



# Full wwPDB X-ray Structure Validation Report ⓘ

Dec 8, 2022 – 04:36 PM EST

PDB ID : 8DR2  
Title : Crystal structure of Neisseria gonorrhoeae carbonic anhydrase with 2-cyclohexyl-N-(5-sulfamoyl-1,3,4-thiadiazol-2-yl)acetamide  
Authors : Marapaka, A.K.; Das, C.; Flaherty, D.P.; Yadav, R.  
Deposited on : 2022-07-20  
Resolution : 2.81 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](mailto: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	:	2.31.2
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)	:	2.31.2

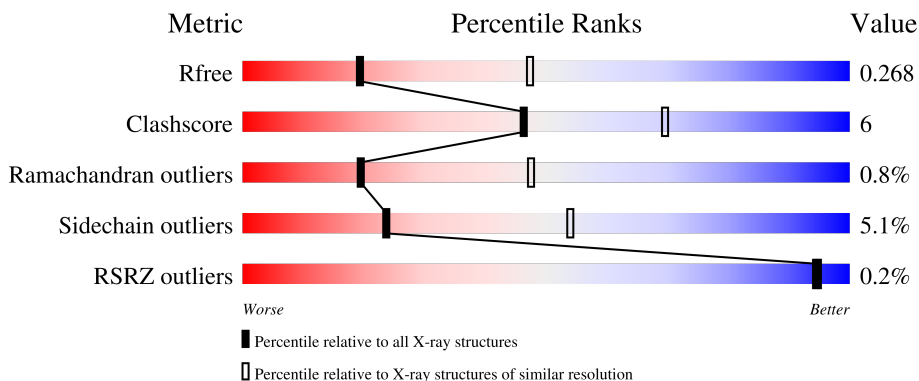
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*





The reported resolution of this entry is 2.81 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	3617 (2.84-2.80)
Clashscore	141614	4060 (2.84-2.80)
Ramachandran outliers	138981	3978 (2.84-2.80)
Sidechain outliers	138945	3980 (2.84-2.80)
RSRZ outliers	127900	3552 (2.84-2.80)

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  $\geq 3$ , 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

There are 5 unique types of molecules in this entry. The entry contains 7130 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
1	A	222	Total	C	N	O	S	0	0	0
			1743	1103	308	326	6			
1	C	222	Total	C	N	O	S	0	0	0
			1740	1102	306	326	6			
1	E	222	Total	C	N	O	S	0	0	0
			1748	1107	307	328	6			
1	G	222	Total	C	N	O	S	0	0	0
			1751	1107	308	330	6			

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...*

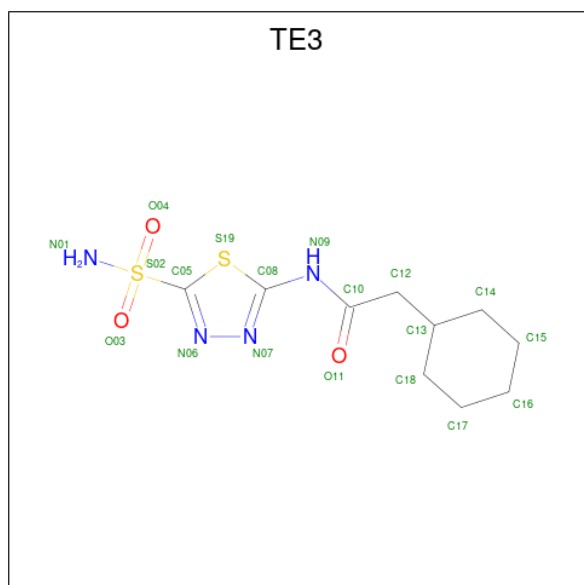
Chain	Residue	Modelled	Actual	Comment	Reference
C	-12	HIS	-	expression tag	UNP Q50940
C	-11	HIS	-	expression tag	UNP Q50940
C	-10	ASP	-	expression tag	UNP Q50940
C	-9	SER	-	expression tag	UNP Q50940
C	-8	GLY	-	expression tag	UNP Q50940
C	-7	LEU	-	expression tag	UNP Q50940
C	-6	VAL	-	expression tag	UNP 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 2-cyclohexyl-N-(5-sulfamoyl-1,3,4-thiadiazol-2-yl)acetamide (three-letter code: TE3) (formula:  $C_{10}H_{16}N_4O_3S_2$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			19	10	4	3	2		
2	C	1	Total	C	N	O	S	0	0
			19	10	4	3	2		
2	E	1	Total	C	N	O	S	0	0
			19	10	4	3	2		
2	G	1	Total	C	N	O	S	0	0
			19	10	4	3	2		

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

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

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

*Continued on next page...*

*Continued from previous page...*

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

- Molecule 5 is water.

