wwPDB X-ray Structure Validation Summary Report (i)

May 23, 2020 – 05:37 pm BST

PDB ID	:	2FGH
Title	:	ATP bound gelsolin
Authors	:	Ma, Q.; Robinson, R.C.; Burtnick, L.D.; Urosev, D.
Deposited on	:	2005-12-22
Resolution	:	2.80 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 (i) symbol.

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

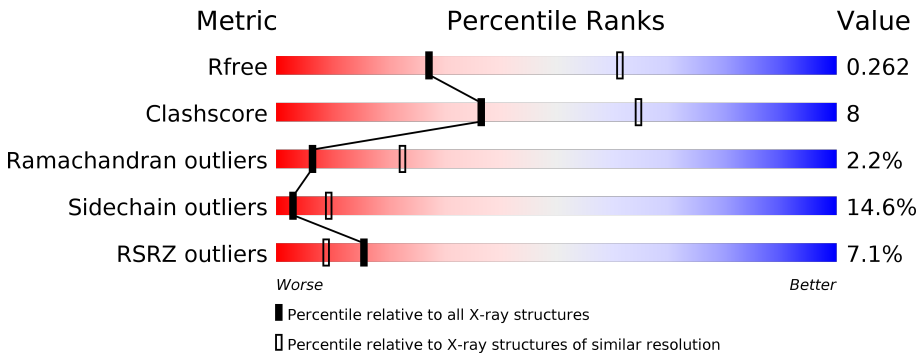
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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}	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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% 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	731	7% 74% 21% • ••
1	B	731	7% 72% 22% • ••

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 11378 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.

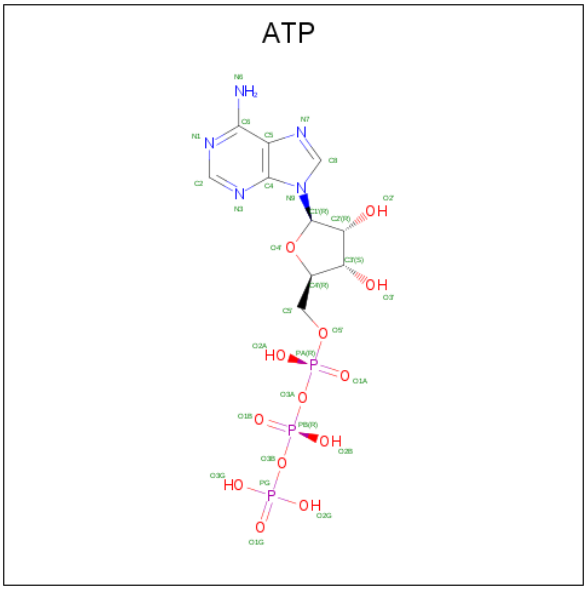
- Molecule 1 is a protein called gelsolin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	721	5633	3564	981	1072	16	0	0	0
1	B	721	5633	3564	981	1072	16	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	25	MET	-	INSERTION	UNP Q28372
B	25	MET	-	INSERTION	UNP Q28372

- Molecule 2 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		
2	A	1	31	10	5	13	3	0	0

Continued on next page...

Continued from previous page...

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

- Molecule 3 is water.

