



# Full wwPDB X-ray Structure Validation Report ⓘ

Dec 2, 2024 – 11:15 AM EST

PDB ID : 9CGR  
Title : Structure of human MAIT A-F7 TCR in complex with human MR1-Pyridoxal  
Authors : Awad, W.; Rossjohn, R.  
Deposited on : 2024-06-30  
Resolution : 2.40 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 : 2022.3.0, CSD as543be (2022)  
Xtrriage (Phenix) : 1.21  
EDS : 3.0  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.004 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.40

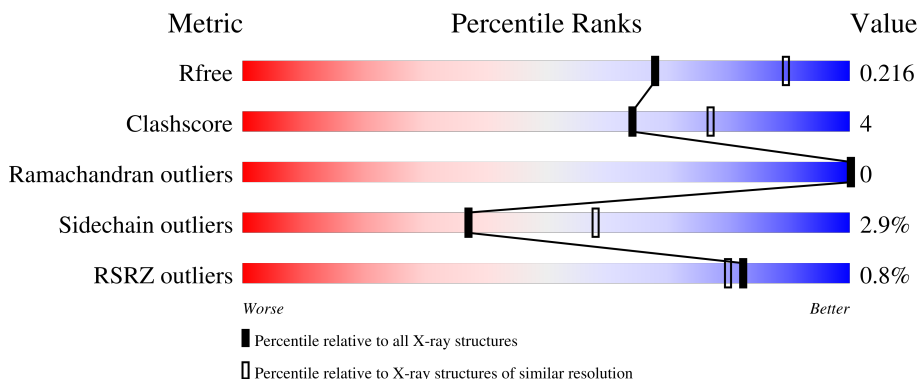
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.






Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	164625	4642 (2.40-2.40)
Clashscore	180529	5218 (2.40-2.40)
Ramachandran outliers	177936	5158 (2.40-2.40)
Sidechain outliers	177891	5159 (2.40-2.40)
RSRZ outliers	164620	4642 (2.40-2.40)

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	271	 86% 11% ..
1	C	271	 86% 11% ..
2	B	100	 86% 12% .
2	F	100	 96% .
3	D	204	 86% 10% .

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	G	204	 90% 8% •
4	E	246	 % 83% 15% ••
4	H	246	 % 91% 7% •

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
7	ACY	C	305	-	-	X	-

## 2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 13293 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 Major histocompatibility complex class I-related gene protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	266	2191	1402	379	398	12	0	5	0
1	C	266	2189	1406	378	394	11	0	3	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP Q95460
A	261	SER	CYS	conflict	UNP Q95460
C	0	MET	-	initiating methionine	UNP Q95460
C	261	SER	CYS	conflict	UNP Q95460

- Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	98	797	511	136	147	3	0	0	0
2	F	100	817	521	138	154	4	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	initiating methionine	UNP P61769
F	0	MET	-	initiating methionine	UNP P61769

- Molecule 3 is a protein called TCR TRAV1-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	D	197	1506	955	238	303	10	0	2	0

*Continued on next page...*

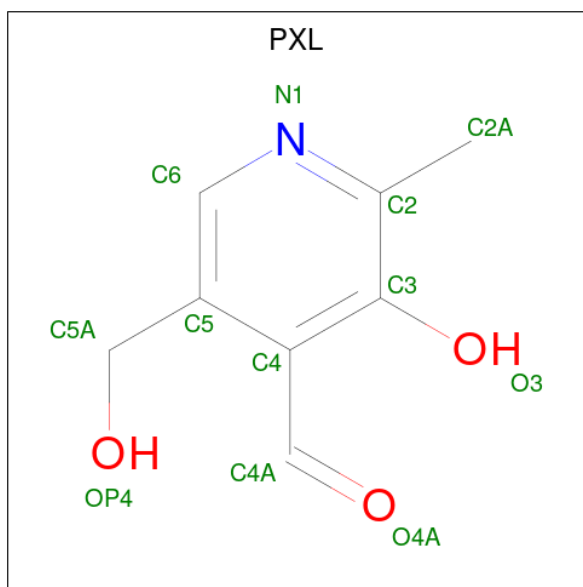
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	G	200	1578	998	252	319	9	1	4	0

