



wwPDB EM Validation Summary Report ⓘ

Nov 20, 2022 – 01:19 AM EST

PDB ID : 7SGL
EMDB ID : EMD-25113
Title : DNA-PK complex of DNA end processing
Authors : Liu, L.; Li, J.; Chen, X.; Yang, W.; Gellert, M.
Deposited on : 2021-10-06
Resolution : 3.00 Å(reported)

This is a wwPDB EM 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/EMValidationReportHelp>

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:

EMDB validation analysis : 0.0.1.dev43
Mogul : 1.8.5 (274361), CSD as541be (2020)
MolProbity : 4.02b-467
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.31.3

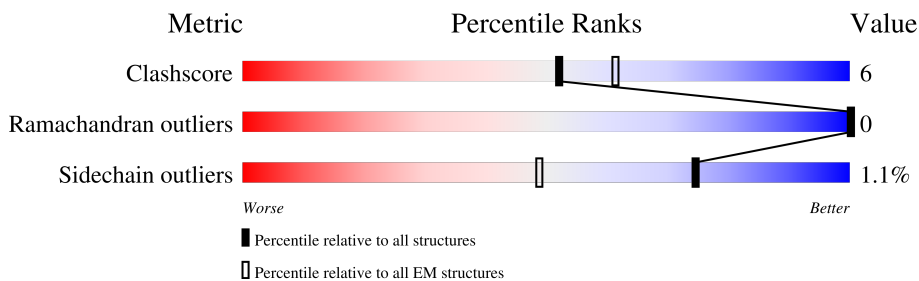
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.00 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)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	4128	81% (green), 12% (yellow), 7% (grey)
2	B	612	70% (green), 12% (yellow), 17% (grey)
3	C	732	70% (green), 16% (yellow), 14% (grey)
4	D	701	50% (green), 8% (yellow), 42% (grey)
5	E	54	59% (yellow), 41% (orange)
5	F	54	11% (yellow), 85% (grey)

2 Entry composition i

There are 9 unique types of molecules in this entry. The entry contains 44704 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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-dependent protein kinase catalytic subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace	
			Total	C	N	O	P			S
1	A	3852	30853	19740	5219	5691	4	199	5	0

- Molecule 2 is a protein called X-ray repair cross-complementing protein 6.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	507	4102	2625	694	765	18	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	GLY	-	expression tag	UNP P12956
B	-1	PRO	-	expression tag	UNP P12956
B	0	VAL	-	expression tag	UNP P12956

- Molecule 3 is a protein called X-ray repair cross-complementing protein 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	631	5099	3267	856	950	26	0	0

- Molecule 4 is a protein called Protein artemis.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	406	3311	2111	578	601	21	1	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	693	ALA	-	expression tag	UNP Q96SD1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	694	ALA	-	expression tag	UNP Q96SD1
D	695	ALA	-	expression tag	UNP Q96SD1
D	696	LEU	-	expression tag	UNP Q96SD1
D	697	GLU	-	expression tag	UNP Q96SD1
D	698	VAL	-	expression tag	UNP Q96SD1
D	699	LEU	-	expression tag	UNP Q96SD1
D	700	PHE	-	expression tag	UNP Q96SD1
D	701	GLN	-	expression tag	UNP Q96SD1

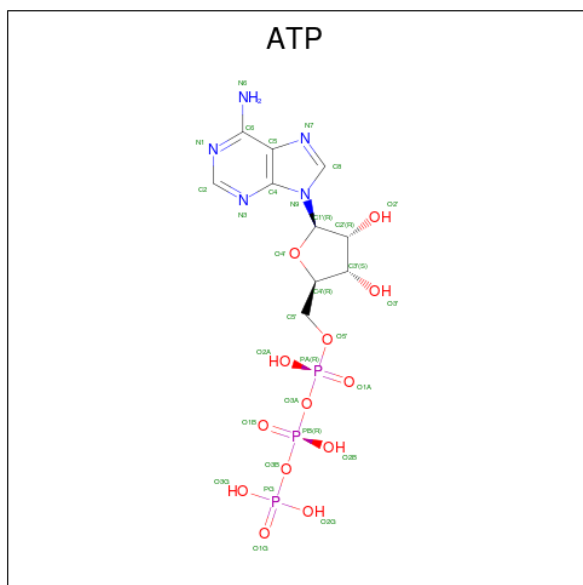
- Molecule 5 is a DNA chain called Hairpin_1.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	54	Total	C	N	O	P	0	0
			1104	528	201	322	53		
5	F	8	Total	C	N	O	P	0	0
			164	78	30	48	8		

