



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 7, 2026 – 01:15 AM UTC

PDB ID : 8PQ9 / pdb_00008pq9
Title : c-KIT kinase domain in complex with avapritinib
Authors : Teuber, A.; Kleinboelting, S.; Mueller, M.P.; Rauh, D.
Deposited on : 2023-07-11
Resolution : 1.70 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 : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Buster-report : wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

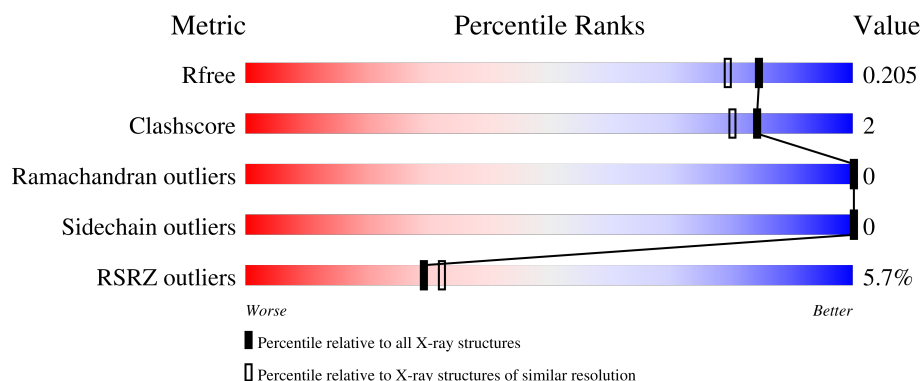
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.70 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}	180053	5551 (1.70-1.70)
Clashscore	190562	5924 (1.70-1.70)
Ramachandran outliers	187476	5846 (1.70-1.70)
Sidechain outliers	187428	5846 (1.70-1.70)
RSRZ outliers	180081	5554 (1.70-1.70)

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 $\leq 5\%$. 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	327	6% 87% 9%
1	C	327	4% 86% 6% 9%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	SO4	C	1002	-	-	X	-

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5232 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 Mast/stem cell growth factor receptor Kit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	296	Total	C	N	O	S	0	9	0
			2327	1496	382	431	18			
1	C	299	Total	C	N	O	S	0	6	0
			2340	1499	390	434	17			

There are 70 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	548	GLY	-	expression tag	UNP P10721
A	549	SER	-	expression tag	UNP P10721
A	550	MET	-	expression tag	UNP P10721
A	563	SER	ILE	engineered mutation	UNP P10721
A	569	SER	VAL	engineered mutation	UNP P10721
A	609	GLN	TYR	engineered mutation	UNP P10721
A	631	SER	LEU	engineered mutation	UNP P10721
A	651	GLU	MET	engineered mutation	UNP P10721
A	662	HIS	ILE	engineered mutation	UNP P10721
A	688	GLU	-	linker	UNP P10721
A	689	PHE	-	linker	UNP P10721
A	690	VAL	-	linker	UNP P10721
A	691	PRO	-	linker	UNP P10721
A	752	TYR	-	linker	UNP P10721
A	753	LYS	-	linker	UNP P10721
A	754	VAL	-	linker	UNP P10721
A	755	ALA	-	linker	UNP P10721
A	756	PRO	-	linker	UNP P10721
A	757	GLU	-	linker	UNP P10721
A	758	ASP	-	linker	UNP P10721
A	759	LEU	-	linker	UNP P10721
A	760	TYR	-	linker	UNP P10721
A	761	LYS	-	linker	UNP P10721
A	762	ASP	-	linker	UNP P10721
A	763	PHE	-	linker	UNP P10721

Continued on next page...

Continued from previous page...

