



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 15, 2020 – 07:11 am BST

PDB ID : 4TLE  
Title : Crystal structure of N-terminal C1 domain of KaiC  
Authors : Abe, J.; Hiyama, T.B.; Mukaiyama, A.; Son, S.; Akiyama, S.  
Deposited on : 2014-05-29  
Resolution : 1.94 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 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

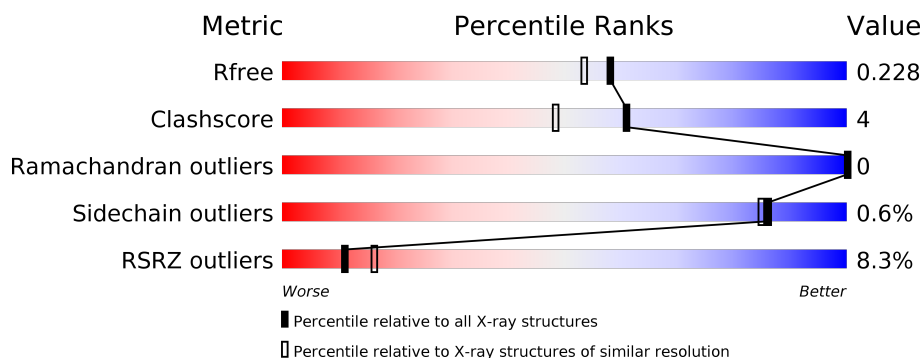
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.94 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	4310 (1.96-1.92)
Clashscore	141614	1023 (1.94-1.94)
Ramachandran outliers	138981	1007 (1.94-1.94)
Sidechain outliers	138945	1007 (1.94-1.94)
RSRZ outliers	127900	4250 (1.96-1.92)

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  $\geq 3$ , 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	253	<div> <div>14%</div> <div> <div></div> <div>75%</div> <div>10%</div> <div>14%</div> </div> </div>
1	B	253	<div> <div>5%</div> <div> <div></div> <div>79%</div> <div>7%</div> <div>14%</div> </div> </div>
1	C	253	<div> <div>5%</div> <div> <div></div> <div>81%</div> <div>7%</div> <div>12%</div> </div> </div>
1	D	253	<div> <div>6%</div> <div> <div></div> <div>79%</div> <div>7%</div> <div>13%</div> </div> </div>
1	E	253	<div> <div>8%</div> <div> <div></div> <div>77%</div> <div>9%</div> <div>14%</div> </div> </div>
1	F	253	<div> <div>5%</div> <div> <div></div> <div>72%</div> <div>13%</div> <div>15%</div> </div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 11457 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 Circadian clock protein kinase KaiC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	218	Total	C	N	O	S	4	6	0
			1766	1122	307	332	5			
1	B	217	Total	C	N	O	S	0	9	0
			1791	1139	313	333	6			
1	C	222	Total	C	N	O	S	0	9	0
			1834	1162	323	343	6			
1	D	219	Total	C	N	O	S	0	6	0
			1770	1124	309	331	6			
1	E	217	Total	C	N	O	S	0	6	0
			1763	1120	309	329	5			
1	F	214	Total	C	N	O	S	0	11	0
			1768	1128	308	326	6			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	157	CYS	SER	engineered mutation	UNP Q79PF4
B	157	CYS	SER	engineered mutation	UNP Q79PF4
C	157	CYS	SER	engineered mutation	UNP Q79PF4
D	157	CYS	SER	engineered mutation	UNP Q79PF4
E	157	CYS	SER	engineered mutation	UNP Q79PF4
F	157	CYS	SER	engineered mutation	UNP Q79PF4

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	D	1	Total	Mg	0	0
			1	1		
2	E	1	Total	Mg	0	0
			1	1		
2	B	1	Total	Mg	0	0
			1	1		

*Continued on next page...*

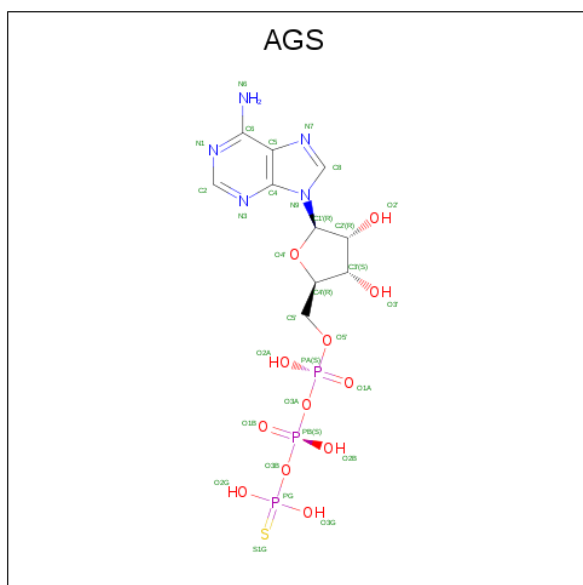
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	C	1	Total	Mg	0	0
			1	1		
2	A	1	Total	Mg	0	0
			1	1		
2	F	1	Total	Mg	0	0
			1	1		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	D	1	Total	Cl	0	0
			1	1		
3	E	1	Total	Cl	0	0
			1	1		
3	B	1	Total	Cl	0	0
			1	1		
3	C	1	Total	Cl	0	0
			1	1		
3	A	1	Total	Cl	0	0
			1	1		
3	F	1	Total	Cl	0	0
			1	1		

- Molecule 4 is PHOSPHOTHIOPHOSPHORIC ACID-ADENYLATE ESTER (three-letter code: AGS) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>5</sub>O<sub>12</sub>P<sub>3</sub>S).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	A	1	Total 31	C 10	N 5	O 12	P 3	S 1	0	0
4	B	1	Total 31	C 10	N 5	O 12	P 3	S 1	0	0
4	C	1	Total 31	C 10	N 5	O 12	P 3	S 1	0	0
4	D	1	Total 31	C 10	N 5	O 12	P 3	S 1	0	0
4	E	1	Total 31	C 10	N 5	O 12	P 3	S 1	0	0
4	F	1	Total 31	C 10	N 5	O 12	P 3	S 1	0	0

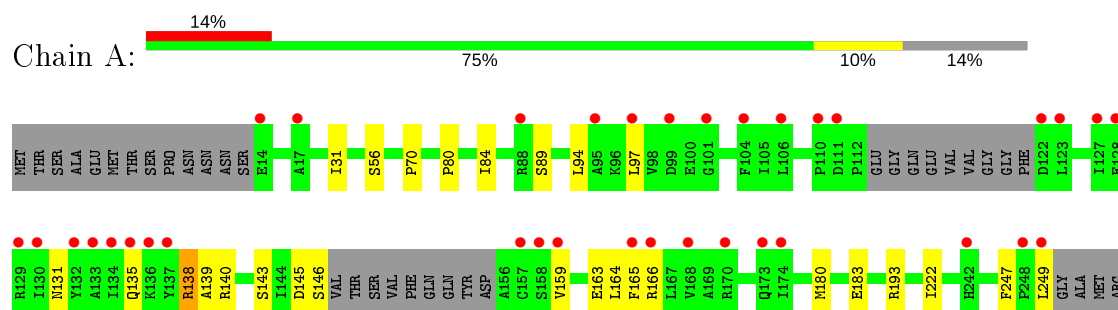
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	62	Total 62	O 62	0	0
5	B	110	Total 110	O 110	0	0
5	C	105	Total 105	O 105	0	0
5	D	112	Total 112	O 112	0	0
5	E	89	Total 89	O 89	0	0
5	F	89	Total 89	O 89	0	0

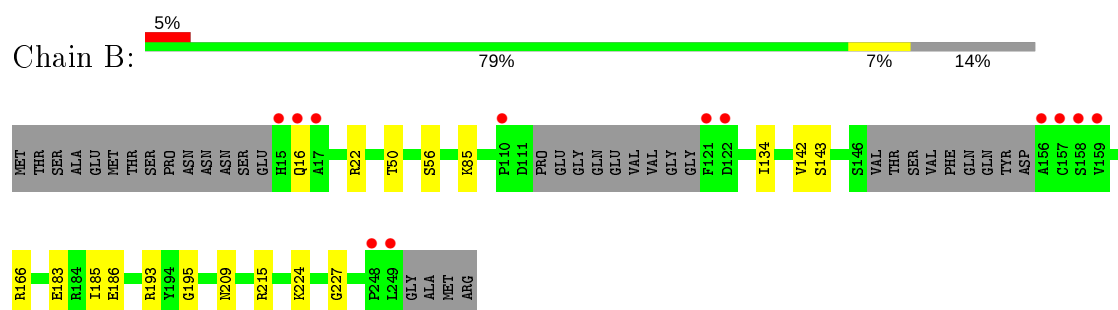
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

