



Full wwPDB X-ray Structure Validation Report ⓘ

Oct 15, 2023 – 06:08 PM EDT

PDB ID : 8DQF
Title : Crystal structure of Neisseria gonorrhoeae carbonic anhydrase with N-(5-sulfamoyl-1,3,4-thiadiazol-2-yl)cyclohexanecarboxamide
Authors : Marapaka, A.K.; Das, C.; Flaherty, D.P.; Yadav, R.
Deposited on : 2022-07-19
Resolution : 2.80 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 : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.36
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) : 2.36

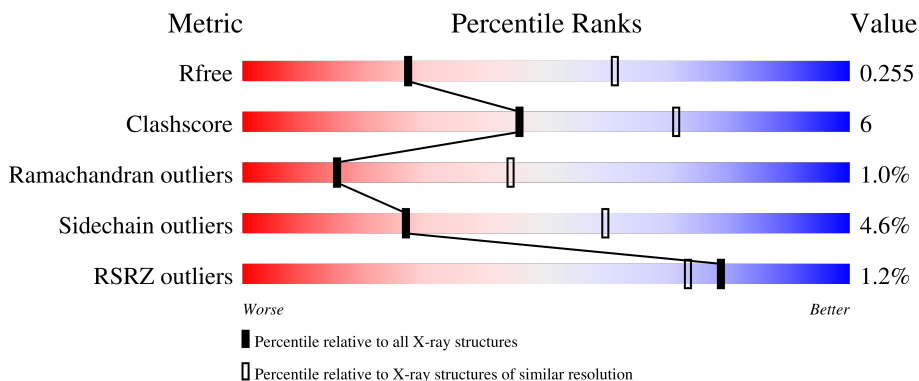
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	243	
1	C	243	
1	E	243	
1	G	243	

2 Entry composition i

There are 5 unique types of molecules in this entry. The entry contains 7114 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 Carbonic anhydrase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	222	Total 1747	C 1105	N 307	O 329	S 6	0	0	0
1	C	222	Total 1732	C 1097	N 305	O 324	S 6	0	0	0
1	E	222	Total 1739	C 1101	N 307	O 325	S 6	0	1	0
1	G	222	Total 1747	C 1106	N 309	O 326	S 6	0	0	0

There are 68 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-16	HIS	-	expression tag	UNP Q50940
A	-15	HIS	-	expression tag	UNP Q50940
A	-14	HIS	-	expression tag	UNP Q50940
A	-13	HIS	-	expression tag	UNP Q50940
A	-12	HIS	-	expression tag	UNP Q50940
A	-11	HIS	-	expression tag	UNP Q50940
A	-10	ASP	-	expression tag	UNP Q50940
A	-9	SER	-	expression tag	UNP Q50940
A	-8	GLY	-	expression tag	UNP Q50940
A	-7	LEU	-	expression tag	UNP Q50940
A	-6	VAL	-	expression tag	UNP Q50940
A	-5	PRO	-	expression tag	UNP Q50940
A	-4	ARG	-	expression tag	UNP Q50940
A	-3	GLY	-	expression tag	UNP Q50940
A	-2	SER	-	expression tag	UNP Q50940
A	-1	HIS	-	expression tag	UNP Q50940
A	0	MET	-	expression tag	UNP Q50940
C	-16	HIS	-	expression tag	UNP Q50940
C	-15	HIS	-	expression tag	UNP Q50940
C	-14	HIS	-	expression tag	UNP Q50940
C	-13	HIS	-	expression tag	UNP Q50940

Continued on next page...

Continued from previous page...

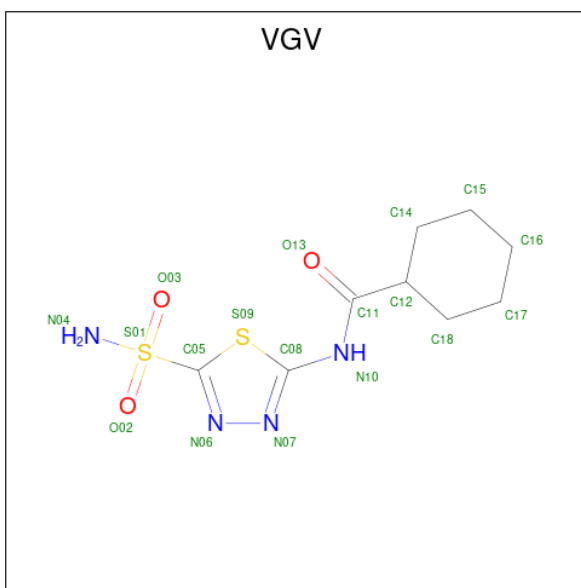
Chain	Residue	Modelled	Actual	Comment	Reference
C	-12	HIS	-	expression tag	UNP Q50940
C	-11	HIS	-	expression tag	UNP Q50940
C	-10	ASP	-	expression tag	UNP Q50940
C	-9	SER	-	expression tag	UNP Q50940
C	-8	GLY	-	expression tag	UNP Q50940
C	-7	LEU	-	expression tag	UNP Q50940
C	-6	VAL	-	expression tag	UNP Q50940
C	-5	PRO	-	expression tag	UNP Q50940
C	-4	ARG	-	expression tag	UNP Q50940
C	-3	GLY	-	expression tag	UNP Q50940
C	-2	SER	-	expression tag	UNP Q50940
C	-1	HIS	-	expression tag	UNP Q50940
C	0	MET	-	expression tag	UNP Q50940
E	-16	HIS	-	expression tag	UNP Q50940
E	-15	HIS	-	expression tag	UNP Q50940
E	-14	HIS	-	expression tag	UNP Q50940
E	-13	HIS	-	expression tag	UNP Q50940
E	-12	HIS	-	expression tag	UNP Q50940
E	-11	HIS	-	expression tag	UNP Q50940
E	-10	ASP	-	expression tag	UNP Q50940
E	-9	SER	-	expression tag	UNP Q50940
E	-8	GLY	-	expression tag	UNP Q50940
E	-7	LEU	-	expression tag	UNP Q50940
E	-6	VAL	-	expression tag	UNP Q50940
E	-5	PRO	-	expression tag	UNP Q50940
E	-4	ARG	-	expression tag	UNP Q50940
E	-3	GLY	-	expression tag	UNP Q50940
E	-2	SER	-	expression tag	UNP Q50940
E	-1	HIS	-	expression tag	UNP Q50940
E	0	MET	-	expression tag	UNP Q50940
G	-16	HIS	-	expression tag	UNP Q50940
G	-15	HIS	-	expression tag	UNP Q50940
G	-14	HIS	-	expression tag	UNP Q50940
G	-13	HIS	-	expression tag	UNP Q50940
G	-12	HIS	-	expression tag	UNP Q50940
G	-11	HIS	-	expression tag	UNP Q50940
G	-10	ASP	-	expression tag	UNP Q50940
G	-9	SER	-	expression tag	UNP Q50940
G	-8	GLY	-	expression tag	UNP Q50940
G	-7	LEU	-	expression tag	UNP Q50940
G	-6	VAL	-	expression tag	UNP Q50940
G	-5	PRO	-	expression tag	UNP Q50940