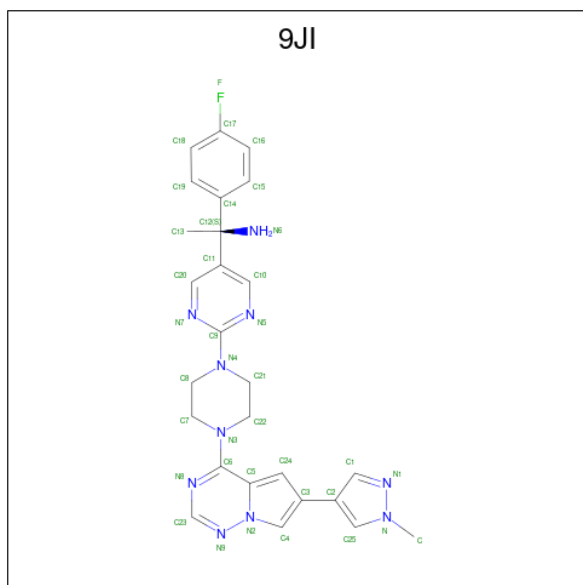
Chain	Residue	Modelled	Actual	Comment	Reference
A	764	LEU	-	linker	UNP P10721
A	765	THR	-	linker	UNP P10721
A	768	HIS	ASP	engineered mutation	UNP P10721
A	804	ASN	ARG	engineered mutation	UNP P10721
A	825	ASP	VAL	engineered mutation	UNP P10721
A	844	SER	CYS	engineered mutation	UNP P10721
A	890	SER	LEU	engineered mutation	UNP P10721
A	894	TYR	HIS	engineered mutation	UNP P10721
A	912	ASP	LEU	engineered mutation	UNP P10721
A	923	ASP	LEU	engineered mutation	UNP P10721
C	548	GLY	-	expression tag	UNP P10721
C	549	SER	-	expression tag	UNP P10721
C	550	MET	-	expression tag	UNP P10721
C	563	SER	ILE	engineered mutation	UNP P10721
C	569	SER	VAL	engineered mutation	UNP P10721
C	609	GLN	TYR	engineered mutation	UNP P10721
C	631	SER	LEU	engineered mutation	UNP P10721
C	651	GLU	MET	engineered mutation	UNP P10721
C	662	HIS	ILE	engineered mutation	UNP P10721
C	688	GLU	-	linker	UNP P10721
C	689	PHE	-	linker	UNP P10721
C	690	VAL	-	linker	UNP P10721
C	691	PRO	-	linker	UNP P10721
C	692	TYR	-	linker	UNP P10721
C	693	LYS	-	linker	UNP P10721
C	754	VAL	-	linker	UNP P10721
C	755	ALA	-	linker	UNP P10721
C	756	PRO	-	linker	UNP P10721
C	757	GLU	-	linker	UNP P10721
C	758	ASP	-	linker	UNP P10721
C	759	LEU	-	linker	UNP P10721
C	760	TYR	-	linker	UNP P10721
C	761	LYS	-	linker	UNP P10721
C	762	ASP	-	linker	UNP P10721
C	763	PHE	-	linker	UNP P10721
C	764	LEU	-	linker	UNP P10721
C	765	THR	-	linker	UNP P10721
C	768	HIS	ASP	engineered mutation	UNP P10721
C	804	ASN	ARG	engineered mutation	UNP P10721
C	825	ASP	VAL	engineered mutation	UNP P10721
C	844	SER	CYS	engineered mutation	UNP P10721
C	890	SER	LEU	engineered mutation	UNP P10721

Continued on next page...

Continued from previous page...

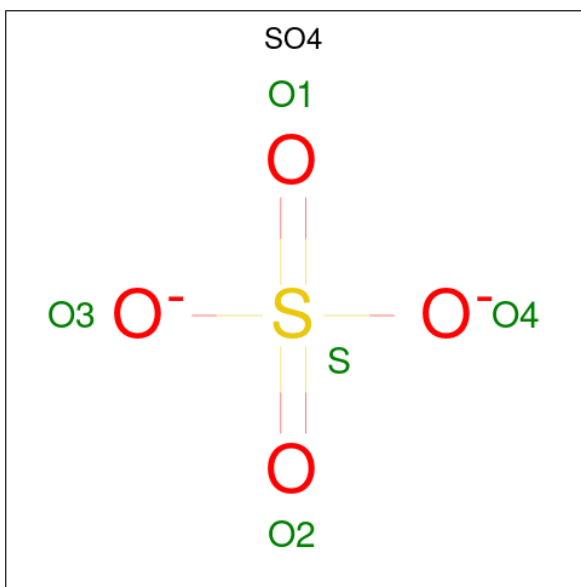
Chain	Residue	Modelled	Actual	Comment	Reference
C	894	TYR	HIS	engineered mutation	UNP P10721
C	912	ASP	LEU	engineered mutation	UNP P10721
C	923	ASP	LEU	engineered mutation	UNP P10721