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

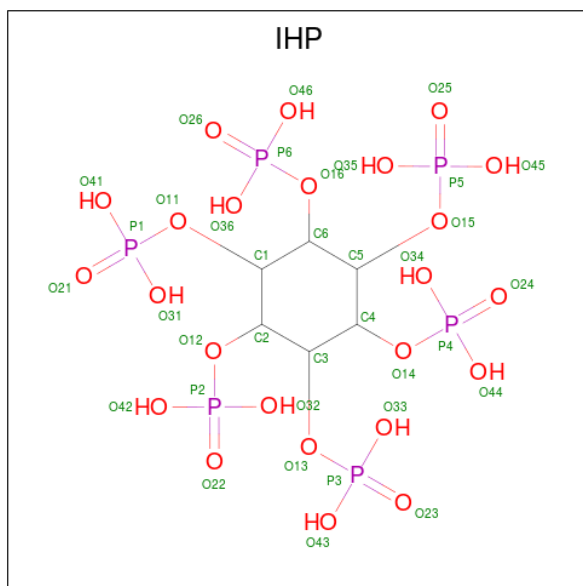
Mol	Chain	Residues	Atoms		AltConf
6	A	1	Total	Mg	0
			1	1	
6	D	2	Total	Mg	0
			2	2	

- Molecule 7 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: C₁₀H₁₆N₅O₁₃P₃) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
7	A	1	31	10	5	13	3	0

- Molecule 8 is INOSITOL HEXAKISPHOSPHATE (three-letter code: IHP) (formula: $C_6H_{18}O_{24}P_6$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				AltConf
			Total	C	O	P	
8	B	1	36	6	24	6	0

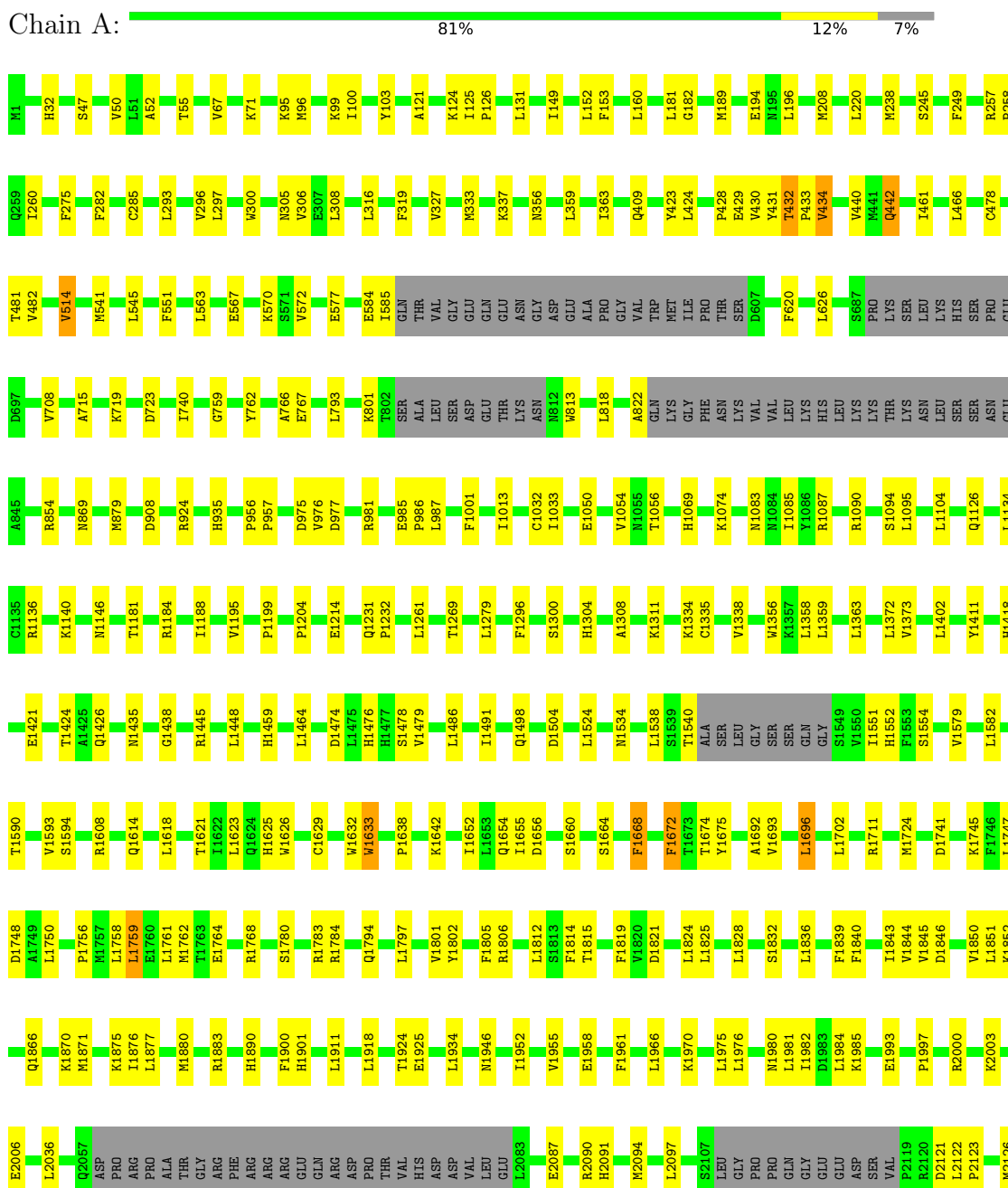
- Molecule 9 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

