



Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 07:41 am BST

PDB ID : 5AAA
Title : Structure of L1198F Mutant Human Anaplastic Lymphoma Kinase in Complex with Crizotinib
Authors : McTigue, M.; Deng, Y.; Liu, W.; Brooun, A.; Stewart, A.
Deposited on : 2015-07-23
Resolution : 1.73 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 following versions of software and data (see [references ⓘ](#)) 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

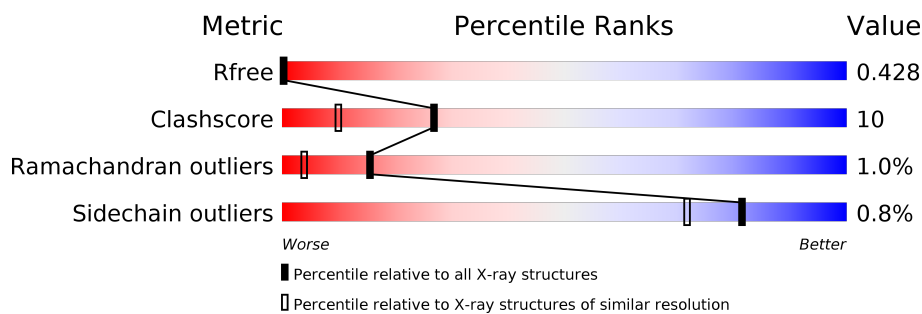
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.73 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	3764 (1.76-1.72)
Clashscore	141614	3923 (1.76-1.72)
Ramachandran outliers	138981	3878 (1.76-1.72)
Sidechain outliers	138945	3878 (1.76-1.72)

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	327	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2607 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.

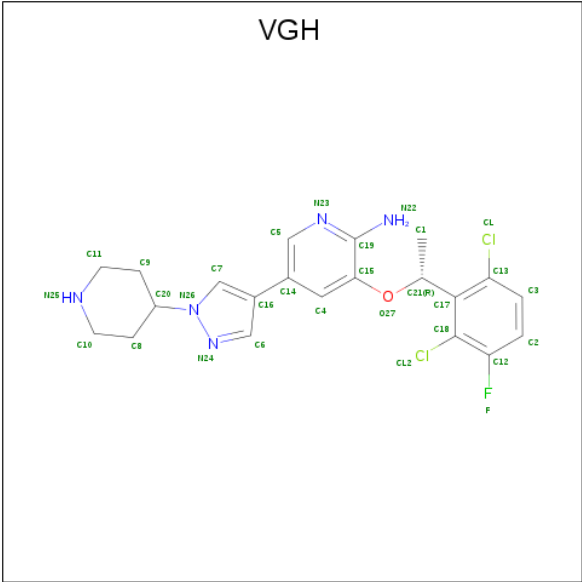
- Molecule 1 is a protein called ALK TYROSINE KINASE RECEPTOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	295	2350	1498	401	428	23	0	4	1

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1085	MET	-	expression tag	UNP Q9UM73
A	1086	ALA	-	expression tag	UNP Q9UM73
A	1087	HIS	-	expression tag	UNP Q9UM73
A	1088	HIS	-	expression tag	UNP Q9UM73
A	1089	HIS	-	expression tag	UNP Q9UM73
A	1090	HIS	-	expression tag	UNP Q9UM73
A	1091	HIS	-	expression tag	UNP Q9UM73
A	1092	HIS	-	expression tag	UNP Q9UM73
A	1198	PHE	LEU	engineered mutation	UNP Q9UM73

- Molecule 2 is 3-[(1R)-1-(2,6-dichloro-3-fluorophenyl)ethoxy]-5-(1-piperidin-4-yl-1H-pyrazol-4-yl)pyridin-2-amine (three-letter code: VGH) (formula: C₂₁H₂₂Cl₂FN₅O).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	Cl	F	N	O	0	0
			30	21	2	1	5	1		

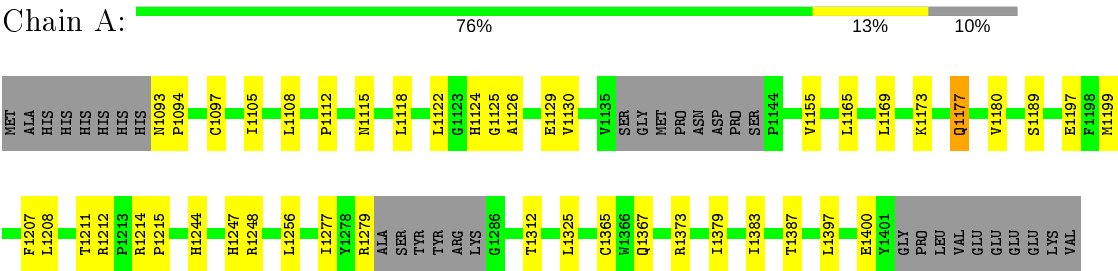
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	227	Total	O	0	0
			227	227		

