Full wwPDB X-ray Structure Validation Report (i)

May 14, 2020 – 11:18 am BST

PDB ID	:	4J0P
Title	:	CRYSTAL STRUCTURE OF BACE-1 IN COMPLEX WITH 5-Cyano-pyridine-2-carboxylic acid [3-((S)-2-amino-4-methyl-5,6-dihydro-4H-[1,3]oxazin-4-yl)-4-fluoro-phenyl]-amide
Authors	:	Kuglstatter, A.; Stihle, M.
Deposited on	:	2013-01-31
Resolution	:	1.97 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 (i) symbol.

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

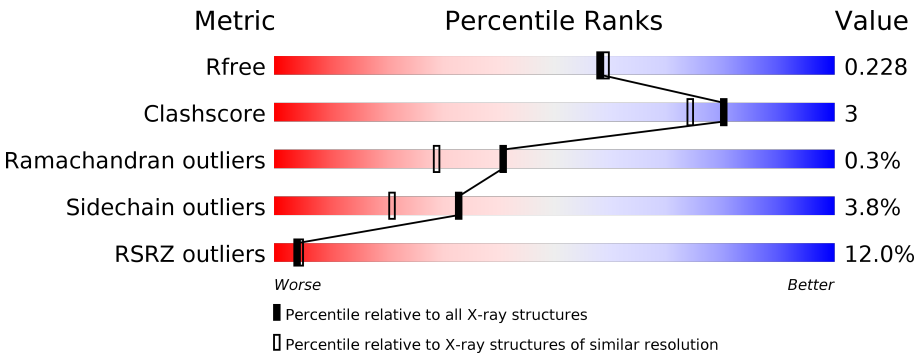
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber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

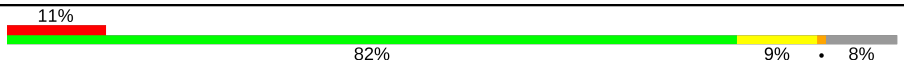
The reported resolution of this entry is 1.97 Å.

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	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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 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	409	

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 3194 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 Beta-secretase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	376	2959	1893	494	558	14	0	0	0

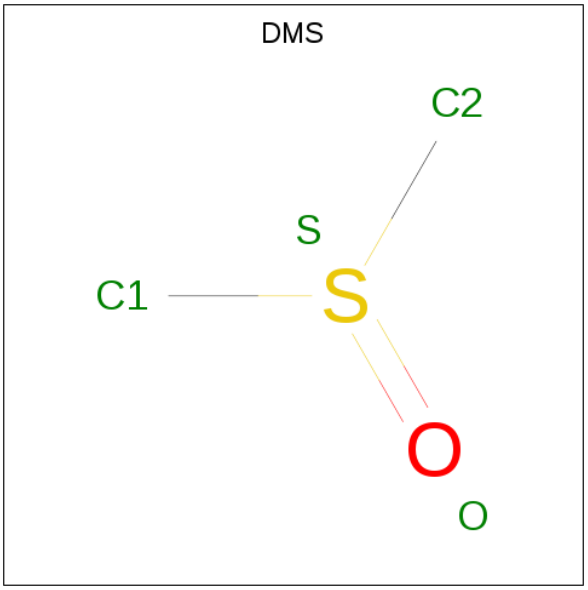
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	307	ALA	LYS	ENGINEERED MUTATION	UNP P56817

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

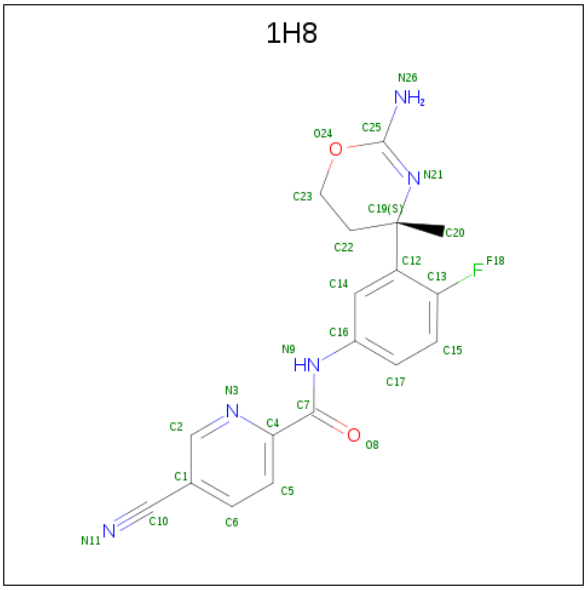
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Na		
2	A	2	2	2	0	0

- Molecule 3 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C_2H_6OS).