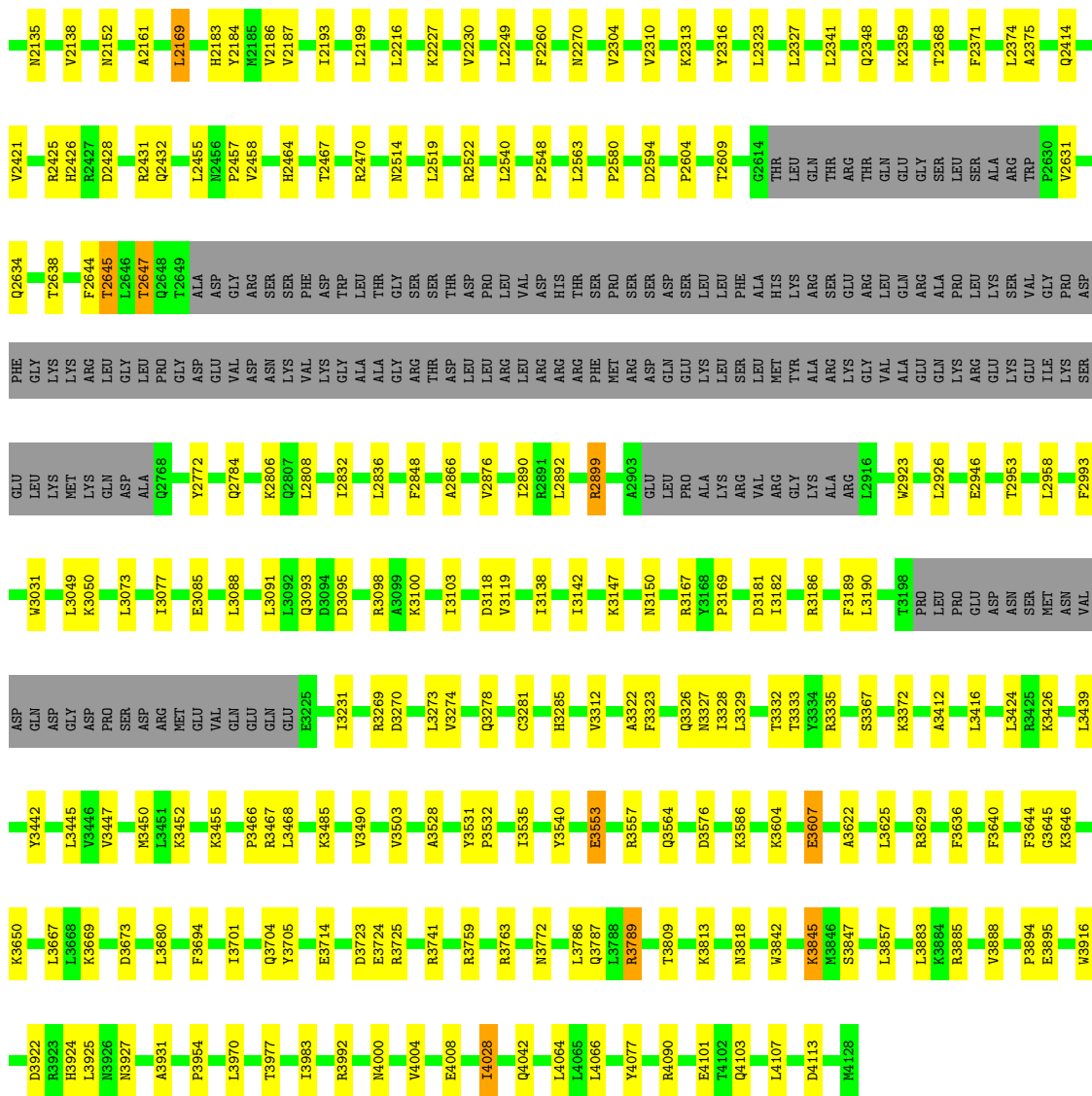
Mol	Chain	Residues	Atoms		AltConf
9	D	1	Total 1	Zn 1	0