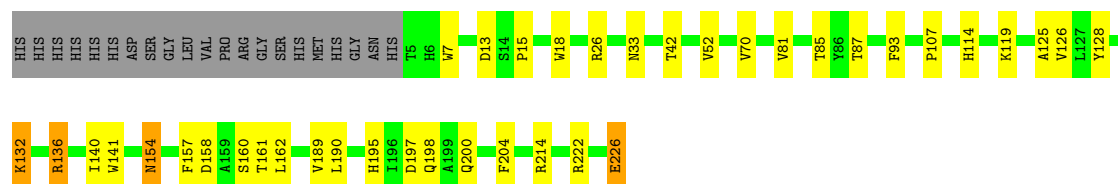
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	9	Total	O	0	0
			9	9		
5	C	2	Total	O	0	0
			2	2		
5	E	2	Total	O	0	0
			2	2		
5	G	5	Total	O	0	0
			5	5		

### 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.

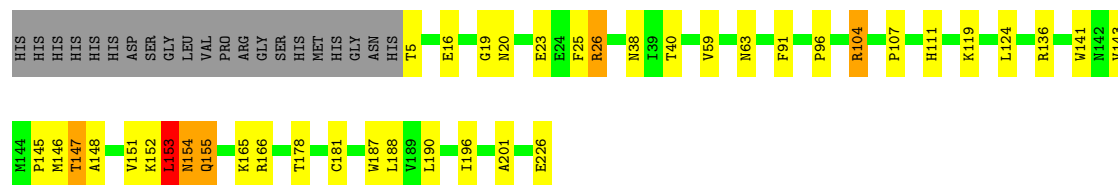
#### • Molecule 1: Carbonic anhydrase

Chain A: 



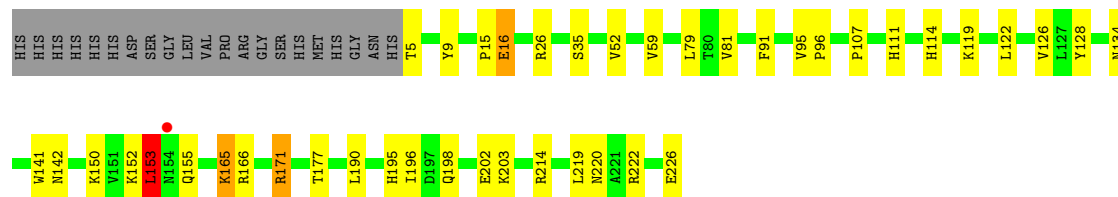
#### • Molecule 1: Carbonic anhydrase

Chain C: 




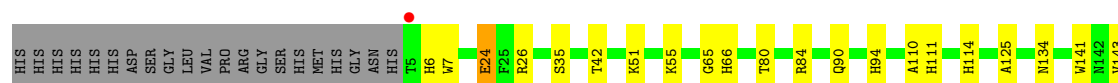
#### • Molecule 1: Carbonic anhydrase

Chain E: 



#### • Molecule 1: Carbonic anhydrase

Chain G: 







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	72.67Å 120.58Å 80.42Å 90.00° 116.62° 90.00°	Depositor
Resolution (Å)	46.20 – 2.81 46.20 – 2.81	Depositor EDS
% Data completeness (in resolution range)	91.0 (46.20-2.81) 91.1 (46.20-2.81)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.70 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.8.0352	Depositor
R, $R_{free}$	0.192 , 0.266 0.197 , 0.268	Depositor DCC
$R_{free}$ test set	1321 reflections (4.81%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.4	Xtriage
Anisotropy	0.114	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 32.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.057 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	7130	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.77% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: TE3, SO4, ZN

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	0.42	0/1789	0.75	0/2436
1	C	0.43	0/1786	0.78	0/2432
1	E	0.42	0/1794	0.78	0/2441
1	G	0.41	0/1797	0.77	0/2446
All	All	0.42	0/7166	0.77	0/9755

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	3
1	C	0	5
1	E	0	3
1	G	0	1
All	All	0	12

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (12) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	136	ARG	Sidechain
1	A	214	ARG	Sidechain
1	A	26	ARG	Sidechain
1	C	104	ARG	Sidechain
1	C	136	ARG	Sidechain

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
1	C	155	GLN	Peptide
1	C	166	ARG	Sidechain
1	C	26	ARG	Sidechain
1	E	153	LEU	Peptide
1	E	171	ARG	Sidechain
1	E	214	ARG	Sidechain
1	G	26	ARG	Sidechain