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	-4	ARG	-	expression tag	UNP Q50940
G	-3	GLY	-	expression tag	UNP Q50940
G	-2	SER	-	expression tag	UNP Q50940
G	-1	HIS	-	expression tag	UNP Q50940
G	0	MET	-	expression tag	UNP Q50940

- Molecule 2 is N-(5-sulfamoyl-1,3,4-thiadiazol-2-yl)cyclohexanecarboxamide (three-letter code: VGV) (formula: C₉H₁₄N₄O₃S₂) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
2	A	1	Total	C	N	O	S	0	0
			18	9	4	3	2		
2	C	1	Total	C	N	O	S	0	0
			18	9	4	3	2		
2	E	1	Total	C	N	O	S	0	0
			18	9	4	3	2		
2	G	1	Total	C	N	O	S	0	0
			18	9	4	3	2		

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
3	A	1	Total	Zn	0	0
			1	1		
3	C	1	Total	Zn	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	E	1	Total Zn 1 1	0	0
3	G	1	Total Zn 1 1	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



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

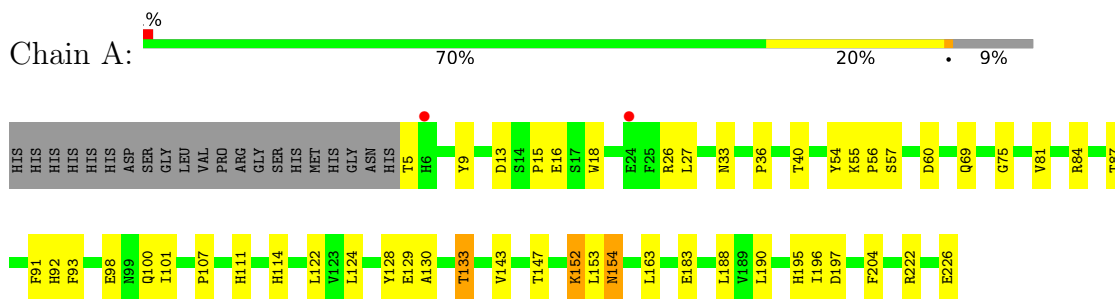
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	13	Total O 13 13	0	0
5	C	6	Total O 6 6	0	0
5	E	3	Total O 3 3	0	0
5	G	6	Total O 6 6	0	0

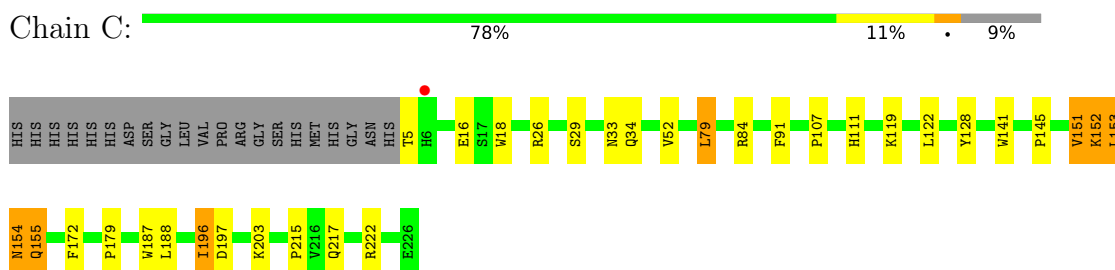
3 Residue-property plots [i](#)

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.

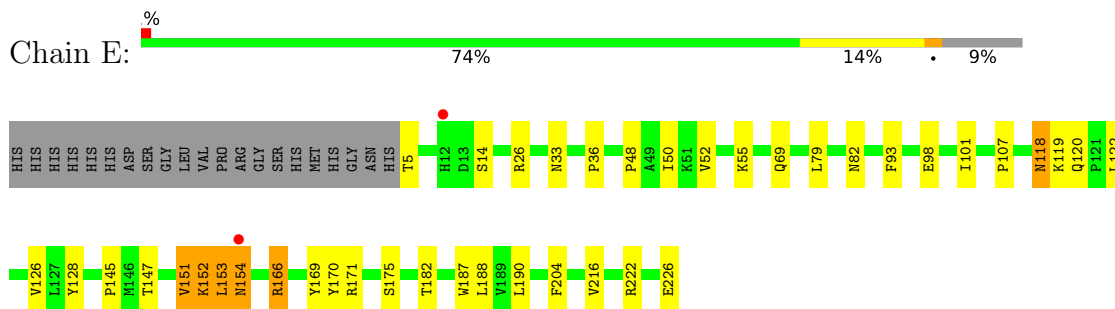
- Molecule 1: Carbonic anhydrase



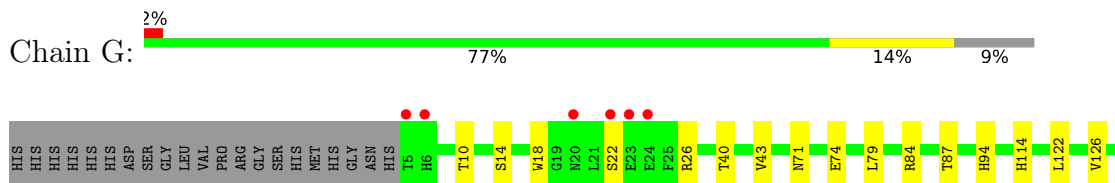
- Molecule 1: Carbonic anhydrase



- Molecule 1: Carbonic anhydrase



- Molecule 1: Carbonic anhydrase





4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	71.91Å 122.82Å 78.53Å 90.00° 117.26° 90.00°	Depositor
Resolution (Å)	46.11 – 2.80 46.11 – 2.80	Depositor EDS
% Data completeness (in resolution range)	98.3 (46.11-2.80) 98.3 (46.11-2.80)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.94 (at 2.81Å)	Xtrriage
Refinement program	REFMAC 5.8.0352	Depositor
R, R_{free}	0.187 , 0.255 0.191 , 0.255	Depositor DCC
R_{free} test set	1405 reflections (4.76%)	wwPDB-VP
Wilson B-factor (Å ²)	30.4	Xtrriage
Anisotropy	0.086	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 31.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.026 for h,-k,-h-l	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	7114	wwPDB-VP
Average B, all atoms (Å ²)	29.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.11% 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: ZN, SO4, VGV