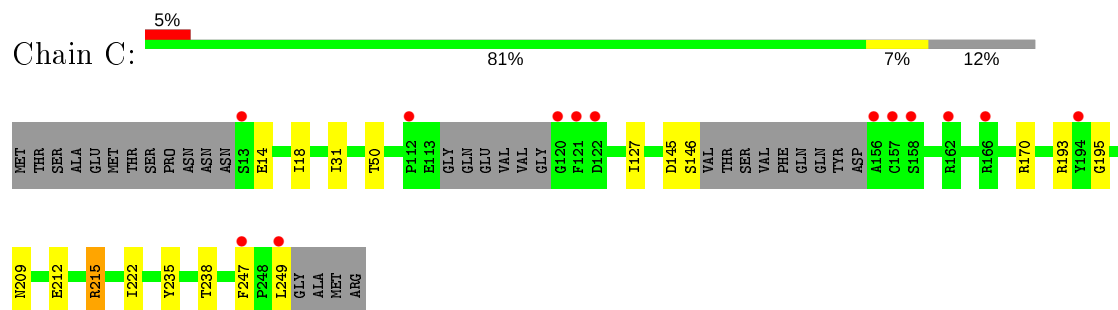
- Molecule 1: Circadian clock protein kinase KaiC



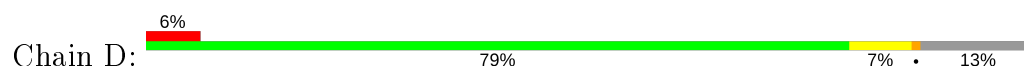
- Molecule 1: Circadian clock protein kinase KaiC

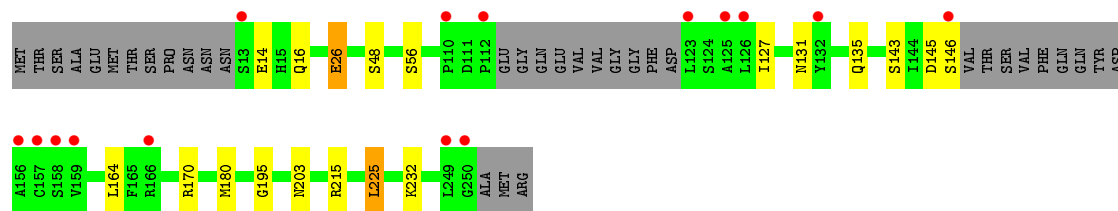


- Molecule 1: Circadian clock protein kinase KaiC

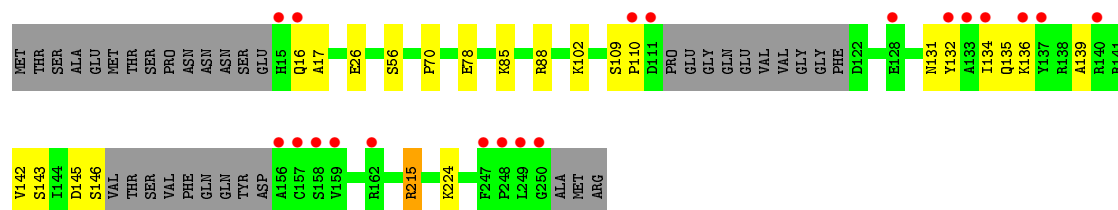
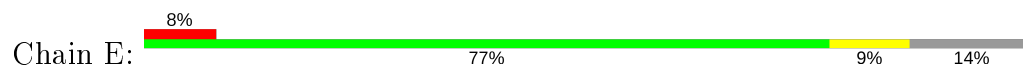


- Molecule 1: Circadian clock protein kinase KaiC

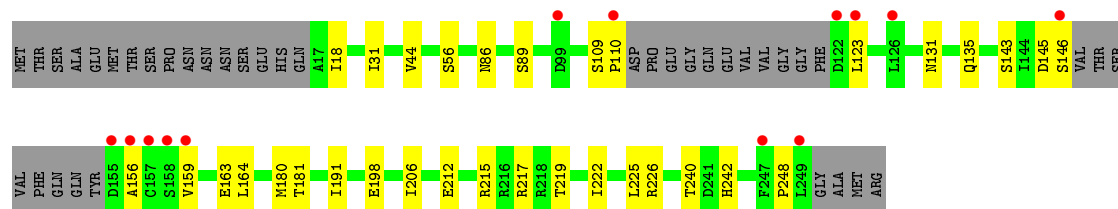
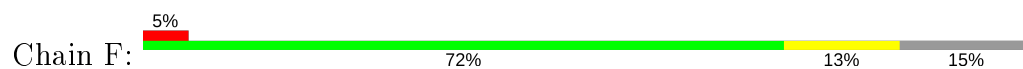




- Molecule 1: Circadian clock protein kinase KaiC



- Molecule 1: Circadian clock protein kinase KaiC



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	79.93Å 133.32Å 151.06Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.60 – 1.94 40.58 – 1.94	Depositor EDS
% Data completeness (in resolution range)	98.9 (42.60-1.94) 94.6 (40.58-1.94)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.78 (at 1.94Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, $R_{free}$	0.189 , 0.229 0.190 , 0.228	Depositor DCC
$R_{free}$ test set	6003 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	27.5	Xtriage
Anisotropy	0.289	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 55.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	11457	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.88% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, AGS, CL

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	0/1803	0.56	0/2429
1	B	0.48	0/1826	0.61	0/2457
1	C	0.48	0/1866	0.60	0/2509
1	D	0.52	0/1807	0.62	1/2431 (0.0%)
1	E	0.50	0/1793	0.61	0/2413
1	F	0.54	0/1812	0.65	1/2438 (0.0%)
All	All	0.49	0/10907	0.61	2/14677 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	225	LEU	CA-CB-CG	5.85	128.75	115.30
1	D	225	LEU	CA-CB-CG	5.04	126.88	115.30

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	1766	0	1786	18	0
1	B	1791	0	1808	18	0
1	C	1834	0	1845	15	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	1770	0	1795	16	0
1	E	1763	0	1784	16	1
1	F	1768	0	1815	25	1
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	1	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
4	A	31	0	12	1	0
4	B	31	0	12	0	0
4	C	31	0	12	0	0
4	D	31	0	12	3	0
4	E	31	0	12	0	0
4	F	31	0	12	0	0
5	A	62	0	0	0	0
5	B	110	0	0	2	0
5	C	105	0	0	3	0
5	D	112	0	0	1	0
5	E	89	0	0	2	0
5	F	89	0	0	1	0
All	All	11457	0	10905	98	1

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 4.

All (98) 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:B:183:GLU:OE1	1:B:193:ARG:NH1	2.12	0.82
1:A:247:PHE:HB3	1:A:249:LEU:HB3	1.62	0.81
1:E:215:ARG:NH2	5:E:401:HOH:O	2.14	0.79
1:C:215:ARG:NH2	5:C:401:HOH:O	2.18	0.76
1:B:215:ARG:NH2	5:B:508:HOH:O	2.20	0.72
4:A:303:AGS:O3G	1:B:224:LYS:NZ	2.27	0.67