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. 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. 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: ALK TYROSINE KINASE RECEPTOR



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	51.47Å 57.28Å 104.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	52.35 – 1.73 50.25 – 2.02	Depositor EDS
% Data completeness (in resolution range)	89.5 (52.35-1.73) 95.7 (50.25-2.02)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.34 (at 2.01Å)	Xtriage
Refinement program	CNX 2005	Depositor
R, R_{free}	0.200 , 0.206 0.420 , 0.428	Depositor DCC
R_{free} test set	591 reflections (2.92%)	wwPDB-VP
Wilson B-factor (Å ²)	34.1	Xtriage
Anisotropy	0.608	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 51.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.76	EDS
Total number of atoms	2607	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 5.33% 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

5.1 Standard geometry

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

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	0.26	0/2408	0.45	0/3263

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	2350	0	2317	47	0
2	A	30	0	22	3	0
3	A	227	0	0	4	0
All	All	2607	0	2339	48	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 10.

All (48) 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:1105:ILE:HD12	1:A:1105:ILE:H	1.20	1.03

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1124:HIS:HB3	1:A:1129:GLU:HA	1.56	0.87
1:A:1105:ILE:HG13	1:A:1165:LEU:HD11	1.61	0.81
1:A:1105:ILE:CD1	1:A:1105:ILE:H	1.94	0.81
1:A:1105:ILE:HD12	1:A:1105:ILE:N	2.00	0.75
1:A:1093:ASN:HB2	1:A:1105:ILE:HD13	1.69	0.74
1:A:1214:ARG:HG3	1:A:1215:PRO:HD2	1.69	0.74
1:A:1165:LEU:O	1:A:1169:LEU:HD13	1.95	0.66
1:A:1248:ARG:CZ	1:A:1277:ILE:HG21	2.25	0.65
1:A:1279:ARG:HE	1:A:1279:ARG:HA	1.60	0.65
1:A:1367:GLN:O	1:A:1373:ARG:HD3	1.98	0.63
1:A:1093:ASN:HB2	1:A:1105:ILE:CD1	2.29	0.62
1:A:1097[A]:CYS:HB3	3:A:2029:HOH:O	2.00	0.62
1:A:1155:VAL:O	1:A:1155:VAL:HG23	2.02	0.59
1:A:1367:GLN:HB2	1:A:1373:ARG:HG2	1.85	0.58
1:A:1244:HIS:HD2	3:A:2008:HOH:O	1.86	0.58
1:A:1279:ARG:NE	1:A:1279:ARG:HA	2.19	0.57
1:A:1248:ARG:NH1	1:A:1277:ILE:HG21	2.20	0.57
1:A:1177:GLN:NE2	1:A:1177:GLN:H	2.03	0.57
1:A:1199:MET:HA	1:A:1199:MET:HE2	1.87	0.57
1:A:1105:ILE:HG13	1:A:1165:LEU:CD1	2.32	0.55
1:A:1112:PRO:HG2	1:A:1115:ASN:OD1	2.06	0.55
1:A:1180:VAL:HG21	1:A:1256:LEU:HD12	1.89	0.55
1:A:1097[B]:CYS:HB2	3:A:2029:HOH:O	2.07	0.54
1:A:1208:LEU:O	1:A:1212:ARG:HG3	2.08	0.54
1:A:1365:CYS:O	1:A:1373:ARG:HD2	2.07	0.53
1:A:1197:GLU:HG2	1:A:1199:MET:HE3	1.91	0.53
1:A:1093:ASN:HD22	1:A:1105:ILE:HD13	1.75	0.52
1:A:1199:MET:CE	1:A:1199:MET:HA	2.42	0.49
1:A:1256:LEU:HD11	2:A:9000:VGH:CL2	2.50	0.49
1:A:1108:LEU:HD11	1:A:1169:LEU:CD1	2.43	0.48
1:A:1247:HIS:O	1:A:1248:ARG:HB2	2.14	0.47
1:A:1197:GLU:HG2	1:A:1199:MET:CE	2.44	0.47
2:A:9000:VGH:CL2	2:A:9000:VGH:H12C	2.52	0.47
1:A:1207:PHE:O	1:A:1211:THR:HG22	2.16	0.46
1:A:1124:HIS:CD2	1:A:1129:GLU:HB3	2.51	0.45
1:A:1383:ILE:O	1:A:1387:THR:HG23	2.18	0.43
1:A:1325:LEU:HD22	1:A:1325:LEU:N	2.34	0.42
1:A:1180:VAL:HG13	1:A:1197:GLU:HB3	2.01	0.41
1:A:1199:MET:O	2:A:9000:VGH:H5	2.20	0.41
1:A:1093:ASN:HA	1:A:1094:PRO:HD3	1.84	0.41
1:A:1169:LEU:O	1:A:1173:LYS:HG2	2.20	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1189:SER:N	3:A:2037:HOH:O	2.50	0.41
1:A:1214:ARG:HG3	1:A:1215:PRO:CD	2.46	0.41
1:A:1122:LEU:HD12	1:A:1130:VAL:HG12	2.03	0.41
1:A:1108:LEU:HD21	1:A:1169:LEU:HD12	2.02	0.41
1:A:1325:LEU:HD21	1:A:1397:LEU:HD22	2.01	0.41
1:A:1312:THR:HG23	1:A:1379:ILE:HD11	2.03	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	293/327 (90%)	281 (96%)	9 (3%)	3 (1%)	15 4

