wwPDB X-ray Structure Validation Summary Report (i)

May 25, 2020 – 11:48 pm BST

PDB ID	:	3PJR
Title	:	HELICASE SUBSTRATE COMPLEX
Authors	:	Velankar, S.S.; Soultanas, P.; Dillingham, M.S.; Subramanya, H.S.; Wigley, D.B.
Deposited on	:	1999-03-12
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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

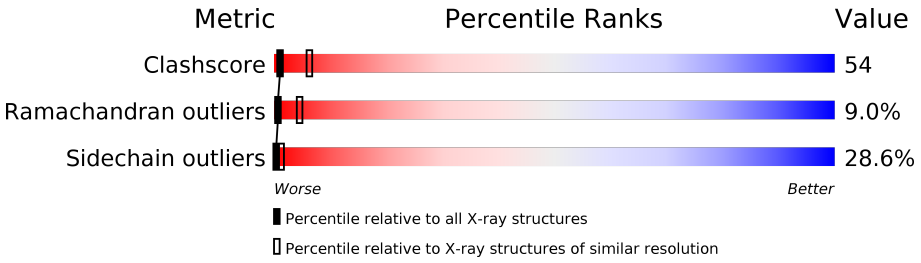
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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)

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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	Y	15	20% 80%
2	Z	10	100%
3	A	724	23% 40% 21% 6% 11%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5765 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(*GP*CP*AP*GP*TP*GP*CP*TP*CP*GP*TP*TP*TP*TP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	Y	15	Total	C	N	O	P	0	0	0
			303	147	48	94	14			

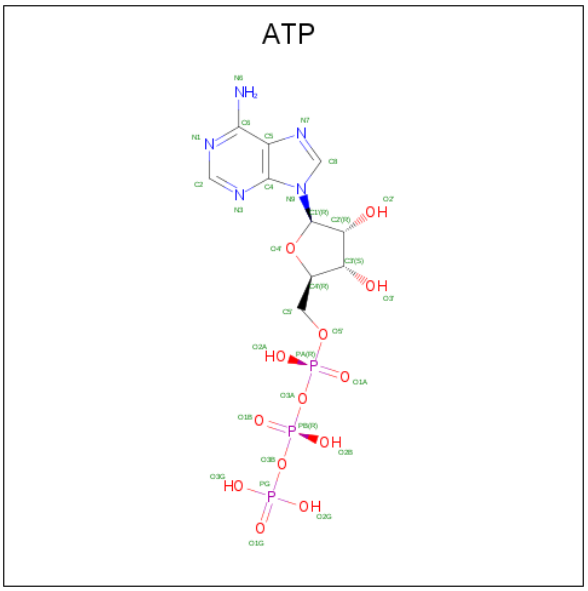
- Molecule 2 is a DNA chain called 5'-D(*CP*GP*AP*GP*CP*AP*CP*TP*GP*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	Z	10	Total	C	N	O	P	0	0	0
			201	96	39	57	9			

- Molecule 3 is a protein called HELICASE PCRA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	646	Total	C	N	O	S	0	0	0
			5230	3304	916	991	19			

- Molecule 4 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: $C_{10}H_{16}N_5O_{13}P_3$).



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

3 Residue-property plots (i)

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.

Note EDS was not executed.

- Molecule 1: 5'-D(*GP*CP*AP*GP*TP*GP*CP*TP*CP*GP*TP*TP*TP*TP*T)-3'