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:193:ARG:NH1	1:B:195:GLY:O	2.26	0.67
1:F:31[B]:ILE:HD11	1:F:248:PRO:HB3	1.76	0.67
1:C:193:ARG:NH1	1:D:195:GLY:O	2.27	0.66
1:C:14:GLU:OE2	5:C:477:HOH:O	2.13	0.66
1:E:131:ASN:OD1	1:E:135:GLN:NE2	2.27	0.66
1:B:193:ARG:NH2	1:C:195:GLY:O	2.29	0.65
1:F:198:GLU:OE2	5:F:482:HOH:O	2.14	0.65
1:F:123:LEU:HD22	1:F:163:GLU:HB3	1.83	0.60
1:A:164:LEU:HD11	1:A:180:MET:HE1	1.84	0.59
1:D:215:ARG:NH1	5:D:469:HOH:O	2.35	0.59
1:A:183:GLU:OE1	1:A:193:ARG:NH2	2.32	0.57
1:B:56[A]:SER:HB2	1:B:143:SER:HB3	1.88	0.56
1:A:138:ARG:O	1:A:138:ARG:HG3	2.05	0.56
1:E:16:GLN:CD	1:E:17:ALA:H	2.09	0.55
3:C:302:CL:CL	1:D:232[B]:LYS:HE2	2.43	0.55
1:D:56[A]:SER:HB2	1:D:143:SER:HB3	1.87	0.55
1:F:131:ASN:O	1:F:135:GLN:HG2	2.06	0.54
1:F:191:ILE:HG21	1:F:198:GLU:HB3	1.88	0.54
1:B:22:ARG:NH1	5:B:506:HOH:O	2.40	0.54
1:F:156:ALA:HB3	1:F:159:VAL:HG12	1.89	0.54
1:C:238:THR:HG23	1:C:247:PHE:HZ	1.72	0.53
1:A:70:PRO:HG2	1:A:139:ALA:HA	1.89	0.53
1:B:166:ARG:HG3	1:B:166:ARG:HH11	1.74	0.52
1:F:56[B]:SER:HB2	1:F:143:SER:HB3	1.91	0.52
1:C:31:ILE:HG22	1:C:222:ILE:HD12	1.91	0.52
1:A:165:PHE:CE1	1:F:110:PRO:HG2	2.44	0.52
1:D:56[B]:SER:HB3	1:D:143:SER:HB3	1.93	0.51
1:B:185:ILE:HG13	1:B:186:GLU:HG3	1.94	0.49
1:E:134:ILE:HD11	1:E:142:VAL:HG21	1.93	0.49
1:F:240:THR:OG1	1:F:242:HIS:CD2	2.66	0.49
1:E:56[B]:SER:HB3	1:E:143:SER:HB3	1.94	0.48
1:B:56[B]:SER:HB3	1:B:143:SER:HB3	1.95	0.48
1:C:238:THR:HG23	1:C:247:PHE:CZ	2.47	0.48
1:D:164:LEU:HD11	1:D:180:MET:HE1	1.95	0.48
1:A:80:PRO:O	1:A:84:ILE:HG13	2.14	0.48
1:D:14:GLU:HG2	1:D:16:GLN:HE21	1.78	0.48
1:A:138:ARG:O	1:A:138:ARG:CG	2.62	0.48
1:A:145[A]:ASP:HA	1:A:146[A]:SER:HA	1.64	0.47
1:B:85:LYS:HE3	1:C:18:ILE:HA	1.95	0.47
1:B:134:ILE:HD11	1:B:142:VAL:HG21	1.97	0.47
1:E:70:PRO:HB2	1:E:139:ALA:HA	1.96	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:109:SER:HB2	1:F:110:PRO:HD2	1.96	0.47
1:E:145[A]:ASP:HA	1:E:146[A]:SER:HA	1.64	0.47
1:B:166:ARG:HG3	1:B:166:ARG:NH1	2.30	0.47
1:F:123:LEU:CD2	1:F:163:GLU:HB3	2.45	0.47
1:C:145[A]:ASP:HA	1:C:146[A]:SER:HA	1.56	0.47
1:E:132:TYR:CE1	1:E:136:LYS:HE3	2.49	0.47
1:D:26:GLU:H	1:D:26:GLU:CD	2.16	0.47
1:D:127:ILE:HG21	1:D:170:ARG:HG2	1.97	0.46
1:A:89:SER:HB2	1:B:227:GLY:O	2.16	0.46
1:D:203:ASN:HB3	1:D:225:LEU:HD23	1.97	0.46
1:C:127:ILE:CG2	1:C:170:ARG:HG2	2.46	0.46
1:B:85:LYS:NZ	1:C:14:GLU:OE1	2.43	0.46
1:A:165:PHE:CZ	1:F:110:PRO:HG2	2.50	0.45
1:C:212:GLU:HB3	5:C:490:HOH:O	2.17	0.45
4:D:303:AGS:O3G	1:E:224:LYS:NZ	2.48	0.45
1:E:88:ARG:HA	1:E:88:ARG:HD2	1.68	0.45
1:F:212:GLU:HG2	1:F:215[A]:ARG:NH2	2.32	0.45
1:F:164[A]:LEU:HD11	1:F:180[A]:MET:SD	2.57	0.44
1:E:78:GLU:OE1	1:F:226:ARG:HD3	2.17	0.44
1:E:56[A]:SER:HB2	1:E:143:SER:HB3	1.99	0.44
1:D:131:ASN:O	1:D:135:GLN:HG2	2.17	0.43
1:A:31:ILE:HG22	1:A:222:ILE:HD12	2.01	0.43
1:B:50:THR:HG22	1:B:209:ASN:HB2	2.00	0.43
1:E:109:SER:HA	1:E:110:PRO:HD3	1.83	0.43
1:A:159:VAL:O	1:A:163:GLU:HG2	2.18	0.42
1:B:215:ARG:HA	1:B:215:ARG:HD3	1.80	0.42
1:E:215:ARG:HA	1:E:215:ARG:HD3	1.75	0.42
1:F:217[B]:ARG:HE	1:F:219:THR:HG23	1.83	0.42
1:F:240:THR:HG1	1:F:242:HIS:CD2	2.38	0.42
1:F:86:ASN:O	1:F:89:SER:HB3	2.20	0.42
1:C:235:TYR:HE2	1:C:249:LEU:HB2	1.85	0.42
1:D:145[A]:ASP:HA	1:D:146[A]:SER:HA	1.63	0.42
1:A:97[B]:LEU:HD12	1:A:97[B]:LEU:HA	1.75	0.42
1:E:102[A]:LYS:NZ	5:E:402:HOH:O	2.25	0.42
1:D:48[B]:SER:HA	4:D:303:AGS:S1G	2.60	0.42
1:D:26:GLU:N	1:D:26:GLU:CD	2.73	0.41
1:D:48[A]:SER:HA	4:D:303:AGS:S1G	2.61	0.41
1:C:50:THR:HG22	1:C:209:ASN:HB2	2.01	0.41
1:F:191:ILE:HG23	1:F:206[B]:ILE:HD11	2.03	0.41
1:F:217[B]:ARG:HE	1:F:219:THR:CG2	2.34	0.41
1:F:31[B]:ILE:CD1	1:F:248:PRO:HB3	2.49	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:85:LYS:HE3	1:F:18:ILE:HG12	2.03	0.41
1:A:131:ASN:O	1:A:135:GLN:HG3	2.21	0.41
1:C:215:ARG:HD3	1:C:215:ARG:HA	1.67	0.41
1:A:84:ILE:HG23	1:A:94:LEU:HB2	2.02	0.41
1:F:145[A]:ASP:HA	1:F:146[A]:SER:HA	1.59	0.41
1:B:16:GLN:OE1	1:B:16:GLN:N	2.54	0.40
1:F:44:VAL:O	1:F:181:THR:HA	2.21	0.40
1:A:56[B]:SER:HB3	1:A:143:SER:HB3	2.03	0.40
1:D:14:GLU:HG3	1:D:16:GLN:HG2	2.02	0.40
1:F:31[A]:ILE:HG22	1:F:222:ILE:HD12	2.02	0.40

All (1) 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)
1:E:102[A]:LYS:NZ	1:F:212:GLU:OE2[4_555]	2.13	0.07

## 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	217/253 (86%)	212 (98%)	5 (2%)	0	100	100
1	B	219/253 (87%)	212 (97%)	7 (3%)	0	100	100
1	C	224/253 (88%)	217 (97%)	7 (3%)	0	100	100
1	D	218/253 (86%)	210 (96%)	8 (4%)	0	100	100
1	E	216/253 (85%)	212 (98%)	4 (2%)	0	100	100
1	F	218/253 (86%)	211 (97%)	7 (3%)	0	100	100
All	All	1312/1518 (86%)	1274 (97%)	38 (3%)	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	194/217 (89%)	191 (98%)	3 (2%)	65	56
1	B	196/217 (90%)	196 (100%)	0	100	100
1	C	200/217 (92%)	199 (100%)	1 (0%)	88	88
1	D	194/217 (89%)	193 (100%)	1 (0%)	88	88
1	E	192/217 (88%)	190 (99%)	2 (1%)	76	71
1	F	195/217 (90%)	195 (100%)	0	100	100
All	All	1171/1302 (90%)	1164 (99%)	7 (1%)	86	85

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	138	ARG
1	A	140	ARG
1	A	166	ARG
1	C	215	ARG
1	D	26	GLU
1	E	26	GLU
1	E	215	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 18 ligands modelled in this entry, 12 are 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
4	AGS	E	303	2	26,33,33	0.73	0	26,52,52	1.34	2 (7%)
4	AGS	A	303	2	26,33,33	0.81	1 (3%)	26,52,52	1.07	2 (7%)
4	AGS	F	303	2	26,33,33	0.79	0	26,52,52	1.00	2 (7%)
4	AGS	D	303	2	26,33,33	0.79	1 (3%)	26,52,52	1.08	2 (7%)
4	AGS	B	303	2	26,33,33	0.78	1 (3%)	26,52,52	1.10	2 (7%)
4	AGS	C	303	2	26,33,33	0.82	1 (3%)	26,52,52	1.12	2 (7%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	AGS	E	303	2	-	8/17/38/38	0/3/3/3
4	AGS	A	303	2	-	7/17/38/38	0/3/3/3
4	AGS	F	303	2	-	6/17/38/38	0/3/3/3
4	AGS	D	303	2	-	9/17/38/38	0/3/3/3
4	AGS	B	303	2	-	7/17/38/38	0/3/3/3
4	AGS	C	303	2	-	6/17/38/38	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	303	AGS	PG-S1G	2.37	1.95	1.90

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	303	AGS	PG-S1G	2.25	1.95	1.90
4	D	303	AGS	PG-S1G	2.11	1.95	1.90
4	A	303	AGS	PG-S1G	2.11	1.95	1.90

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	E	303	AGS	O2G-PG-O3B	4.27	118.89	104.64
4	C	303	AGS	O2G-PG-O3B	3.82	117.38	104.64
4	D	303	AGS	O2G-PG-O3B	3.19	115.30	104.64
4	B	303	AGS	O2G-PG-O3B	3.05	114.82	104.64
4	A	303	AGS	O2G-PG-O3B	3.00	114.64	104.64
4	F	303	AGS	O2G-PG-O3B	2.79	113.96	104.64
4	B	303	AGS	C5-C6-N6	2.58	124.27	120.35
4	A	303	AGS	C5-C6-N6	2.41	124.02	120.35
4	D	303	AGS	C5-C6-N6	2.28	123.82	120.35
4	F	303	AGS	C5-C6-N6	2.27	123.80	120.35
4	C	303	AGS	C5-C6-N6	2.24	123.75	120.35
4	E	303	AGS	C5-C6-N6	2.21	123.71	120.35

There are no chirality outliers.