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

- Molecule 4 is N-{3-[(4S)-2-amino-4-methyl-5,6-dihydro-4H-1,3-oxazin-4-yl]-4-fluorophenyl}-5-cyanopyridine-2-carboxamide (three-letter code: 1H8) (formula: $C_{18}H_{16}FN_5O_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	F	N	O		
4	A	1	26	18	1	5	2	0	0

- Molecule 5 is water.

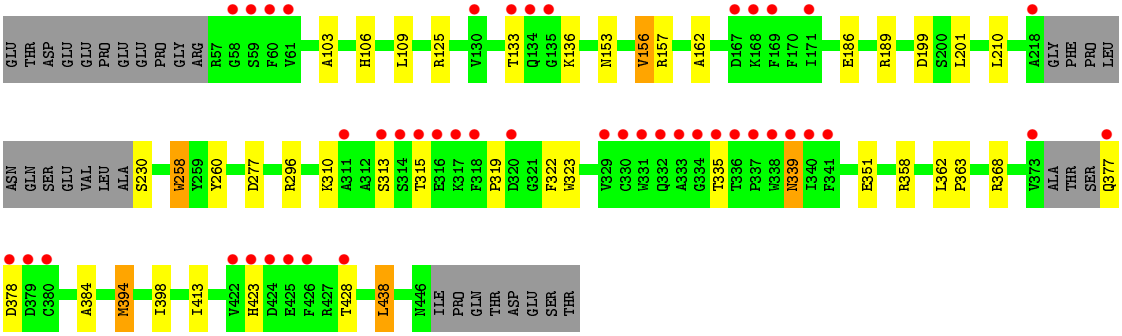
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	203	203	203	0	0

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: Beta-secretase 1

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α, β, γ	102.54Å 102.54Å 169.79Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.08 – 1.97 49.08 – 1.97	Depositor EDS
% Data completeness (in resolution range)	100.0 (49.08-1.97) 100.0 (49.08-1.97)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$< I/\sigma(I) >$ [1]	1.26 (at 1.97Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.176 , 0.231 0.173 , 0.228	Depositor DCC
R_{free} test set	1898 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	41.6	Xtriage
Anisotropy	0.134	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 51.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	3194	wwPDB-VP
Average B, all atoms (Å^2)	47.0	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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: NA, 1H8, DMS

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	1.02	1/3033 (0.0%)	1.01	10/4120 (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	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	351	GLU	CD-OE1	5.21	1.31	1.25

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	157	ARG	NE-CZ-NH1	6.99	123.79	120.30
1	A	157	ARG	NE-CZ-NH2	-6.80	116.90	120.30
1	A	277	ASP	CB-CG-OD1	5.89	123.60	118.30
1	A	358	ARG	NE-CZ-NH2	-5.71	117.44	120.30
1	A	394	MET	CG-SD-CE	-5.53	91.35	100.20
1	A	438	LEU	CB-CG-CD1	5.51	120.36	111.00
1	A	189	ARG	NE-CZ-NH1	5.13	122.86	120.30
1	A	358	ARG	CG-CD-NE	-5.04	101.21	111.80
1	A	277	ASP	CB-CG-OD2	-5.01	113.79	118.30
1	A	199	ASP	CB-CG-OD2	5.01	122.81	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	258	TRP	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	2959	0	2865	16	0
2	A	2	0	0	0	0
3	A	4	0	6	0	0
4	A	26	0	16	0	0
5	A	203	0	0	1	2
All	All	3194	0	2887	16	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 3.

All (16) 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:368:ARG:HG2	1:A:384:ALA:HB2	1.80	0.62
1:A:362:LEU:HB3	1:A:363:PRO:HD2	1.84	0.58
1:A:394:MET:CE	1:A:398:ILE:HG21	2.34	0.57
1:A:339:ASN:N	1:A:339:ASN:OD1	2.40	0.53
1:A:296:ARG:NH1	5:A:775:HOH:O	2.42	0.52
1:A:394:MET:HE2	1:A:398:ILE:HG21	1.92	0.51
1:A:368:ARG:CG	1:A:384:ALA:HB2	2.41	0.51
1:A:319:PRO:O	1:A:322:PHE:HB3	2.12	0.50
1:A:260:TYR:HB3	1:A:413:ILE:HD11	1.95	0.48
1:A:310:LYS:HD3	1:A:323:TRP:CD1	2.48	0.48
1:A:106:HIS:HB3	1:A:109:LEU:HG	1.99	0.45
1:A:186:GLU:O	1:A:186:GLU:HG3	2.15	0.45
1:A:156:VAL:HG21	1:A:201:LEU:HA	2.00	0.43
1:A:313:SER:C	1:A:315:THR:H	2.24	0.41
1:A:103:ALA:CB	1:A:162:ALA:HB1	2.50	0.41
1:A:423:HIS:CG	1:A:428:THR:HG22	2.56	0.41