- Molecule 4 is a protein called TCR TRBV 6-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	E	242	1885	1191	324	359	11	0	4	0
4	H	244	1895	1198	328	359	10	0	3	0

- Molecule 5 is 3-HYDROXY-5-(HYDROXYMETHYL)-2-METHYLISONICOTINALDEHYDE (three-letter code: PXL) (formula: C<sub>8</sub>H<sub>9</sub>NO<sub>3</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
5	A	1	11	8	1	2	0	0
5	C	1	11	8	1	2	0	0

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C O 6 3 3	0	0
6	A	1	Total C O 6 3 3	0	0
6	C	1	Total C O 6 3 3	0	0
6	C	1	Total C O 6 3 3	0	0
6	C	1	Total C O 6 3 3	0	0
6	C	1	Total C O 6 3 3	0	0
6	D	1	Total C O 6 3 3	0	0
6	F	1	Total C O 6 3 3	0	0
6	F	1	Total C O 6 3 3	0	0
6	G	1	Total C O 6 3 3	0	0

- Molecule 7 is ACETIC ACID (three-letter code: ACY) (formula: C<sub>2</sub>H<sub>4</sub>O<sub>2</sub>).



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

- Molecule 8 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	1	Total Cl 1 1	0	0
8	C	1	Total Cl 1 1	0	0

- Molecule 9 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	H	1	Total Ca 1 1	0	0

- Molecule 10 is water.

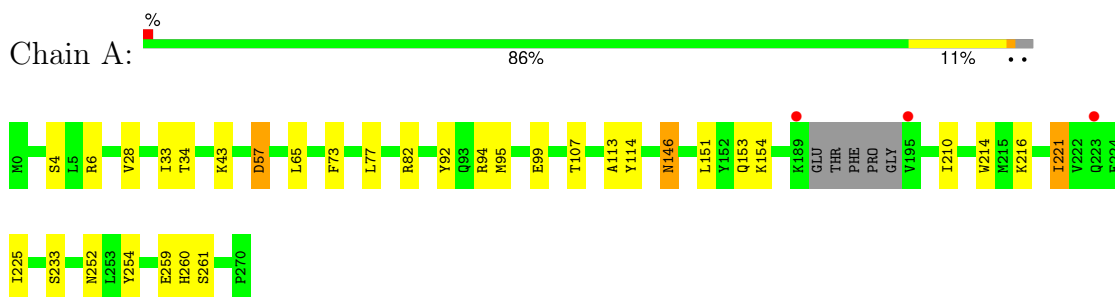
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
10	A	67	Total 67	O 67	0	0
10	B	18	Total 18	O 18	0	0
10	C	57	Total 57	O 57	0	0
10	D	22	Total 22	O 22	0	0
10	E	27	Total 27	O 27	0	0
10	F	22	Total 22	O 22	0	0
10	G	55	Total 55	O 55	0	0
10	H	62	Total 62	O 62	0	0



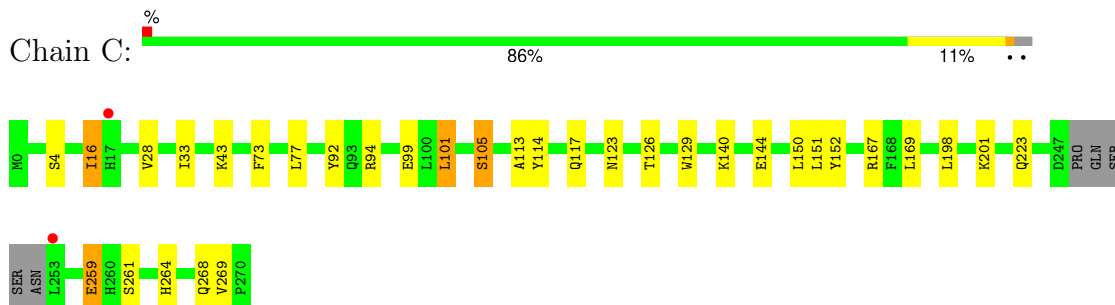
### 3 Residue-property plots [i](#)

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.