- Molecule 2 is Avapritinib (CCD ID: 9JI) (formula: C₂₆H₂₇FN₁₀) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	F	N	0	0
			37	26	1	10		
2	C	1	Total	C	F	N	0	0
			37	26	1	10		

- Molecule 3 is SULFATE ION (CCD ID: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	C	1	Total	O	S	0	0
			5	4	1		

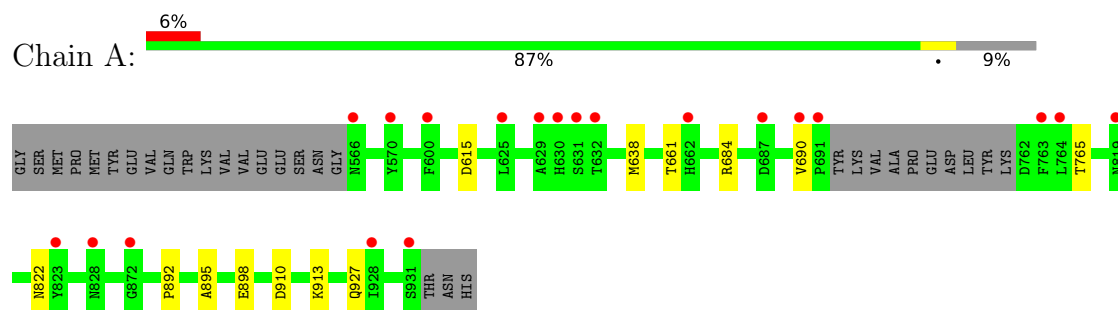
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	229	Total	O	0	0
			229	229		
4	C	257	Total	O	0	0
			257	257		

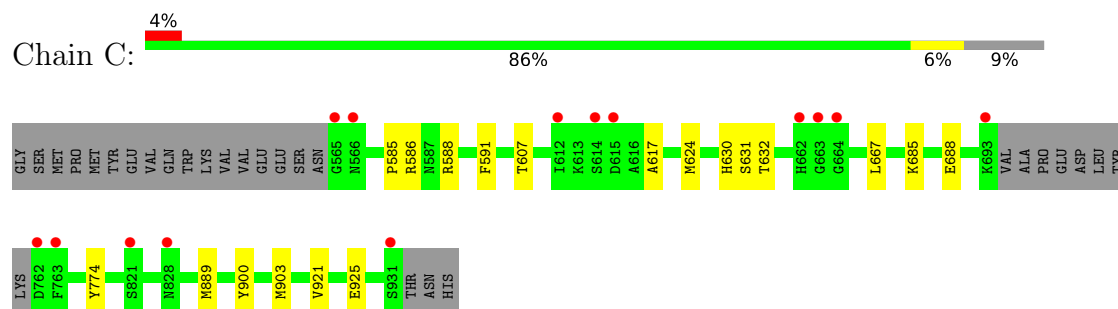
3 Residue-property plots [i](#)

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: Mast/stem cell growth factor receptor Kit



- Molecule 1: Mast/stem cell growth factor receptor Kit



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	59.02Å 59.30Å 191.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.00 – 1.70 35.00 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.9 (35.00-1.70) 99.9 (35.00-1.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.66 (at 1.70Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.191 , 0.204 0.191 , 0.205	Depositor DCC
R_{free} test set	3750 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	34.2	Xtriage
Anisotropy	0.527	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 38.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.017 for k,h,-l	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	5232	wwPDB-VP
Average B, all atoms (Å ²)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.87% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: SO4, 9JI

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	0.24	0/2409	0.48	0/3267
1	C	0.28	0/2417	0.51	0/3277
All	All	0.26	0/4826	0.49	0/6544

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 [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	2327	0	2253	8	0
1	C	2340	0	2258	13	0
2	A	37	0	0	0	0
2	C	37	0	0	0	0
3	C	5	0	0	2	0
4	A	229	0	0	1	0
4	C	257	0	0	1	0
All	All	5232	0	4511	21	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 2.

