Full wwPDB X-ray Structure Validation Report ⓘ

Sep 8, 2022 – 01:09 pm BST

PDB ID	:	7OVI
Title	:	Protein kinase MKK7 in complex with phenethyltriazole-substituted pyrazolopyrimidine
Authors	:	Kleinboelting, S.; Buehrmann, M.; Mueller, M.P.; Rauh, D.
Deposited on	:	2021-06-15
Resolution	:	1.95 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ 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:

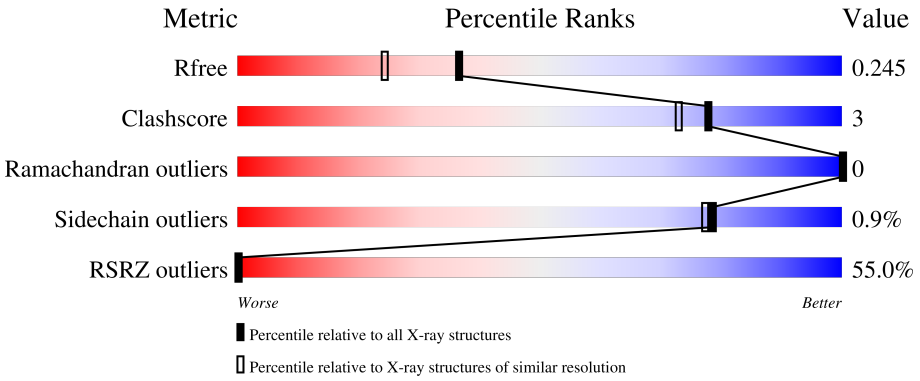
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.30
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.30

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 1.95 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	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (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	318	45% 75% 8% 18%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2178 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.

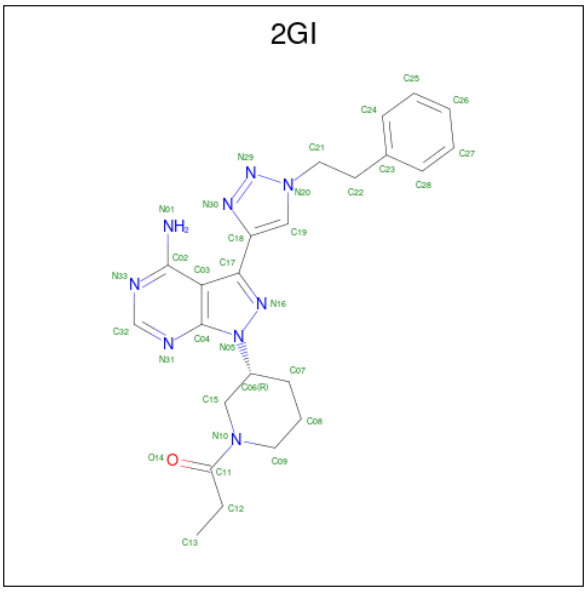
- Molecule 1 is a protein called Dual specificity mitogen-activated protein kinase kinase 7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	262	Total	C	N	O	S	0	0	0
			2062	1318	355	373	16			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	107	MET	-	initiating methionine	UNP O14733
A	108	GLY	-	expression tag	UNP O14733
A	109	HIS	-	expression tag	UNP O14733
A	110	HIS	-	expression tag	UNP O14733
A	111	HIS	-	expression tag	UNP O14733
A	112	HIS	-	expression tag	UNP O14733
A	113	HIS	-	expression tag	UNP O14733
A	114	HIS	-	expression tag	UNP O14733
A	115	SER	-	expression tag	UNP O14733
A	116	ALA	-	expression tag	UNP O14733

- Molecule 2 is 1-[(3 {R})-3-[4-azanyl-3-[1-(2-phenylethyl)-1,2,3-triazol-4-yl]pyrazolo[3,4-d]pyrimidin-1-yl]piperidin-1-yl]propan-1-one (three-letter code: 2GI) (formula: $C_{23}H_{27}N_9O$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C N O 33 23 9 1	0	0

- Molecule 3 is water.