## 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	1743	0	1702	23	0
1	C	1740	0	1696	24	0
1	E	1748	0	1711	19	0
1	G	1751	0	1710	18	0
2	A	19	0	0	1	0
2	C	19	0	0	1	0
2	E	19	0	0	0	0
2	G	19	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	10	0	0	0	0
4	E	15	0	0	0	0
4	G	15	0	0	0	0
5	A	9	0	0	0	0
5	C	2	0	0	0	0
5	E	2	0	0	0	0
5	G	5	0	0	0	0
All	All	7130	0	6819	87	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 (87) 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:152:LYS:O	1:C:154:ASN:N	2.00	0.94
1:A:154:ASN:H	1:A:154:ASN:HD22	1.22	0.87
1:G:143:VAL:HG11	1:G:153:LEU:HD13	1.60	0.84
1:C:38:ASN:ND2	1:C:40:THR:HG23	1.95	0.82
1:G:158:ASP:OD1	1:G:160:SER:HB2	1.82	0.79
1:C:153:LEU:O	1:C:155:GLN:N	2.20	0.75
1:E:122:LEU:C	1:E:122:LEU:HD23	2.11	0.71
1:C:152:LYS:C	1:C:154:ASN:H	1.96	0.66
1:G:65:GLY:O	1:G:208:VAL:HG13	1.96	0.65
1:E:79:LEU:HD13	1:E:114:HIS:NE2	2.14	0.63
1:C:196:ILE:HD12	1:C:201:ALA:HB2	1.80	0.63
1:C:187:TRP:C	1:C:188:LEU:HD12	2.20	0.62
1:C:63:ASN:O	1:C:148:ALA:HA	1.99	0.61
1:C:38:ASN:HD21	1:C:40:THR:HG23	1.61	0.61
1:E:107:PRO:HB2	1:E:196:ILE:HD11	1.84	0.60
1:A:154:ASN:H	1:A:154:ASN:ND2	1.99	0.58
2:G:301:TE3:O11	2:G:301:TE3:N07	2.32	0.58
1:C:153:LEU:C	1:C:155:GLN:N	2.56	0.57
1:C:151:VAL:HG12	1:C:154:ASN:CB	2.34	0.57
1:C:143:VAL:O	1:C:151:VAL:HG21	2.05	0.56
1:A:70:VAL:HG21	1:A:140:ILE:HD13	1.88	0.56
1:A:132:LYS:O	1:A:195:HIS:HD2	1.89	0.56
1:A:158:ASP:OD1	1:A:160:SER:HB2	2.06	0.55
1:G:158:ASP:OD1	1:G:160:SER:CB	2.54	0.55
2:A:301:TE3:N07	2:A:301:TE3:O11	2.36	0.54
1:G:143:VAL:HG11	1:G:153:LEU:CD1	2.33	0.54
2:C:301:TE3:O11	2:C:301:TE3:N07	2.37	0.53
1:E:35:SER:C	1:E:177:THR:HG22	2.28	0.53
1:C:188:LEU:HD12	1:C:188:LEU:N	2.24	0.53
1:C:152:LYS:C	1:C:154:ASN:N	2.54	0.53
1:E:141:TRP:O	1:E:203:LYS:HE2	2.10	0.52
1:A:7:TRP:HA	1:A:13:ASP:HB3	1.93	0.51
1:C:178:THR:O	1:C:181:CYS:N	2.43	0.51
1:A:126:VAL:HG22	1:A:190:LEU:HD12	1.92	0.51
1:E:107:PRO:HD2	1:E:128:TYR:O	2.11	0.51
1:C:196:ILE:CD1	1:C:201:ALA:HB2	2.42	0.49
1:A:197:ASP:OD1	1:A:200:GLN:HG3	2.12	0.49
1:C:23:GLU:O	1:C:26:ARG:HB2	2.13	0.49
1:G:188:LEU:HD12	1:G:188:LEU:N	2.28	0.49
1:G:7:TRP:O	1:G:66:HIS:HE1	1.96	0.49
1:A:107:PRO:HD2	1:A:128:TYR:O	2.13	0.48
1:A:87:THR:O	1:A:114:HIS:HA	2.13	0.48

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:94:HIS:HA	1:G:212:ASN:HD21	1.79	0.48
1:A:154:ASN:HD22	1:A:154:ASN:N	2.02	0.48
1:E:198:GLN:O	1:E:202:GLU:HG3	2.14	0.48
1:C:38:ASN:HD21	1:C:40:THR:CG2	2.24	0.47
1:C:124:LEU:HD11	1:C:190:LEU:HG	1.97	0.47
1:E:9:TYR:HA	1:E:15:PRO:HD3	1.96	0.47
1:C:19:GLY:HA2	1:C:25:PHE:O	2.15	0.47
1:C:96:PRO:HA	1:C:107:PRO:O	2.14	0.47
1:E:153:LEU:HD22	1:E:155:GLN:CB	2.45	0.46
1:A:136:ARG:HD2	1:A:161:THR:O	2.14	0.46
1:E:134:ASN:HB2	1:E:195:HIS:O	2.15	0.46
1:A:158:ASP:OD1	1:A:160:SER:CB	2.63	0.46
1:A:33:ASN:HA	1:A:222:ARG:CZ	2.45	0.45
1:G:94:HIS:CD2	1:G:111:HIS:CE1	3.06	0.44
1:E:59:VAL:O	1:E:153:LEU:HB2	2.17	0.44
1:G:187:TRP:C	1:G:188:LEU:HD12	2.38	0.44
1:E:165:LYS:HD2	1:E:166:ARG:HD3	2.00	0.44
1:E:91:PHE:HA	1:E:111:HIS:O	2.18	0.44
1:A:93:PHE:HB2	1:A:204:PHE:CE2	2.53	0.43
1:E:122:LEU:C	1:E:122:LEU:CD2	2.83	0.43
1:E:219:LEU:HA	1:E:222:ARG:HH21	1.83	0.43
1:C:91:PHE:HA	1:C:111:HIS:O	2.18	0.43
1:G:176:LEU:HD22	2:G:301:TE3:S19	2.59	0.43
1:G:94:HIS:CD2	1:G:111:HIS:ND1	2.87	0.43
1:E:26:ARG:HH11	1:E:26:ARG:HA	1.84	0.42
1:E:126:VAL:HG22	1:E:190:LEU:HD12	2.02	0.42
1:G:24:GLU:O	1:G:180:PRO:HB3	2.19	0.42
1:A:33:ASN:HD22	1:A:222:ARG:CG	2.33	0.42
1:A:125:ALA:HB3	1:A:189:VAL:HG22	2.02	0.42
1:A:33:ASN:HD22	1:A:222:ARG:HG3	1.85	0.42
1:A:42:THR:OG1	1:A:226:GLU:HG2	2.19	0.42
1:A:126:VAL:CG2	1:A:190:LEU:HD12	2.49	0.42
1:G:134:ASN:HB2	1:G:195:HIS:O	2.20	0.42
1:A:15:PRO:HA	1:A:18:TRP:CD2	2.55	0.42
1:C:146:MET:O	1:C:147:THR:HG23	2.20	0.42
1:G:80:THR:HA	1:G:84:ARG:O	2.19	0.42
1:E:171:ARG:HG3	1:E:171:ARG:HH11	1.85	0.42
1:A:157:PHE:HZ	1:A:162:LEU:HD11	1.85	0.41
1:C:145:PRO:HG3	1:C:151:VAL:HG23	2.03	0.41
1:C:59:VAL:HG23	1:C:155:GLN:CB	2.50	0.41
1:G:35:SER:O	1:G:214:ARG:NH1	2.53	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:95:VAL:HA	1:E:96:PRO:HA	1.88	0.40
1:A:33:ASN:HD22	1:A:222:ARG:CD	2.34	0.40
1:G:110:ALA:O	1:G:125:ALA:HA	2.21	0.40
1:G:42:THR:OG1	1:G:226:GLU:HG2	2.21	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%)	206 (94%)	14 (6%)	0	100	100
1	C	220/243 (90%)	200 (91%)	18 (8%)	2 (1%)	17	44
1	E	220/243 (90%)	209 (95%)	8 (4%)	3 (1%)	11	32
1	G	220/243 (90%)	207 (94%)	11 (5%)	2 (1%)	17	44
All	All	880/972 (90%)	822 (93%)	51 (6%)	7 (1%)	19	47