All (21) 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:C:889:MET:HG2	3:C:1002:SO4:O2	1.71	0.89
1:C:921:VAL:O	1:C:925:GLU:HG3	1.92	0.70
1:A:684:ARG:NH2	4:A:1101:HOH:O	2.25	0.69
1:A:615:ASP:OD1	1:A:615:ASP:N	2.37	0.57
1:C:607:THR:CG2	1:C:617:ALA:HB1	2.36	0.56
1:C:591:PHE:CZ	1:C:624:MET:HE1	2.42	0.55
1:A:910:ASP:HB3	1:A:913:LYS:HD2	1.89	0.54
1:C:685:LYS:HE3	1:C:688:GLU:OE2	2.10	0.52
1:C:774:TYR:CE1	1:C:925:GLU:HG2	2.44	0.51
1:A:892:PRO:HG2	1:A:895:ALA:HB2	1.92	0.51
1:C:624:MET:HE2	1:C:667:LEU:HD22	1.94	0.50
1:C:630:HIS:CE1	1:C:632:THR:HG23	2.46	0.50
1:C:630:HIS:ND1	1:C:631:SER:N	2.63	0.46
1:C:889:MET:HA	3:C:1002:SO4:O2	2.15	0.46
1:C:900:TYR:O	1:C:903:MET:HB2	2.15	0.46
1:C:586:ARG:HB2	4:C:1220:HOH:O	2.16	0.45
1:A:638[A]:MET:HE1	1:A:661:THR:HG21	1.98	0.45
1:A:690:VAL:O	1:A:765:THR:HA	2.17	0.44
1:A:822:ASN:OD1	1:A:822:ASN:C	2.63	0.42
1:A:898[A]:GLU:OE1	1:A:927:GLN:HG2	2.20	0.41
1:C:585:PRO:HG2	1:C:588:ARG:CG	2.50	0.41

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	301/327 (92%)	297 (99%)	4 (1%)	0	100	100
1	C	301/327 (92%)	297 (99%)	4 (1%)	0	100	100
All	All	602/654 (92%)	594 (99%)	8 (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	248/285 (87%)	248 (100%)	0	100	100
1	C	248/285 (87%)	248 (100%)	0	100	100
All	All	496/570 (87%)	496 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	649	ASN
1	C	787	ASN

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

3 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	9JI	A	1001	-	40,42,42	0.35	0	51,62,62	0.58	0
3	SO4	C	1002	-	4,4,4	0.70	0	6,6,6	0.79	0
2	9JI	C	1001	-	40,42,42	0.30	0	51,62,62	0.57	0

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	9JI	A	1001	-	-	3/24/34/34	0/6/6/6
2	9JI	C	1001	-	-	4/24/34/34	0/6/6/6

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (7) torsion outliers are listed below:

