Full wwPDB X-ray Structure Validation Report ⓘ

Feb 6, 2025 – 10:31 AM EST

PDB ID	:	7HZM
Title	:	Group deposition of Coxsackievirus A16 (G-10) 2A protease in complex with inhibitors from the ASAP AViDD centre – Crystal structure of Coxsackievirus A16 (G-10) 2A protease in complex with ASAP-0030611-001 (A71EV2A-x2351)
Authors	:	Lithgo, R.M.; Fairhead, M.; Koekemoer, L.; Balcomb, B.H.; Capkin, E.; Chandran, A.V.; Golding, M.; Godoy, A.S.; Aschenbrenner, J.C.; Marples, P.G.; Ni, X.; Thompson, W.; Tomlinson, C.W.E.; Wild, C.; Winokan, M.; Xavier, M.-A.E.; Kenton, N.; Tucker, J.; DiPoto, M.; Lee, A.; Fearon, D.; von Delft, F.
Deposited on	:	2025-01-14
Resolution	:	1.19 Å(reported)

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

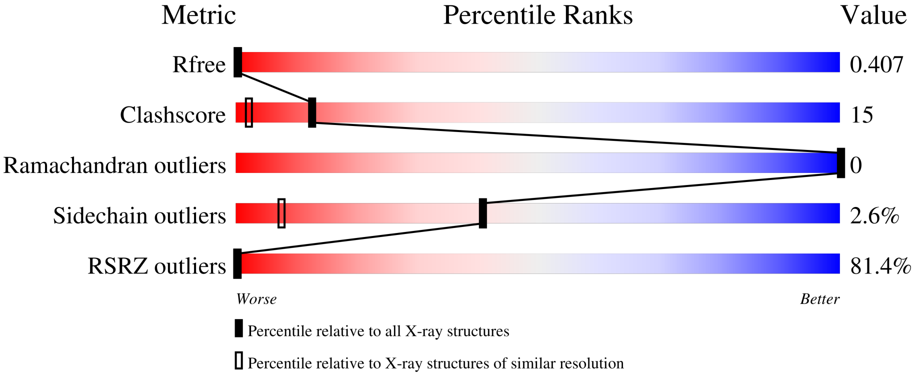
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.21
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

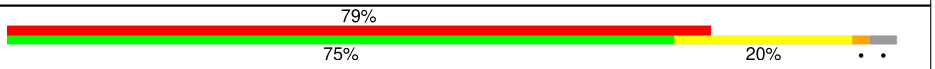
The reported resolution of this entry is 1.19 Å.

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}	164625	1079 (1.20-1.20)
Clashscore	180529	1183 (1.20-1.20)
Ramachandran outliers	177936	1146 (1.20-1.20)
Sidechain outliers	177891	1146 (1.20-1.20)
RSRZ outliers	164620	1078 (1.20-1.20)

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	144	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

---

CCP4 : 9.0.004 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.40

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	DMS	A	202	-	-	-	X
3	DMS	A	204	-	-	-	X

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 1367 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.

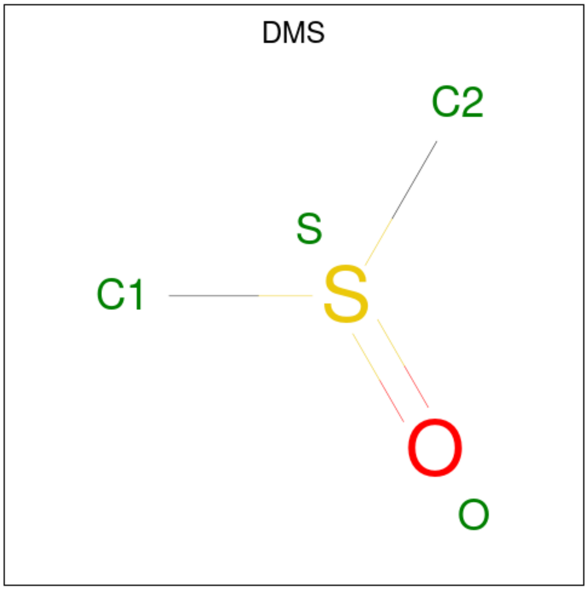
- Molecule 1 is a protein called Protease 2A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	140	1103	686	196	214	7	0	3	0