3 Residue-property plots

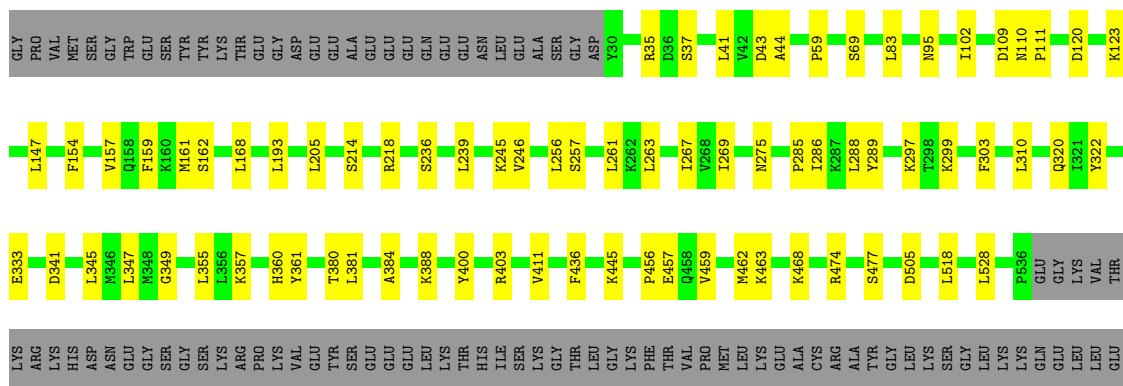
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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.