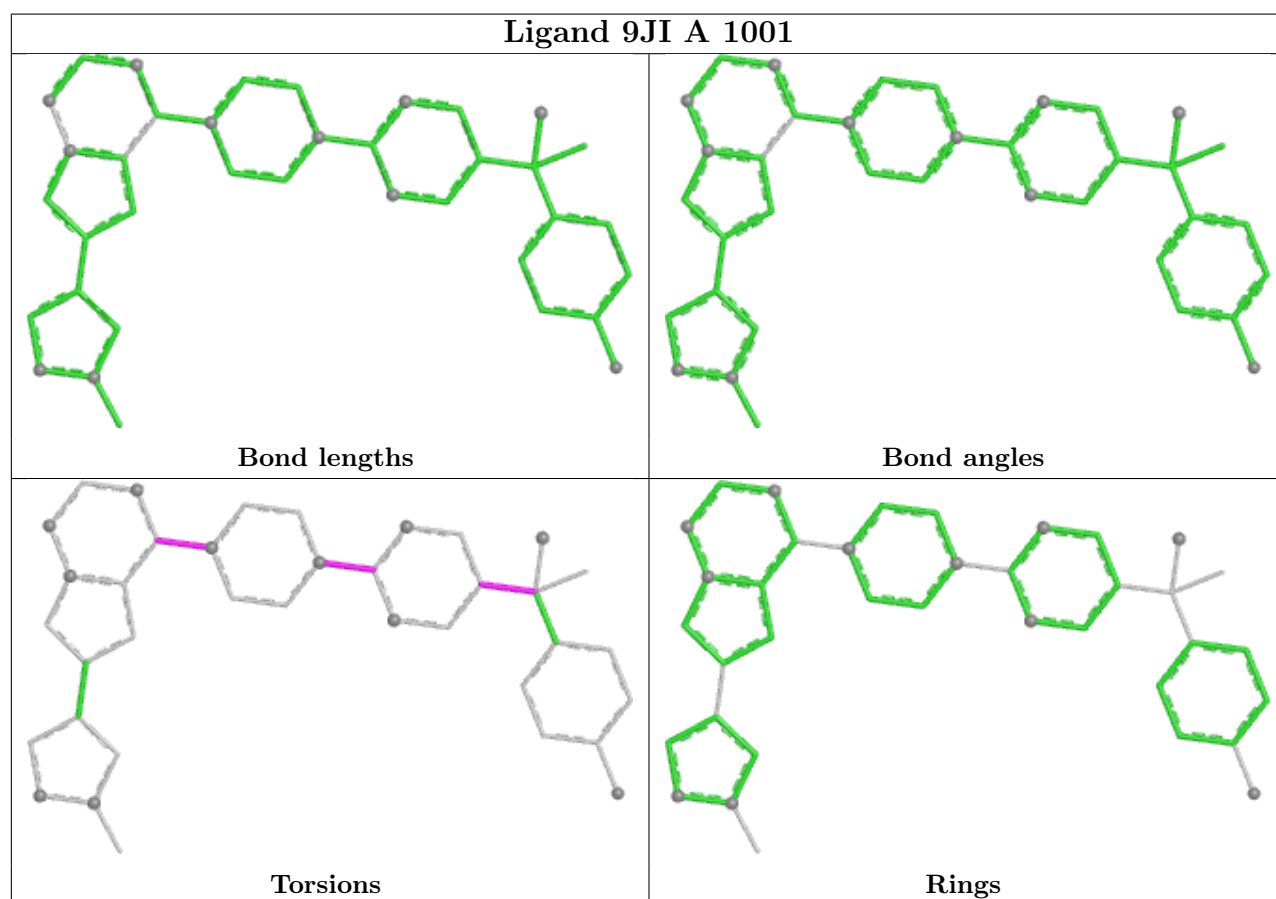
Mol	Chain	Res	Type	Atoms
2	A	1001	9JI	N8-C6-N3-C22
2	C	1001	9JI	N8-C6-N3-C7
2	C	1001	9JI	N5-C9-N4-C8
2	C	1001	9JI	N7-C9-N4-C8
2	A	1001	9JI	C20-C11-C12-N6
2	C	1001	9JI	C10-C11-C12-N6
2	A	1001	9JI	N7-C9-N4-C21