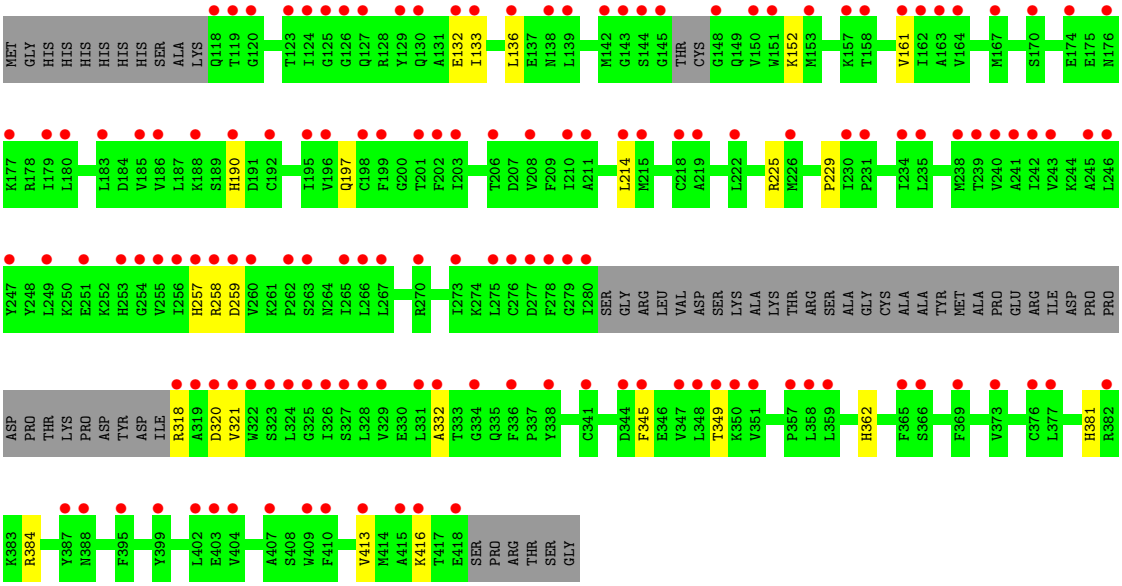
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	83	Total O 83 83	0	0

3 Residue-property plots (i)

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: Dual specificity mitogen-activated protein kinase kinase 7

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	61.72Å 68.49Å 83.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.74 – 1.95 49.74 – 1.95	Depositor EDS
% Data completeness (in resolution range)	99.5 (49.74-1.95) 99.5 (49.74-1.95)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.21 (at 1.95Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158, PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.209 , 0.250 0.205 , 0.245	Depositor DCC
R_{free} test set	1324 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	51.0	Xtriage
Anisotropy	0.111	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	2178	wwPDB-VP
Average B, all atoms (Å^2)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.27% 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: 2GI

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	0.47	0/2101	0.58	0/2828

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	2062	0	2058	12	0
2	A	33	0	0	0	0
3	A	83	0	0	0	0
All	All	2178	0	2058	12	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.

All (12) 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:257:HIS:NE2	1:A:259:ASP:OD1	2.34	0.59

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:381:HIS:HA	1:A:384:ARG:HG3	1.85	0.59
1:A:152:LYS:HE3	1:A:161:VAL:HG11	1.90	0.54
1:A:225:ARG:HD3	1:A:416:LYS:O	2.12	0.48
1:A:190:HIS:HA	1:A:197:GLN:OE1	2.14	0.48
1:A:161:VAL:HG12	1:A:214:LEU:HD22	1.95	0.47
1:A:345:PHE:O	1:A:349:THR:HG23	2.17	0.44
1:A:318:ARG:O	1:A:321:VAL:HG12	2.18	0.43
1:A:229:PRO:HB3	1:A:332:ALA:O	2.18	0.43
1:A:413:VAL:HA	1:A:416:LYS:HD2	2.00	0.42
1:A:133:ILE:HD12	1:A:136:LEU:HD12	2.01	0.42
1:A:258:ARG:HA	1:A:320:ASP:HA	2.02	0.42

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	
1	A	256/318 (80%)	254 (99%)	2 (1%)	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 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	221/277 (80%)	219 (99%)	2 (1%)	78 77

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	132	GLU
1	A	362	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

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	$\#\|Z\| > 2$
2	2GI	A	501	1	34,37,37	1.79	7 (20%)	34,52,52	1.62	5 (14%)

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	2GI	A	501	1	-	2/11/29/29	0/5/5/5

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	2GI	C11-N10	5.42	1.46	1.35
2	A	501	2GI	C17-N16	-3.81	1.32	1.35
2	A	501	2GI	C02-N01	3.18	1.45	1.34
2	A	501	2GI	C19-N20	-2.93	1.32	1.35
2	A	501	2GI	C15-N10	2.62	1.49	1.46
2	A	501	2GI	N16-N05	-2.61	1.34	1.37
2	A	501	2GI	C17-C18	2.49	1.52	1.50