All (43) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	303	AGS	PB-O3B-PG-O2G
4	A	303	AGS	PB-O3B-PG-O3G
4	A	303	AGS	C5'-O5'-PA-O2A
4	F	303	AGS	PB-O3B-PG-O2G
4	F	303	AGS	PB-O3B-PG-O3G
4	E	303	AGS	PB-O3B-PG-O2G
4	E	303	AGS	PB-O3B-PG-O3G
4	E	303	AGS	C5'-O5'-PA-O1A
4	D	303	AGS	PB-O3B-PG-O2G
4	D	303	AGS	PB-O3B-PG-O3G
4	D	303	AGS	C5'-O5'-PA-O1A
4	B	303	AGS	PB-O3B-PG-O2G
4	B	303	AGS	PB-O3B-PG-O3G
4	C	303	AGS	PB-O3B-PG-O2G
4	C	303	AGS	PB-O3B-PG-O3G
4	E	303	AGS	C5'-O5'-PA-O3A
4	D	303	AGS	C5'-O5'-PA-O3A
4	C	303	AGS	PA-O3A-PB-O1B

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
4	E	303	AGS	C5'-O5'-PA-O2A
4	D	303	AGS	C5'-O5'-PA-O2A
4	B	303	AGS	C5'-O5'-PA-O2A
4	A	303	AGS	PA-O3A-PB-O2B
4	E	303	AGS	PA-O3A-PB-O1B
4	E	303	AGS	PA-O3A-PB-O2B
4	D	303	AGS	PA-O3A-PB-O2B
4	B	303	AGS	PA-O3A-PB-O1B
4	E	303	AGS	PG-O3B-PB-O2B
4	A	303	AGS	PG-O3B-PB-O1B
4	A	303	AGS	PG-O3B-PB-O2B
4	D	303	AGS	PG-O3B-PB-O2B
4	A	303	AGS	C5'-O5'-PA-O3A
4	F	303	AGS	C5'-O5'-PA-O3A
4	B	303	AGS	C5'-O5'-PA-O3A
4	C	303	AGS	C5'-O5'-PA-O3A
4	D	303	AGS	C3'-C4'-C5'-O5'
4	F	303	AGS	PA-O3A-PB-O2B
4	D	303	AGS	PA-O3A-PB-O1B
4	B	303	AGS	PA-O3A-PB-O2B
4	C	303	AGS	PA-O3A-PB-O2B
4	F	303	AGS	C5'-O5'-PA-O1A
4	F	303	AGS	C5'-O5'-PA-O2A
4	B	303	AGS	C5'-O5'-PA-O1A
4	C	303	AGS	C5'-O5'-PA-O1A

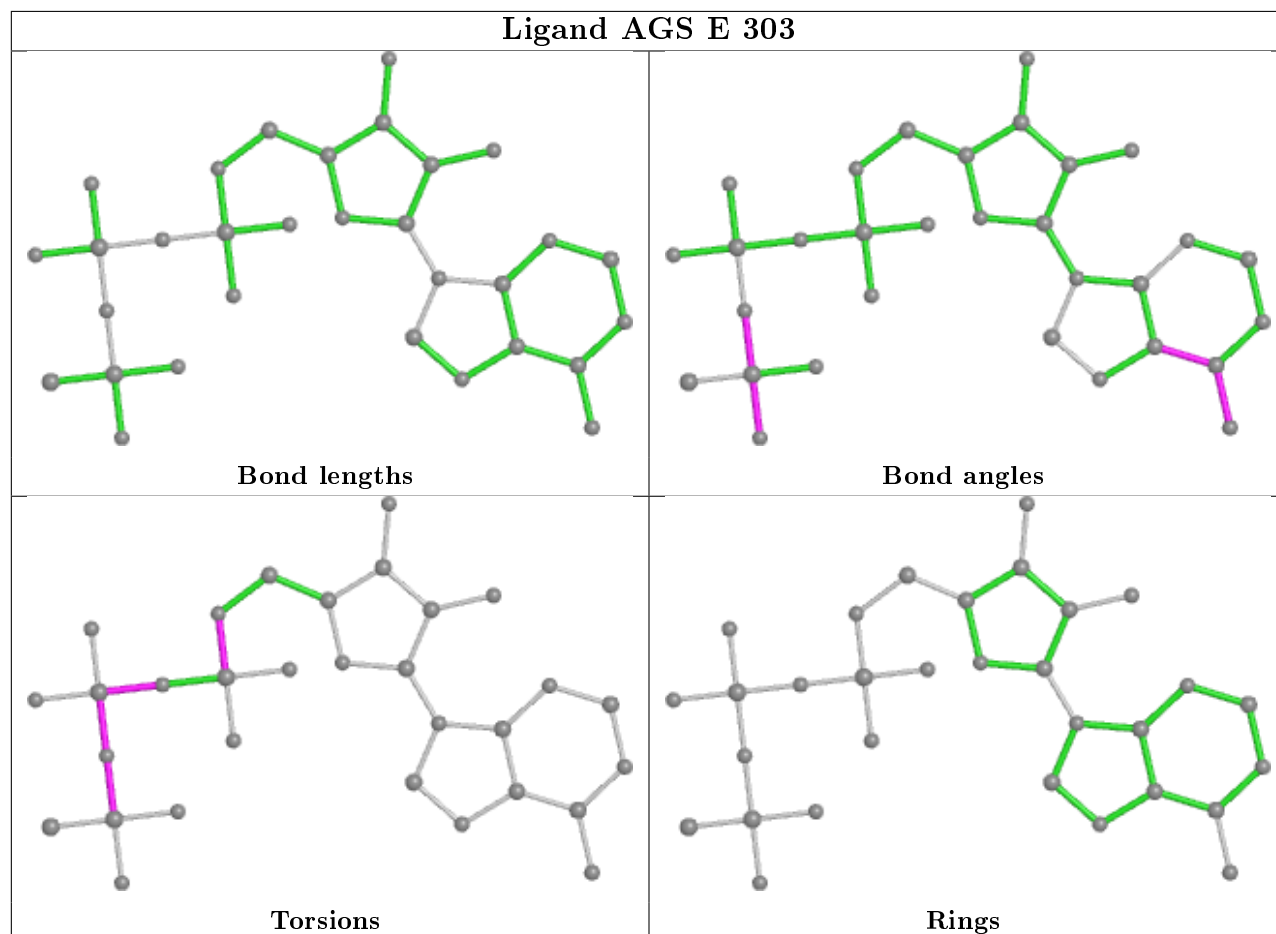
There are no ring outliers.