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1126	ALA
1	A	1400	GLU
1	A	1125	GLY

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	260/285 (91%)	258 (99%)	2 (1%)	81	72

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

Mol	Chain	Res	Type
1	A	1118	LEU
1	A	1177	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1093	ASN
1	A	1115	ASN
1	A	1124	HIS
1	A	1177	GLN
1	A	1188	GLN
1	A	1217	GLN
1	A	1238	GLN
1	A	1243	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

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	VGH	A	9000	-	31,33,33	2.34	13 (41%)	36,47,47	1.64	7 (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	VGH	A	9000	-	-	0/12/24/24	0/4/4/4

All (13) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	9000	VGH	C7-N26	5.35	1.41	1.35
2	A	9000	VGH	C18-C12	4.87	1.44	1.38
2	A	9000	VGH	C18-C17	4.75	1.47	1.39
2	A	9000	VGH	C13-C17	4.74	1.46	1.39
2	A	9000	VGH	C4-C15	3.22	1.44	1.38
2	A	9000	VGH	C19-N22	3.00	1.41	1.34
2	A	9000	VGH	C17-C21	2.42	1.57	1.53
2	A	9000	VGH	C5-C14	2.27	1.43	1.39
2	A	9000	VGH	C4-C14	2.23	1.43	1.39
2	A	9000	VGH	C2-C12	2.20	1.42	1.37
2	A	9000	VGH	C2-C3	2.04	1.42	1.38
2	A	9000	VGH	C3-C13	2.03	1.43	1.38
2	A	9000	VGH	O27-C15	2.02	1.41	1.37

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	9000	VGH	C15-O27-C21	4.52	126.12	118.17
2	A	9000	VGH	O27-C21-C17	4.44	114.20	107.64
2	A	9000	VGH	C17-C13-CL	3.07	123.46	120.49
2	A	9000	VGH	C13-C17-C18	2.77	118.09	114.90
2	A	9000	VGH	C17-C18-CL2	2.38	123.06	120.32
2	A	9000	VGH	C1-C21-C17	-2.31	110.76	113.48
2	A	9000	VGH	C4-C15-C19	-2.06	115.93	119.25

There are no chirality outliers.

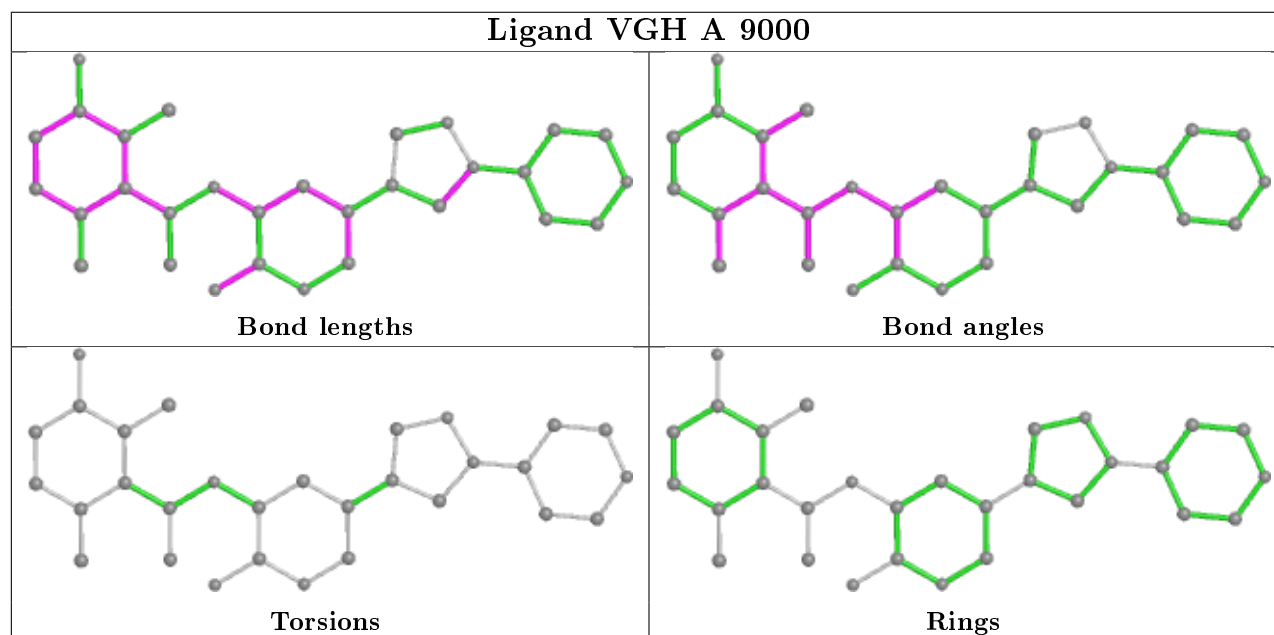
There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	9000	VGH	3	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