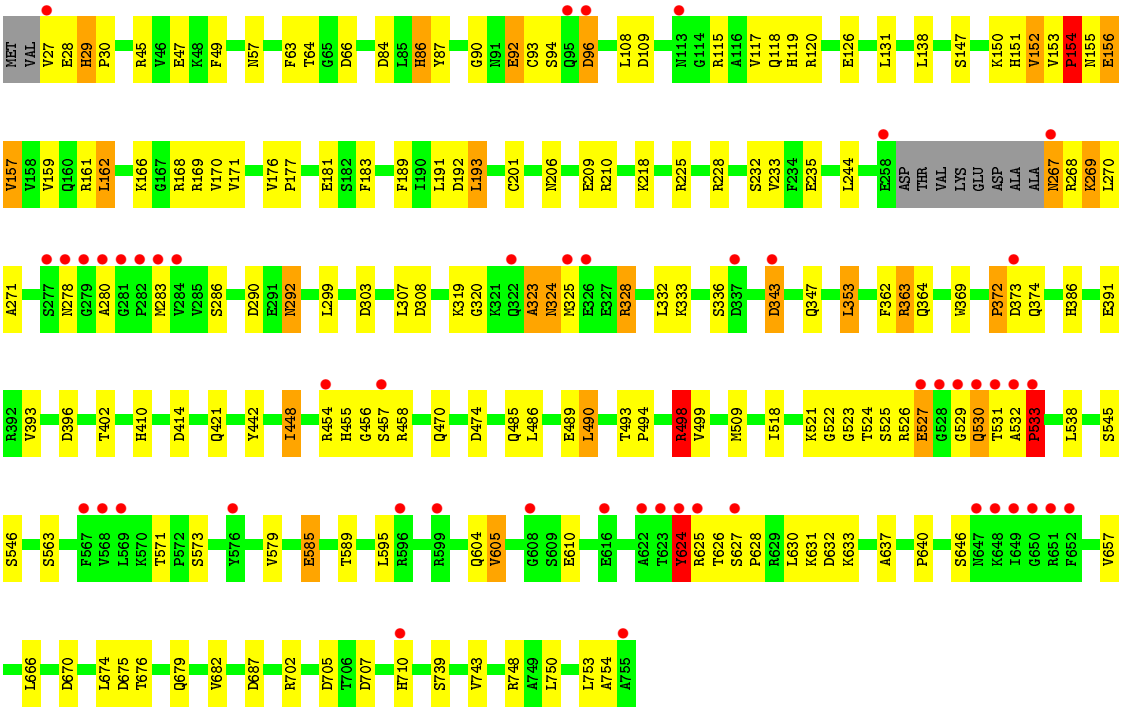
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	27	27	27	0	0
			Total	O		
3	B	23	23	23	0	0

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 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.

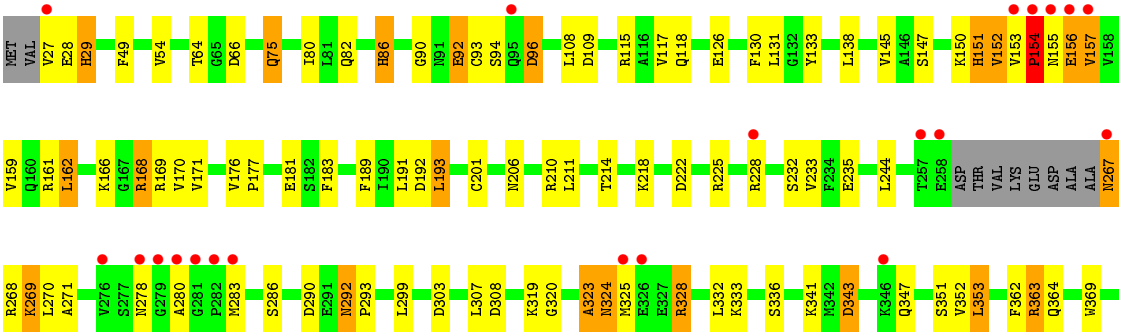
- Molecule 1: gelsolin

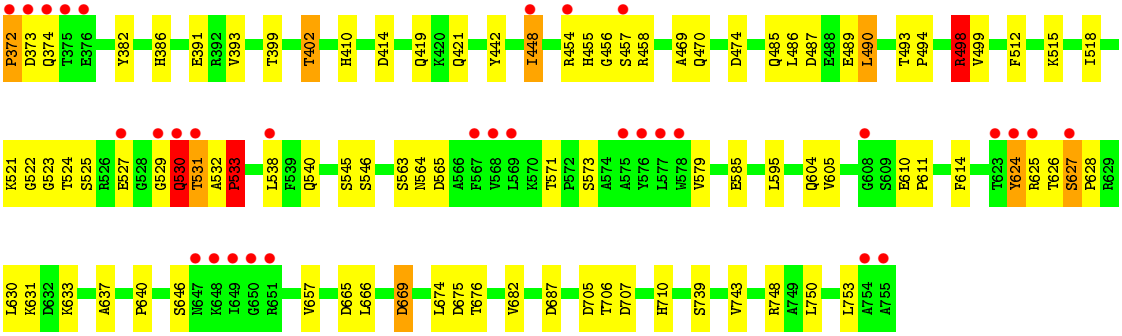
Chain A:



- Molecule 1: gelsolin

Chain B:




P372
D373
Q374
T375
E376
Y382
H386
E391
R392
V393
T399
T402
H410
D414
Q419
K420
Q421
Y442
I448
R454
H455
G456
S457
R458
A469
Q470
D474
Q485
L486
D487
E488
E489
L490
T493
P494
R498
V499
F512
K515
I518
K521
G522
G523
T524
S525
R526
E527
G528
G529
Q530
T531
A532
P533
L538
F539
Q540
S545
S546
S563
N564
D565
A566
F567
V568
L569
K570
T571
P572
S573
A574
A575
Y576
L577
W578
V579
E585
L595
Q604
V605
G608
S609
E610
P611
F614
T623
Y624
R625
T626
S627
P628
R629
L630
K631
D632
K633
A637
P640
S646
M647
K648
I649
G650
R651
V657
D665
L666
D669
L674
D675
T676
V682
D687
D705
T706
D707
H710
S739
V743
R748
A749
L750
L753
A754
A755

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 4 21 2	Depositor
Cell constants a, b, c, α, β, γ	167.34Å 167.34Å 149.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.80 19.99 – 2.80	Depositor EDS
% Data completeness (in resolution range)	98.6 (30.00-2.80) 98.9 (19.99-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.49 (at 2.79Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.242 , 0.265 0.241 , 0.262	Depositor DCC
R_{free} test set	2650 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	56.7	Xtriage
Anisotropy	0.069	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 46.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	11378	wwPDB-VP
Average B, all atoms (Å^2)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.87% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: 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	A	0.50	0/5766	0.81	15/7812 (0.2%)
1	B	0.51	0/5766	0.81	17/7812 (0.2%)
All	All	0.50	0/11532	0.81	32/15624 (0.2%)

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
1	A	0	2
1	B	0	1
All	All	0	3

There are no bond length outliers.

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	414	ASP	CB-CG-OD2	7.35	124.92	118.30
1	B	498	ARG	NE-CZ-NH1	6.92	123.76	120.30
1	A	474	ASP	CB-CG-OD2	6.54	124.18	118.30
1	A	498	ARG	NE-CZ-NH1	6.15	123.37	120.30
1	A	414	ASP	CB-CG-OD2	6.13	123.82	118.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	624	TYR	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	86	HIS	Sidechain
1	B	624	TYR	Peptide

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	5633	0	5493	94	1
1	B	5633	0	5493	84	1
2	A	31	0	12	1	0
2	B	31	0	12	1	0
3	A	27	0	0	22	0
3	B	23	0	0	9	0
All	All	11378	0	11010	177	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 8.

The worst 5 of 177 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:152:VAL:O	1:B:153:VAL:HG13	1.66	0.94
1:A:152:VAL:O	1:A:153:VAL:HG13	1.70	0.92
1:B:75:GLN:HG3	3:B:9387:HOH:O	1.69	0.92
1:B:627:SER:HA	3:B:9398:HOH:O	1.72	0.90
1:A:624:TYR:HA	3:A:5397:HOH:O	1.71	0.88

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:632:ASP:OD2	1:B:419:GLN:OE1[5_545]	2.18	0.02

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	717/731 (98%)	664 (93%)	37 (5%)	16 (2%)	6 22
1	B	717/731 (98%)	664 (93%)	38 (5%)	15 (2%)	7 23
All	All	1434/1462 (98%)	1328 (93%)	75 (5%)	31 (2%)	6 22

5 of 31 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	29	HIS
1	A	280	ALA
1	A	324	ASN
1	A	456	GLY
1	A	527	GLU

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
1	A	591/599 (99%)	506 (86%)	85 (14%)	3 10
1	B	591/599 (99%)	503 (85%)	88 (15%)	3 9
All	All	1182/1198 (99%)	1009 (85%)	173 (15%)	3 9

5 of 173 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	674	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	154	PRO
1	B	624	TYR
1	A	739	SER
1	B	86	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	564	ASN
1	B	91	ASN
1	B	507	HIS
1	A	647	ASN
1	B	118	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)

2 ligands 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	$\#\|Z\| > 2$
2	ATP	B	9380	-	26,33,33	1.22	3 (11%)	31,52,52	1.95	4 (12%)
2	ATP	A	5380	-	26,33,33	1.01	2 (7%)	31,52,52	1.94	4 (12%)

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
2	ATP	B	9380	-	-	5/18/38/38	0/3/3/3
2	ATP	A	5380	-	-	4/18/38/38	0/3/3/3

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	9380	ATP	C2-N3	3.65	1.38	1.32
2	A	5380	ATP	C2-N3	3.25	1.37	1.32
2	B	9380	ATP	C2-N1	2.84	1.39	1.33
2	B	9380	ATP	O4'-C1'	2.69	1.44	1.41
2	A	5380	ATP	C2-N1	2.24	1.38	1.33