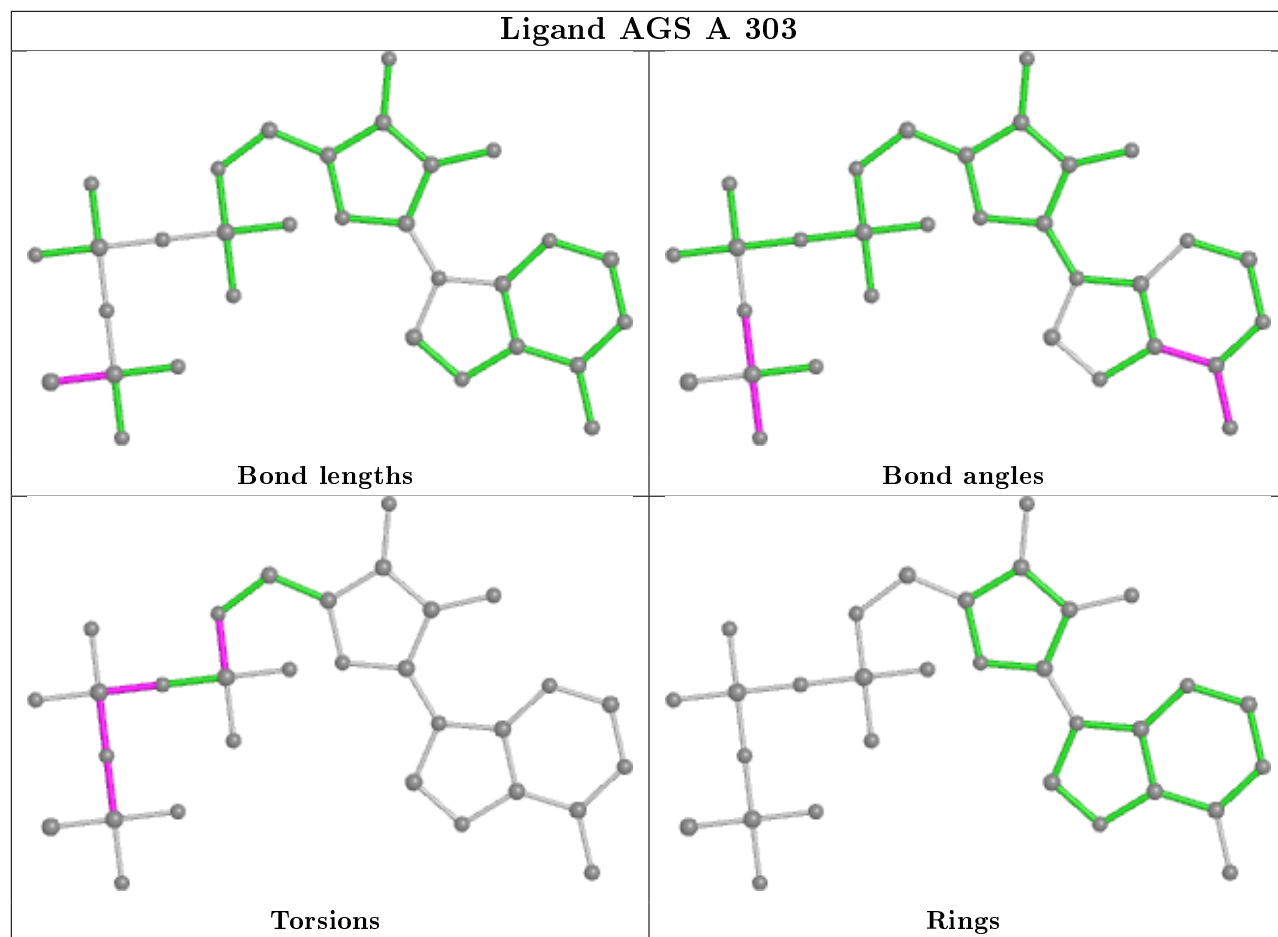
2 monomers are involved in 4 short contacts:

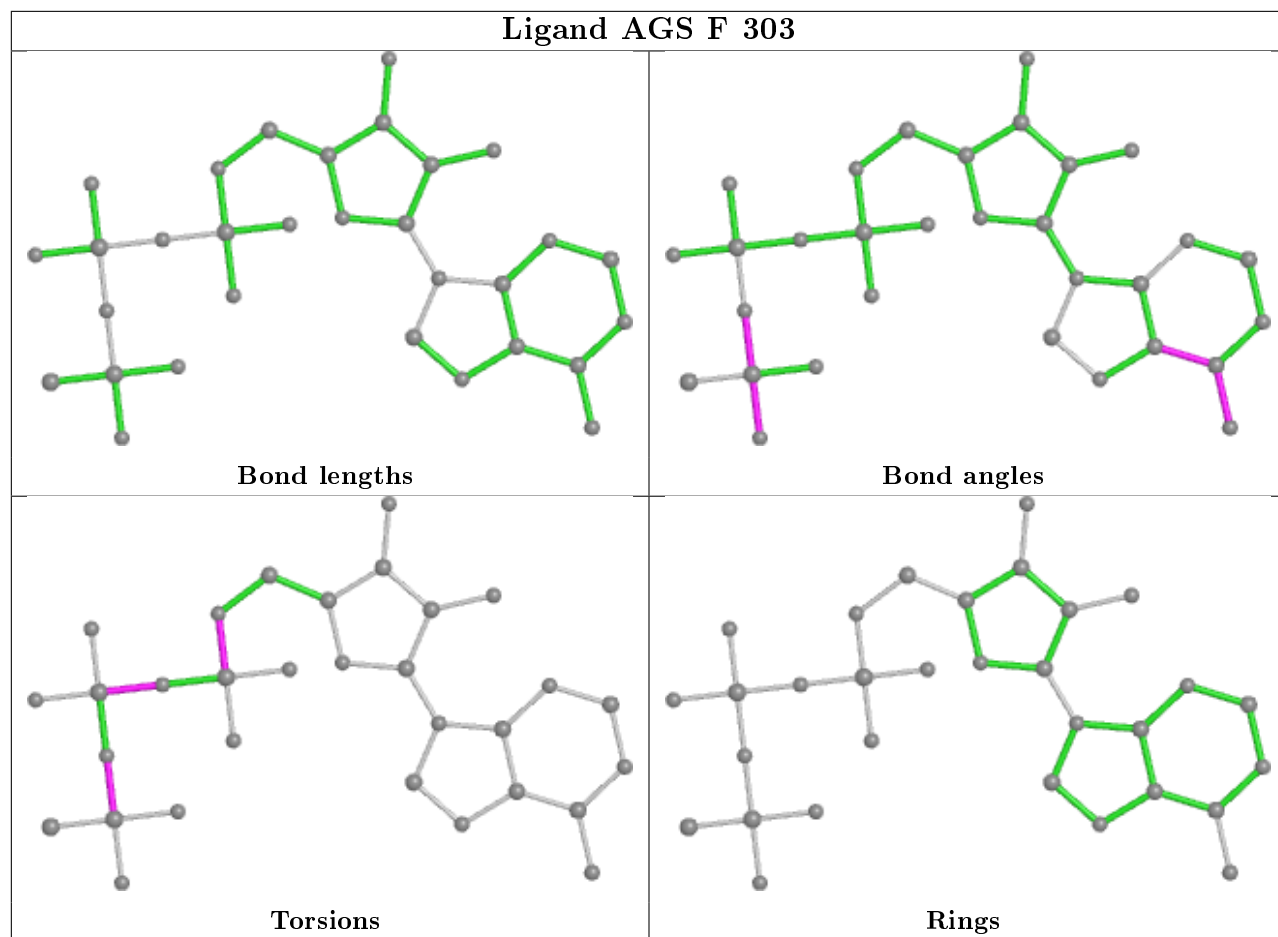
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	303	AGS	1	0
4	D	303	AGS	3	0

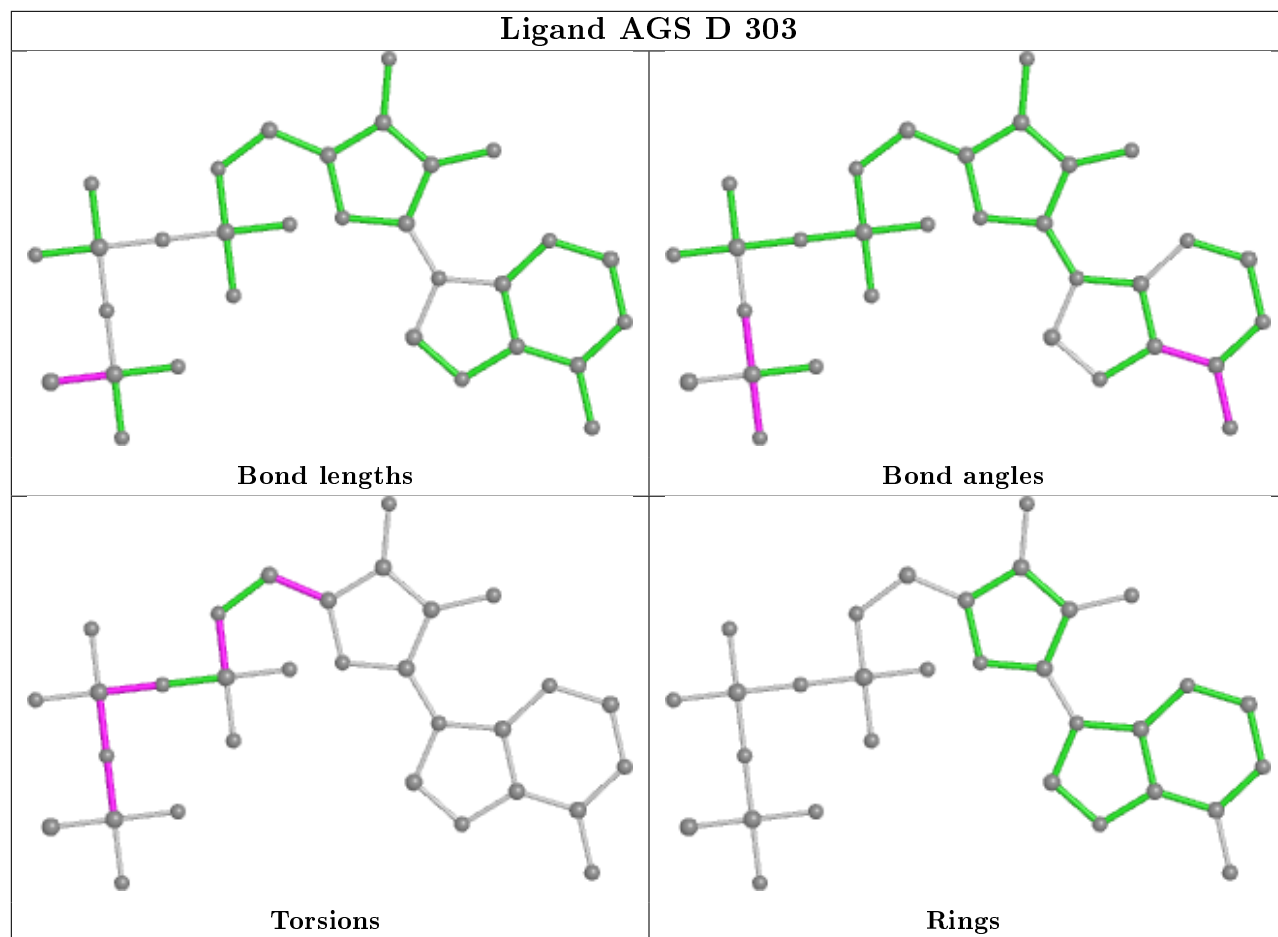
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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