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
2	A	1	1	1	0	0

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



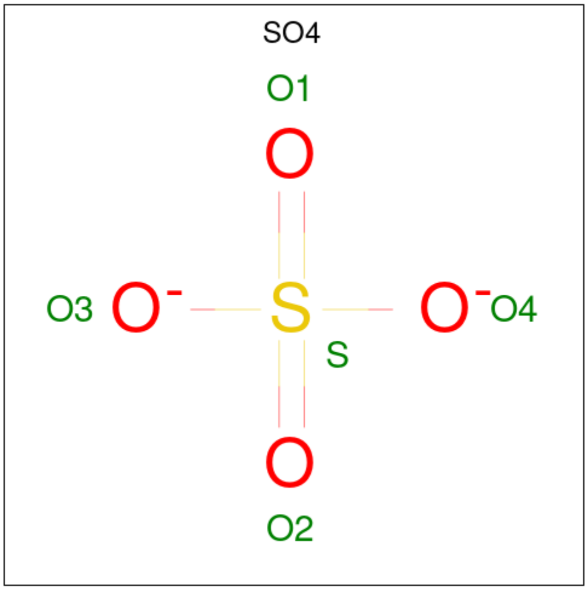
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	S		
3	A	1	4	2	1	1	0	0
3	A	1	4	2	1	1	0	0

Continued on next page...

Continued from previous page...

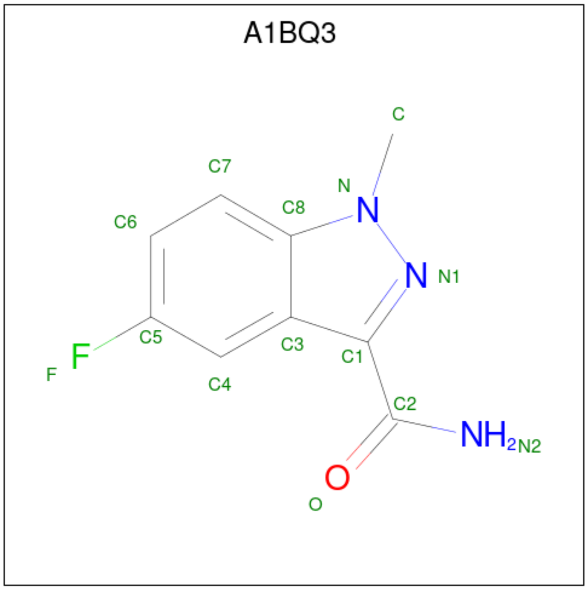
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O S 4 2 1 1	0	0
3	A	1	Total C O S 4 2 1 1	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 5 is 5-fluoro-1-methyl-1H-indazole-3-carboxamide (three-letter code: A1BQ3) (formula: $C_9H_8FN_3O$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C F N O 14 9 1 3 1	0	0

- Molecule 6 is water.

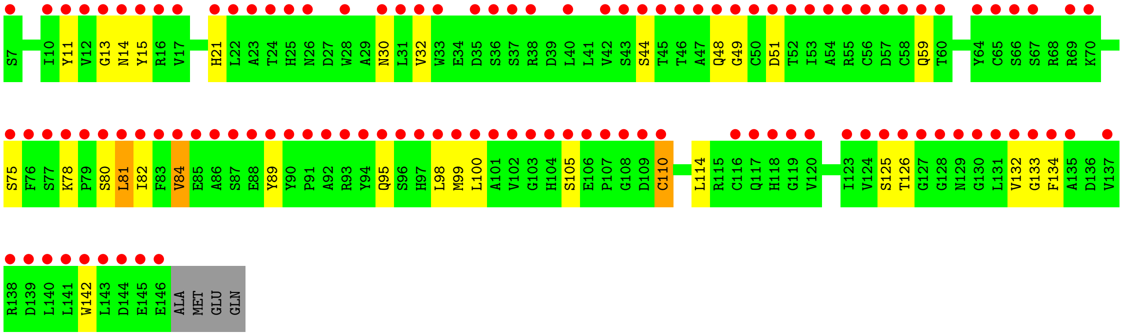
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	228	Total O 228 228	0	0

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: Protease 2A

Chain A:



4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	86.47Å 57.05Å 32.38Å 90.00° 94.82° 90.00°	Depositor
Resolution (Å)	47.57 – 1.19 47.57 – 1.19	Depositor EDS
% Data completeness (in resolution range)	81.1 (47.57-1.19) 81.7 (47.57-1.19)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.37 (at 1.19Å)	Xtriage
Refinement program	REFMAC 5.8.0267, REFMAC5	Depositor
R, R_{free}	0.330 , 0.381 0.360 , 0.407	Depositor DCC
R_{free} test set	2460 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	12.8	Xtriage
Anisotropy	0.011	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 43.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.86	EDS
Total number of atoms	1367	wwPDB-VP
Average B, all atoms (Å^2)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.98% 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: SO4, ZN, DMS, A1BQ3

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	0.80	0/1129	0.95	1/1535 (0.1%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	110	CYS	CA-CB-SG	-5.59	103.93	114.00

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	1103	0	1034	33	0
2	A	1	0	0	0	0
3	A	16	0	24	0	0
4	A	5	0	0	0	0
5	A	14	0	0	0	0
6	A	228	0	0	12	2
All	All	1367	0	1058	33	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.

All (33) 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:99:MET:O	1:A:133:GLY:HA2	1.60	1.02
1:A:98:LEU:HD21	1:A:126:THR:OG1	1.81	0.79
1:A:32:VAL:O	6:A:301:HOH:O	2.11	0.67
1:A:105:SER:N	6:A:303:HOH:O	2.14	0.67
1:A:100:LEU:HA	1:A:132:VAL:O	1.99	0.62
1:A:81:LEU:HD21	6:A:523:HOH:O	2.03	0.58
1:A:13:GLY:O	1:A:48:GLN:HA	2.05	0.57
1:A:95:GLN:NE2	1:A:125:SER:OG	2.38	0.55
1:A:98:LEU:HD11	1:A:133:GLY:HA3	1.88	0.55
1:A:114:LEU:HD22	1:A:134:PHE:CD2	2.42	0.54
1:A:110:CYS:N	6:A:313:HOH:O	2.42	0.53
1:A:51:ASP:N	6:A:309:HOH:O	2.40	0.53
1:A:142:TRP:N	6:A:311:HOH:O	2.41	0.53
1:A:98:LEU:HD22	1:A:126:THR:HG21	1.93	0.51
1:A:78:LYS:HB2	6:A:384:HOH:O	2.10	0.51
1:A:98:LEU:HD13	1:A:126:THR:HG23	1.92	0.50
1:A:82:ILE:HG21	1:A:100:LEU:HD11	1.94	0.49
1:A:15:TYR:OH	6:A:302:HOH:O	2.13	0.48
1:A:84:VAL:HA	6:A:386:HOH:O	2.14	0.47
1:A:30:ASN:HB3	1:A:44:SER:OG	2.15	0.47
1:A:82:ILE:HB	1:A:98:LEU:HD23	1.98	0.46
1:A:11:TYR:HB3	6:A:415:HOH:O	2.16	0.45
1:A:14:ASN:OD1	6:A:304:HOH:O	2.21	0.43
1:A:114:LEU:HD22	1:A:134:PHE:CG	2.54	0.43
1:A:21:HIS:CD2	1:A:89:TYR:CD1	3.07	0.43
1:A:98:LEU:HD11	1:A:126:THR:OG1	2.19	0.42
1:A:59:GLN:O	1:A:75:SER:HA	2.20	0.42
1:A:78:LYS:CB	6:A:384:HOH:O	2.67	0.42
1:A:99:MET:HE2	1:A:134:PHE:O	2.20	0.42
1:A:11:TYR:HA	1:A:15:TYR:O	2.20	0.41
1:A:11:TYR:HB3	1:A:49:GLY:HA3	2.03	0.41
1:A:13:GLY:O	1:A:49:GLY:N	2.54	0.40
1:A:99:MET:HE3	1:A:134:PHE:HB2	2.02	0.40

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)
6:A:338:HOH:O	6:A:433:HOH:O[1_556]	2.03	0.17

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:A:328:HOH:O	6:A:329:HOH:O[2_556]	2.16	0.04

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	141/144 (98%)	135 (96%)	6 (4%)	0	100 100

There are no Ramachandran outliers to report.