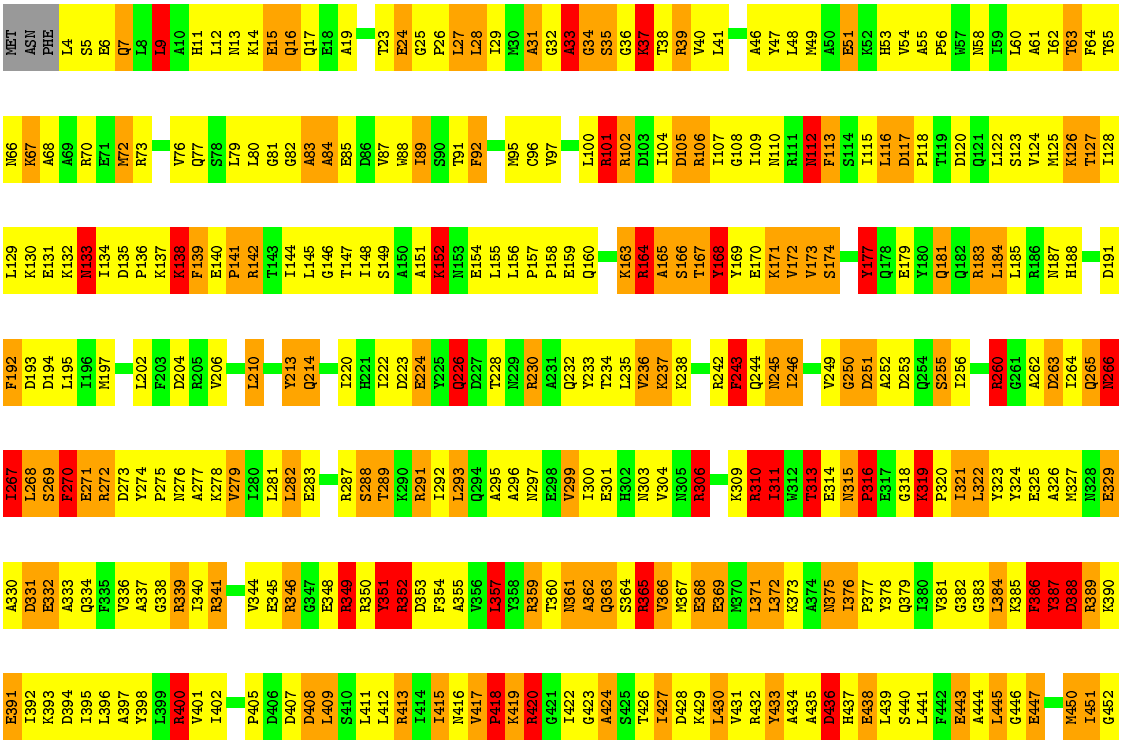
Chain Y:

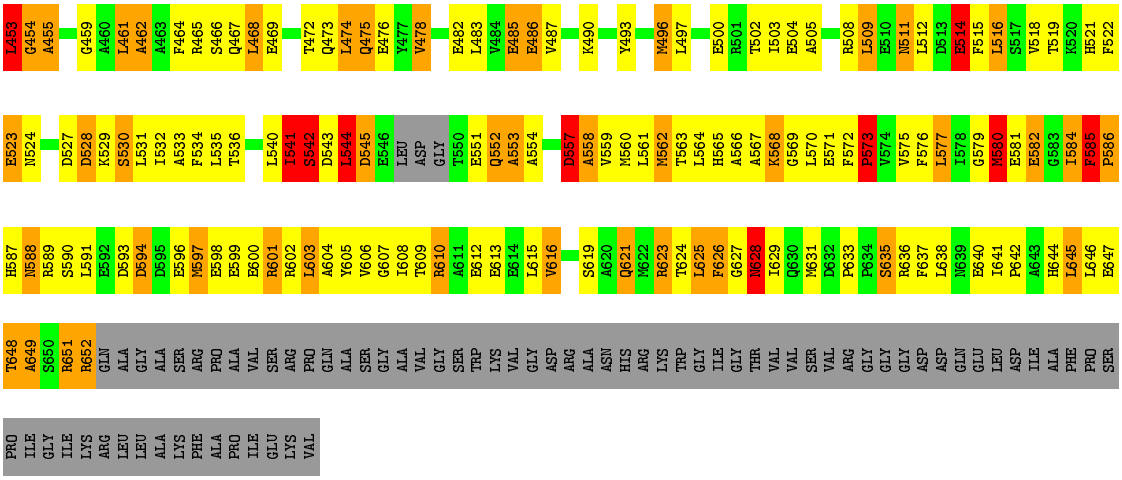
- Molecule 2: 5'-D(*CP*GP*AP*GP*CP*AP*CP*TP*GP*C)-3'

Chain Z:

- Molecule 3: HELICASE PCRA

Chain A:




L453 G454 A455 G459 A460 L461 A462 A463 F464 R465 S466 Q467 L468 E469 T472 Q473 L474 Q475 E476 Y477 V478 E482 L483 V484 E485 E486 V487 K490 Y493 M496 L497 E500 R501 T502 I503 E504 A505 R508 L509 E510 N511 L512 D513 E514 F515 L516 S517 V518 T519 K520 H521 F522
E523 N524 D527 D528 K529 S530 L531 I532 A533 F534 L535 T536 L540 I541 S542 D543 L544 D545 E546 LEU ASP GLY T550 E551 Q552 A553 A554 D587 A558 V559 M560 L561 M562 T563 L564 H565 A566 A567 K568 G569 L570 E571 F572 P573 V574 V575 F576 L577 I578 G579 M580 E581 E582 G583 I584 F585 P586
H587 N588 R589 S590 L591 E592 D593 D594 D595 E596 M597 E598 E599 E600 R601 R602 L603 A604 Y605 V606 G607 I608 T609 R610 A611 E612 E613 E614 L615 V616 S619 A620 Q621 M622 R623 T624 L625 F626 G627 N628 I629 Q630 M631 D632 P633 P634 S635 R636 F637 L638 N639 E640 I641 P642 A643 H644 L645 L646 E647
T648 A649 S650 R651 R652 GLN ALA GLY ALA SER ARG PRO ALA VAL SER ARG PRO GLN ALA SER GLY ALA VAL GLY SER TRP LYS VAL GLY ASP ARG ALA ASN HIS ARG LYS TRP GLY ILE GLY THR VAL VAL SER VAL ARG GLY GLY GLY ASP ASP GLN GLU LEU ASP ILE ALA PHE PRO SER
PRO ILE GLY ILE LYS ARG LEU LEU ALA LYS PHE ALA PRO ILE GLU LYS VAL