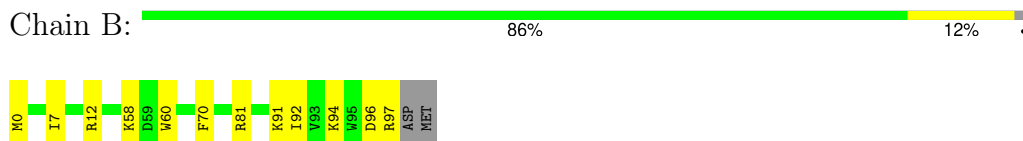
- Molecule 1: Major histocompatibility complex class I-related gene protein



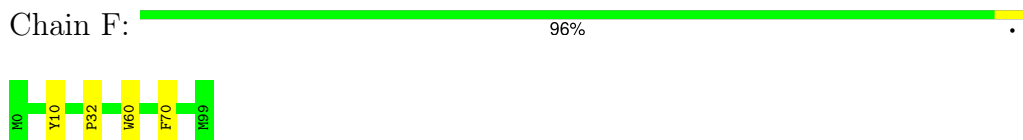
- Molecule 1: Major histocompatibility complex class I-related gene protein



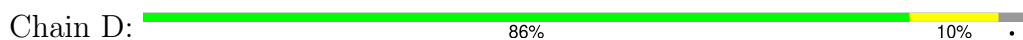
- Molecule 2: Beta-2-microglobulin

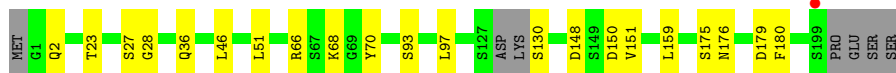


- Molecule 2: Beta-2-microglobulin

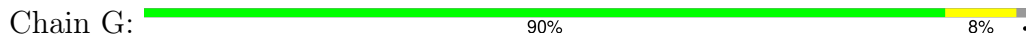


- Molecule 3: TCR TRAV1-2

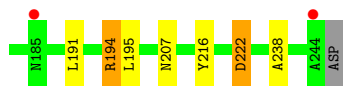
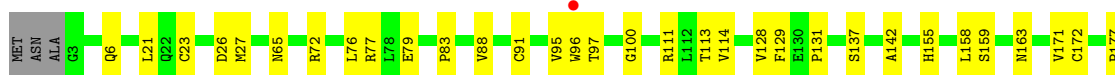
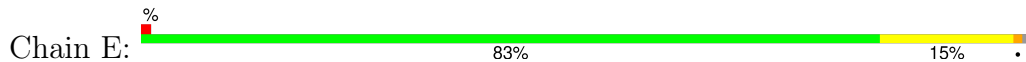




• Molecule 3: TCR TRAV1-2



• Molecule 4: TCR TRBV 6-1



• Molecule 4: TCR TRBV 6-1



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	216.07Å 69.90Å 142.44Å 90.00° 103.83° 90.00°	Depositor
Resolution (Å)	46.15 – 2.40 46.15 – 2.40	Depositor EDS
% Data completeness (in resolution range)	98.1 (46.15-2.40) 91.0 (46.15-2.40)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.26 (at 2.39Å)	Xtrriage
Refinement program	PHENIX 1.21.1_5286	Depositor
R, $R_{free}$	0.173 , 0.216 0.173 , 0.216	Depositor DCC
$R_{free}$ test set	77648 reflections (2.51%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	53.7	Xtrriage
Anisotropy	0.107	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 46.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	13293	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	66.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.56% 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: ACY, CA, CL, PXL, GOL