- Molecule 1: DNA-dependent protein kinase catalytic subunit



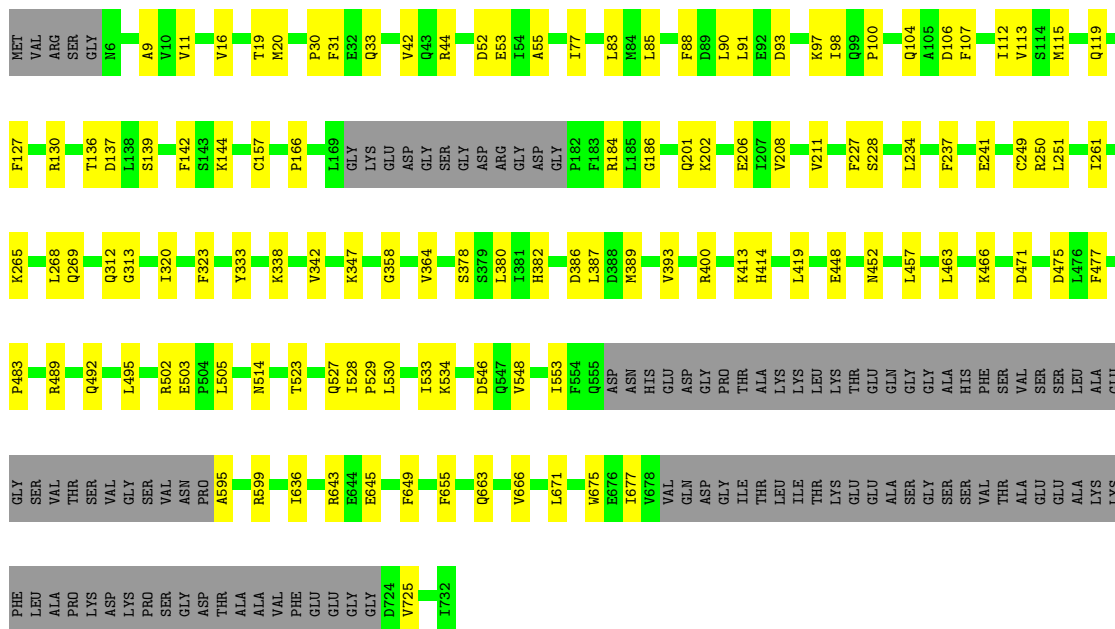


● Molecule 2: X-ray repair cross-complementing protein 6



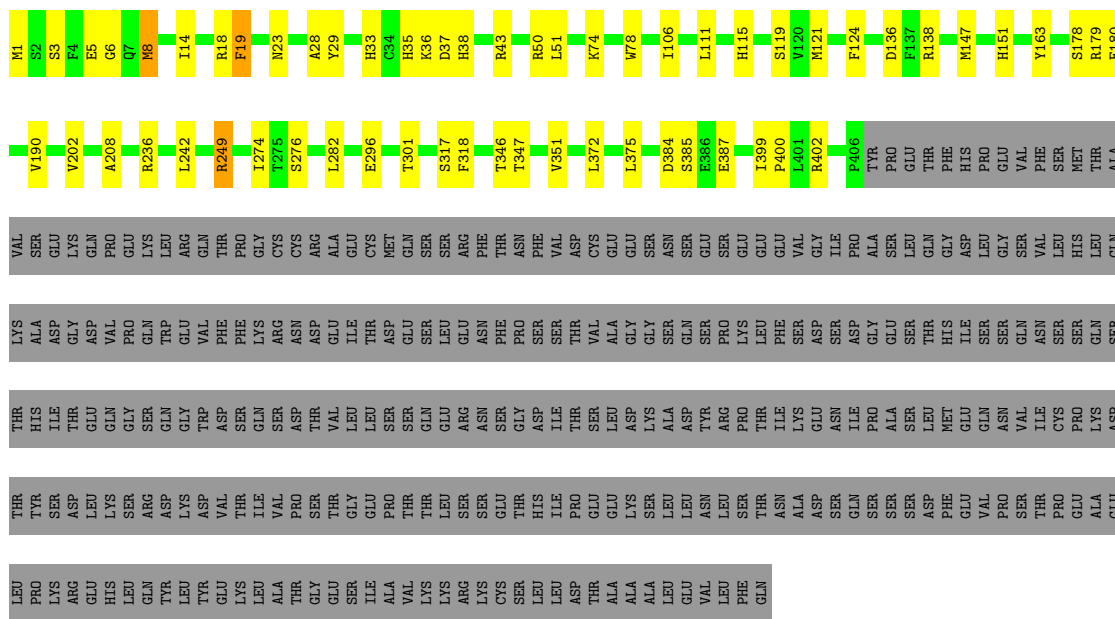
● Molecule 3: X-ray repair cross-complementing protein 5

Chain C: 70% 16% 14%



● Molecule 4: Protein artemis

Chain D: 50% 8% 42%



● Molecule 5: Hairpin_1

Chain E: 59% 41%



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	161380	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	55.9	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor

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, ATP, MG, TPO, IHP

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	0.24	0/31449	0.38	1/42499 (0.0%)
2	B	0.24	0/4184	0.40	0/5638
3	C	0.24	0/5200	0.40	0/7007
4	D	0.25	0/3397	0.41	0/4600
5	E	3.58	213/1238 (17.2%)	1.32	9/1909 (0.5%)
5	F	1.55	1/183 (0.5%)	1.23	3/280 (1.1%)
All	All	0.64	214/45651 (0.5%)	0.46	13/61933 (0.0%)

The worst 5 of 214 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	E	37	DC	N1-C6	-10.16	1.31	1.37
5	E	34	DC	N1-C6	-9.99	1.31	1.37
5	E	2	DC	N1-C6	-9.77	1.31	1.37
5	E	22	DT	C3'-O3'	-9.37	1.31	1.44
5	E	40	DC	N1-C6	-9.14	1.31	1.37

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	E	50	DT	O4'-C1'-N1	6.51	112.56	108.00
5	F	50	DT	O4'-C4'-C3'	-5.70	102.22	104.50
5	E	25	DA	OP1-P-OP2	-5.68	111.09	119.60
5	F	52	DT	N3-C4-O4	5.62	123.28	119.90
1	A	434	VAL	CB-CA-C	5.46	121.77	111.40

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	30853	0	31141	307	0
2	B	4102	0	4174	47	0
3	C	5099	0	5128	76	0
4	D	3311	0	3298	42	0
5	E	1104	0	607	49	0
5	F	164	0	91	22	0
6	A	1	0	0	0	0
6	D	2	0	0	0	0
7	A	31	0	12	0	0
8	B	36	0	6	2	0
9	D	1	0	0	0	0
All	All	44704	0	44457	510	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 6.