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	153	LEU
1	C	154	ASN
1	E	153	LEU
1	G	154	ASN
1	G	160	SER
1	E	16	GLU
1	E	220	ASN

### 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	192/213 (90%)	183 (95%)	9 (5%)	26	57
1	C	191/213 (90%)	181 (95%)	10 (5%)	23	53
1	E	193/213 (91%)	183 (95%)	10 (5%)	23	53
1	G	194/213 (91%)	184 (95%)	10 (5%)	23	53
All	All	770/852 (90%)	731 (95%)	39 (5%)	24	54

All (39) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	52	VAL
1	A	81	VAL
1	A	85	THR
1	A	119	LYS
1	A	132	LYS
1	A	141	TRP
1	A	154	ASN
1	A	198	GLN
1	A	226	GLU
1	C	5	THR
1	C	16	GLU
1	C	20	ASN
1	C	104	ARG
1	C	119	LYS
1	C	141	TRP
1	C	147	THR
1	C	153	LEU
1	C	165	LYS
1	C	226	GLU
1	E	5	THR
1	E	16	GLU
1	E	52	VAL
1	E	81	VAL
1	E	119	LYS

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	E	142	ASN
1	E	150	LYS
1	E	152	LYS
1	E	165	LYS
1	E	226	GLU
1	G	6	HIS
1	G	24	GLU
1	G	51	LYS
1	G	55	LYS
1	G	90	GLN
1	G	114	HIS
1	G	141	TRP
1	G	203	LYS
1	G	223	VAL
1	G	226	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	33	ASN
1	A	154	ASN
1	A	195	HIS
1	C	33	ASN
1	C	66	HIS
1	E	66	HIS
1	G	33	ASN
1	G	66	HIS
1	G	69	GLN
1	G	142	ASN
1	G	195	HIS
1	G	198	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 [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 18 ligands modelled in this entry, 4 are monoatomic - leaving 14 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
2	TE3	E	301	3	15,20,20	2.46	3 (20%)	19,28,28	3.17	7 (36%)
4	SO4	G	305	-	4,4,4	0.35	0	6,6,6	0.11	0
2	TE3	A	301	3	15,20,20	2.46	5 (33%)	19,28,28	2.95	8 (42%)
4	SO4	E	303	-	4,4,4	0.28	0	6,6,6	0.15	0
4	SO4	A	303	-	4,4,4	0.35	0	6,6,6	0.16	0
4	SO4	G	304	-	4,4,4	0.29	0	6,6,6	0.12	0
4	SO4	G	303	-	4,4,4	0.37	0	6,6,6	0.16	0
2	TE3	G	301	3	15,20,20	2.37	4 (26%)	19,28,28	5.18	9 (47%)
4	SO4	A	304	-	4,4,4	0.28	0	6,6,6	0.12	0
4	SO4	C	303	-	4,4,4	0.30	0	6,6,6	0.16	0
4	SO4	C	304	-	4,4,4	0.31	0	6,6,6	0.19	0
2	TE3	C	301	3	15,20,20	2.69	5 (33%)	19,28,28	3.32	10 (52%)
4	SO4	E	305	-	4,4,4	0.24	0	6,6,6	0.16	0
4	SO4	E	304	-	4,4,4	0.31	0	6,6,6	0.09	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	TE3	E	301	3	-	1/6/22/22	0/2/2/2
2	TE3	G	301	3	-	1/6/22/22	0/2/2/2
2	TE3	A	301	3	-	2/6/22/22	0/2/2/2
2	TE3	C	301	3	-	2/6/22/22	0/2/2/2

