wwPDB X-ray Structure Validation Summary Report (i)

Jan 27, 2022 – 06:16 pm GMT

PDB ID	:	7P3V
Title	:	B-Raf V600E structure bound to a new inhibitor
Authors	:	Schneider, M.; Gelin, M.; Cohen-Gonsaud, M.; Labesse, G.
Deposited on	:	2021-07-08
Resolution	:	2.37 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:

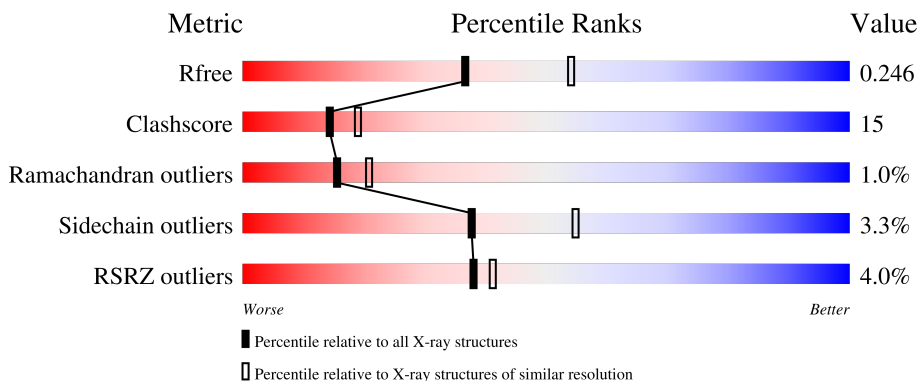|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.26
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.26

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 2.37 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		0.246
Clashscore		15
Ramachandran outliers		1.0%
Sidechain outliers		3.3%
RSRZ outliers		4.0%

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}	130704	5509 (2.40-2.36)
Clashscore	141614	6082 (2.40-2.36)
Ramachandran outliers	138981	5973 (2.40-2.36)
Sidechain outliers	138945	5975 (2.40-2.36)
RSRZ outliers	127900	5397 (2.40-2.36)

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	275	5% 67% 25% • 6%
1	B	275	3% 69% 28% ••

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 4428 atoms, of which 38 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 Serine/threonine-protein kinase B-raf.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	258	Total 2046	C 1298	N 359	O 376	S 13	0	0	0
1	B	272	Total 2145	C 1358	N 380	O 394	S 13	0	0	0

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	447	PHE	-	expression tag	UNP P15056
A	543	ALA	ILE	conflict	UNP P15056
A	544	SER	ILE	conflict	UNP P15056
A	551	LYS	ILE	conflict	UNP P15056
A	562	ARG	GLN	conflict	UNP P15056
A	588	ASN	LEU	conflict	UNP P15056
A	600	GLU	VAL	engineered mutation	UNP P15056
A	630	SER	LYS	conflict	UNP P15056
A	667	GLU	PHE	conflict	UNP P15056
A	673	SER	TYR	conflict	UNP P15056
A	688	ARG	ALA	conflict	UNP P15056
A	706	SER	LEU	conflict	UNP P15056
A	709	ARG	GLN	conflict	UNP P15056
A	713	GLU	SER	conflict	UNP P15056
A	716	GLU	LEU	conflict	UNP P15056
A	720	GLU	-	expression tag	UNP P15056
A	721	ALA	-	expression tag	UNP P15056
B	447	PHE	-	expression tag	UNP P15056
B	543	ALA	ILE	conflict	UNP P15056
B	544	SER	ILE	conflict	UNP P15056
B	551	LYS	ILE	conflict	UNP P15056
B	562	ARG	GLN	conflict	UNP P15056
B	588	ASN	LEU	conflict	UNP P15056
B	600	GLU	VAL	engineered mutation	UNP P15056
B	630	SER	LYS	conflict	UNP P15056

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	667	GLU	PHE	conflict	UNP P15056
B	673	SER	TYR	conflict	UNP P15056
B	688	ARG	ALA	conflict	UNP P15056
B	706	SER	LEU	conflict	UNP P15056
B	709	ARG	GLN	conflict	UNP P15056
B	713	GLU	SER	conflict	UNP P15056
B	716	GLU	LEU	conflict	UNP P15056
B	720	GLU	-	expression tag	UNP P15056
B	721	ALA	-	expression tag	UNP P15056

