WORLDWIDE PDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Dec 10, 2024 – 12:03 PM EST

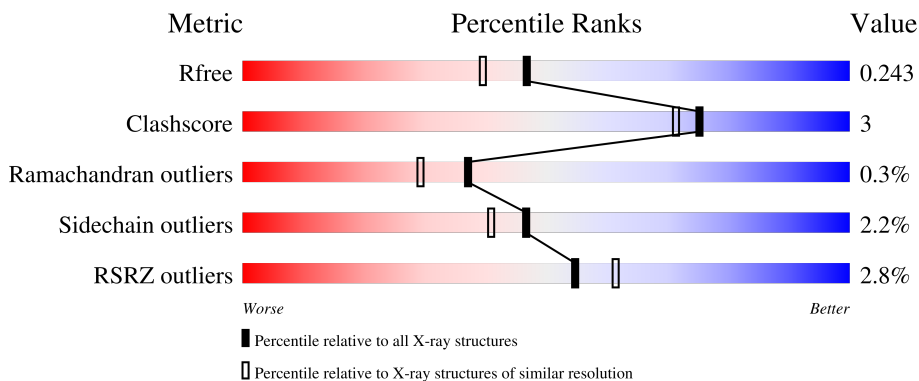|		
PDB ID	:	9E31
Title	:	Discovery of Potent, Highly Selective and Efficacious SMARCA2 Degraders - Compound 6
Authors	:	Strickland, C.; Rice, C.
Deposited on	:	2024-10-23
Resolution	:	1.96 Å(reported)

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 1.96 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	Percentile Ranks	Value
Rfree		0.243
Clashscore		3
Ramachandran outliers		0.3%
Sidechain outliers		2.2%
RSRZ outliers		2.8%

Worse ← → *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	3187 (1.96-1.96)
Clashscore	180529	3412 (1.96-1.96)
Ramachandran outliers	177936	3390 (1.96-1.96)
Sidechain outliers	177891	3390 (1.96-1.96)
RSRZ outliers	164620	3186 (1.96-1.96)

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 of 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	123	2% / 85% / 10% · ·
1	B	123	3% / 87% / 6% · 7%
1	C	123	2% / 90% / 6% · ·

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 3167 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.
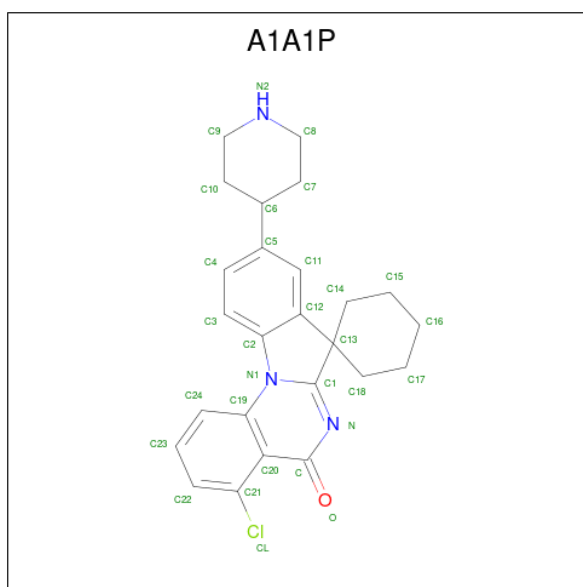
- Molecule 1 is a protein called Isoform Short of Probable global transcription activator SNF2L2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	118	Total	C	N	O	S	0	0	0
			960	609	167	181	3			
1	B	115	Total	C	N	O	S	0	0	0
			937	596	164	174	3			
1	C	119	Total	C	N	O	S	0	0	0
			973	616	169	185	3			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1371	SER	-	expression tag	UNP P51531
A	1372	MET	-	expression tag	UNP P51531
B	1371	SER	-	expression tag	UNP P51531
B	1372	MET	-	expression tag	UNP P51531
C	1371	SER	-	expression tag	UNP P51531
C	1372	MET	-	expression tag	UNP P51531

- Molecule 2 is (12'R)-4'-chloro-9'-(piperidin-4-yl)-5'H-spiro[cyclohexane-1,7'-indolo[1,2-a]quinazolin]-5'-one (three-letter code: A1A1P) (formula: $C_{25}H_{26}ClN_3O$) (labeled as "Ligand of Interest" by depositor).

