	1/1	0.90	0.25	0.79	112,112,112,112	0
4	ADP	D	701	27/27	0.91	0.23	0.30	110,129,134,135	0
4	ADP	B	701	27/27	0.93	0.22	0.05	109,129,135,136	0
6	MN	B	703	1/1	0.98	0.19	-0.01	100,100,100,100	0
4	ADP	F	702	27/27	0.92	0.21	-0.11	111,127,136,136	0
4	ADP	C	701	27/27	0.89	0.20	-0.26	150,157,166,167	0
4	ADP	A	701	27/27	0.92	0.17	-0.36	150,157,164,165	0
4	ADP	E	701	27/27	0.94	0.19	-0.56	152,161,169,169	0
6	MN	F	701	1/1	0.93	0.13	-1.14	103,103,103,103	0
6	MN	D	703	1/1	0.96	0.15	-1.42	103,103,103,103	0
5	ZN	C	702	1/1	0.89	0.23	-	160,160,160,160	0
5	ZN	F	703	1/1	0.86	0.23	-	109,109,109,109	0
5	ZN	A	702	1/1	0.94	0.13	-	162,162,162,162	0
5	ZN	B	702	1/1	0.96	0.22	-	112,112,112,112	0

6.5 Other polymers [i](#)

There are no such residues in this entry.