- Molecule 2 is {N}-[3-[5-(2-azanylpyrimidin-4-yl)-2-[(3 {S})-morpholin-3-yl]-1,3-thiazol-4-yl]-2-fluoranyl-phenyl]-2,5-bis(fluoranyl)benzenesulfonamide (three-letter code: 5I4) (formula: $C_{23}H_{19}F_3N_6O_3S_2$).
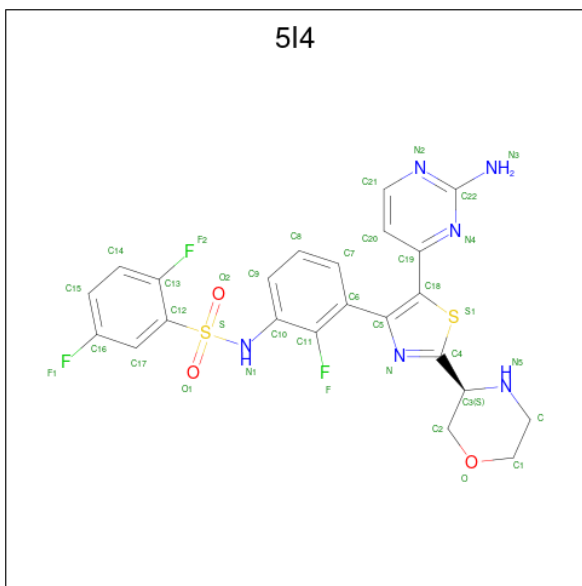
Mol	Chain	Residues	Atoms							ZeroOcc	AltConf
2	A	1	Total	C	F	H	N	O	S	0	0
			56	23	3	19	6	3	2		
2	B	1	Total	C	F	H	N	O	S	0	0
			56	23	3	19	6	3	2		

- Molecule 3 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	74	Total	O	0	0
			74	74		
3	B	51	Total	O	0	0
			51	51		

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: Serine/threonine-protein kinase B-raf
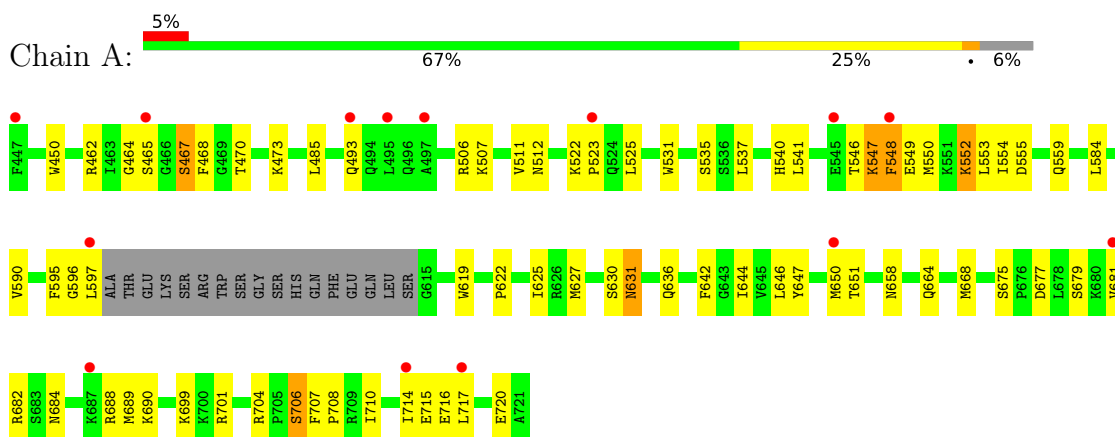
• Molecule 1: Serine/threonine-protein kinase B-raf
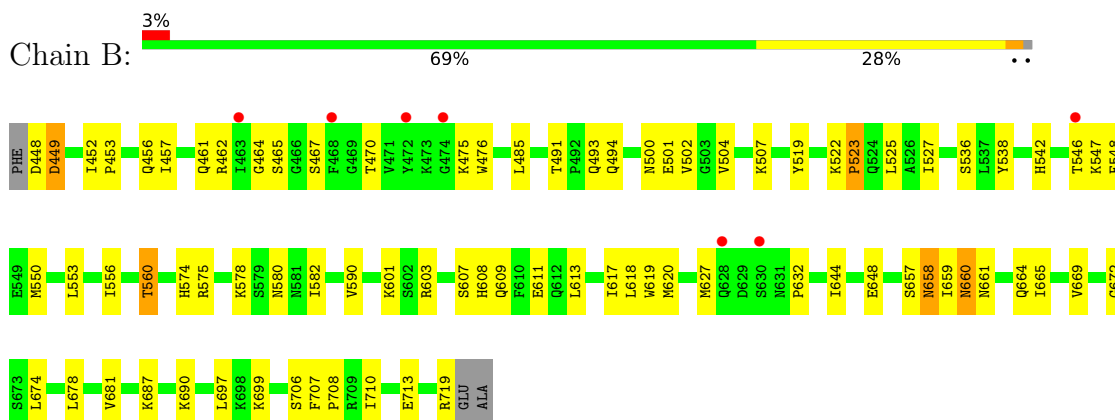
4   Data and refinement statistics   ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	52.63Å 104.85Å 109.56Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.55 – 2.37 48.55 – 2.37	Depositor EDS
% Data completeness (in resolution range)	99.7 (48.55-2.37) 99.8 (48.55-2.37)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.09 (at 2.37Å)	Xtriage
Refinement program	PHENIX 1.19.1_4122, PHENIX 1.19.1_4122	Depositor
R, R_{free}	0.195 , 0.247 0.194 , 0.246	Depositor DCC
R_{free} test set	2002 reflections (7.92%)	wwPDB-VP
Wilson B-factor (Å^2)	56.4	Xtriage
Anisotropy	0.423	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.021 for -h,l,k	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	4428	wwPDB-VP
Average B, all atoms (Å^2)	69.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.48% 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 ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 5I4