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	0.43	0/1793	0.78	1/2441 (0.0%)
1	C	0.43	0/1778	0.80	1/2423 (0.0%)
1	E	0.45	0/1789	0.80	1/2439 (0.0%)
1	G	0.42	0/1793	0.77	2/2440 (0.1%)
All	All	0.43	0/7153	0.79	5/9743 (0.1%)

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	2
1	C	0	3
1	E	0	1
1	G	0	5
All	All	0	11

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	171	ARG	NE-CZ-NH1	6.23	123.42	120.30
1	G	197	ASP	CB-CA-C	-6.03	98.34	110.40
1	A	154	ASN	CB-CA-C	-5.75	98.89	110.40
1	C	197	ASP	CB-CA-C	-5.63	99.14	110.40
1	G	129	GLU	CB-CA-C	-5.37	99.66	110.40

There are no chirality outliers.

All (11) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	26	ARG	Sidechain
1	A	84	ARG	Sidechain
1	C	155	GLN	Peptide
1	C	26	ARG	Sidechain
1	C	84	ARG	Sidechain
1	E	151	VAL	Peptide
1	G	136	ARG	Sidechain
1	G	166	ARG	Sidechain
1	G	206	ARG	Sidechain
1	G	26	ARG	Sidechain
1	G	84	ARG	Sidechain

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	1747	0	1704	24	0
1	C	1732	0	1681	21	0
1	E	1739	0	1683	26	0
1	G	1747	0	1713	15	0
2	A	18	0	0	1	0
2	C	18	0	0	1	0
2	E	18	0	0	1	0
2	G	18	0	0	2	0
3	A	1	0	0	0	0
3	C	1	0	0	0	0
3	E	1	0	0	0	0
3	G	1	0	0	0	0
4	A	10	0	0	0	0
4	C	15	0	0	0	0
4	E	15	0	0	1	0
4	G	5	0	0	0	0
5	A	13	0	0	0	0
5	C	6	0	0	2	0
5	E	3	0	0	1	0
5	G	6	0	0	0	0
All	All	7114	0	6781	89	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 6.