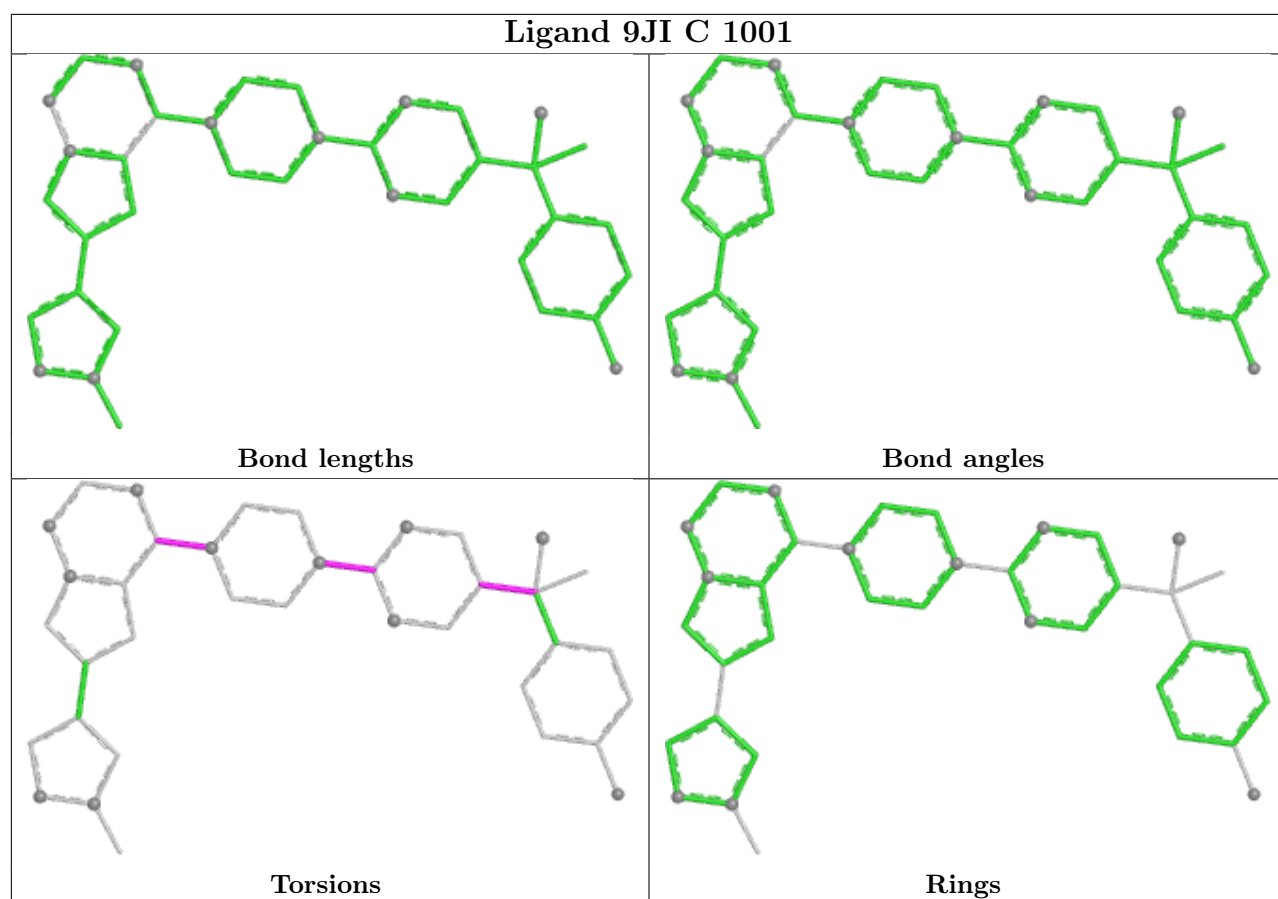
There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	1002	SO4	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 than 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 [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	296/327 (90%)	0.45	20 (6%)	23 25	18, 38, 69, 92	9 (3%)
1	C	299/327 (91%)	0.27	14 (4%)	36 40	23, 38, 67, 84	6 (2%)
All	All	595/654 (90%)	0.36	34 (5%)	29 32	18, 38, 69, 92	15 (2%)