All (17) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	301	TE3	S02-N01	5.71	1.71	1.60
2	G	301	TE3	S02-N01	5.44	1.71	1.60
2	E	301	TE3	S02-N01	5.32	1.70	1.60
2	E	301	TE3	C10-N09	5.22	1.47	1.35
2	A	301	TE3	S02-N01	4.95	1.70	1.60
2	G	301	TE3	C10-N09	4.89	1.46	1.35
2	A	301	TE3	C10-N09	4.82	1.46	1.35
2	E	301	TE3	C08-N09	4.78	1.45	1.36
2	C	301	TE3	C10-N09	4.56	1.45	1.35
2	A	301	TE3	C08-N09	4.50	1.44	1.36
2	C	301	TE3	C08-N09	4.38	1.44	1.36
2	C	301	TE3	O03-S02	3.72	1.50	1.43
2	G	301	TE3	C08-N09	3.70	1.43	1.36
2	C	301	TE3	O04-S02	3.19	1.49	1.43
2	A	301	TE3	O03-S02	3.18	1.49	1.43
2	G	301	TE3	O04-S02	2.84	1.49	1.43
2	A	301	TE3	O04-S02	2.61	1.48	1.43

All (34) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	G	301	TE3	O04-S02-N01	14.96	129.55	107.36
2	G	301	TE3	O04-S02-O03	-9.76	102.72	118.76
2	C	301	TE3	O03-S02-C05	-9.09	97.91	108.59
2	E	301	TE3	O03-S02-N01	7.37	118.29	107.36
2	G	301	TE3	O04-S02-C05	-7.14	100.20	108.59
2	G	301	TE3	O03-S02-C05	-6.61	100.83	108.59
2	E	301	TE3	O04-S02-O03	-6.55	107.98	118.76
2	G	301	TE3	C08-N09-C10	-6.08	113.12	129.54
2	A	301	TE3	C08-N09-C10	-5.84	113.78	129.54
2	C	301	TE3	C12-C10-N09	5.77	122.42	114.50
2	A	301	TE3	O03-S02-C05	-5.54	102.08	108.59
2	A	301	TE3	O04-S02-O03	-5.37	109.94	118.76
2	E	301	TE3	O03-S02-C05	-5.33	102.33	108.59
2	C	301	TE3	C08-N09-C10	-5.18	115.56	129.54
2	A	301	TE3	C12-C10-N09	4.85	121.15	114.50
2	G	301	TE3	C12-C10-N09	4.45	120.61	114.50
2	C	301	TE3	O03-S02-N01	4.34	113.80	107.36
2	E	301	TE3	C08-N09-C10	-4.05	118.61	129.54
2	A	301	TE3	O03-S02-N01	3.95	113.22	107.36
2	E	301	TE3	C12-C10-N09	3.90	119.85	114.50
2	G	301	TE3	O03-S02-N01	3.85	113.07	107.36

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	301	TE3	C05-S02-N01	3.38	114.48	108.25
2	C	301	TE3	O04-S02-O03	-3.09	113.68	118.76
2	C	301	TE3	C18-C13-C12	3.00	115.00	111.52
2	C	301	TE3	C14-C13-C12	-2.78	108.29	111.52
2	C	301	TE3	O11-C10-C12	-2.60	117.69	121.50
2	E	301	TE3	O11-C10-C12	-2.58	117.72	121.50
2	G	301	TE3	C13-C12-C10	-2.54	107.11	112.87
2	A	301	TE3	O11-C10-N09	-2.54	119.00	123.63
2	A	301	TE3	C14-C13-C12	-2.39	108.74	111.52
2	C	301	TE3	O04-S02-C05	2.20	111.18	108.59
2	G	301	TE3	O11-C10-N09	-2.10	119.81	123.63
2	E	301	TE3	O04-S02-N01	2.05	110.41	107.36
2	C	301	TE3	O11-C10-N09	-2.05	119.89	123.63

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	301	TE3	C10-C12-C13-C14
2	A	301	TE3	C10-C12-C13-C18
2	C	301	TE3	C10-C12-C13-C14
2	C	301	TE3	C10-C12-C13-C18
2	E	301	TE3	C10-C12-C13-C18
2	G	301	TE3	C10-C12-C13-C18