All (89) 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:E:152:LYS:O	1:E:154:ASN:CB	2.29	0.81
1:E:50:ILE:HG23	1:E:79:LEU:HD21	1.66	0.75
1:C:152:LYS:O	1:C:154:ASN:N	2.18	0.71
1:E:152:LYS:O	1:E:154:ASN:N	2.26	0.69
1:C:33:ASN:HD22	1:C:222:ARG:HD3	1.60	0.65
1:E:33:ASN:HD22	1:E:222:ARG:HD3	1.61	0.65
1:C:79:LEU:C	1:C:79:LEU:HD23	2.18	0.64
2:E:301:VGV:O13	2:E:301:VGV:N07	2.28	0.62
1:C:141:TRP:O	1:C:203:LYS:HE2	2.00	0.61
1:E:36:PRO:HG3	1:E:98:GLU:HB3	1.82	0.61
1:C:145:PRO:HG3	1:C:151:VAL:HG23	1.82	0.60
1:C:29:SER:HB2	5:C:404:HOH:O	2.03	0.59
1:E:188:LEU:N	1:E:188:LEU:HD12	2.17	0.59
1:E:107:PRO:HD2	1:E:128:TYR:O	2.04	0.57
1:A:130:ALA:HB3	1:G:10:THR:HB	1.86	0.57
1:A:143:VAL:HG21	1:A:153:LEU:HD13	1.89	0.55
1:E:126:VAL:HG22	1:E:190:LEU:HD12	1.87	0.55
1:A:101:ILE:HD13	1:A:101:ILE:N	2.21	0.55
1:A:100:GLN:C	1:A:101:ILE:HD13	2.27	0.54
1:A:27:LEU:HD21	1:A:183:GLU:HG3	1.88	0.53
1:A:91:PHE:HA	1:A:111:HIS:O	2.09	0.53
1:G:177:THR:OG1	2:G:301:VGV:N04	2.43	0.52
1:C:172:PHE:HB2	5:C:402:HOH:O	2.09	0.52
1:E:145:PRO:HG3	1:E:151:VAL:HG23	1.92	0.52
2:G:301:VGV:O13	2:G:301:VGV:N07	2.38	0.52
1:A:107:PRO:HD2	1:A:128:TYR:O	2.10	0.51
2:A:301:VGV:O13	2:A:301:VGV:N07	2.41	0.51
1:A:69:GLN:HG3	1:A:92:HIS:HB3	1.93	0.51
1:G:79:LEU:C	1:G:79:LEU:HD23	2.32	0.50
1:E:50:ILE:HG23	1:E:79:LEU:CD2	2.41	0.50
1:A:55:LYS:HB2	1:A:56:PRO:HD2	1.93	0.50
1:C:33:ASN:HD22	1:C:222:ARG:CD	2.23	0.50
1:E:93:PHE:HB2	1:E:204:PHE:CE2	2.46	0.49
1:E:152:LYS:HB2	1:E:152:LYS:NZ	2.26	0.49
1:G:122:LEU:C	1:G:122:LEU:HD23	2.33	0.49
1:A:87:THR:O	1:A:114:HIS:HA	2.13	0.49
1:E:152:LYS:O	1:E:154:ASN:CA	2.60	0.49
1:A:36:PRO:HG3	1:A:98:GLU:HB3	1.96	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:18:TRP:CZ3	1:G:179:PRO:HD3	2.49	0.47
1:G:126:VAL:HG22	1:G:190:LEU:HD12	1.97	0.47
1:E:118:ASN:O	1:E:119:LYS:CB	2.63	0.47
1:A:107:PRO:HB2	1:A:196:ILE:HD11	1.95	0.47
1:C:153:LEU:O	1:C:155:GLN:N	2.39	0.46
1:E:187:TRP:C	1:E:188:LEU:HD12	2.35	0.46
1:C:34:GLN:HB3	1:C:217:GLN:HG3	1.98	0.46
1:A:93:PHE:HB2	1:A:204:PHE:CE1	2.51	0.46
1:C:153:LEU:C	1:C:155:GLN:N	2.70	0.46
1:C:107:PRO:HB2	1:C:196:ILE:HD11	1.98	0.45
1:A:124:LEU:HD11	1:A:190:LEU:HD11	1.98	0.45
1:C:122:LEU:C	1:C:122:LEU:HD23	2.37	0.45
1:A:9:TYR:HA	1:A:15:PRO:HD3	1.99	0.45
1:C:153:LEU:C	1:C:155:GLN:H	2.17	0.45
1:G:94:HIS:HA	1:G:212:ASN:OD1	2.17	0.45
1:E:48:PRO:HB3	1:E:82:ASN:HD22	1.82	0.45
1:E:170:TYR:O	1:E:188:LEU:HA	2.17	0.45
1:C:215:PRO:O	1:C:217:GLN:NE2	2.47	0.44
1:C:18:TRP:CZ3	1:C:179:PRO:HD3	2.53	0.44
1:E:152:LYS:NZ	1:E:152:LYS:CB	2.80	0.44
1:C:91:PHE:HA	1:C:111:HIS:O	2.18	0.44
1:G:187:TRP:C	1:G:188:LEU:HD12	2.38	0.44
1:A:54:TYR:CE2	1:A:163:LEU:HD11	2.53	0.43
1:A:60:ASP:CB	1:A:152:LYS:HA	2.49	0.43
1:G:71:ASN:HD22	1:G:150:LYS:HE2	1.83	0.43
1:E:169:TYR:HB3	1:E:190:LEU:HD23	2.01	0.43
1:G:226:GLU:O	1:G:226:GLU:HG2	2.18	0.43
1:E:188:LEU:N	1:E:188:LEU:CD1	2.82	0.43
1:C:152:LYS:HB3	1:C:152:LYS:NZ	2.34	0.43
1:A:129:GLU:O	1:A:195:HIS:HA	2.19	0.42
1:G:132:LYS:O	1:G:195:HIS:HD2	2.03	0.42
1:E:151:VAL:HA	5:E:403:HOH:O	2.18	0.42
1:G:188:LEU:HD12	1:G:188:LEU:N	2.34	0.42
1:E:122:LEU:C	1:E:122:LEU:HD23	2.40	0.42
1:G:43:VAL:O	1:G:225:ILE:HA	2.20	0.42
1:A:122:LEU:HD11	1:A:188:LEU:HD13	2.02	0.41
1:A:27:LEU:CD2	1:A:183:GLU:HG3	2.50	0.41
1:A:33:ASN:HD22	1:A:222:ARG:HD3	1.86	0.41
1:E:26:ARG:HD2	1:E:26:ARG:HA	1.89	0.41
1:C:33:ASN:HD22	1:C:222:ARG:CG	2.34	0.41
1:G:209:GLY:O	1:G:210:SER:HB3	2.21	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:301:VGV:O13	2:C:301:VGV:N07	2.51	0.41
1:E:69:GLN:HE21	1:E:69:GLN:HB2	1.67	0.41
1:C:187:TRP:C	1:C:188:LEU:HD12	2.41	0.41
1:A:133:THR:HA	1:A:197:ASP:HB2	2.02	0.41
1:A:60:ASP:HB2	1:A:152:LYS:HA	2.03	0.40
1:C:107:PRO:HD2	1:C:128:TYR:O	2.21	0.40
1:E:166:ARG:HD3	4:E:304:SO4:O1	2.22	0.40
1:E:175:SER:HA	1:E:182:THR:O	2.20	0.40
1:A:15:PRO:HG3	1:A:18:TRP:CZ2	2.56	0.40
1:G:87:THR:O	1:G:114:HIS:HA	2.22	0.40

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	220/243 (90%)	209 (95%)	10 (4%)	1 (0%)	29	61
1	C	220/243 (90%)	204 (93%)	14 (6%)	2 (1%)	17	46
1	E	221/243 (91%)	208 (94%)	10 (4%)	3 (1%)	11	34
1	G	220/243 (90%)	206 (94%)	11 (5%)	3 (1%)	11	34
All	All	881/972 (91%)	827 (94%)	45 (5%)	9 (1%)	15	44

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	153	LEU
1	E	153	LEU
1	A	75	GLY
1	C	154	ASN
1	E	154	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	G	22	SER
1	G	14	SER
1	G	40	THR
1	E	14	SER

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	193/213 (91%)	182 (94%)	11 (6%)	20	50
1	C	189/213 (89%)	181 (96%)	8 (4%)	30	63
1	E	190/213 (89%)	178 (94%)	12 (6%)	18	46
1	G	193/213 (91%)	189 (98%)	4 (2%)	53	84
All	All	765/852 (90%)	730 (95%)	35 (5%)	27	60

All (35) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	5	THR
1	A	13	ASP
1	A	16	GLU
1	A	40	THR
1	A	57	SER
1	A	81	VAL
1	A	133	THR
1	A	147	THR
1	A	152	LYS
1	A	154	ASN
1	A	226	GLU
1	C	5	THR
1	C	16	GLU
1	C	52	VAL
1	C	79	LEU
1	C	119	LYS
1	C	151	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	152	LYS
1	C	196	ILE
1	E	5	THR
1	E	52	VAL
1	E	55	LYS
1	E	101	ILE
1	E	118	ASN
1	E	120	GLN
1	E	147	THR
1	E	152	LYS
1	E	153	LEU
1	E	166	ARG
1	E	216	VAL
1	E	226	GLU
1	G	74	GLU
1	G	154	ASN
1	G	166	ARG
1	G	182	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (16) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	33	ASN
1	A	66	HIS
1	A	155	GLN
1	A	195	HIS
1	C	33	ASN
1	C	53	ASN
1	C	66	HIS
1	C	69	GLN
1	C	118	ASN
1	E	20	ASN
1	E	33	ASN
1	E	69	GLN
1	E	82	ASN
1	E	142	ASN
1	G	71	ASN
1	G	195	HIS