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	9380	ATP	PA-O3A-PB	-6.52	110.44	132.83
2	A	5380	ATP	PA-O3A-PB	-5.75	113.11	132.83
2	A	5380	ATP	N3-C2-N1	-4.99	120.88	128.68
2	B	9380	ATP	N3-C2-N1	-4.69	121.35	128.68
2	A	5380	ATP	C1'-N9-C4	-4.28	119.12	126.64

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	9380	ATP	C3'-C4'-C5'-O5'
2	B	9380	ATP	O4'-C4'-C5'-O5'
2	A	5380	ATP	O4'-C4'-C5'-O5'
2	A	5380	ATP	C3'-C4'-C5'-O5'
2	B	9380	ATP	PB-O3B-PG-O1G

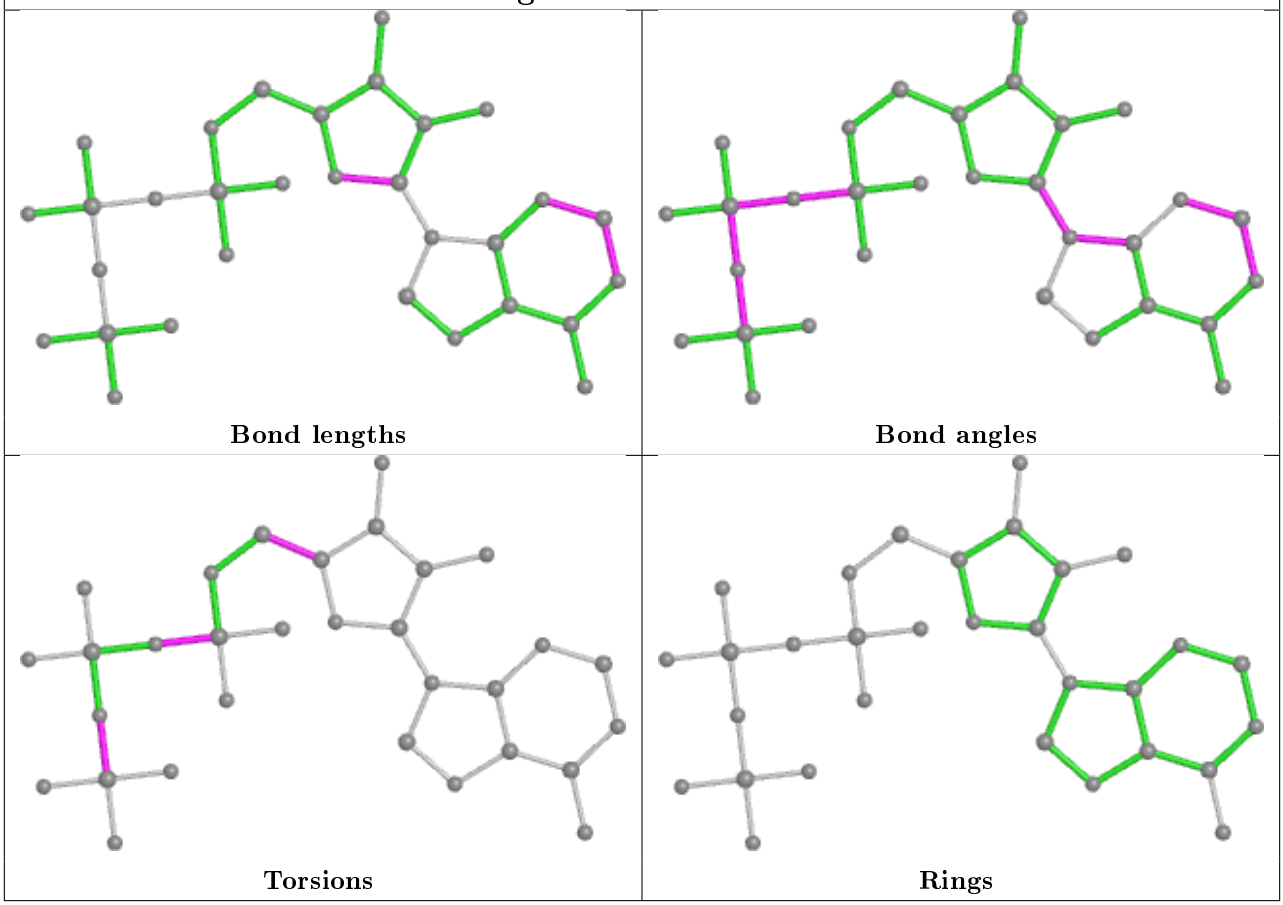
There are no ring outliers.