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	501	2GI	N31-C32-N33	-5.22	120.52	128.68
2	A	501	2GI	C17-N16-N05	3.96	108.33	105.17
2	A	501	2GI	C22-C21-N20	-2.40	107.27	111.35
2	A	501	2GI	C12-C11-N10	2.06	121.29	117.95
2	A	501	2GI	C28-C23-C24	2.02	121.35	118.17

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	501	2GI	N10-C11-C12-C13
2	A	501	2GI	O14-C11-C12-C13

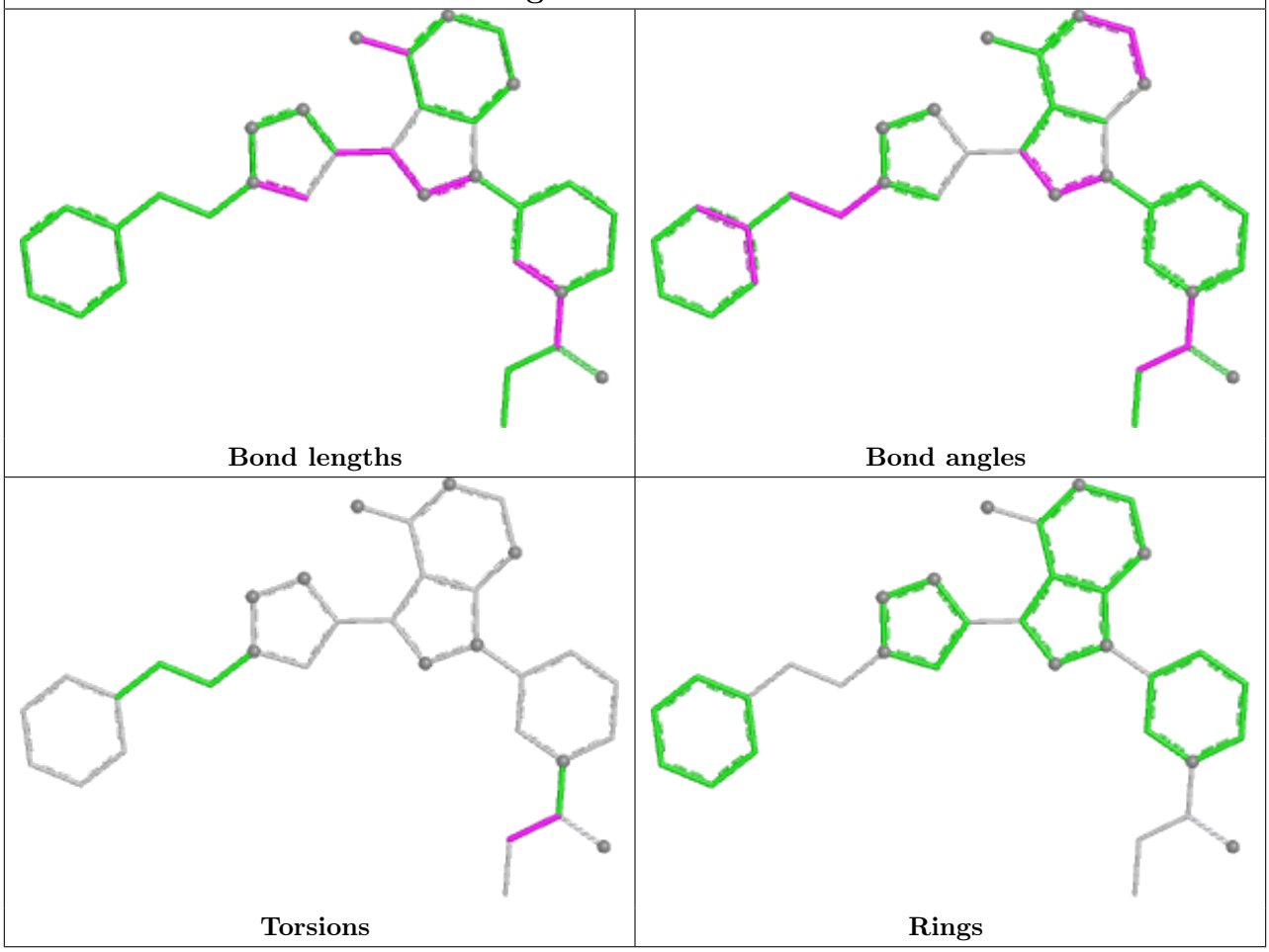
There are no ring outliers.