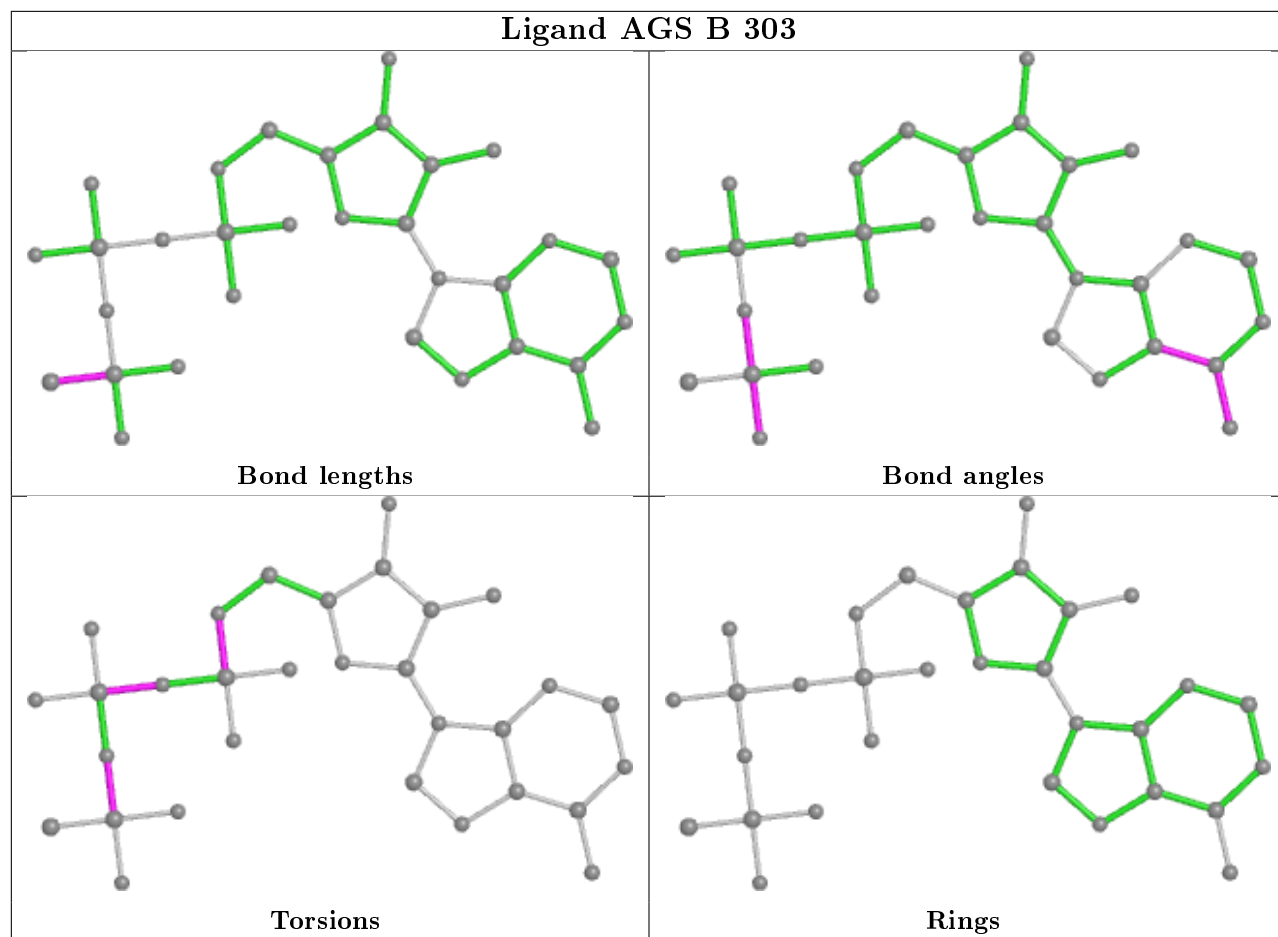
any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

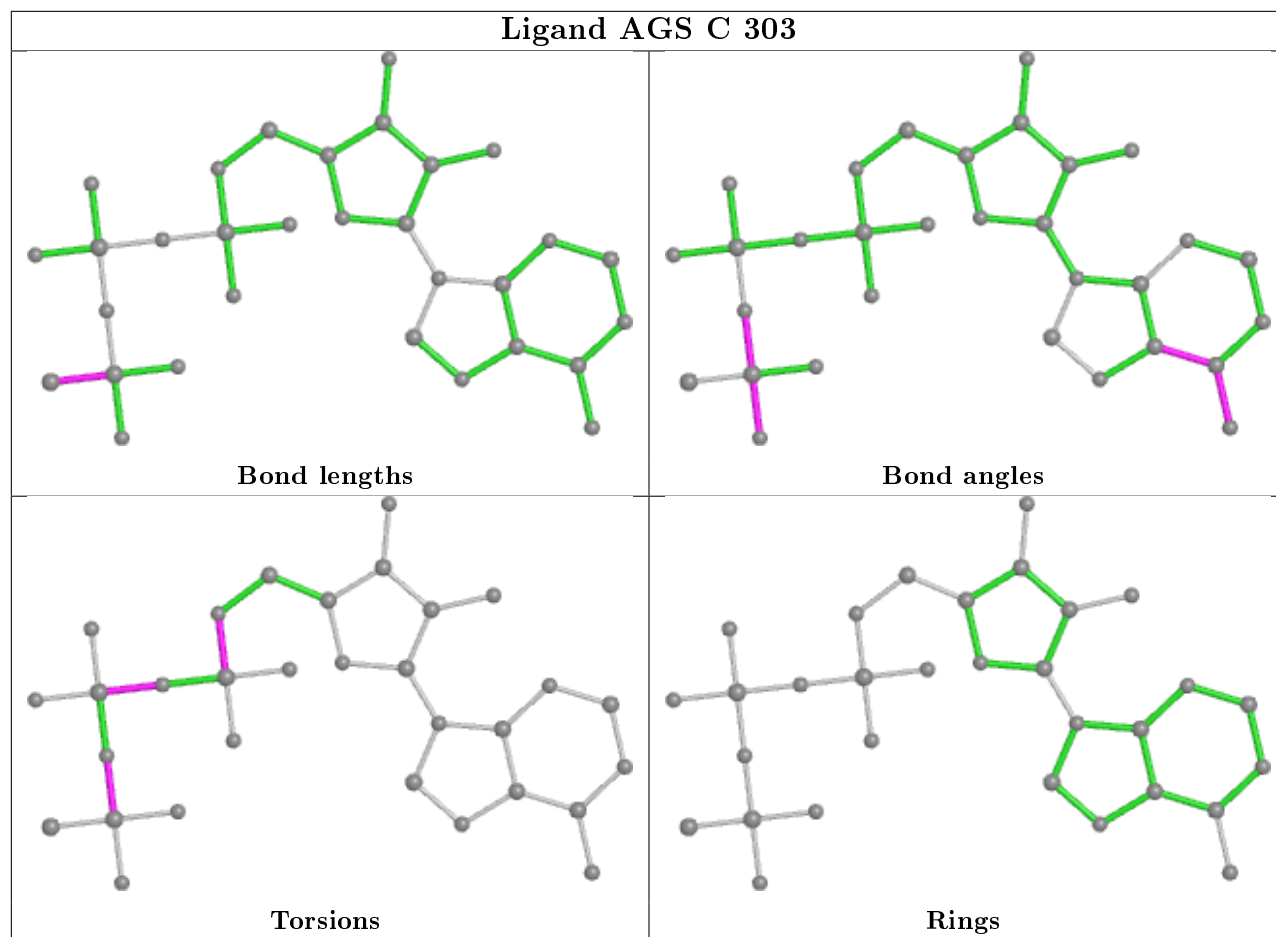












## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [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, 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	218/253 (86%)	0.76	35 (16%) 1 2	23, 45, 84, 118	0
1	B	217/253 (85%)	0.17	12 (5%) 25 31	23, 34, 70, 129	0
1	C	222/253 (87%)	0.18	13 (5%) 22 28	21, 35, 73, 109	0
1	D	219/253 (86%)	0.27	15 (6%) 17 23	19, 33, 75, 104	0
1	E	217/253 (85%)	0.36	20 (9%) 9 13	19, 35, 78, 117	0
1	F	214/253 (84%)	0.28	13 (6%) 21 27	19, 36, 73, 116	0
All	All	1307/1518 (86%)	0.34	108 (8%) 11 16	19, 36, 80, 129	0