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	120/120 (100%)	116 (97%)	4 (3%)	33 4

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	80[A]	SER
1	A	80[B]	SER
1	A	81	LEU
1	A	84	VAL

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

Mol	Chain	Res	Type
1	A	71	HIS
1	A	95	GLN

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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 7 ligands modelled in this entry, 1 is monoatomic - leaving 6 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	$\#\|Z\| > 2$
5	A1BQ3	A	207	-	10,15,15	0.97	0	13,22,22	0.81	0
3	DMS	A	205	-	3,3,3	0.25	0	3,3,3	0.14	0
3	DMS	A	204	-	3,3,3	0.22	0	3,3,3	0.10	0
4	SO4	A	206	-	4,4,4	0.35	0	6,6,6	0.07	0
3	DMS	A	203	-	3,3,3	0.22	0	3,3,3	0.10	0
3	DMS	A	202	-	3,3,3	0.19	0	3,3,3	0.11	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
5	A1BQ3	A	207	-	-	0/0/4/4	0/2/2/2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

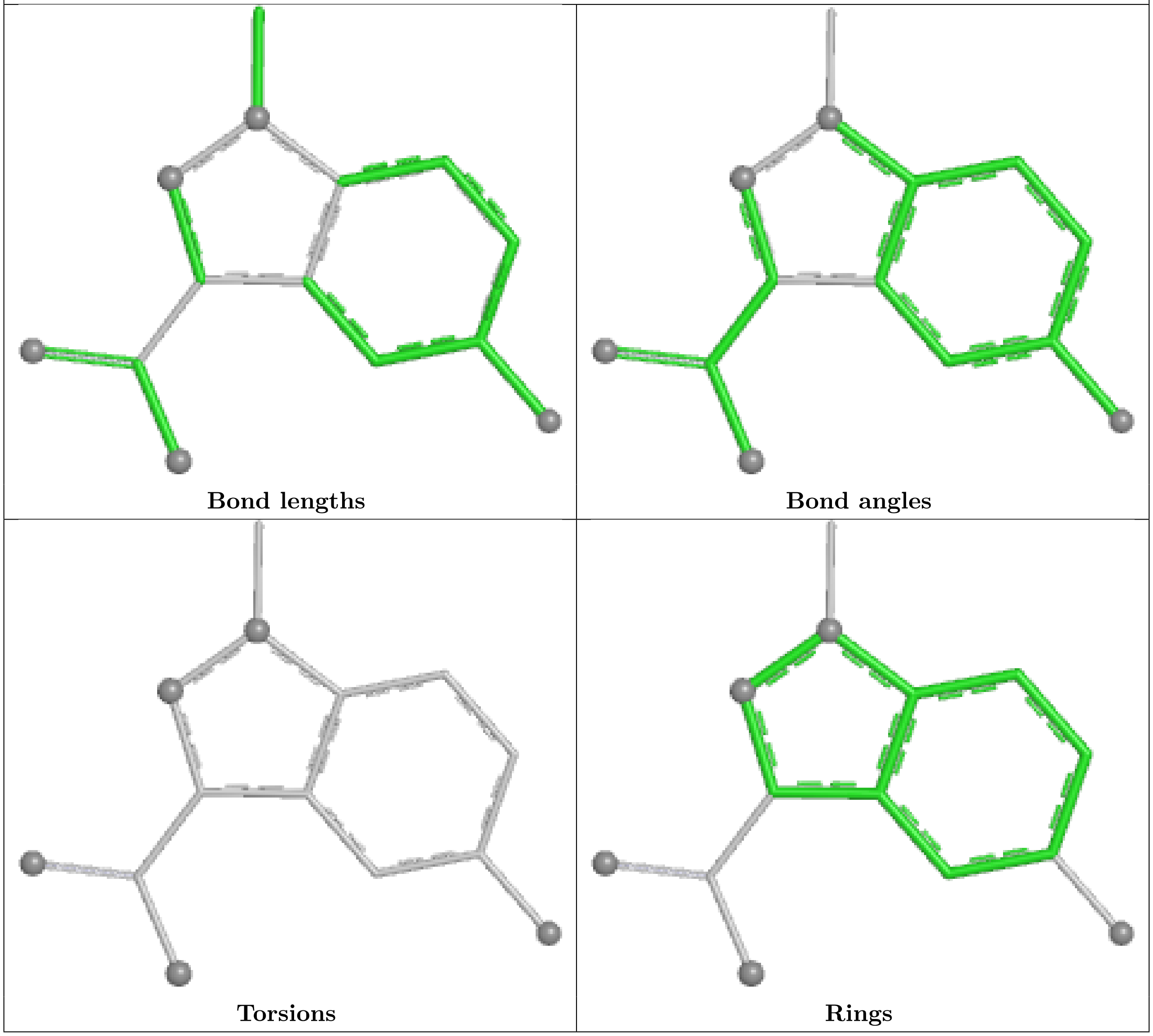
There are no torsion outliers.