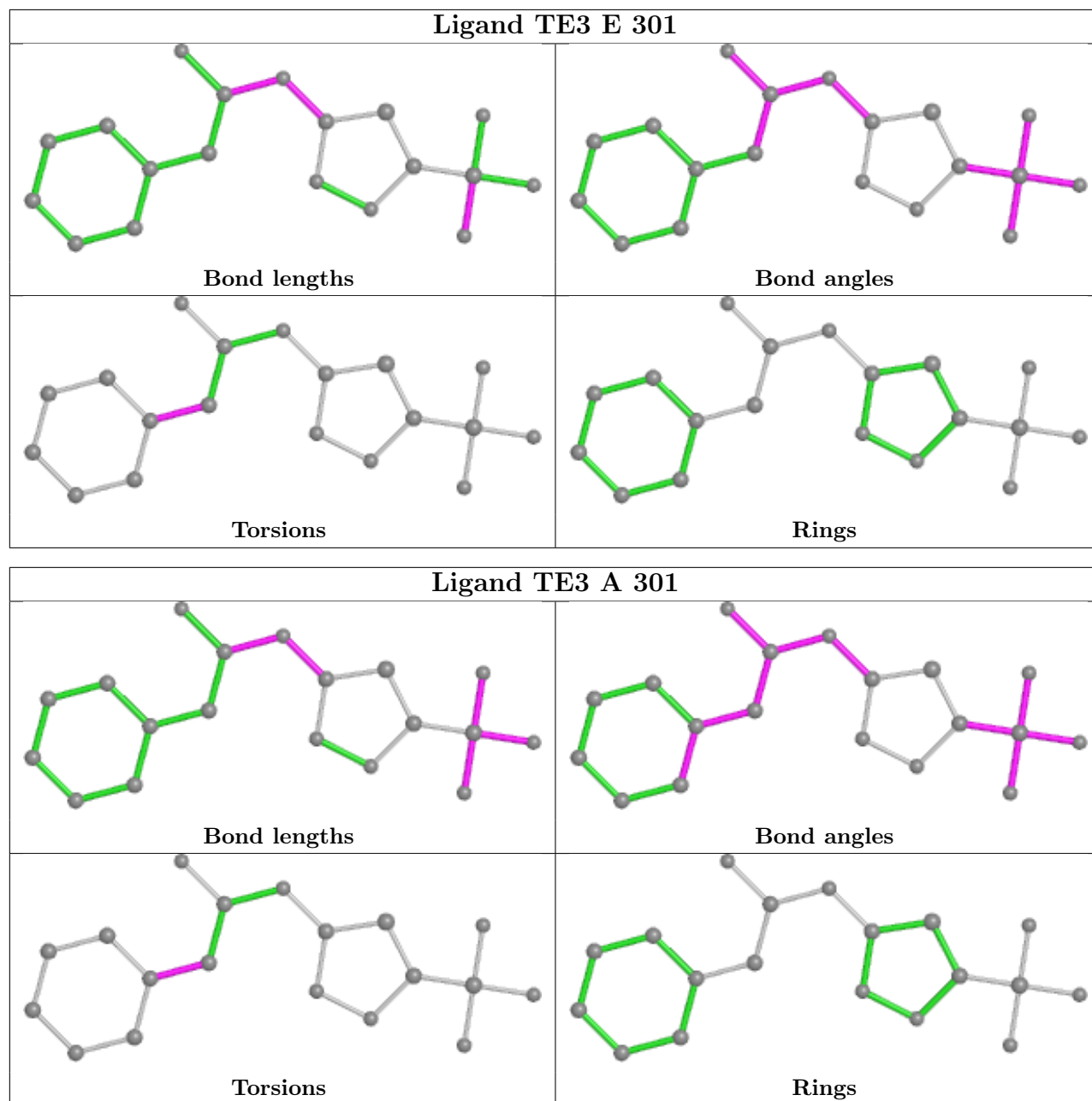
There are no ring outliers.

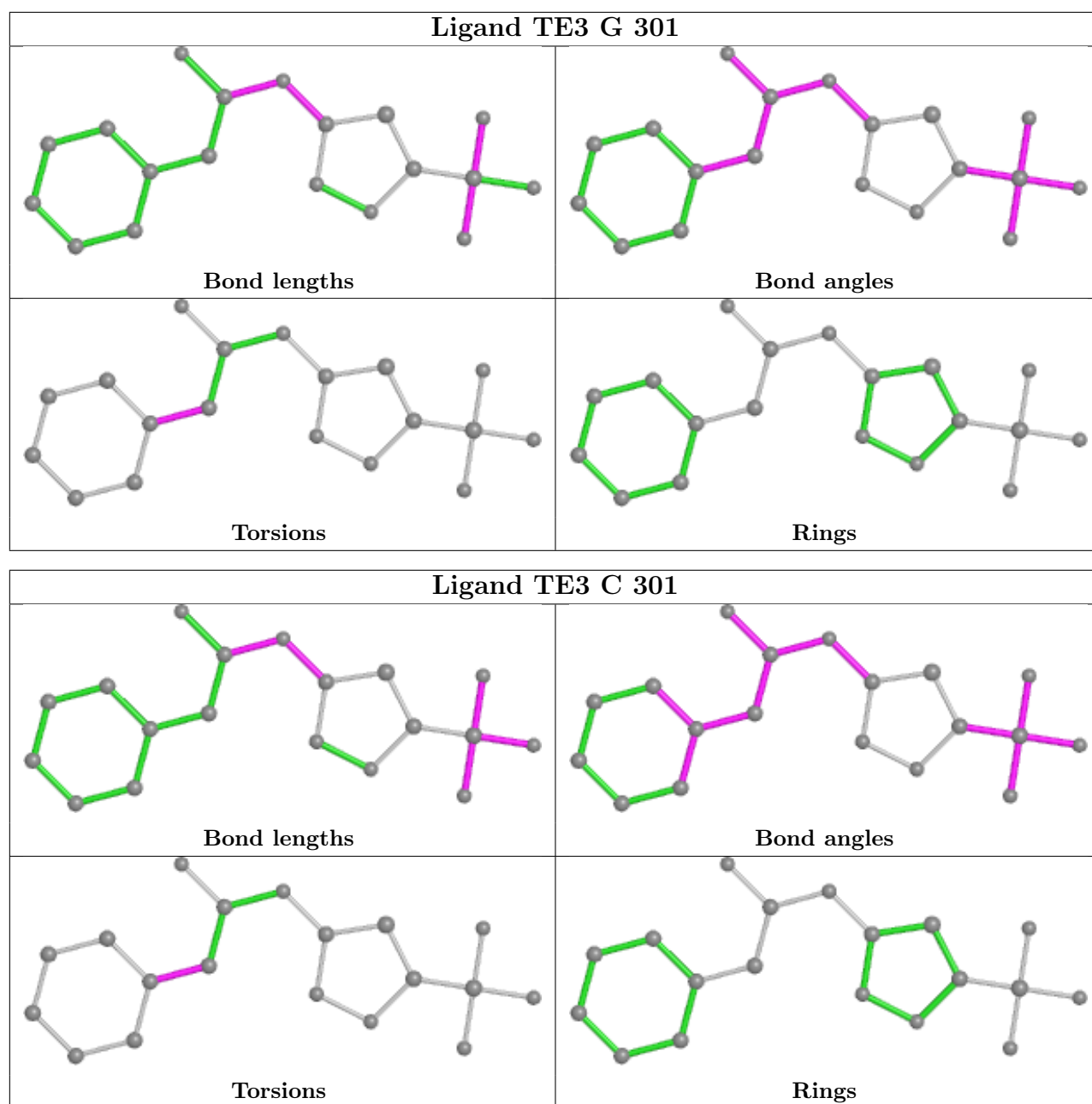
3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	301	TE3	1	0
2	G	301	TE3	2	0
2	C	301	TE3	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	222/243 (91%)	-0.59	0	100 100	24, 37, 55, 77	0
1	C	222/243 (91%)	-0.55	0	100 100	24, 36, 55, 88	0
1	E	222/243 (91%)	-0.54	1 (0%)	91 88	23, 36, 60, 105	0
1	G	222/243 (91%)	-0.51	1 (0%)	91 88	24, 37, 55, 75	0
All	All	888/972 (91%)	-0.55	2 (0%)	95 94	23, 37, 56, 105	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	154	ASN	5.7
1	G	5	THR	4.0