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	0.24	0/2271	0.48	0/3087
1	C	0.24	0/2264	0.48	0/3078
2	B	0.24	0/820	0.48	0/1114
2	F	0.25	0/840	0.49	0/1141
3	D	0.26	0/1545	0.47	0/2100
3	G	0.25	0/1625	0.47	0/2205
4	E	0.25	0/1948	0.50	0/2656
4	H	0.26	0/1955	0.51	0/2666
All	All	0.25	0/13268	0.49	0/18047

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [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	2191	0	2076	22	0
1	C	2189	0	2084	21	0
2	B	797	0	752	5	0
2	F	817	0	764	3	0
3	D	1506	0	1390	12	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	G	1578	0	1495	10	0
4	E	1885	0	1756	24	0
4	H	1895	0	1777	10	0
5	A	11	0	7	1	0
5	C	11	0	7	1	0
6	A	12	0	16	2	0
6	C	24	0	32	3	0
6	D	6	0	8	1	0
6	F	12	0	16	0	0
6	G	6	0	8	0	0
7	A	8	0	6	1	0
7	C	12	0	9	4	0
8	A	1	0	0	0	0
8	C	1	0	0	0	0
9	H	1	0	0	0	0
10	A	67	0	0	2	0
10	B	18	0	0	0	0
10	C	57	0	0	1	0
10	D	22	0	0	0	0
10	E	27	0	0	0	0
10	F	22	0	0	0	0
10	G	55	0	0	0	0
10	H	62	0	0	1	0
All	All	13293	0	12203	96	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 4.

All (96) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:43:LYS:NZ	5:A:301:PXL:O3	2.13	0.81
1:A:94:ARG:HH12	6:A:302:GOL:H32	1.52	0.74
3:G:159:LEU:HB3	4:H:172:CYS:HB2	1.71	0.70
3:D:2:GLN:HE21	3:D:27:SER:H	1.41	0.68
1:C:77:LEU:HD13	1:C:92:TYR:HB2	1.75	0.68
3:D:28:GLY:HA3	3:D:93[B]:SER:HB3	1.77	0.66
1:A:77:LEU:HD13	1:A:92:TYR:HB2	1.78	0.65
1:A:146:ASN:ND2	10:A:402:HOH:O	2.29	0.65
1:C:94:ARG:HH12	6:C:303:GOL:H11	1.61	0.63
3:D:130:SER:HB2	4:E:129:PHE:HE2	1.64	0.63