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 A1BQ3 A 207



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	140/144 (97%)	5.59	114 (81%)	0	0	3, 14, 36, 62	38 (27%)

All (114) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	89	TYR	15.3
1	A	81	LEU	14.9
1	A	135	ALA	14.9
1	A	82	ILE	14.8
1	A	92	ALA	14.7
1	A	123	ILE	14.5
1	A	90	TYR	14.4
1	A	94	TYR	13.6
1	A	84	VAL	13.6
1	A	134	PHE	13.4
1	A	124	VAL	13.1
1	A	103	GLY	12.5
1	A	50	CYS	12.5
1	A	47	ALA	12.3
1	A	132	VAL	12.3
1	A	131	LEU	12.0
1	A	108	GLY	12.0
1	A	83	PHE	11.7
1	A	91	PRO	11.2
1	A	107	PRO	10.3
1	A	86	ALA	10.0
1	A	129	ASN	9.9
1	A	49	GLY	9.6
1	A	15	TYR	9.4
1	A	130	GLY	9.2
1	A	45	THR	9.2
1	A	80[A]	SER	9.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	133	GLY	9.0
1	A	46	THR	9.0
1	A	36[A]	SER	8.9
1	A	104	HIS	8.8
1	A	32	VAL	8.7
1	A	109	ASP	8.4
1	A	48	GLN	8.0
1	A	57[A]	ASP	7.7
1	A	142	TRP	7.7
1	A	102	VAL	7.7
1	A	13	GLY	7.6
1	A	144	ASP	7.6
1	A	87	SER	7.6
1	A	12	VAL	7.6
1	A	93	ARG	7.6
1	A	126	THR	7.4
1	A	95	GLN	7.1
1	A	88	GLU	6.8
1	A	53	ILE	6.7
1	A	125	SER	6.6
1	A	11	TYR	6.6
1	A	143	LEU	6.5
1	A	145	GLU	6.5
1	A	33	TRP	6.5
1	A	141	LEU	6.4
1	A	54	ALA	6.2
1	A	44	SER	6.1
1	A	100	LEU	6.0
1	A	42	VAL	5.9
1	A	105	SER	5.7
1	A	128	GLY	5.7
1	A	51	ASP	5.6
1	A	146	GLU	5.5
1	A	14	ASN	5.1
1	A	30	ASN	5.1
1	A	127	GLY	4.8
1	A	52	THR	4.7
1	A	106	GLU	4.5
1	A	137	VAL	4.5
1	A	119	GLY	4.5
1	A	101	ALA	4.4
1	A	58	CYS	4.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	26	ASN	4.3
1	A	31	LEU	4.1
1	A	98	LEU	4.1
1	A	21	HIS	4.1
1	A	97	HIS	4.1
1	A	66	SER	4.1
1	A	64	TYR	4.1
1	A	140	LEU	3.9
1	A	79	PRO	3.9
1	A	40	LEU	3.9
1	A	67	SER	3.8
1	A	70	LYS	3.8
1	A	22	LEU	3.8
1	A	65	CYS	3.8
1	A	43	SER	3.8
1	A	69	ARG	3.6
1	A	120	VAL	3.5
1	A	25	HIS	3.3
1	A	37	SER	3.3
1	A	7	SER	3.3
1	A	117	GLN	3.2
1	A	96	SER	3.2
1	A	56	CYS	3.2
1	A	78	LYS	3.1
1	A	10	ILE	3.1
1	A	55	ARG	3.1
1	A	85	GLU	3.0
1	A	60	THR	3.0
1	A	118	HIS	2.9
1	A	99	MET	2.9
1	A	116	CYS	2.9
1	A	35	ASP	2.8
1	A	139	ASP	2.8
1	A	38	ARG	2.7
1	A	24	THR	2.7
1	A	110	CYS	2.6
1	A	77	SER	2.6
1	A	23	ALA	2.6
1	A	16	ARG	2.5
1	A	75	SER	2.4
1	A	59	GLN	2.4
1	A	76	PHE	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	138	ARG	2.2
1	A	28	TRP	2.1
1	A	17	VAL	2.0