### 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 [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, 95<sup>th</sup> 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.

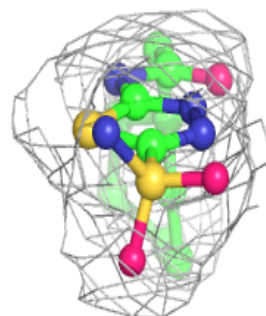
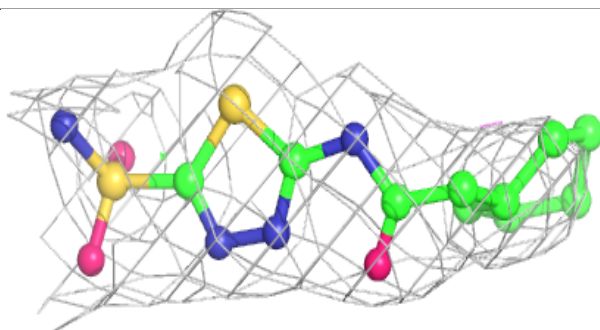
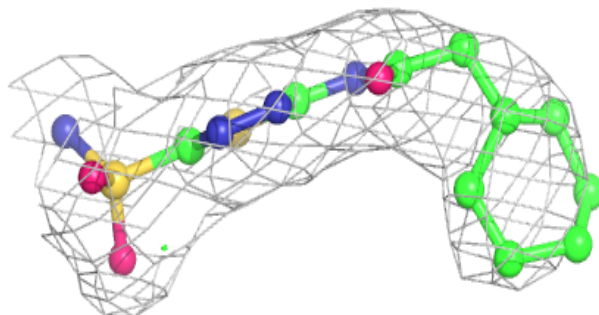
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	SO4	G	305	5/5	0.90	0.78	149,152,159,165	0
4	SO4	C	304	5/5	0.91	0.23	77,80,90,97	0
4	SO4	G	304	5/5	0.92	0.19	70,80,84,85	0
4	SO4	E	305	5/5	0.93	0.26	59,65,73,74	0
4	SO4	E	304	5/5	0.96	0.17	66,67,74,76	0
4	SO4	A	304	5/5	0.96	0.14	64,68,73,79	0
2	TE3	A	301	19/19	0.97	0.15	28,38,50,50	0
2	TE3	G	301	19/19	0.97	0.17	29,34,41,41	0
2	TE3	E	301	19/19	0.98	0.15	34,39,71,73	0
2	TE3	C	301	19/19	0.98	0.13	30,34,45,45	0
3	ZN	E	302	1/1	0.98	0.08	30,30,30,30	0
4	SO4	E	303	5/5	0.99	0.11	31,32,36,36	0
4	SO4	A	303	5/5	0.99	0.10	43,43,44,53	0
3	ZN	A	302	1/1	0.99	0.10	34,34,34,34	0
4	SO4	G	303	5/5	0.99	0.08	34,34,37,38	0
4	SO4	C	303	5/5	0.99	0.09	36,42,44,47	0
3	ZN	G	302	1/1	0.99	0.10	30,30,30,30	0
3	ZN	C	302	1/1	1.00	0.09	29,29,29,29	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.