The worst 5 of 510 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:A:1924:THR:CG2	1:A:1975:LEU:CD1	2.26	1.13
1:A:1924:THR:HG22	1:A:1975:LEU:CD1	1.80	1.11
1:A:1924:THR:CG2	1:A:1975:LEU:HD11	1.85	1.03
1:A:1924:THR:HG22	1:A:1975:LEU:HD12	1.37	1.02
1:A:1924:THR:HG21	1:A:1975:LEU:HD11	1.47	0.93

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 PDB entries followed by that with respect to all EM entries.

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	3829/4128 (93%)	3731 (97%)	98 (3%)	0	100	100
2	B	505/612 (82%)	497 (98%)	8 (2%)	0	100	100
3	C	623/732 (85%)	605 (97%)	18 (3%)	0	100	100
4	D	405/701 (58%)	394 (97%)	11 (3%)	0	100	100
All	All	5362/6173 (87%)	5227 (98%)	135 (2%)	0	100	100

There are no Ramachandran outliers to report.

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 PDB entries followed by that with respect to all EM entries.

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	3431/3667 (94%)	3389 (99%)	42 (1%)	71	90
2	B	461/550 (84%)	458 (99%)	3 (1%)	84	94
3	C	571/649 (88%)	567 (99%)	4 (1%)	84	94
4	D	372/639 (58%)	366 (98%)	6 (2%)	62	86
All	All	4835/5505 (88%)	4780 (99%)	55 (1%)	74	90

5 of 55 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2899	ARG
1	A	3694	PHE
4	D	372	LEU
4	D	8	MET
1	A	3426	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 62 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	2508	GLN
3	C	423	GLN
1	A	3166	ASN
3	C	269	GLN
3	C	514	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](#)

4 non-standard protein/DNA/RNA residues are modelled in this entry.

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
1	TPO	A	2645	1	8,10,11	1.08	0	10,14,16	1.94	1 (10%)
1	TPO	A	2647	1	8,10,11	0.77	0	10,14,16	1.01	1 (10%)
1	TPO	A	2638	1	8,10,11	1.58	1 (12%)	10,14,16	1.90	1 (10%)
1	TPO	A	2609	1	8,10,11	1.59	1 (12%)	10,14,16	1.89	1 (10%)

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
1	TPO	A	2645	1	-	0/9/11/13	-
1	TPO	A	2647	1	-	1/9/11/13	-
1	TPO	A	2638	1	-	2/9/11/13	-
1	TPO	A	2609	1	-	2/9/11/13	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	2609	TPO	P-O1P	3.40	1.61	1.50
1	A	2638	TPO	P-O1P	3.37	1.61	1.50

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	2645	TPO	P-OG1-CB	-5.53	106.50	123.21
1	A	2609	TPO	P-OG1-CB	-5.38	106.94	123.21
1	A	2638	TPO	P-OG1-CB	-5.36	107.01	123.21
1	A	2647	TPO	O-C-CA	-2.46	118.33	124.78

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	2609	TPO	CB-OG1-P-O2P
1	A	2638	TPO	O-C-CA-CB
1	A	2609	TPO	CB-OG1-P-O1P
1	A	2638	TPO	CB-OG1-P-O1P
1	A	2647	TPO	O-C-CA-CB

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	2645	TPO	1	0
1	A	2647	TPO	2	0

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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	$\# Z > 2$
8	IHP	B	701	-	36,36,36	1.48	6 (16%)	54,60,60	0.50	0
7	ATP	A	4201	6	26,33,33	0.67	0	31,52,52	0.73	1 (3%)

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
8	IHP	B	701	-	-	3/30/54/54	0/1/1/1
7	ATP	A	4201	6	-	4/18/38/38	0/3/3/3

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	B	701	IHP	P4-O14	3.22	1.65	1.59
8	B	701	IHP	P5-O15	3.20	1.65	1.59
8	B	701	IHP	P3-O13	3.18	1.65	1.59
8	B	701	IHP	P6-O16	3.18	1.65	1.59
8	B	701	IHP	P1-O11	3.18	1.65	1.59

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	A	4201	ATP	C5-C6-N6	2.28	123.81	120.35