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](#)

Of 17 ligands modelled in this entry, 4 are monoatomic - leaving 13 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
4	SO4	C	303	-	4,4,4	0.26	0	6,6,6	0.06	0
4	SO4	A	304	-	4,4,4	0.31	0	6,6,6	0.12	0
2	VGW	G	301	3	14,19,19	2.47	5 (35%)	18,27,27	5.47	11 (61%)
4	SO4	G	303	-	4,4,4	0.37	0	6,6,6	0.38	0
4	SO4	E	303	-	4,4,4	0.42	0	6,6,6	0.21	0
2	VGW	A	301	3	14,19,19	2.24	4 (28%)	18,27,27	5.74	15 (83%)
4	SO4	C	304	-	4,4,4	0.29	0	6,6,6	0.14	0
4	SO4	E	304	-	4,4,4	0.29	0	6,6,6	0.11	0
4	SO4	E	305	-	4,4,4	0.32	0	6,6,6	0.13	0
2	VGW	E	301	3	14,19,19	2.96	5 (35%)	18,27,27	6.01	12 (66%)
4	SO4	A	303	-	4,4,4	0.37	0	6,6,6	0.16	0
2	VGW	C	301	3	14,19,19	2.63	6 (42%)	18,27,27	3.37	10 (55%)
4	SO4	C	305	-	4,4,4	0.33	0	6,6,6	0.17	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	VGV	A	301	3	-	0/6/22/22	0/2/2/2
2	VGV	C	301	3	-	4/6/22/22	0/2/2/2
2	VGV	G	301	3	-	0/6/22/22	0/2/2/2
2	VGV	E	301	3	-	0/6/22/22	0/2/2/2

All (20) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	301	VGV	S01-N04	6.12	1.72	1.60
2	C	301	VGV	S01-N04	6.08	1.72	1.60
2	G	301	VGV	S01-N04	5.93	1.72	1.60
2	E	301	VGV	C11-N10	5.68	1.48	1.35
2	E	301	VGV	C08-N10	5.27	1.45	1.36
2	A	301	VGV	S01-N04	5.21	1.70	1.60
2	C	301	VGV	C11-N10	4.72	1.46	1.35
2	G	301	VGV	C11-N10	4.32	1.45	1.35
2	C	301	VGV	C08-N10	4.01	1.43	1.36
2	G	301	VGV	C08-N10	3.85	1.43	1.36
2	A	301	VGV	C11-N10	3.71	1.44	1.35
2	A	301	VGV	C08-N10	3.71	1.43	1.36
2	E	301	VGV	O03-S01	2.77	1.48	1.43
2	C	301	VGV	O03-S01	2.77	1.48	1.43
2	A	301	VGV	O03-S01	2.31	1.48	1.43
2	E	301	VGV	O13-C11	-2.30	1.18	1.23
2	G	301	VGV	O13-C11	-2.25	1.18	1.23
2	C	301	VGV	O13-C11	-2.22	1.19	1.23
2	C	301	VGV	O02-S01	2.02	1.47	1.43
2	G	301	VGV	O03-S01	2.00	1.47	1.43

All (48) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	301	VGV	O03-S01-N04	14.00	128.12	107.36
2	E	301	VGV	O03-S01-N04	13.96	128.06	107.36
2	E	301	VGV	O03-S01-O02	-12.51	98.21	118.76
2	G	301	VGV	O03-S01-O02	-12.41	98.36	118.76
2	G	301	VGV	O03-S01-N04	11.99	125.14	107.36
2	A	301	VGV	O03-S01-O02	-10.85	100.93	118.76
2	G	301	VGV	O02-S01-C05	-10.53	96.22	108.59
2	A	301	VGV	O03-S01-C05	-10.24	96.56	108.59
2	E	301	VGV	C12-C11-N10	8.80	127.95	115.02

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	301	VG	O02-S01-C05	-8.39	98.73	108.59
2	A	301	VG	O02-S01-C05	-7.78	99.45	108.59
2	C	301	VG	O02-S01-C05	-7.50	99.78	108.59
2	E	301	VG	C08-N10-C11	-6.79	113.92	129.02
2	C	301	VG	O02-S01-N04	6.31	116.71	107.36
2	G	301	VG	C08-N10-C11	-6.23	115.17	129.02
2	C	301	VG	O03-S01-O02	-6.06	108.80	118.76
2	E	301	VG	O13-C11-C12	-5.74	114.65	122.12
2	A	301	VG	C08-N10-C11	-5.51	116.76	129.02
2	C	301	VG	C08-N10-C11	-4.91	118.09	129.02
2	A	301	VG	C05-S01-N04	4.45	116.44	108.25
2	E	301	VG	C16-C15-C14	4.44	120.47	111.42
2	G	301	VG	C12-C11-N10	3.99	120.88	115.02
2	E	301	VG	O02-S01-N04	3.94	113.20	107.36
2	G	301	VG	C05-S01-N04	3.80	115.24	108.25
2	G	301	VG	O03-S01-C05	-3.58	104.38	108.59
2	G	301	VG	O02-S01-N04	3.55	112.62	107.36
2	C	301	VG	C12-C11-N10	3.41	120.03	115.02
2	G	301	VG	C14-C12-C11	-3.25	103.54	110.69
2	A	301	VG	C16-C17-C18	3.17	117.87	111.42
2	A	301	VG	C14-C12-C11	-3.09	103.89	110.69
2	E	301	VG	O13-C11-N10	-3.04	117.39	123.93
2	G	301	VG	C16-C17-C18	3.02	117.57	111.42
2	A	301	VG	C17-C18-C12	2.81	116.66	111.33
2	A	301	VG	C17-C16-C15	2.69	119.45	111.18
2	C	301	VG	C16-C17-C18	2.65	116.81	111.42
2	E	301	VG	C15-C14-C12	2.63	116.31	111.33
2	C	301	VG	O13-C11-C12	-2.60	118.74	122.12
2	A	301	VG	C12-C11-N10	2.50	118.69	115.02
2	E	301	VG	O03-S01-C05	-2.46	105.70	108.59
2	C	301	VG	C05-S01-N04	2.39	112.66	108.25
2	G	301	VG	O13-C11-N10	-2.37	118.83	123.93
2	A	301	VG	O02-S01-N04	2.33	110.82	107.36
2	A	301	VG	C16-C15-C14	2.31	116.12	111.42
2	C	301	VG	C15-C14-C12	2.22	115.53	111.33
2	C	301	VG	O03-S01-C05	2.21	111.19	108.59
2	E	301	VG	C17-C18-C12	2.19	115.48	111.33
2	A	301	VG	O13-C11-N10	-2.12	119.37	123.93
2	A	301	VG	C15-C14-C12	2.08	115.26	111.33

There are no chirality outliers.