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	0.45	0/2089	0.63	0/2819
1	B	0.43	0/2192	0.62	0/2961
All	All	0.44	0/4281	0.62	0/5780

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	2046	0	2032	72	2
1	B	2145	0	2116	66	2
2	A	37	19	0	1	0
2	B	37	19	0	1	0
3	A	74	0	0	3	0
3	B	51	0	0	3	0
All	All	4390	38	4148	130	2

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 15.

The worst 5 of 130 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:548:PHE:HB3	1:A:552:LYS:HG3	1.21	1.12
1:B:681:VAL:HG21	1:B:690:LYS:HD2	1.33	1.08
1:A:664:GLN:O	1:A:668:MET:HG3	1.59	1.01
1:A:710:ILE:O	1:A:714:ILE:HG22	1.68	0.93
1:A:714:ILE:HA	1:A:717:LEU:HD21	1.54	0.90

All (2) 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:715:GLU:OE1	1:B:475:LYS:NZ[2_555]	2.12	0.08
1:A:507:LYS:NZ	1:B:449:ASP:OD2[2_555]	2.15	0.05

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	254/275 (92%)	233 (92%)	19 (8%)	2 (1%)	19	27
1	B	270/275 (98%)	248 (92%)	19 (7%)	3 (1%)	14	18
All	All	524/550 (95%)	481 (92%)	38 (7%)	5 (1%)	15	21

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	658	ASN
1	B	523	PRO
1	B	657	SER
1	A	631	ASN
1	A	548	PHE

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	222/241 (92%)	217 (98%)	5 (2%)	50	68
1	B	231/241 (96%)	221 (96%)	10 (4%)	29	43
All	All	453/482 (94%)	438 (97%)	15 (3%)	38	55

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	467	SER
1	B	660	ASN
1	B	536	SER
1	B	687	LYS
1	B	607	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	500	ASN
1	A	581	ASN
1	B	608	HIS

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

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	5I4	B	801	-	37,41,41	4.22	19 (51%)	42,60,60	3.11	14 (33%)
2	5I4	A	801	-	37,41,41	4.26	18 (48%)	42,60,60	2.83	13 (30%)

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	5I4	B	801	-	-	2/13/31/31	1/5/5/5
2	5I4	A	801	-	-	3/13/31/31	1/5/5/5

The worst 5 of 37 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	801	5I4	C13-C12	-11.74	1.31	1.39
2	B	801	5I4	C13-C12	-9.87	1.32	1.39
2	B	801	5I4	C17-C12	8.77	1.52	1.39
2	A	801	5I4	C15-C16	8.35	1.53	1.37
2	A	801	5I4	C17-C12	8.04	1.51	1.39