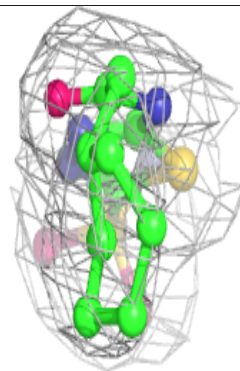
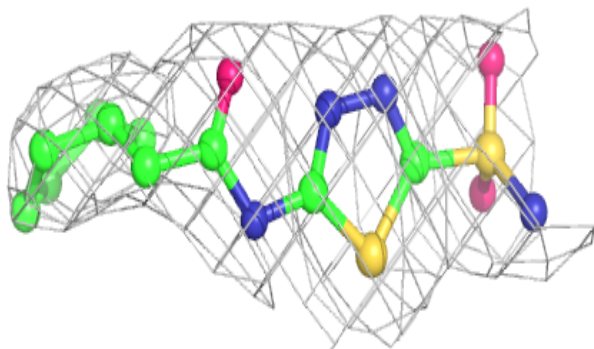
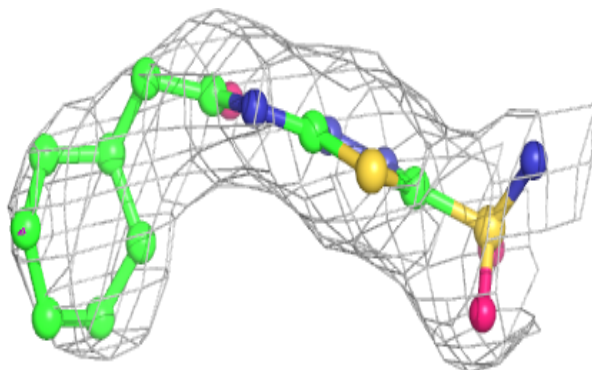


**Electron density around TE3 A 301:**

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

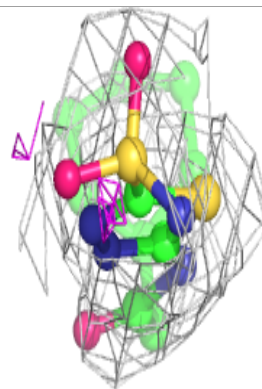
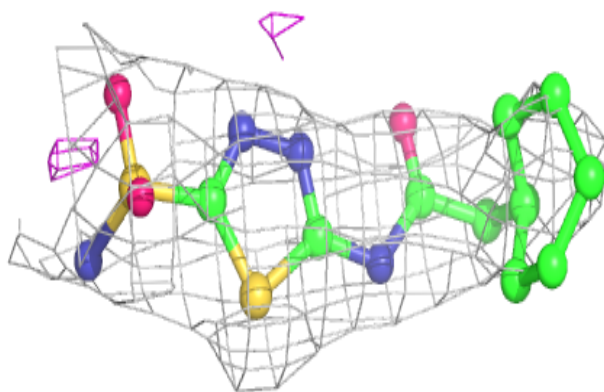
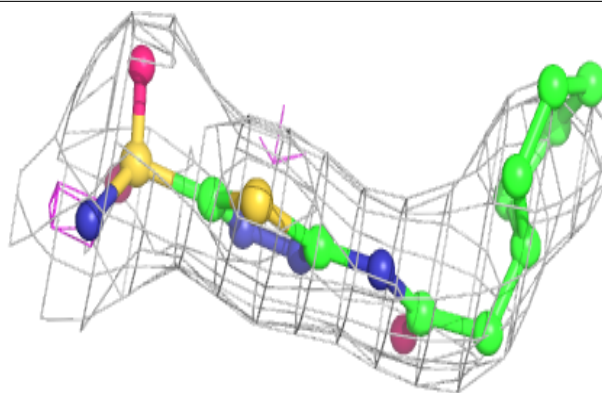
**Electron density around TE3 G 301:**

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

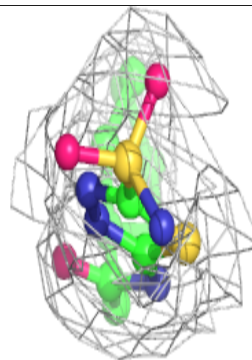
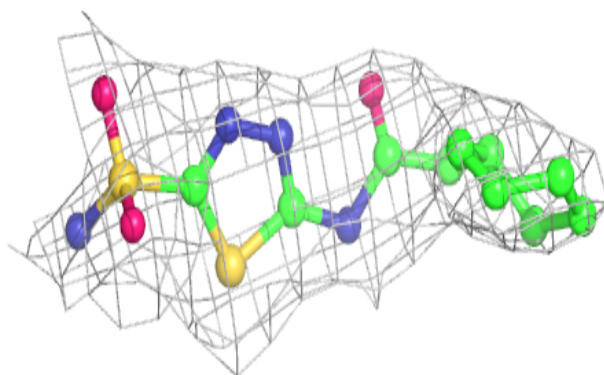
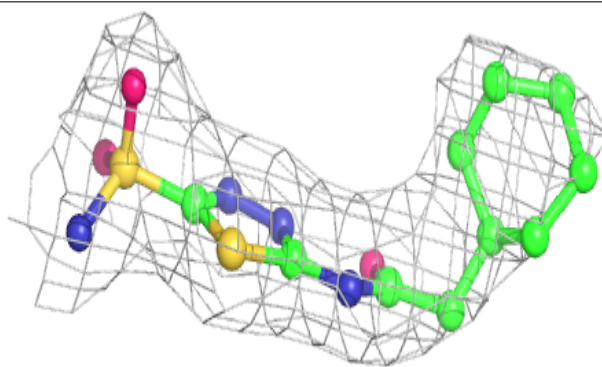


**Electron density around TE3 E 301:**

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

**Electron density around TE3 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)



## 6.5 Other polymers [i](#)

There are no such residues in this entry.