All (4) torsion outliers are listed below:

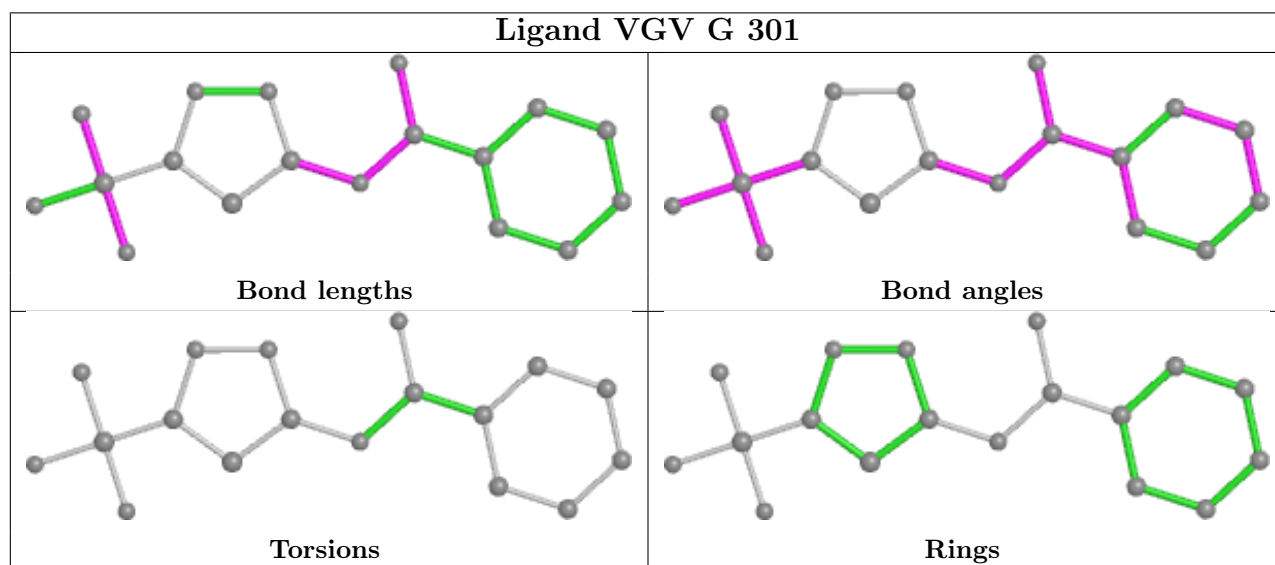
Mol	Chain	Res	Type	Atoms
2	C	301	VGW	O13-C11-C12-C14
2	C	301	VGW	N10-C11-C12-C14
2	C	301	VGW	O13-C11-C12-C18
2	C	301	VGW	N10-C11-C12-C18