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	801	5I4	O2-S-O1	-12.18	104.58	119.55
2	B	801	5I4	O2-S-O1	-12.01	104.79	119.55
2	B	801	5I4	C5-C6-C11	-9.03	112.26	123.56
2	A	801	5I4	C5-C6-C11	-7.33	114.38	123.56
2	B	801	5I4	C12-S-N1	5.63	113.85	107.27

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	801	5I4	N-C5-C6-C7
2	B	801	5I4	N-C5-C6-C7
2	A	801	5I4	C13-C12-S-O1
2	A	801	5I4	C10-N1-S-O1
2	B	801	5I4	C13-C12-S-O1

All (2) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	801	5I4	C-C1-C2-C3-N5-O
2	B	801	5I4	C-C1-C2-C3-N5-O

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	801	5I4	1	0
2	A	801	5I4	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 5I4 B 801

Bond lengths

Bond angles

Torsions

Rings

Ligand 5I4 A 801

Bond lengths

Bond angles

Torsions

Rings
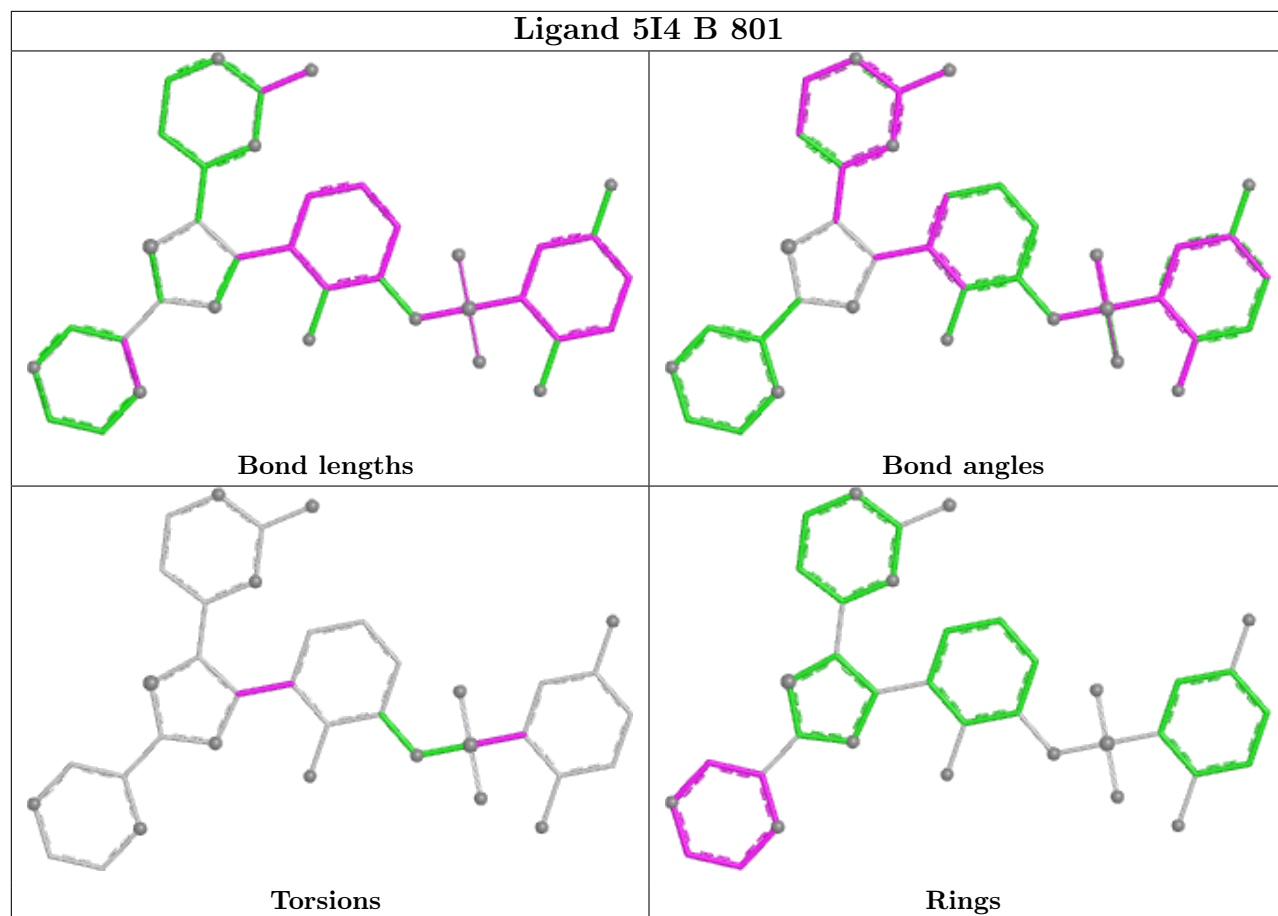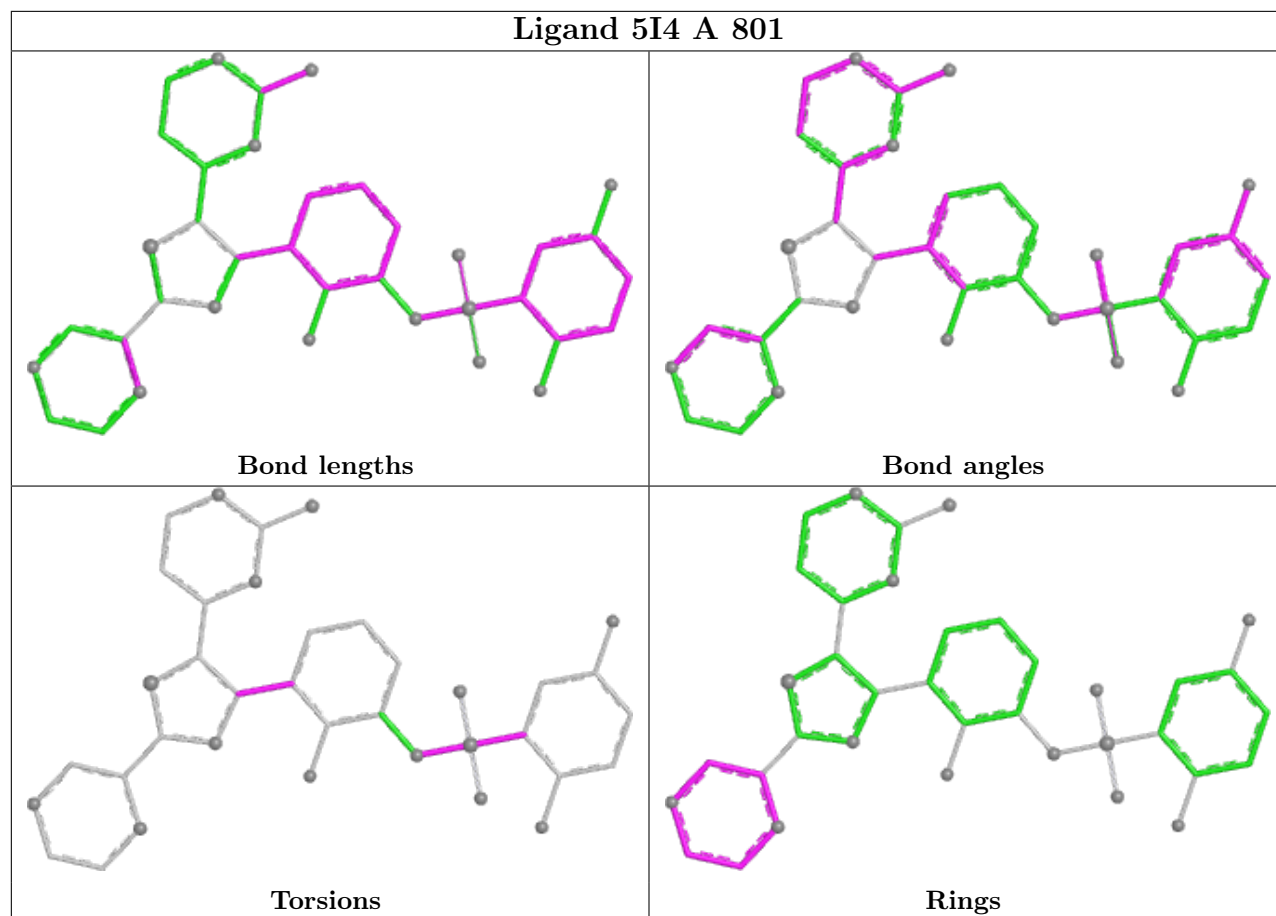
5.7 Other polymers (i)