All (34) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	663	GLY	4.1
1	A	630	HIS	4.0
1	A	600	PHE	4.0
1	A	632	THR	3.9
1	C	566	ASN	3.9
1	C	614	SER	3.6
1	C	615	ASP	3.3
1	C	612	ILE	3.3
1	C	828	ASN	3.2
1	C	565	GLY	3.2
1	A	629	ALA	3.2
1	C	763	PHE	3.1
1	A	566	ASN	3.0
1	A	570	TYR	3.0
1	A	872	GLY	3.0
1	A	931	SER	2.9
1	A	928	ILE	2.8
1	C	762	ASP	2.6
1	A	625	LEU	2.6
1	C	821	SER	2.5
1	A	690	VAL	2.5
1	A	819	ASN	2.5
1	C	931	SER	2.5
1	C	693	LYS	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	687	ASP	2.4
1	A	764	LEU	2.3
1	A	763	PHE	2.3
1	C	662	HIS	2.3
1	A	631	SER	2.2
1	A	828	ASN	2.2
1	A	691	PRO	2.1
1	A	662	HIS	2.0
1	C	664	GLY	2.0
1	A	823	TYR	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 oligosaccharides in this entry.

6.4 Ligands [i](#)

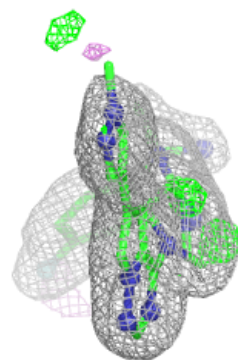
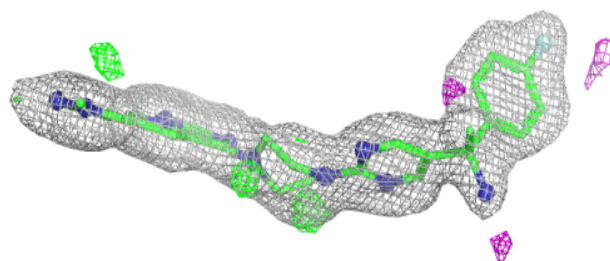
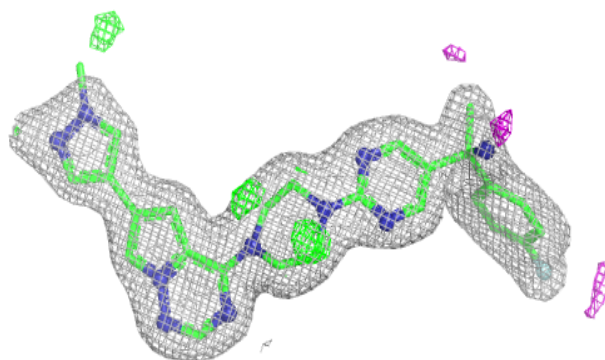
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
2	9JI	A	1001	37/37	0.93	0.09	33,37,48,58	0
2	9JI	C	1001	37/37	0.95	0.07	27,35,42,43	0
3	SO4	C	1002	5/5	0.97	0.12	23,40,53,55	0

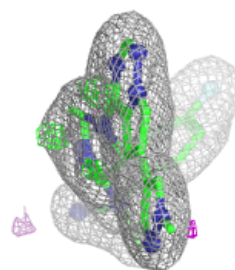
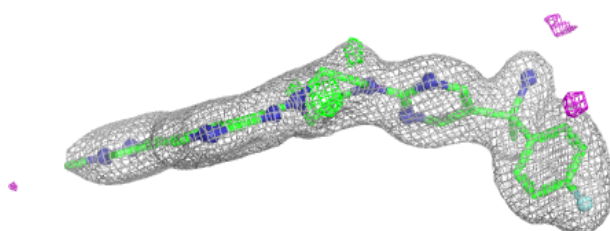
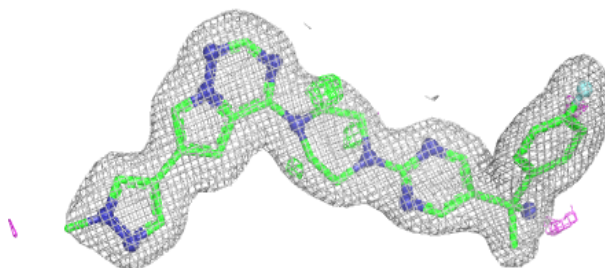
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around 9JI A 1001:

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

**Electron density around 9JI C 1001:**

$2mF_o-DF_c$ (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.