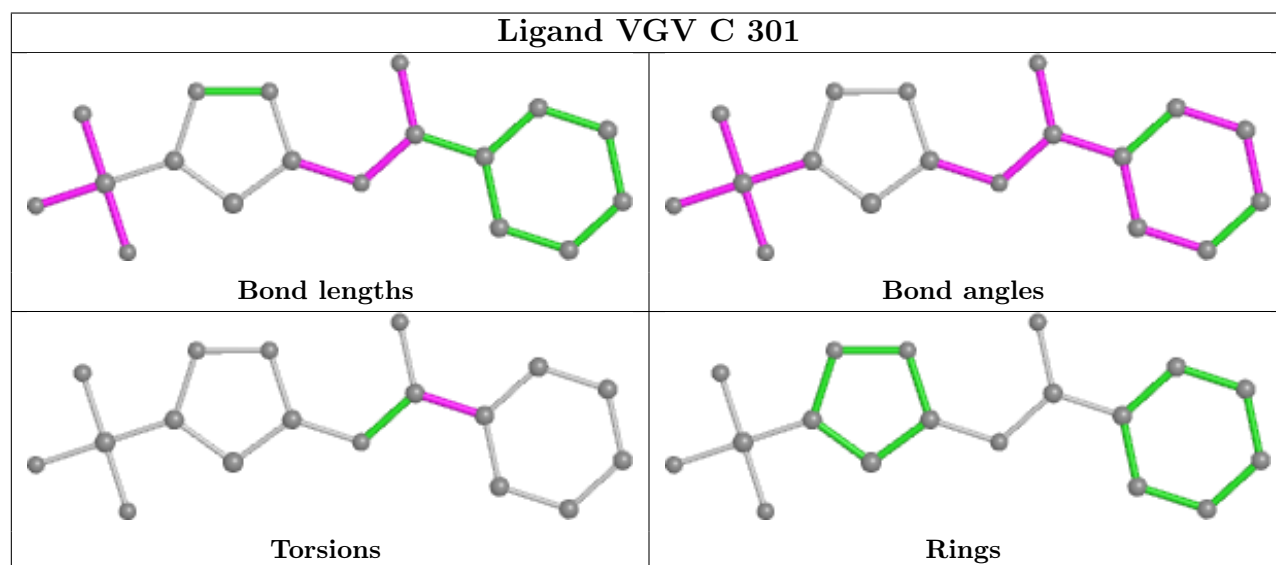
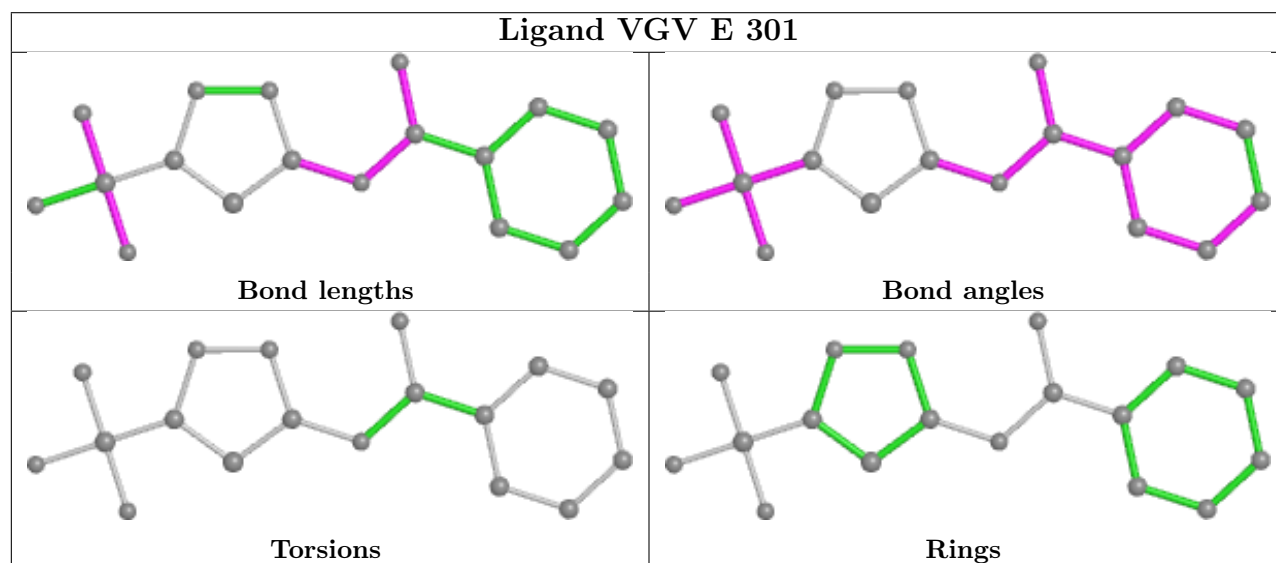
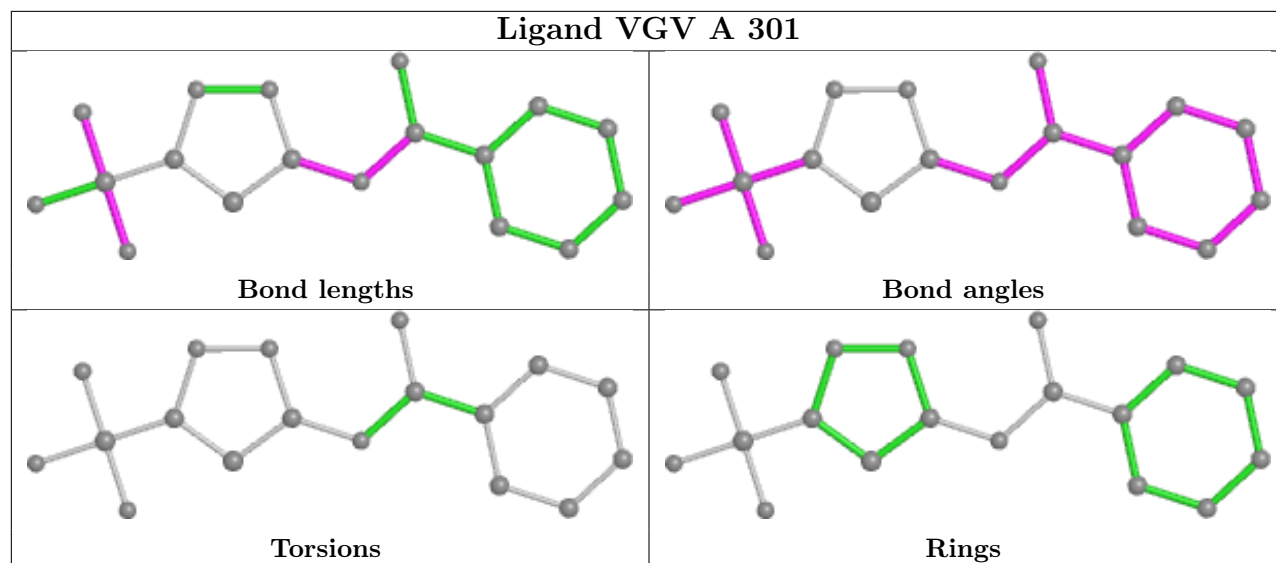
There are no ring outliers.

5 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	G	301	VGW	2	0
2	A	301	VGW	1	0
4	E	304	SO4	1	0
2	E	301	VGW	1	0
2	C	301	VGW	1	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 [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	222/243 (91%)	-0.52	2 (0%) 84 80	17, 27, 43, 64	0
1	C	222/243 (91%)	-0.52	1 (0%) 91 88	15, 26, 48, 79	0
1	E	222/243 (91%)	-0.42	2 (0%) 84 80	17, 27, 43, 72	0
1	G	222/243 (91%)	-0.39	6 (2%) 54 44	19, 30, 47, 65	0
All	All	888/972 (91%)	-0.46	11 (1%) 79 73	15, 27, 47, 79	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	24	GLU	3.4
1	E	154	ASN	3.1
1	G	22	SER	2.9
1	G	20	ASN	2.8
1	G	6	HIS	2.5
1	G	5	THR	2.5
1	C	6	HIS	2.2
1	A	6	HIS	2.1
1	E	12	HIS	2.1
1	A	24	GLU	2.1
1	G	23	GLU	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