A1A1P

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	Cl	N	O	0	0
			30	25	1	3	1		
2	B	1	Total	C	Cl	N	O	0	0
			30	25	1	3	1		
2	C	1	Total	C	Cl	N	O	0	0
			30	25	1	3	1		

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Zn	0	0
			1	1		
3	B	1	Total	Zn	0	0
			1	1		
3	C	2	Total	Zn	0	0
			2	2		

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).

ACT

OXT

O⁻

C

CH3

O

O

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 4	C 2	O 2	0	0
4	B	1	Total 4	C 2	O 2	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	76	Total 76	O 76	0	0
5	B	62	Total 62	O 62	0	0
5	C	57	Total 57	O 57	0	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 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.
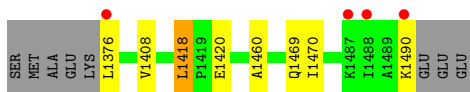
• Molecule 1: Isoform Short of Probable global transcription activator SNF2L2
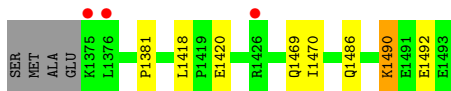
Chain A:

• Molecule 1: Isoform Short of Probable global transcription activator SNF2L2

Chain B:

• Molecule 1: Isoform Short of Probable global transcription activator SNF2L2

Chain C:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α, β, γ	66.03Å 66.03Å 88.81Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.08 – 1.96 48.08 – 1.96	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.08-1.96) 99.8 (48.08-1.96)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.55 (at 1.97Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, R_{free}	0.195 , 0.238 0.203 , 0.243	Depositor DCC
R_{free} test set	1359 reflections (4.36%)	wwPDB-VP
Wilson B-factor (Å^2)	25.5	Xtriage
Anisotropy	0.034	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 26.7	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.002 for -h,-k,l 0.039 for h,-h-k,-l 0.020 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	3167	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.42% 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: ZN, ACT, A1A1P

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	0.44	0/976	0.91	5/1313 (0.4%)
1	B	0.44	0/953	0.83	0/1282
1	C	0.43	0/989	0.76	0/1328
All	All	0.44	0/2918	0.84	5/3923 (0.1%)

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	1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1439	ARG	NE-CZ-NH1	10.04	125.32	120.30
1	A	1415	ARG	NE-CZ-NH1	-6.54	117.03	120.30
1	A	1439	ARG	CD-NE-CZ	5.66	131.53	123.60
1	A	1415	ARG	NE-CZ-NH2	5.51	123.06	120.30
1	A	1485	ARG	NE-CZ-NH2	-5.50	117.55	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1415	ARG	Sidechain

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	960	0	973	5	0
1	B	937	0	959	7	0
1	C	973	0	988	5	0
2	A	30	0	0	0	0
2	B	30	0	0	2	0
2	C	30	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	2	0	0	0	0
4	A	4	0	3	0	0
4	B	4	0	3	0	0
5	A	76	0	0	2	0
5	B	62	0	0	2	0
5	C	57	0	0	2	0
All	All	3167	0	2926	17	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 3.

The worst 5 of 17 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:1408:VAL:HG11	1:B:1470:ILE:CD1	2.16	0.76
1:B:1469:GLN:HG2	5:B:1657:HOH:O	1.92	0.69
1:B:1408:VAL:HG11	1:B:1470:ILE:HD13	1.85	0.57
1:A:1444:ARG:HD2	5:A:1640:HOH:O	2.03	0.57
1:C:1470:ILE:HD13	5:C:1631:HOH:O	2.08	0.53

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	116/123 (94%)	115 (99%)	1 (1%)	0	100	100
1	B	113/123 (92%)	113 (100%)	0	0	100	100
1	C	117/123 (95%)	116 (99%)	0	1 (1%)	14	7
All	All	346/369 (94%)	344 (99%)	1 (0%)	1 (0%)	37	29

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	1490	LYS

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	108/115 (94%)	104 (96%)	4 (4%)	29	19
1	B	106/115 (92%)	104 (98%)	2 (2%)	52	47
1	C	110/115 (96%)	109 (99%)	1 (1%)	75	75
All	All	324/345 (94%)	317 (98%)	7 (2%)	47	41