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)
5:A:774:HOH:O	5:A:774:HOH:O[10_444]	1.48	0.72
5:A:715:HOH:O	5:A:749:HOH:O[5_554]	2.01	0.19

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	370/409 (90%)	358 (97%)	11 (3%)	1 (0%)	41 29

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	378	ASP

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	320/349 (92%)	308 (96%)	12 (4%)	33 21

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	125	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	133	THR
1	A	136	LYS
1	A	153	ASN
1	A	156	VAL
1	A	210	LEU
1	A	230	SER
1	A	258	TRP
1	A	335	THR
1	A	339	ASN
1	A	377	GLN
1	A	438	LEU

Some 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 carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	$\#\|Z\| > 2$
3	DMS	A	503	-	3,3,3	0.50	0	3,3,3	0.82	0
4	1H8	A	504	-	26,28,28	1.15	2 (7%)	30,40,40	2.20	8 (26%)

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
4	1H8	A	504	-	-	0/16/28/28	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	504	1H8	C15-C13	2.42	1.43	1.37
4	A	504	1H8	C12-C13	2.37	1.42	1.38

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	504	1H8	C1-C2-N3	-5.07	115.98	123.51
4	A	504	1H8	C2-N3-C4	4.78	124.78	117.36
4	A	504	1H8	O24-C23-C22	-4.63	102.52	112.11
4	A	504	1H8	O24-C25-N26	4.34	118.31	111.64
4	A	504	1H8	C6-C1-C2	3.26	121.68	117.26
4	A	504	1H8	C23-C22-C19	-2.90	107.91	111.59
4	A	504	1H8	C16-C14-C12	2.41	123.77	120.91
4	A	504	1H8	C15-C17-C16	-2.30	117.64	120.30

There are no chirality outliers.

There are no torsion outliers.

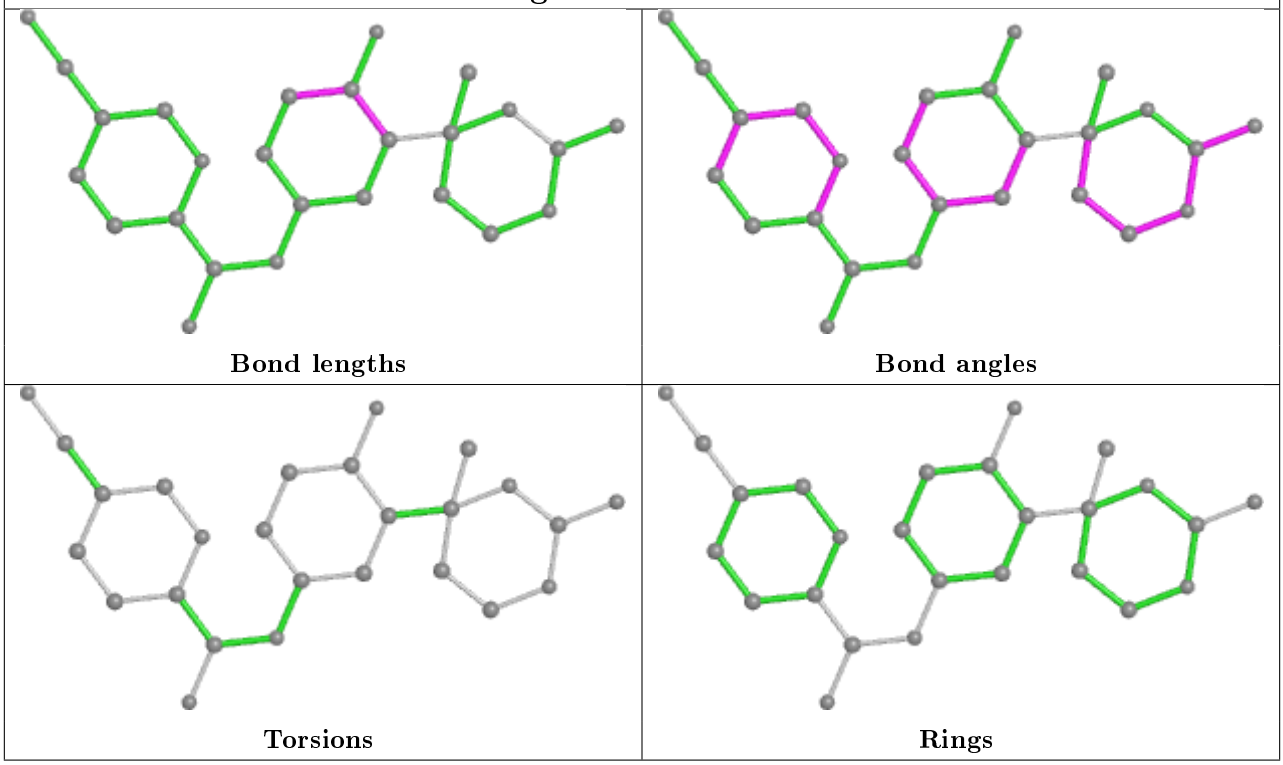
There are no ring outliers.