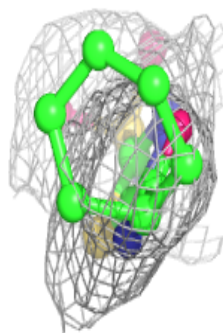
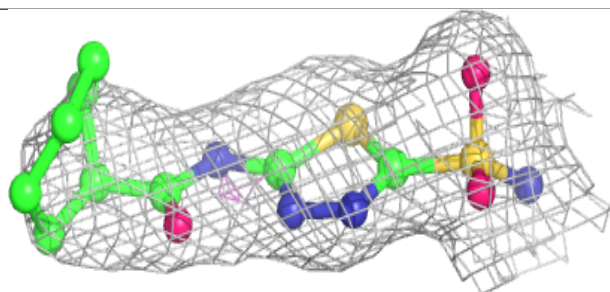
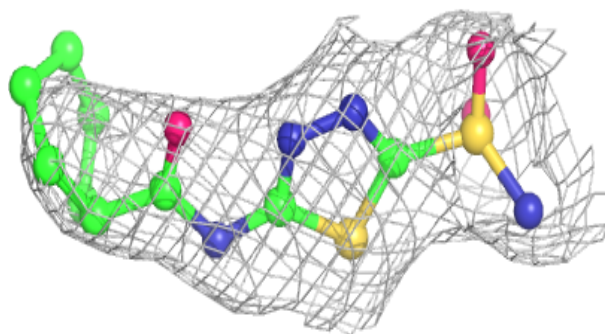
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
4	SO4	A	304	5/5	0.95	0.15	51,54,59,64	0
4	SO4	C	305	5/5	0.95	0.25	51,52,56,61	0
4	SO4	E	304	5/5	0.95	0.18	48,56,61,63	0
4	SO4	C	304	5/5	0.97	0.16	44,50,54,55	0
2	VGW	G	301	18/18	0.97	0.21	28,38,60,60	0
2	VGW	E	301	18/18	0.97	0.17	25,35,45,45	0
4	SO4	E	305	5/5	0.97	0.17	51,53,55,57	0
4	SO4	E	303	5/5	0.98	0.12	36,39,43,44	0
2	VGW	C	301	18/18	0.98	0.15	20,31,50,51	0
2	VGW	A	301	18/18	0.98	0.17	27,37,54,54	0
4	SO4	G	303	5/5	0.98	0.10	29,29,31,34	0
3	ZN	C	302	1/1	0.99	0.07	20,20,20,20	0
3	ZN	E	302	1/1	0.99	0.09	19,19,19,19	0
3	ZN	G	302	1/1	0.99	0.09	25,25,25,25	0
4	SO4	A	303	5/5	0.99	0.08	25,27,31,33	0
3	ZN	A	302	1/1	0.99	0.08	22,22,22,22	0
4	SO4	C	303	5/5	0.99	0.09	23,24,26,26	0

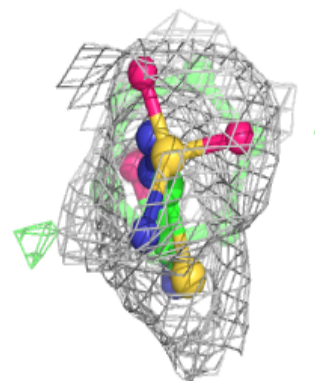
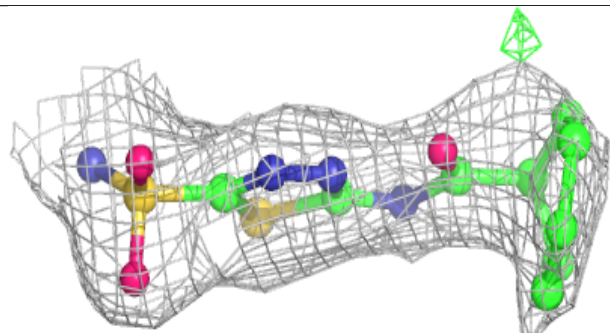
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around VGV G 301:

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

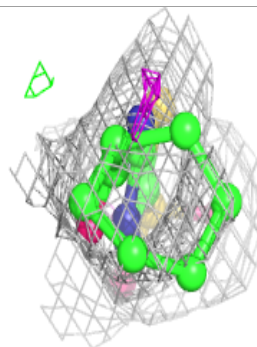
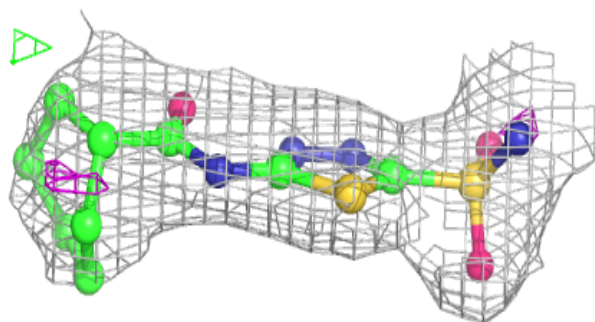
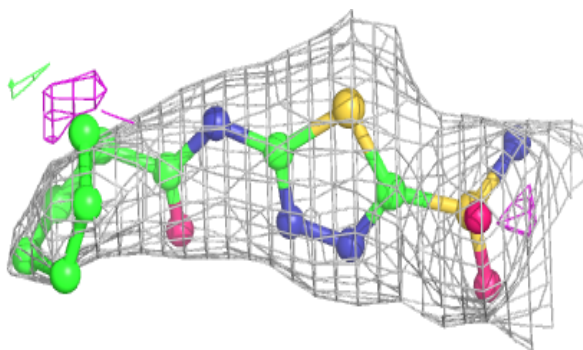
**Electron density around VGV E 301:**

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

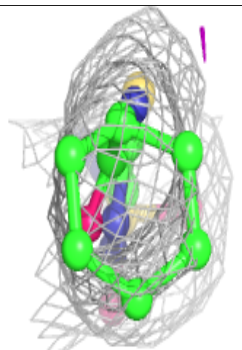
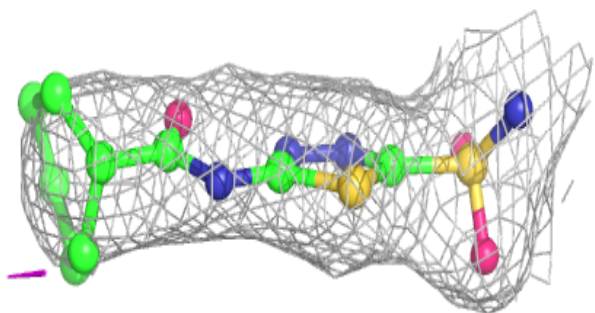
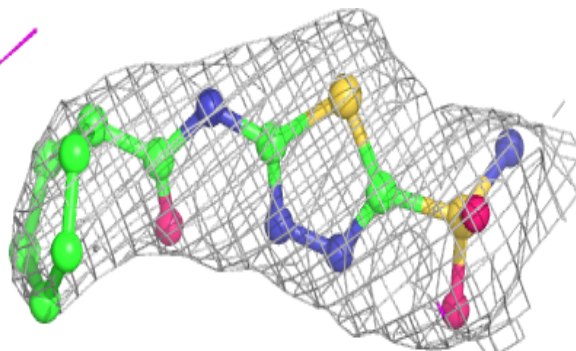


Electron density around VGV C 301:

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

**Electron density around VGV A 301:**

$2mF_o-DF_c$ (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.