All (108) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	121	PHE	12.0
1	A	249	LEU	9.5
1	F	122	ASP	8.9
1	F	157	CYS	8.2
1	F	110	PRO	7.8
1	A	132	TYR	7.6
1	C	157	CYS	7.3
1	C	156	ALA	7.0
1	E	132	TYR	7.0
1	F	123	LEU	6.4
1	F	249	LEU	6.1
1	E	15	HIS	5.9
1	E	157	CYS	5.8
1	F	155	ASP	5.6
1	C	121	PHE	5.2
1	D	13	SER	5.2
1	D	156	ALA	5.2
1	A	174	ILE	5.0
1	E	159	VAL	5.0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	156	ALA	4.8
1	D	249	LEU	4.7
1	B	16	GLN	4.7
1	E	158	SER	4.5
1	C	112	PRO	4.5
1	F	156	ALA	4.5
1	D	123	LEU	4.5
1	B	15	HIS	4.4
1	C	120	GLY	4.4
1	A	158	SER	4.3
1	A	157	CYS	4.3
1	F	158	SER	4.1
1	B	157	CYS	4.0
1	A	170	ARG	4.0
1	E	16	GLN	3.9
1	A	14	GLU	3.9
1	D	157	CYS	3.9
1	A	173	GLN	3.7
1	A	248	PRO	3.7
1	D	132	TYR	3.6
1	A	106	LEU	3.6
1	A	165	PHE	3.6
1	B	158	SER	3.6
1	C	249	LEU	3.6
1	B	248	PRO	3.5
1	A	99	ASP	3.5
1	C	158	SER	3.4
1	A	135	GLN	3.4
1	A	130	ILE	3.4
1	C	13	SER	3.3
1	A	166	ARG	3.3
1	F	126	LEU	3.2
1	B	249	LEU	3.2
1	B	156	ALA	3.1
1	E	111	ASP	3.1
1	D	250	GLY	3.1
1	A	110	PRO	3.1
1	A	134	ILE	3.1
1	A	88	ARG	3.1
1	E	110	PRO	3.1
1	B	122	ASP	3.0
1	A	137	TYR	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	247	PHE	3.0
1	E	250	GLY	3.0
1	A	133	ALA	3.0
1	E	248	PRO	3.0
1	F	159	VAL	2.9
1	B	159	VAL	2.9
1	E	136	LYS	2.9
1	D	126	LEU	2.8
1	C	247	PHE	2.7
1	A	129	ARG	2.7
1	A	159	VAL	2.7
1	D	112	PRO	2.7
1	E	140	ARG	2.7
1	A	111	ASP	2.6
1	E	249	LEU	2.6
1	A	136	LYS	2.6
1	D	158	SER	2.5
1	F	146[A]	SER	2.5
1	E	133	ALA	2.5
1	D	110	PRO	2.5
1	C	194	TYR	2.5
1	A	95	ALA	2.5
1	A	17	ALA	2.4
1	D	125	ALA	2.4
1	F	99	ASP	2.4
1	C	166[A]	ARG	2.4
1	A	123	LEU	2.4
1	F	247	PHE	2.4
1	A	122	ASP	2.3
1	B	17	ALA	2.3
1	E	128	GLU	2.3
1	A	104	PHE	2.3
1	A	168	VAL	2.3
1	D	159	VAL	2.2
1	A	127	ILE	2.2
1	C	162	ARG	2.2
1	A	128	GLU	2.2
1	B	110	PRO	2.2
1	D	146[A]	SER	2.2
1	E	134	ILE	2.2
1	A	101	GLY	2.2
1	C	122	ASP	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	137	TYR	2.1
1	A	97[A]	LEU	2.1
1	D	166	ARG	2.0
1	E	162	ARG	2.0
1	A	242	HIS	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 carbohydrates in this entry.

## 6.4 Ligands [i](#)

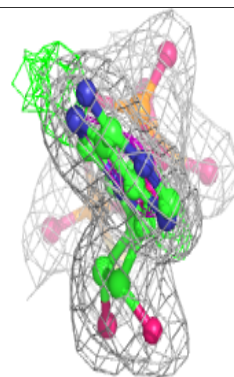
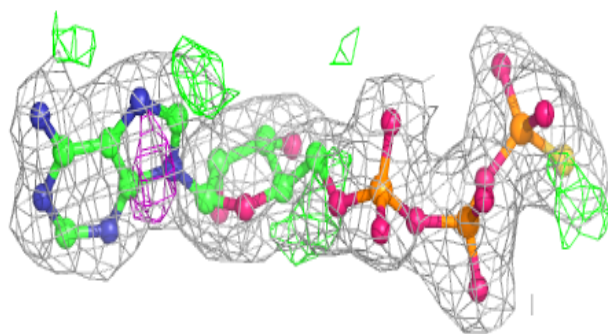
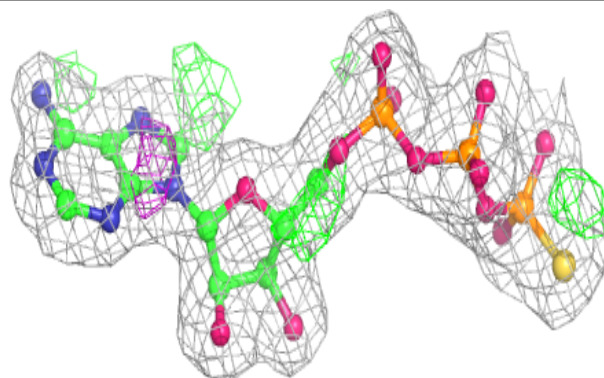
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	MG	C	301	1/1	0.95	0.05	30,30,30,30	0
4	AGS	E	303	31/31	0.96	0.11	19,25,35,41	0
2	MG	B	301	1/1	0.96	0.03	30,30,30,30	0
2	MG	D	301	1/1	0.96	0.12	29,29,29,29	0
4	AGS	D	303	31/31	0.97	0.11	21,25,37,43	0
2	MG	E	301	1/1	0.98	0.03	30,30,30,30	0
2	MG	A	301	1/1	0.98	0.10	35,35,35,35	0
4	AGS	B	303	31/31	0.98	0.08	21,27,35,42	0
4	AGS	C	303	31/31	0.98	0.08	20,25,41,42	0
4	AGS	A	303	31/31	0.98	0.12	29,33,37,40	0
2	MG	F	301	1/1	0.98	0.05	28,28,28,28	0
4	AGS	F	303	31/31	0.98	0.08	22,28,38,59	0
3	CL	B	302	1/1	0.99	0.05	32,32,32,32	0
3	CL	A	302	1/1	0.99	0.07	34,34,34,34	0
3	CL	F	302	1/1	0.99	0.04	28,28,28,28	0
3	CL	E	302	1/1	0.99	0.04	27,27,27,27	0
3	CL	C	302	1/1	1.00	0.04	28,28,28,28	0
3	CL	D	302	1/1	1.00	0.06	22,22,22,22	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.