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 1H8 A 504



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	376/409 (91%)	0.62	45 (11%) 4 5	26, 41, 91, 132	19 (5%)

All (45) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	335	THR	9.0
1	A	315	THR	7.4
1	A	59	SER	7.0
1	A	331	TRP	6.6
1	A	134	GLN	5.9
1	A	380	CYS	5.3
1	A	314	SER	5.2
1	A	336	THR	5.1
1	A	378	ASP	5.1
1	A	330	CYS	5.1
1	A	426	PHE	4.9
1	A	334	GLY	4.8
1	A	340	ILE	4.8
1	A	338	TRP	4.6
1	A	333	ALA	4.3
1	A	337	PRO	4.3
1	A	332	GLN	4.3
1	A	61	VAL	4.1
1	A	339	ASN	3.9
1	A	425	GLU	3.9
1	A	422	VAL	3.9
1	A	316	GLU	3.8
1	A	318	PHE	3.8
1	A	58	GLY	3.7
1	A	317	LYS	3.6
1	A	423	HIS	3.5
1	A	320	ASP	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	313	SER	3.3
1	A	373	VAL	3.2
1	A	424	ASP	3.0
1	A	379	ASP	2.9
1	A	135	GLY	2.9
1	A	133	THR	2.7
1	A	60	PHE	2.7
1	A	329	VAL	2.7
1	A	428	THR	2.6
1	A	171	ILE	2.6
1	A	169	PHE	2.6
1	A	218	ALA	2.6
1	A	341	PHE	2.5
1	A	311	ALA	2.5
1	A	167	ASP	2.3
1	A	168	LYS	2.2
1	A	130	VAL	2.1
1	A	377	GLN	2.1

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)

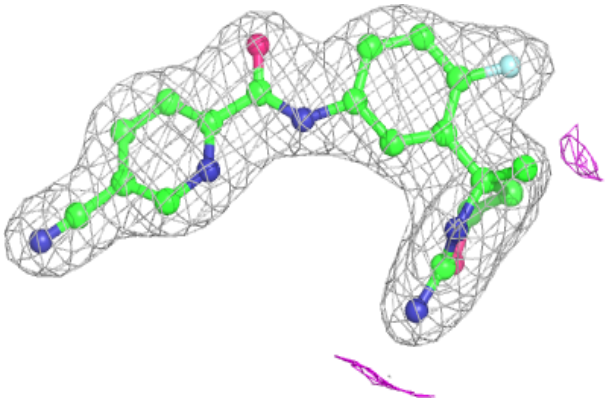
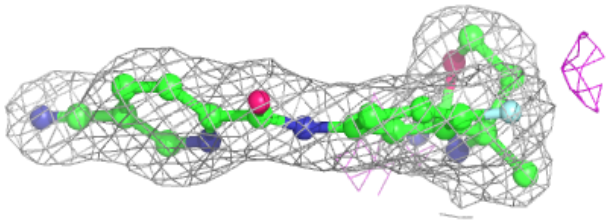
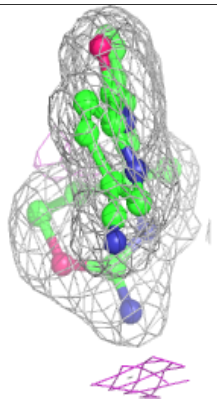
There are no carbohydrates in this entry.

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	NA	A	501	1/1	0.93	0.13	56,56,56,56	0
2	NA	A	502	1/1	0.97	0.26	62,62,62,62	0
4	1H8	A	504	26/26	0.98	0.16	28,35,38,42	0
3	DMS	A	503	4/4	0.98	0.15	75,77,77,77	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 1H8 A 504: $2mF_o$-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.