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1492	GLU
1	B	1376	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	1492	GLU
1	B	1418	LEU
1	A	1490	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	1486	GLN
1	C	1478	GLN
1	B	1478	GLN
1	B	1469	GLN
1	B	1486	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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 9 ligands modelled in this entry, 4 are monoatomic - leaving 5 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	$\#|Z| > 2$
2	A1A1P	B	1501	-	34,35,35	0.57	0	42,53,53	1.51	6 (14%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	ACT	B	1503	3	3,3,3	1.27	0	3,3,3	0.69	0
2	A1A1P	C	1501	-	34,35,35	0.71	1 (2%)	42,53,53	1.44	7 (16%)
4	ACT	A	1503	3	3,3,3	1.57	0	3,3,3	0.68	0
2	A1A1P	A	1501	-	34,35,35	0.68	0	42,53,53	1.50	8 (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
2	A1A1P	A	1501	-	-	0/4/38/38	0/6/6/6
2	A1A1P	C	1501	-	-	0/4/38/38	0/6/6/6
2	A1A1P	B	1501	-	-	2/4/38/38	0/6/6/6

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1501	A1A1P	C1-N1	2.20	1.39	1.37

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1501	A1A1P	N1-C1-N	-4.72	122.46	125.74
2	A	1501	A1A1P	C19-C20-C	-4.52	115.59	119.84
2	C	1501	A1A1P	C19-C20-C	-3.24	116.80	119.84
2	B	1501	A1A1P	C19-C20-C	-3.13	116.89	119.84
2	A	1501	A1A1P	O-C-N	-3.07	115.56	120.23

There are no chirality outliers.
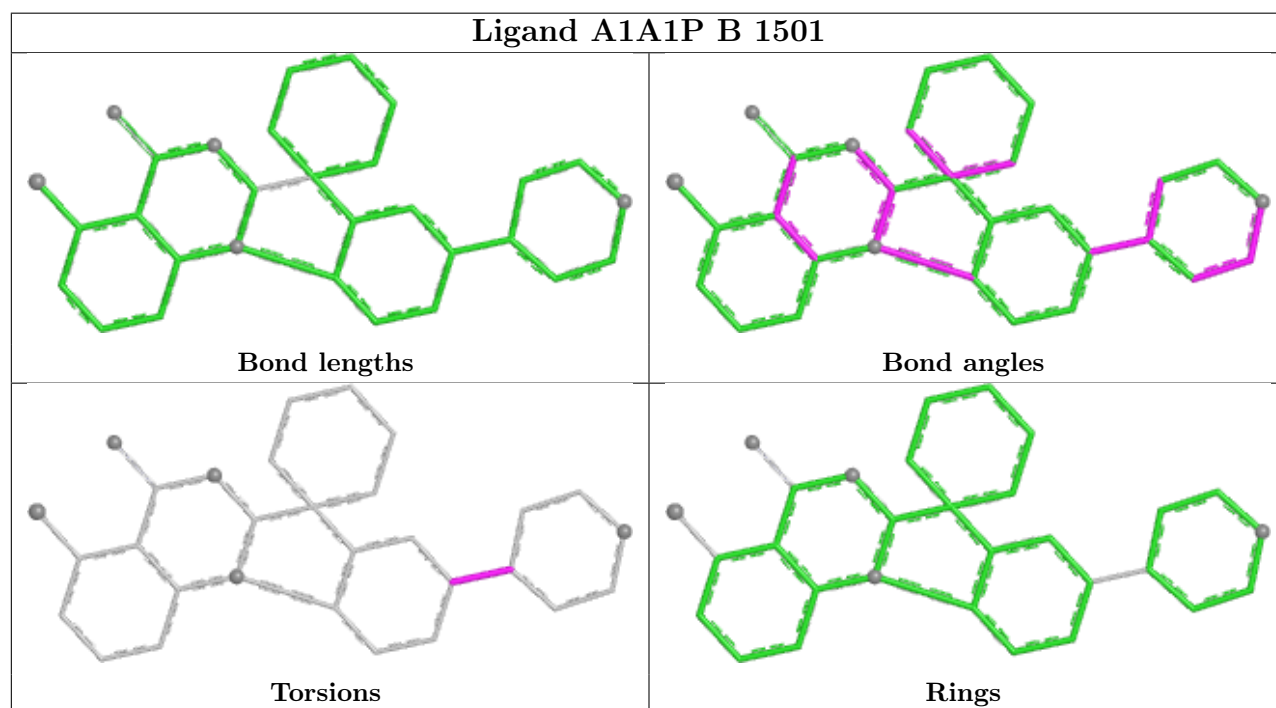
All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1501	A1A1P	C4-C5-C6-C10
2	B	1501	A1A1P	C11-C5-C6-C10