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α, β, γ	105.12Å 105.12Å 380.80Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	10.00 – 3.30	Depositor
% Data completeness (in resolution range)	92.0 (10.00-3.30)	Depositor
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	REFMAC	Depositor
R, R_{free}	0.236 , 0.315	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5765	wwPDB-VP
Average B, all atoms (Å^2)	55.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ATP

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	Y	1.56	4/337 (1.2%)	3.06	50/519 (9.6%)
2	Z	2.27	6/225 (2.7%)	3.84	39/345 (11.3%)
3	A	0.68	7/5319 (0.1%)	1.60	69/7184 (1.0%)
All	All	0.87	17/5881 (0.3%)	1.88	158/8048 (2.0%)

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	Z	1	0
3	A	0	63
All	All	1	63

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	Z	28	DG	N9-C8	12.27	1.46	1.37
2	Z	28	DG	C2'-C1'	11.86	1.64	1.52
2	Z	28	DG	C8-N7	-11.06	1.24	1.30
1	Y	11	DT	C3'-O3'	-7.58	1.34	1.44
2	Z	27	DA	C2'-C1'	7.16	1.59	1.52

The worst 5 of 158 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	652	ARG	CD-NE-CZ	38.94	178.12	123.60
3	A	310	ARG	CD-NE-CZ	36.72	175.00	123.60
2	Z	28	DG	C4-N9-C1'	21.13	153.97	126.50

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	Z	28	DG	C8-N9-C1'	-20.87	99.87	127.00
2	Z	28	DG	N3-C4-C5	-20.27	118.47	128.60

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	Z	33	DG	C3'

5 of 63 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	A	16	GLN	Mainchain
3	A	24	GLU	Mainchain
3	A	27	LEU	Mainchain
3	A	31	ALA	Mainchain
3	A	9	LEU	Mainchain

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	Y	303	0	174	56	0
2	Z	201	0	113	36	0
3	A	5230	0	5230	530	0
4	A	31	0	12	7	0
All	All	5765	0	5529	607	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 54.

The worst 5 of 607 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:Y:12:DT:H5"	1:Y:12:DT:C6	1.65	1.31
1:Y:12:DT:C5'	1:Y:12:DT:H6	1.47	1.27
1:Y:5:DT:H2'	1:Y:6:DG:C8	1.77	1.19

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:31:ALA:HB3	3:A:251:ASP:HB2	1.36	1.08
1:Y:12:DT:C6	1:Y:12:DT:C5'	2.31	1.00

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	642/724 (89%)	471 (73%)	113 (18%)	58 (9%)	1 4

5 of 58 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	33	ALA
3	A	35	SER
3	A	83	ALA
3	A	84	ALA
3	A	113	PHE

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	560/618 (91%)	400 (71%)	160 (29%)	0 1

5 of 160 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	329	GLU
3	A	375	ASN
3	A	599	GLU
3	A	339	ARG
3	A	352	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
3	A	187	ASN
3	A	245	ASN
3	A	511	ASN
3	A	214	GLN
3	A	232	GLN

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

There are no carbohydrates in this entry.

5.6 Ligand geometry

1 ligand is 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	$\#\|Z\|>2$
4	ATP	A	725	-	26,33,33	1.54	6 (23%)	31,52,52	1.60	6 (19%)

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	ATP	A	725	-	-	11/18/38/38	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)
4	A	725	ATP	C8-N7	-3.51	1.28	1.34
4	A	725	ATP	C2-N1	2.31	1.38	1.33
4	A	725	ATP	PG-O2G	-2.28	1.46	1.54
4	A	725	ATP	O4'-C1'	-2.27	1.37	1.41
4	A	725	ATP	C5-N7	-2.21	1.31	1.39