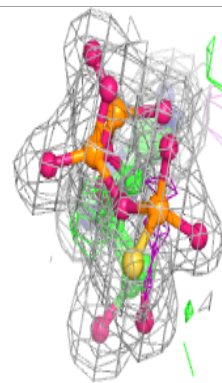
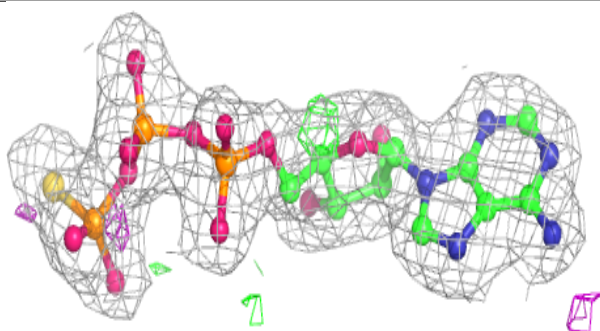
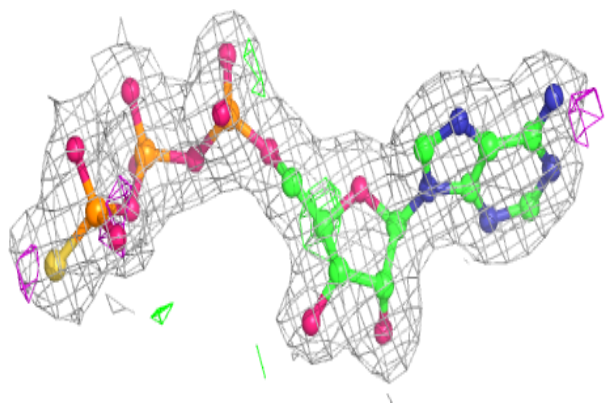
**Electron density around AGS E 303:**

2mF<sub>o</sub>-DF<sub>c</sub> (at 0.7 rmsd) in gray  
mF<sub>o</sub>-DF<sub>c</sub> (at 3 rmsd) in purple (negative)  
and green (positive)

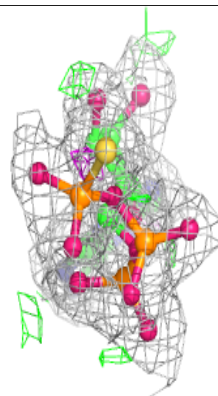
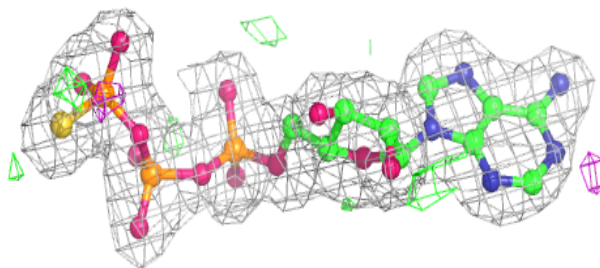
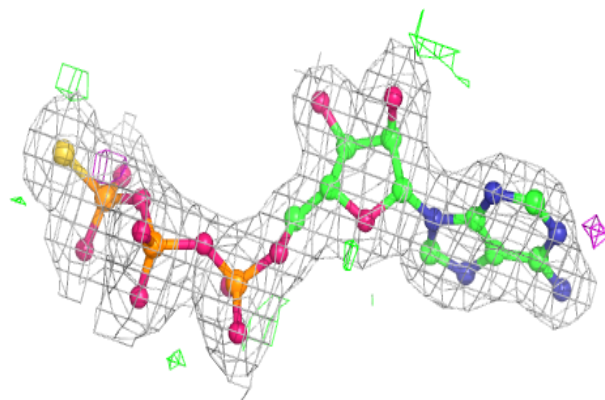


**Electron density around AGS D 303:**

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

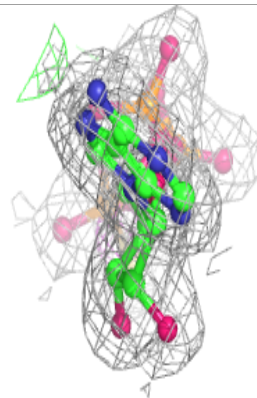
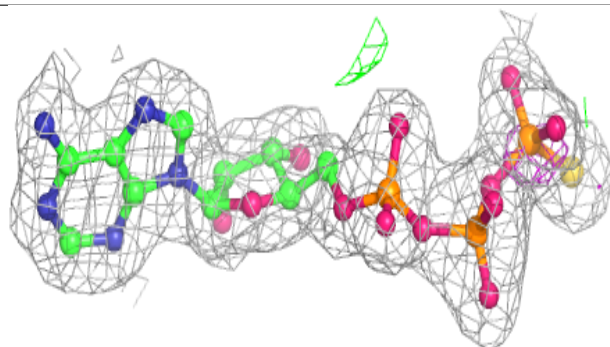
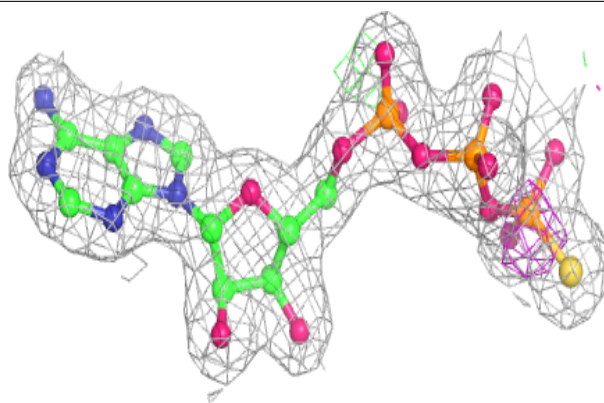
**Electron density around AGS B 303:**

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

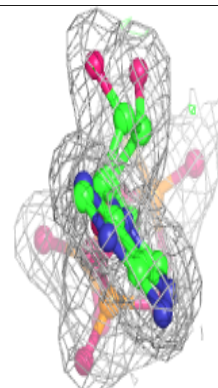
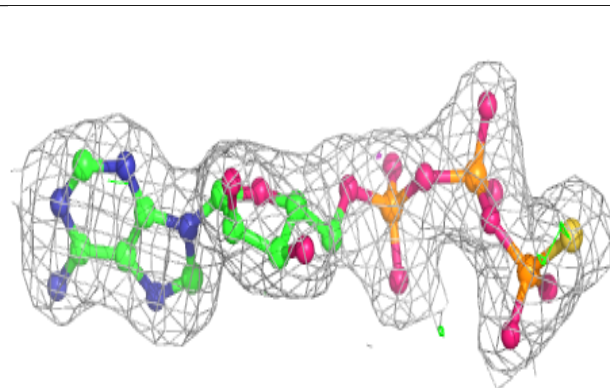
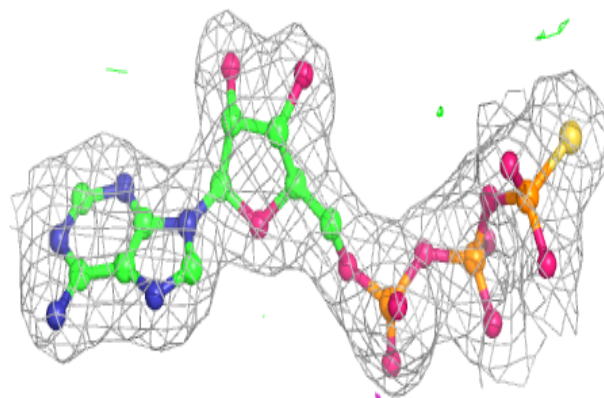


**Electron density around AGS C 303:**

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

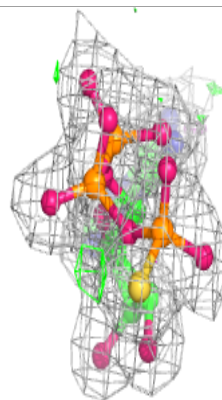
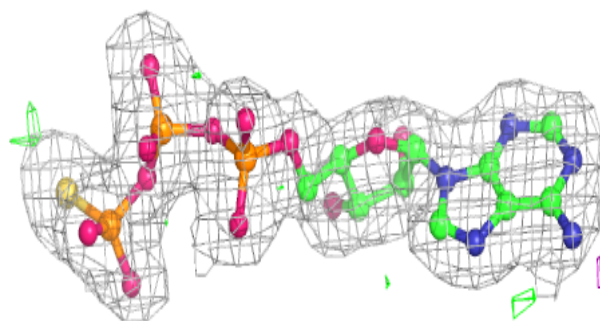
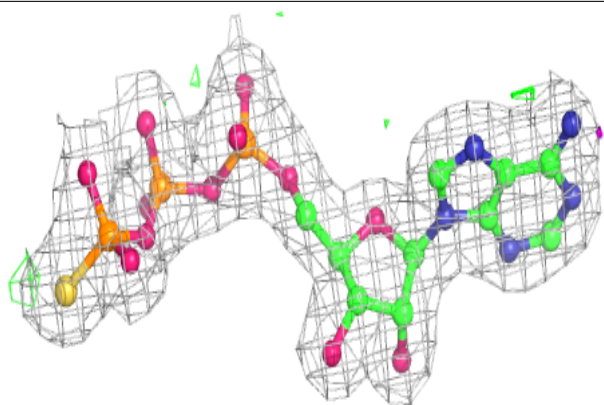
**Electron density around AGS A 303:**

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



**Electron density around AGS F 303:**

$2mF_o-DF_c$  (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.