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1501	A1A1P	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 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 A1A1P B 1501

Bond lengths

Bond angles

Torsions

Rings

Ligand A1A1P C 1501

Bond lengths

Bond angles

Torsions

Rings

Ligand A1A1P A 1501

Bond lengths

Bond angles

Torsions

Rings
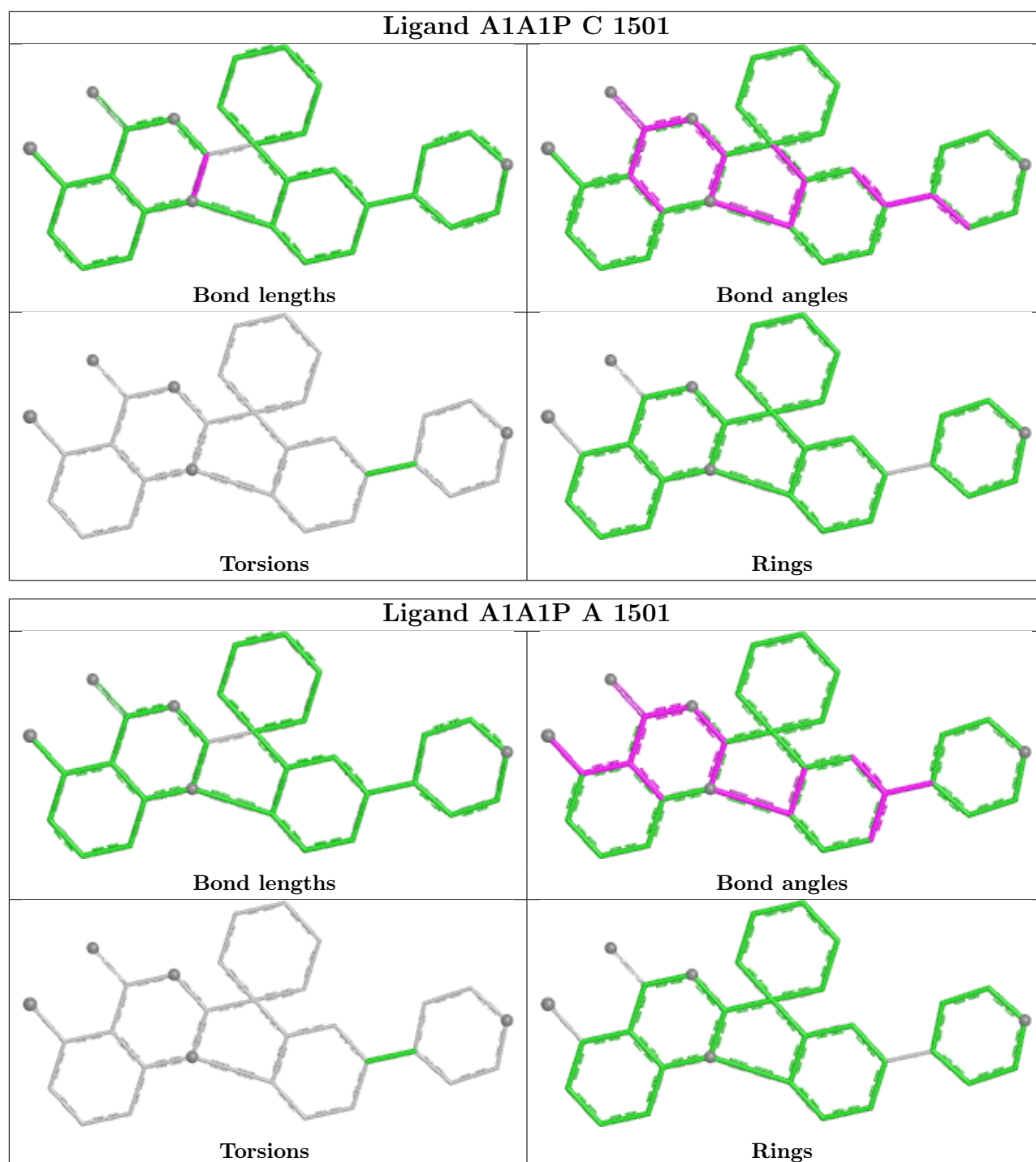
5.7 Other polymers ⓘ

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 ⓘ

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(Å²)	Q<0.9