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

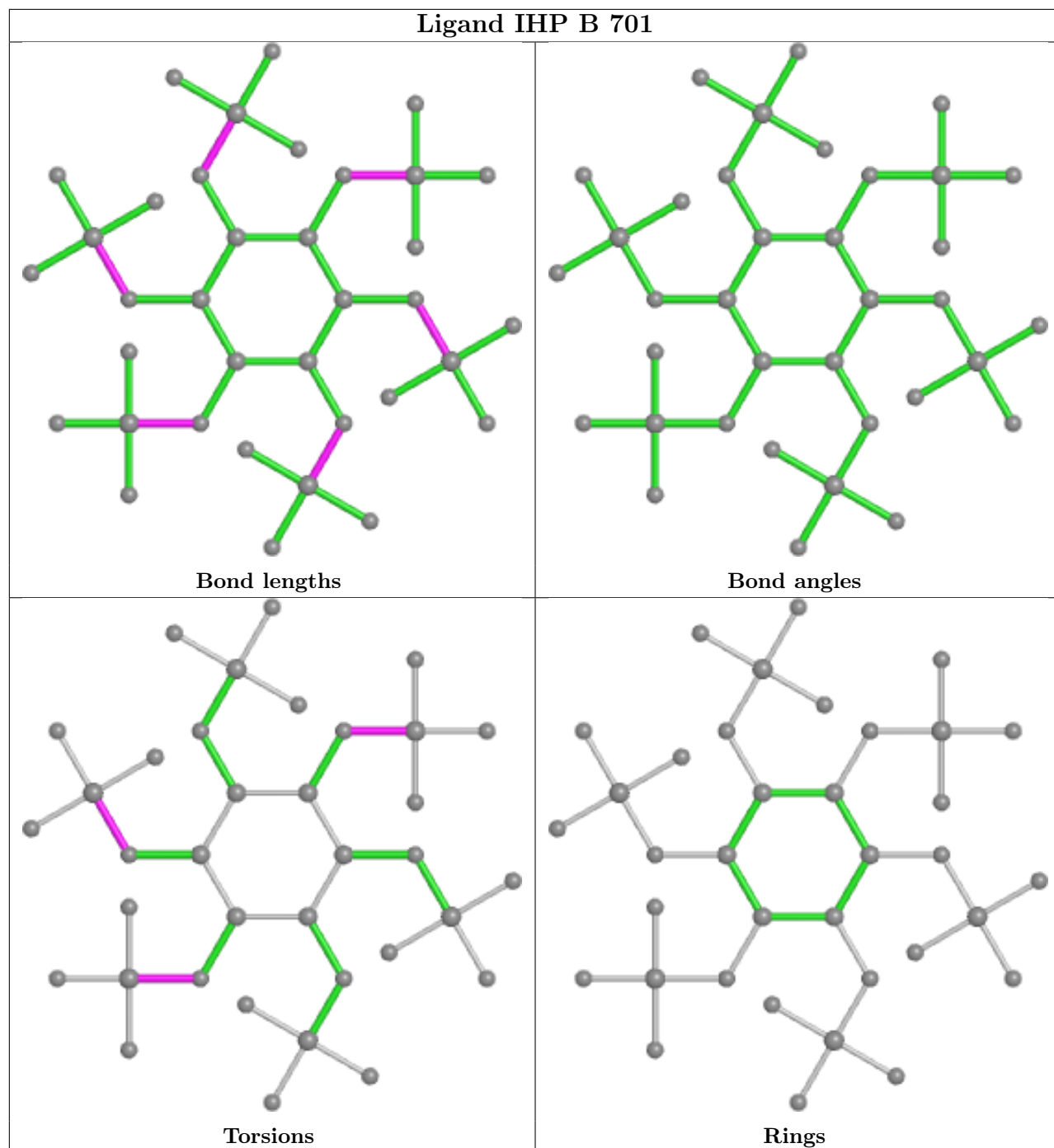
Mol	Chain	Res	Type	Atoms
7	A	4201	ATP	C5'-O5'-PA-O1A
7	A	4201	ATP	C5'-O5'-PA-O2A
7	A	4201	ATP	C5'-O5'-PA-O3A
8	B	701	IHP	C2-O12-P2-O22
8	B	701	IHP	C4-O14-P4-O24