2 monomers are involved in 2 short contacts:

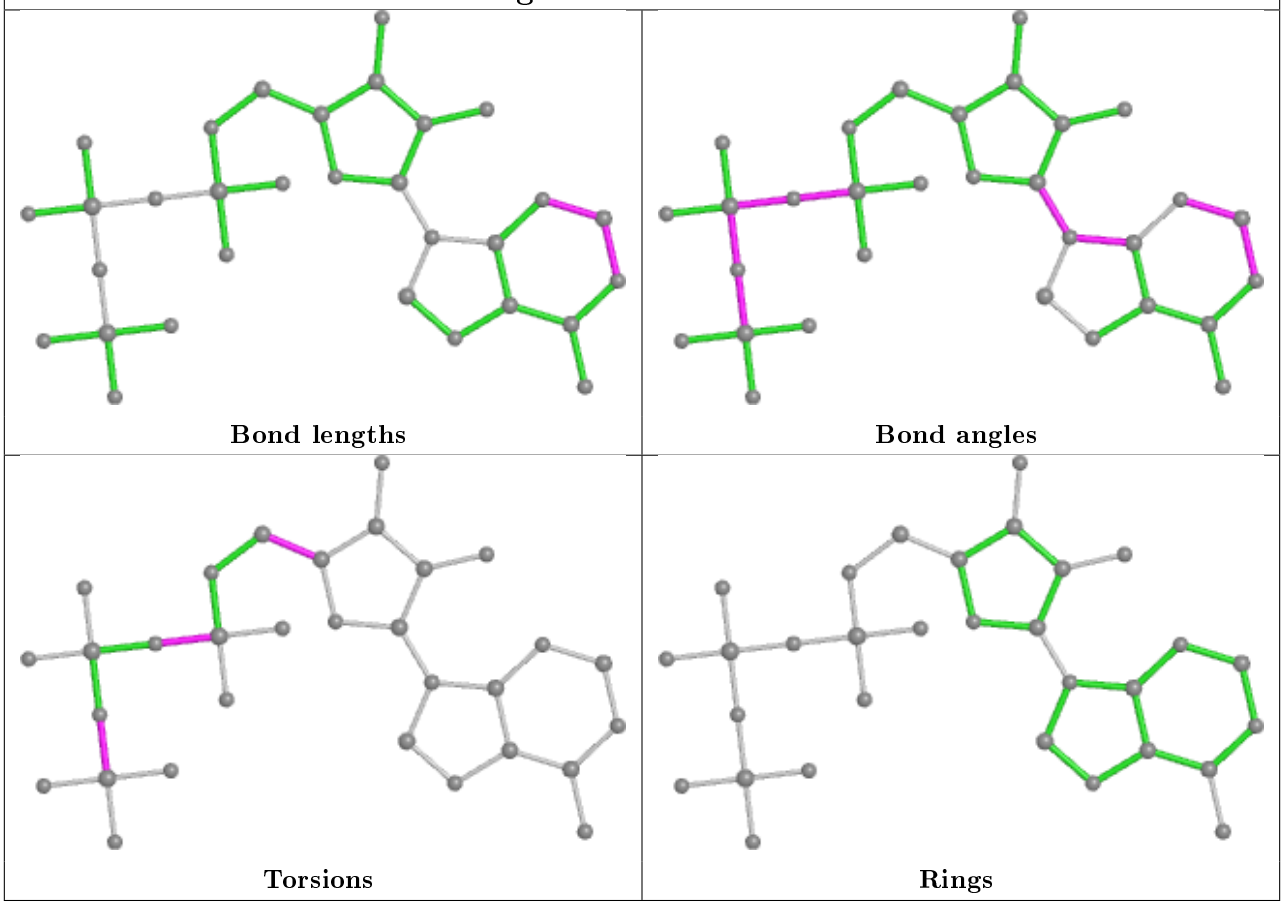
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	9380	ATP	1	0
2	A	5380	ATP	1	0

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.