1	A	118/123 (95%)	-0.12	3 (2%)	58	65	18, 25, 41, 83	0
1	B	115/123 (93%)	0.05	4 (3%)	47	54	20, 28, 49, 60	0
1	C	119/123 (96%)	0.16	3 (2%)	58	65	22, 29, 47, 82	0
All	All	352/369 (95%)	0.03	10 (2%)	55	61	18, 28, 47, 83	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	1376	LEU	4.0
1	C	1375	LYS	3.7
1	B	1376	LEU	3.6
1	B	1488	ILE	3.4
1	B	1490	LYS	3.0

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

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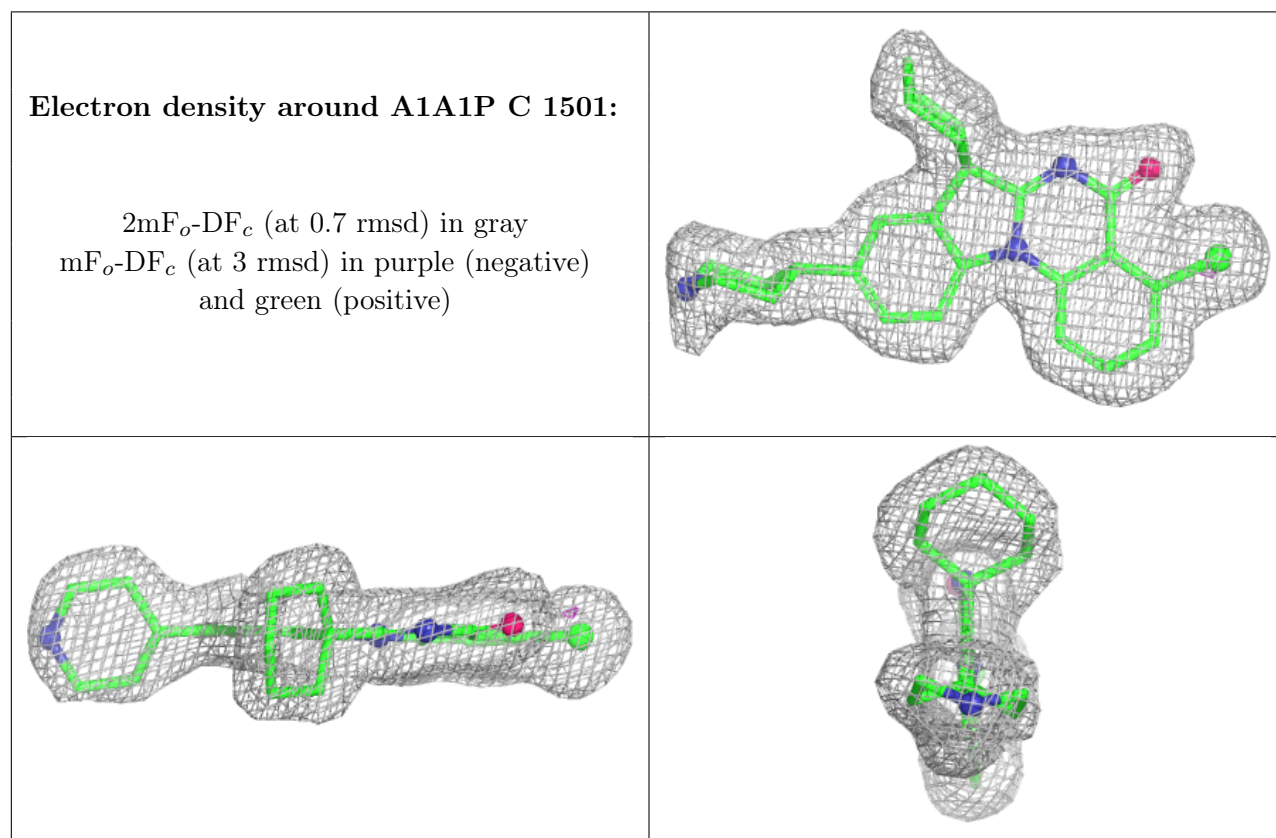
6.3   Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4   Ligands ⓘ