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:E:26:ASP:OD1	4:E:72:ARG:NH1	2.33	0.62
7:C:306:ACY:H1	2:F:32:PRO:HB2	1.80	0.62
4:E:111[B]:ARG:NH1	4:E:113:THR:OG1	2.34	0.61
3:D:28:GLY:HA3	3:D:93[A]:SER:HB3	1.82	0.61
1:A:94:ARG:HH22	6:A:302:GOL:H11	1.67	0.59
4:E:88:VAL:HG22	4:E:111[A]:ARG:HD3	1.83	0.59
4:E:222:ASP:OD1	4:E:222:ASP:N	2.35	0.59
3:G:2:GLN:HE22	3:G:27:SER:H	1.51	0.59
1:C:28:VAL:HG23	1:C:33:ILE:HD13	1.83	0.59
1:C:167:ARG:HH11	6:C:304:GOL:H31	1.69	0.58
3:G:70:TYR:HH	3:G:72:TYR:HD2	1.52	0.57
4:H:224:TRP:HB2	4:H:230:LYS:HD3	1.87	0.57
1:A:28:VAL:HG23	1:A:33:ILE:HD13	1.86	0.56
3:G:4[B]:ILE:HD11	3:G:90:VAL:HB	1.87	0.56
4:H:128:VAL:HG23	4:H:238:ALA:HB3	1.88	0.56
1:C:129:TRP:O	1:C:140:LYS:NZ	2.38	0.55
7:C:308:ACY:H3	2:F:10:TYR:HB3	1.88	0.55
4:E:171:VAL:HG23	4:E:195:LEU:HD23	1.88	0.54
1:C:144:GLU:HA	1:C:150:LEU:HD11	1.90	0.53
1:C:101:LEU:HD23	1:C:105:SER:HB3	1.91	0.53
4:H:86:THR:HG23	4:H:113:THR:HA	1.90	0.53
3:D:51:LEU:HA	3:D:66:ARG:NH1	2.25	0.52
1:A:34:THR:HB	1:A:43:LYS:HE2	1.91	0.52
4:H:11:GLN:HG2	4:H:19:MET:SD	2.50	0.52
4:E:163:ASN:HD21	4:E:207:ASN:HD22	1.58	0.51
4:E:172[B]:CYS:SG	4:E:194:ARG:HD2	2.51	0.51
4:E:65:ASN:OD1	4:E:77:ARG:NH1	2.42	0.51
1:A:153[B]:GLN:OE1	10:A:401:HOH:O	2.18	0.50
3:D:36:GLN:HB2	3:D:46:LEU:HD11	1.92	0.50
3:G:2:GLN:NE2	3:G:27:SER:H	2.09	0.50
1:A:210:ILE:HG13	1:A:260:HIS:HB2	1.94	0.49
1:C:259:GLU:HG2	1:C:264:HIS:HD2	1.77	0.49
4:E:6:GLN:HB3	4:E:91[B]:CYS:SG	2.52	0.49
1:A:151:LEU:HD22	3:G:51:LEU:HD12	1.95	0.49
4:H:230:LYS:HG2	4:H:232:VAL:HG13	1.94	0.49
1:C:43:LYS:NZ	5:C:301:PXL:O3	2.33	0.49
1:A:113:ALA:HB2	2:B:60:TRP:CE2	2.47	0.49
1:C:151:LEU:HD22	3:D:51:LEU:HD12	1.94	0.49
1:A:154:LYS:HD3	3:G:51:LEU:HD11	1.94	0.48
1:A:252:ASN:HB2	1:A:254:TYR:CE1	2.48	0.48
1:A:233:SER:HB3	2:B:12:ARG:HG3	1.96	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:117:GLN:HB2	7:C:305:ACY:H2	1.95	0.47
4:E:83:PRO:HA	4:E:114:VAL:HB	1.95	0.47
1:C:152:TYR:CD1	4:E:100:GLY:HA3	2.50	0.47
4:E:21:LEU:HD12	4:E:76:LEU:HD23	1.97	0.46
6:D:401:GOL:H12	4:H:229:ALA:HA	1.96	0.46
1:C:201:LYS:NZ	10:C:401:HOH:O	2.47	0.46
1:C:123:ASN:ND2	1:C:126:THR:OG1	2.46	0.45
1:C:113:ALA:HB2	2:F:60:TRP:CE2	2.51	0.45
4:E:23:CYS:SG	4:E:91[B]:CYS:HB2	2.56	0.45
2:B:81:ARG:HG3	2:B:92:ILE:HG12	1.99	0.44
4:E:128:VAL:HG23	4:E:238:ALA:HB3	1.99	0.44
1:A:57:ASP:OD1	1:A:57:ASP:N	2.44	0.44
2:B:7:ILE:HD12	2:B:91:LYS:HD2	1.99	0.44
3:D:23:THR:HG22	3:D:70:TYR:HB2	2.00	0.44
1:A:114:TYR:HB3	7:A:303:ACY:H3	1.98	0.44