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	258/275 (93%)	0.36	14 (5%)	25	28	42, 64, 109, 124	12 (4%)
1	B	272/275 (98%)	0.44	7 (2%)	56	57	44, 68, 107, 148	16 (5%)
All	All	530/550 (96%)	0.40	21 (3%)	38	41	42, 65, 109, 148	28 (5%)

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	447	PHE	3.7
1	A	495	LEU	3.6
1	A	465	SER	2.8
1	B	474	GLY	2.8
1	A	545	GLU	2.7

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

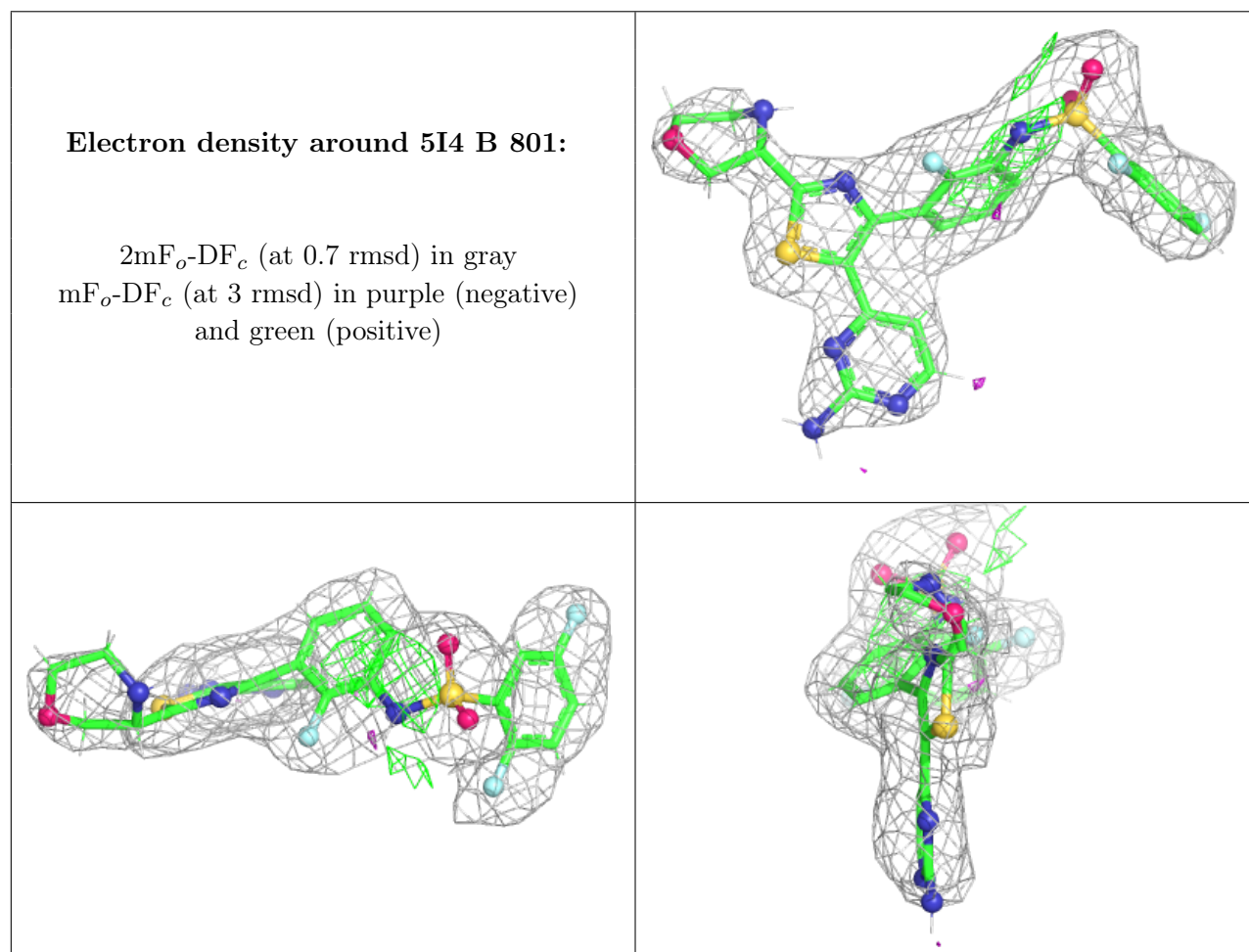There are no non-standard protein/DNA/RNA residues in this entry.