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	A1A1P	C	1501	30/30	0.95	0.08	22,28,40,41	0
4	ACT	A	1503	4/4	0.95	0.08	23,24,26,28	0
2	A1A1P	A	1501	30/30	0.96	0.08	17,21,38,40	0
2	A1A1P	B	1501	30/30	0.96	0.09	18,23,48,52	0
4	ACT	B	1503	4/4	0.97	0.08	26,30,31,33	0
3	ZN	C	1503	1/1	0.99	0.03	48,48,48,48	0
3	ZN	B	1502	1/1	0.99	0.02	26,26,26,26	0
3	ZN	C	1502	1/1	0.99	0.03	35,35,35,35	0
3	ZN	A	1502	1/1	1.00	0.01	24,24,24,24	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.
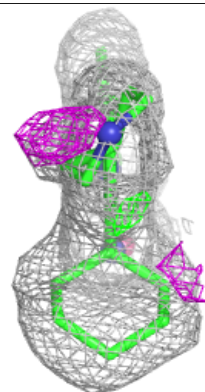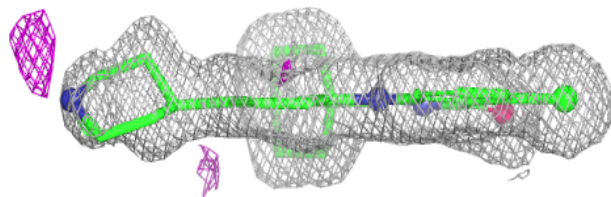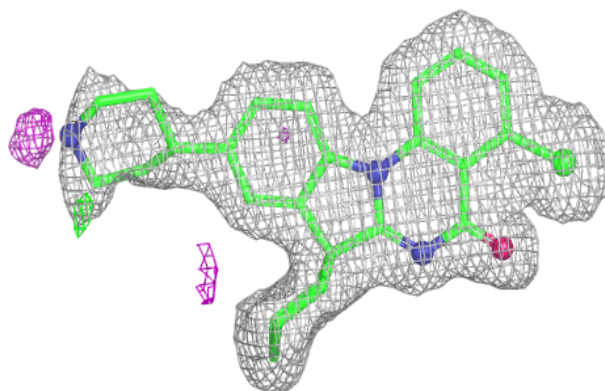
Electron density around A1A1P C 1501:

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 A1A1P A 1501:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
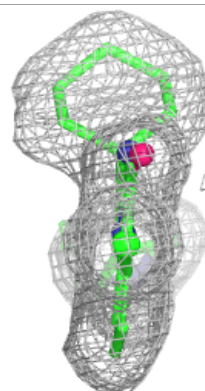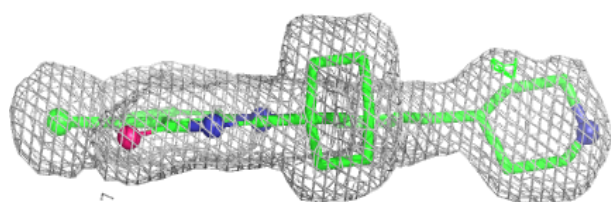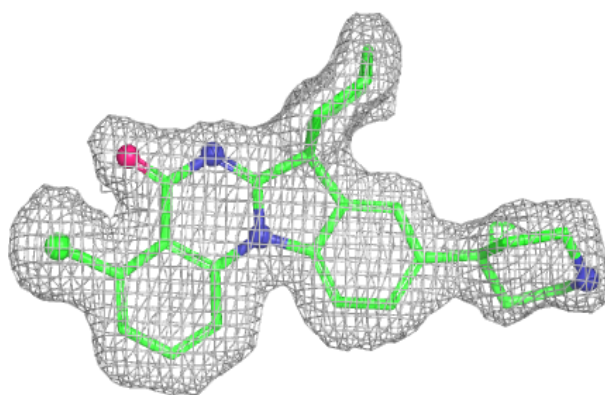
Electron density around A1A1P B 1501:

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

6.5 Other polymers ⓘ

There are no such residues in this entry.