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

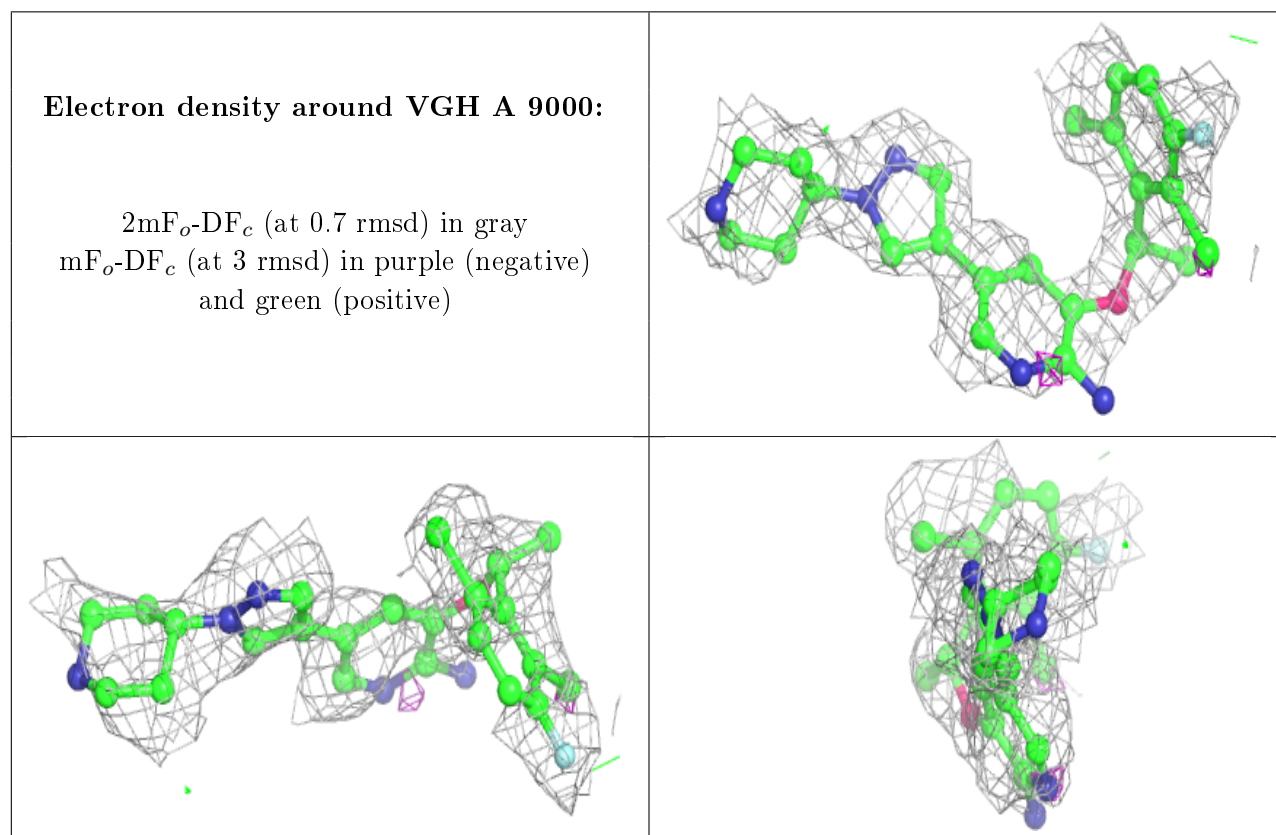
6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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.



6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.