No monomer is involved in short contacts.

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 2GI A 501



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, 95^{th} 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	262/318 (82%)	2.54	144 (54%)	0	0	38, 54, 86, 111	0

All (144) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	125	GLY	16.9
1	A	280	ILE	11.8
1	A	318	ARG	10.4
1	A	259	ASP	8.4
1	A	145	GLY	8.1
1	A	258	ARG	7.9
1	A	183	LEU	7.8
1	A	320	ASP	7.5
1	A	278	PHE	6.9
1	A	277	ASP	6.8
1	A	124	ILE	6.5
1	A	319	ALA	6.4
1	A	118	GLN	6.1
1	A	126	GLY	5.5
1	A	255	VAL	5.4
1	A	150	VAL	5.4
1	A	119	THR	5.3
1	A	345	PHE	5.3
1	A	331	LEU	5.1
1	A	257	HIS	4.8
1	A	358	LEU	4.4
1	A	321	VAL	4.3
1	A	328	LEU	4.2
1	A	418	GLU	4.2
1	A	179	ILE	4.1
1	A	262	PRO	4.1
1	A	242	ILE	4.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	410	PHE	4.0
1	A	210	ILE	4.0
1	A	260	VAL	3.9
1	A	351	VAL	3.9
1	A	185	VAL	3.8
1	A	243	VAL	3.8
1	A	192	CYS	3.8
1	A	347	VAL	3.8
1	A	186	VAL	3.7
1	A	230	ILE	3.7
1	A	136	LEU	3.6
1	A	267	LEU	3.6
1	A	334	GLY	3.6
1	A	129	TYR	3.6
1	A	276	CYS	3.6
1	A	222	LEU	3.6
1	A	256	ILE	3.5
1	A	338	TYR	3.5
1	A	158	THR	3.4
1	A	234	ILE	3.4
1	A	388	ASN	3.4
1	A	247	TYR	3.4
1	A	133	ILE	3.3
1	A	251	GLU	3.3
1	A	404	VAL	3.3
1	A	148	GLY	3.2
1	A	373	VAL	3.2
1	A	127	GLN	3.2
1	A	202	PHE	3.2
1	A	254	GLY	3.2
1	A	365	PHE	3.2
1	A	214	LEU	3.2
1	A	211	ALA	3.2
1	A	246	LEU	3.1
1	A	359	LEU	3.1
1	A	399	TYR	3.1
1	A	198	CYS	3.1
1	A	413	VAL	3.1
1	A	196	VAL	3.1
1	A	382	ARG	3.1
1	A	415	ALA	3.0
1	A	130	GLN	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	402	LEU	2.9
1	A	199	PHE	2.9
1	A	327	SER	2.9
1	A	322	TRP	2.9
1	A	170	SER	2.9
1	A	249	LEU	2.9
1	A	273	ILE	2.9
1	A	188	LYS	2.9
1	A	266	LEU	2.9
1	A	153	MET	2.9
1	A	403	GLU	2.8
1	A	151	TRP	2.8
1	A	208	VAL	2.8
1	A	231	PRO	2.8
1	A	279	GLY	2.8
1	A	180	LEU	2.8
1	A	161	VAL	2.8
1	A	143	GLY	2.8
1	A	357	PRO	2.8
1	A	174	GLU	2.8
1	A	215	MET	2.7
1	A	123	THR	2.7
1	A	241	ALA	2.7
1	A	348	LEU	2.7
1	A	177	LYS	2.6
1	A	336	PHE	2.6
1	A	132	GLU	2.6
1	A	218	CYS	2.6
1	A	350	LYS	2.6
1	A	341	CYS	2.6
1	A	120	GLY	2.6
1	A	139	LEU	2.5
1	A	144	SER	2.5
1	A	238	MET	2.5
1	A	409	TRP	2.5
1	A	376	CYS	2.5
1	A	263	SER	2.5
1	A	265	ILE	2.5
1	A	329	VAL	2.5
1	A	395	PHE	2.5
1	A	324	LEU	2.5
1	A	349	THR	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	332	ALA	2.5
1	A	387	TYR	2.5
1	A	195	ILE	2.4
1	A	157	LYS	2.4
1	A	253	HIS	2.4
1	A	201	THR	2.3
1	A	416	LYS	2.3
1	A	240	VAL	2.3
1	A	323	SER	2.3
1	A	366	SER	2.3
1	A	344	ASP	2.3
1	A	407	ALA	2.2
1	A	226	MET	2.2
1	A	219	ALA	2.2
1	A	245	ALA	2.2
1	A	206	THR	2.2
1	A	203	ILE	2.2
1	A	377	LEU	2.2
1	A	176	ASN	2.1
1	A	162	ILE	2.1
1	A	326	ILE	2.1
1	A	163	ALA	2.1
1	A	275	LEU	2.1
1	A	369	PHE	2.1
1	A	164	VAL	2.1
1	A	239	THR	2.1
1	A	235	LEU	2.1
1	A	138	ASN	2.1
1	A	270	ARG	2.1
1	A	325	GLY	2.0
1	A	142	MET	2.0
1	A	167	MET	2.0
1	A	190	HIS	2.0