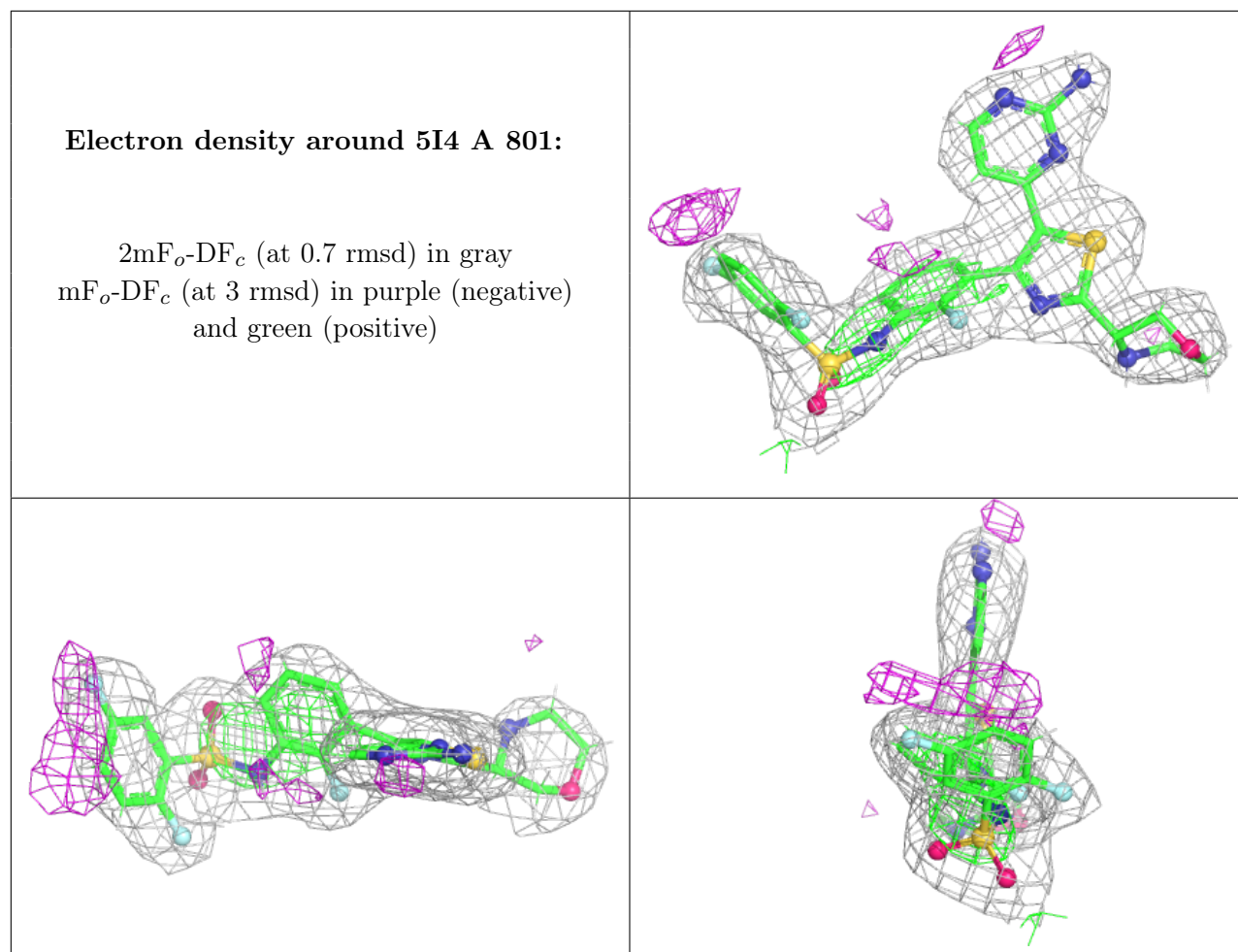## 6.3 Carbohydrates ⓘ

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	5I4	B	801	37/37	0.92	0.19	50,64,88,94	5
2	5I4	A	801	37/37	0.94	0.17	37,50,71,73	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 5I4 B 801:

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

Electron density around 5I4 A 801:

$2mF_o$-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 ⓘ

There are no such residues in this entry.