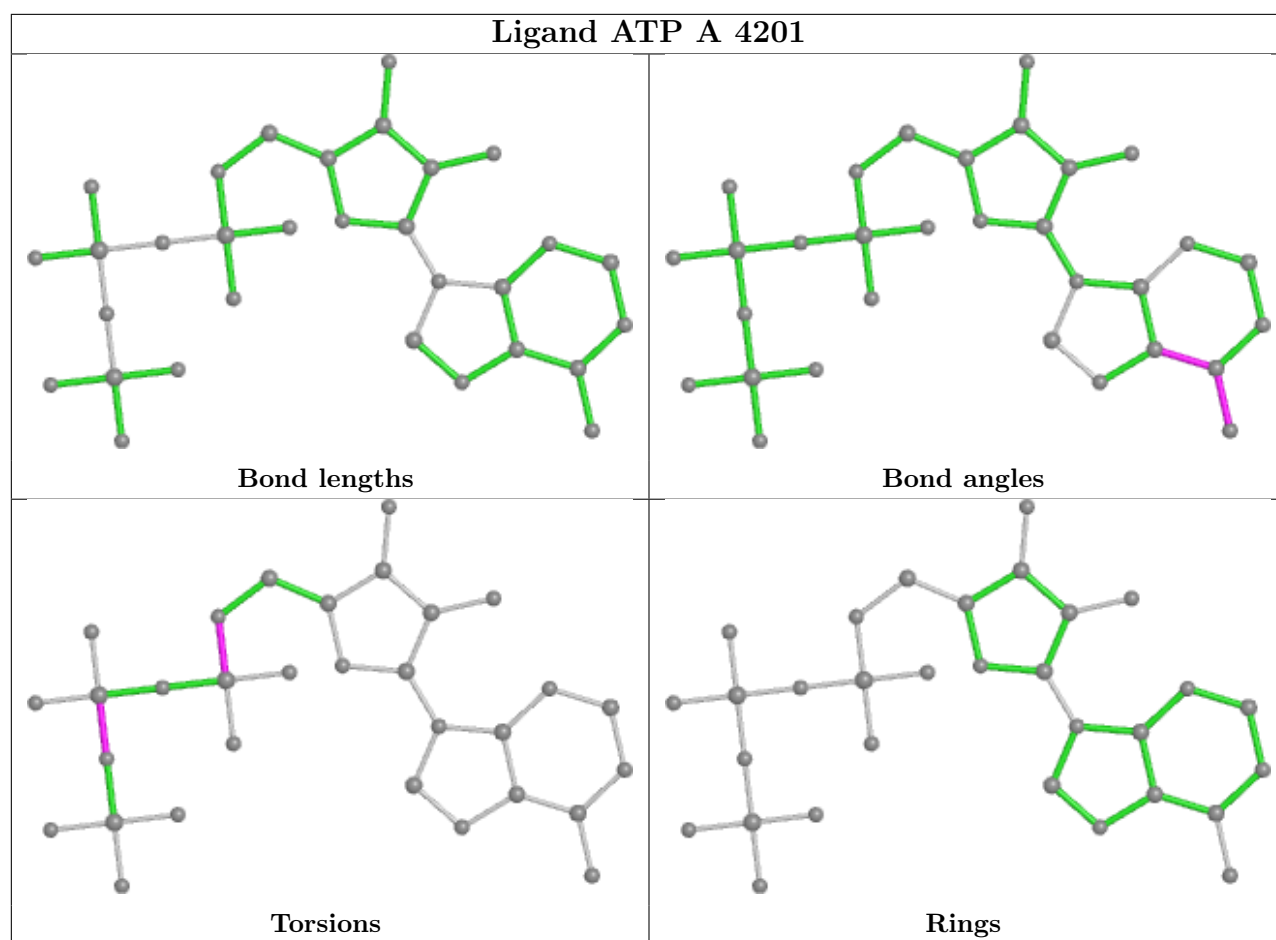
There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	B	701	IHP	2	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.





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 Map visualisation

This section contains visualisations of the EMDB entry EMD-25113. These allow visual inspection of the internal detail of the map and identification of artifacts.

Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

This section was not generated.

6.2 Central slices

This section was not generated.

6.3 Largest variance slices

This section was not generated.

6.4 Orthogonal surface views

This section was not generated.

6.5 Mask visualisation

This section was not generated. No masks/segmentation were deposited.

7 Map analysis

This section contains the results of statistical analysis of the map.

7.1 Map-value distribution

This section was not generated.

7.2 Volume estimate versus contour level

This section was not generated.

7.3 Rotationally averaged power spectrum

This section was not generated. The rotationally averaged power spectrum had issues being displayed.

8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit

This section was not generated.