Ligand ATP B 9380



Ligand ATP A 5380



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	721/731 (98%)	0.14	50 (6%)	16	10	28, 54, 110, 114	0
1	B	721/731 (98%)	0.21	53 (7%)	14	8	28, 54, 110, 114	0
All	All	1442/1462 (98%)	0.18	103 (7%)	16	9	28, 54, 110, 114	0

The worst 5 of 103 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	280	ALA	11.0
1	A	531	THR	7.9
1	B	281	GLY	7.4
1	B	27	VAL	7.4
1	A	529	GLY	7.4

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)

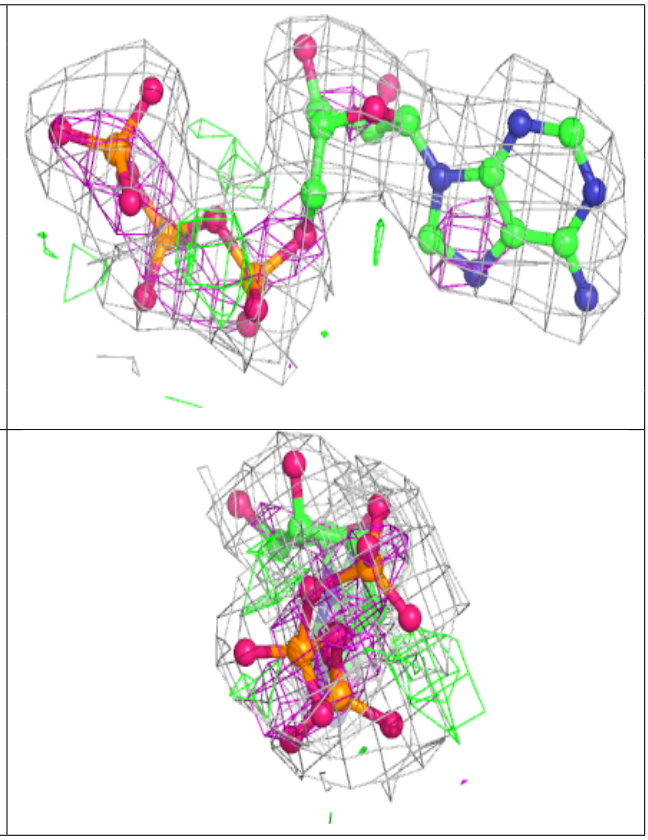
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	ATP	B	9380	31/31	0.96	0.15	20,26,33,34	0
2	ATP	A	5380	31/31	0.96	0.15	20,26,33,34	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

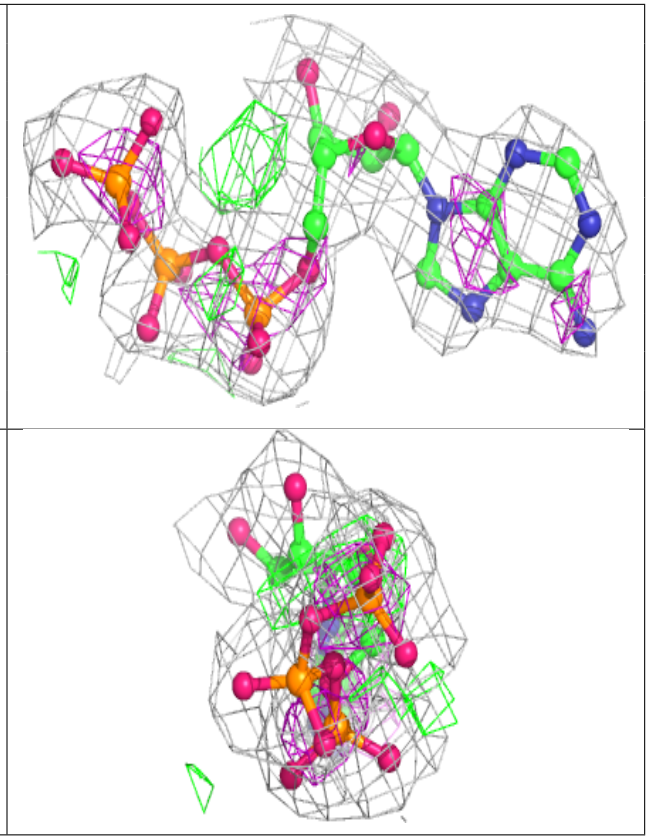
Electron density around ATP B 9380:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around ATP A 5380:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



6.5 Other polymers (i)

There are no such residues in this entry.