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	725	ATP	PB-O3B-PG	-4.26	118.21	132.83
4	A	725	ATP	C4-C5-N7	3.76	113.31	109.40
4	A	725	ATP	PA-O3A-PB	-3.12	122.12	132.83
4	A	725	ATP	C5-C6-N1	-2.40	114.92	120.35
4	A	725	ATP	O2'-C2'-C3'	-2.25	104.54	111.82

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

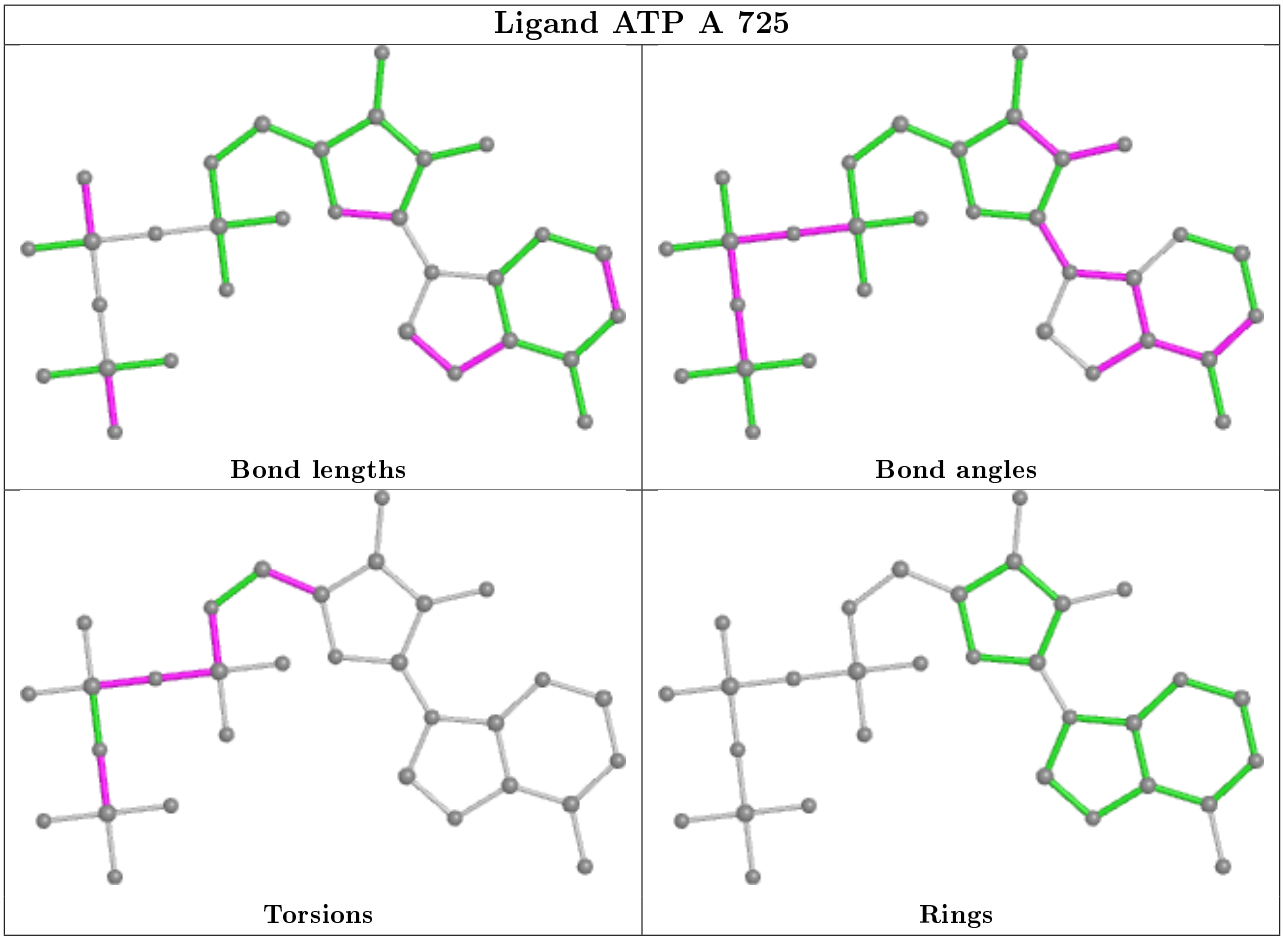
Mol	Chain	Res	Type	Atoms
4	A	725	ATP	PB-O3B-PG-O3G
4	A	725	ATP	C5'-O5'-PA-O1A
4	A	725	ATP	C5'-O5'-PA-O2A
4	A	725	ATP	C5'-O5'-PA-O3A
4	A	725	ATP	O4'-C4'-C5'-O5'

There are no ring outliers.

1 monomer is involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	725	ATP	7	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 then 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

There are no such residues in this entry.

5.8 Polymer linkage issues

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.