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 ⓘ

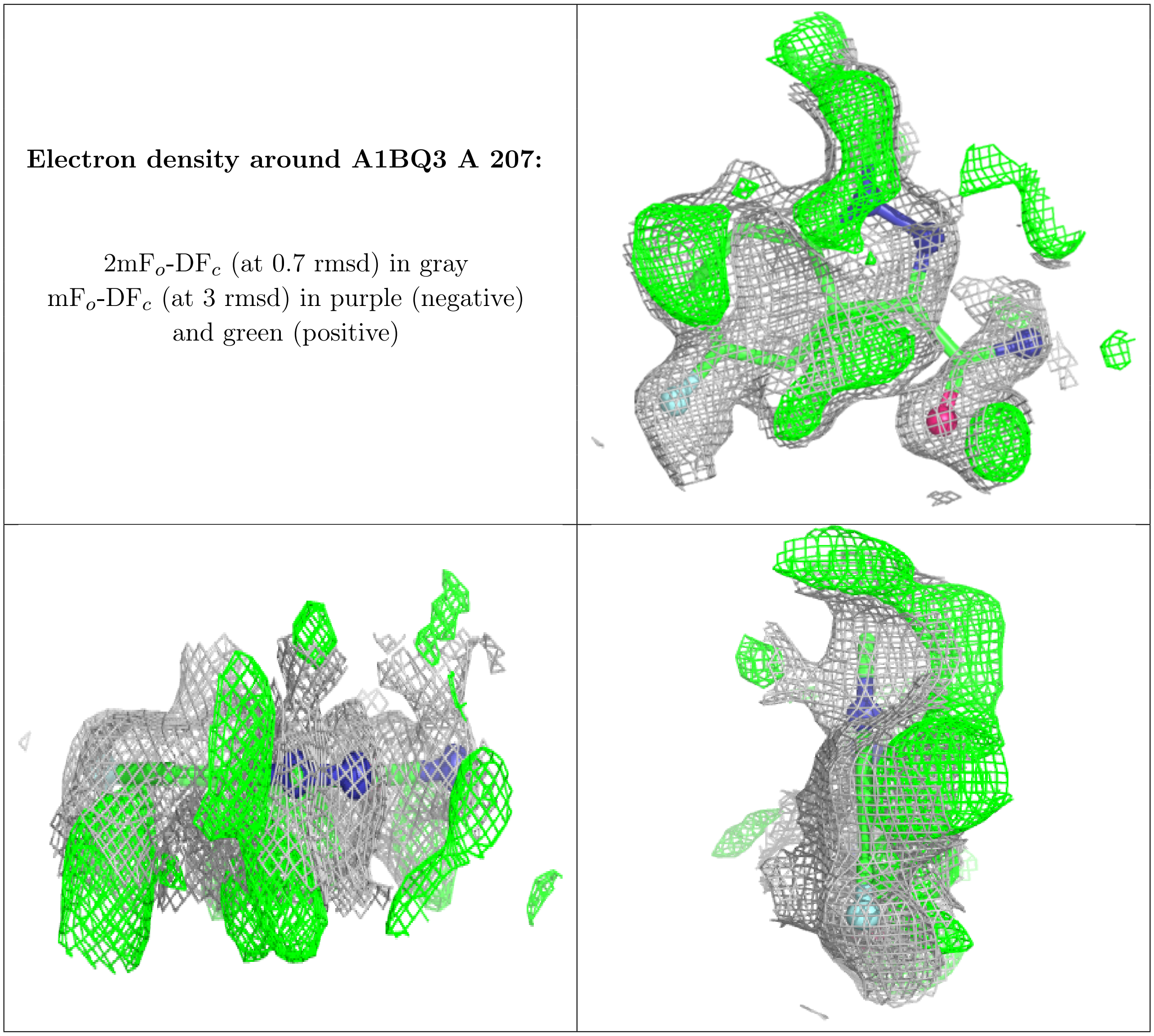
There are no monosaccharides in this entry.

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, 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
4	SO4	A	206	5/5	0.37	0.29	101,101,103,106	0
3	DMS	A	202	4/4	0.41	0.44	103,105,105,107	0
3	DMS	A	204	4/4	0.50	0.43	87,89,90,90	0
5	A1BQ3	A	207	14/14	0.54	0.23	12,12,13,14	14
3	DMS	A	203	4/4	0.57	0.37	79,79,81,82	0
3	DMS	A	205	4/4	0.75	0.25	48,48,49,50	0
2	ZN	A	201	1/1	0.96	0.10	13,13,13,13	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 A1BQ3 A 207:

$2mF_o$-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.