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 monosaccharides in this entry.

6.4 Ligands

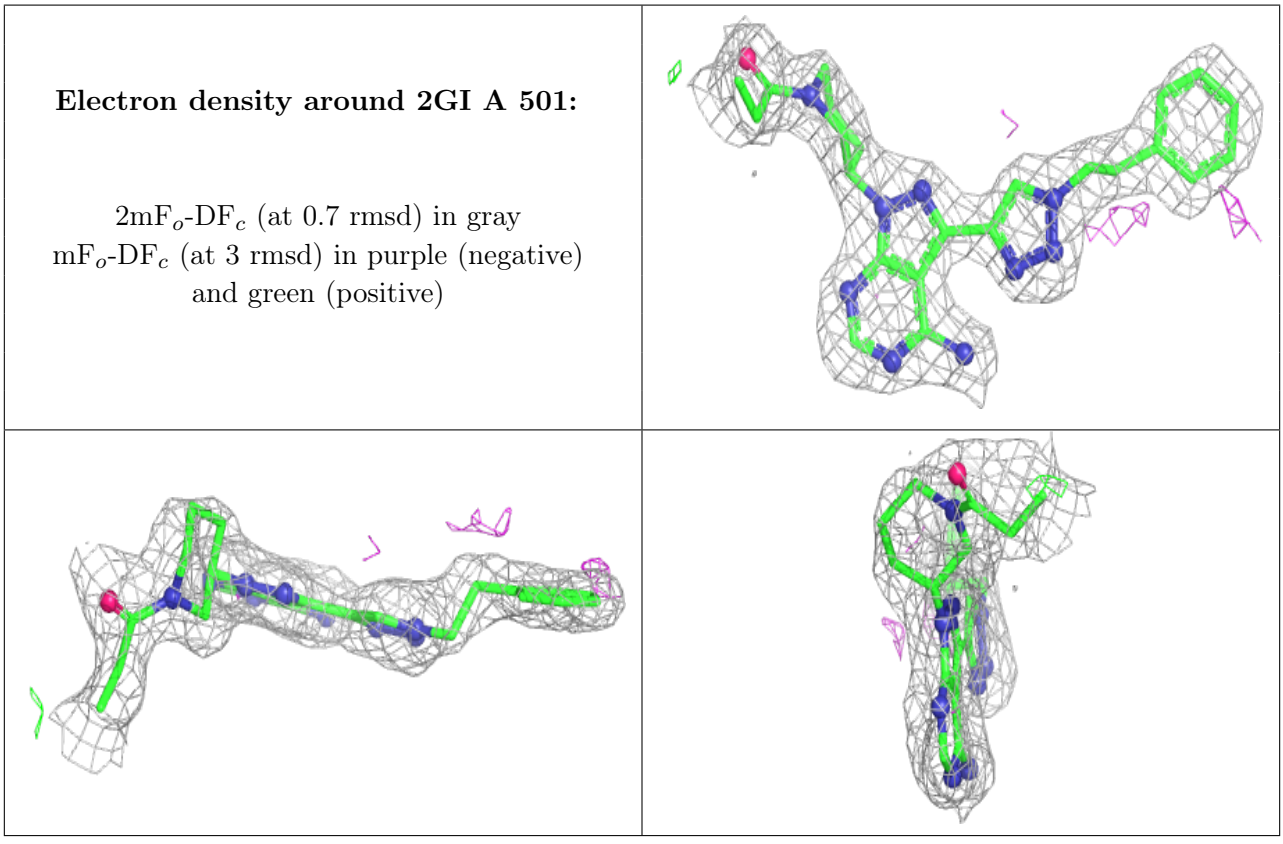
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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	2GI	A	501	33/33	0.77	0.24	43,48,57,58	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 2GI A 501:

$2mF_o$-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.