4:H:174:ASP:HB2	4:H:191:LEU:HD12	2.00	0.43
3:D:130:SER:HB2	4:E:129:PHE:CE2	2.49	0.43
4:E:27:MET:HE3	4:E:27:MET:HB2	1.92	0.43
3:D:148:ASP:HB3	3:D:151:VAL:HB	2.01	0.43
4:E:163:ASN:HD21	4:E:207:ASN:ND2	2.17	0.43
4:E:95:VAL:HG12	4:E:96:TRP:CD1	2.55	0.42
1:A:99:GLU:HB2	1:A:107:THR:OG1	2.19	0.42
4:H:120:ASN:HB2	10:H:420:HOH:O	2.18	0.42
1:A:214:TRP:H	1:A:225:ILE:HD13	1.85	0.42
1:A:6:ARG:NH1	2:B:58:LYS:HG2	2.34	0.42
4:E:155:HIS:HB3	4:E:216:TYR:HB2	2.00	0.42
1:A:221:ILE:O	1:A:225:ILE:HG13	2.19	0.42
1:C:4:SER:HB3	1:C:99:GLU:HG2	2.02	0.42
3:D:175:SER:HB3	3:D:180:PHE:CG	2.54	0.42
4:E:88:VAL:HG22	4:E:111[B]:ARG:HG2	2.01	0.42
1:C:198:LEU:HD13	1:C:269:VAL:HG21	2.01	0.41
4:H:95:VAL:HG12	4:H:96:TRP:CD1	2.55	0.41
1:C:16:ILE:H	1:C:16:ILE:HD13	1.85	0.41
1:A:4:SER:HB3	1:A:99:GLU:HG2	2.01	0.41
1:C:114:TYR:HB3	7:C:305:ACY:H3	2.00	0.41
3:G:91:LYS:HG2	3:G:95:TYR:HA	2.01	0.41
3:G:162:ARG:NH2	3:G:165:ASP:OD1	2.47	0.41
3:D:97:LEU:HD23	3:D:97:LEU:HA	1.97	0.41
4:E:177:PRO:HG3	4:E:191:LEU:HD13	2.03	0.41
1:C:167:ARG:NH1	6:C:304:GOL:H31	2.32	0.41
1:C:169[A]:LEU:HD23	1:C:169[A]:LEU:HA	1.98	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:E:158:LEU:HD23	4:E:159:SER:N	2.36	0.41
1:A:216:LYS:N	1:A:221:ILE:HD11	2.36	0.40
4:E:131:PRO:HG2	4:E:142:ALA:HB1	2.02	0.40
3:G:60:PHE:CD1	3:G:75:LEU:HD22	2.56	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	267/271 (98%)	261 (98%)	6 (2%)	0	100	100
1	C	265/271 (98%)	259 (98%)	6 (2%)	0	100	100
2	B	96/100 (96%)	96 (100%)	0	0	100	100
2	F	98/100 (98%)	98 (100%)	0	0	100	100
3	D	195/204 (96%)	190 (97%)	5 (3%)	0	100	100
3	G	202/204 (99%)	195 (96%)	7 (4%)	0	100	100
4	E	244/246 (99%)	236 (97%)	8 (3%)	0	100	100
4	H	245/246 (100%)	240 (98%)	5 (2%)	0	100	100
All	All	1612/1642 (98%)	1575 (98%)	37 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	232/241 (96%)	222 (96%)	10 (4%)	25	42
1	C	231/241 (96%)	223 (96%)	8 (4%)	31	51
2	B	87/95 (92%)	82 (94%)	5 (6%)	17	29
2	F	90/95 (95%)	89 (99%)	1 (1%)	70	84
3	D	163/181 (90%)	158 (97%)	5 (3%)	35	56
3	G	178/181 (98%)	175 (98%)	3 (2%)	56	75
4	E	199/212 (94%)	194 (98%)	5 (2%)	42	63
4	H	200/212 (94%)	194 (97%)	6 (3%)	36	57
All	All	1380/1458 (95%)	1337 (97%)	43 (3%)	37	56

All (43) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	57	ASP
1	A	65	LEU
1	A	73	PHE
1	A	82	ARG
1	A	95[A]	MET
1	A	95[B]	MET
1	A	146	ASN
1	A	221	ILE
1	A	259	GLU
1	A	261	SER
2	B	0	MET
2	B	70	PHE
2	B	94	LYS
2	B	96	ASP
2	B	97	ARG
1	C	16	ILE
1	C	73	PHE
1	C	101	LEU
1	C	105	SER
1	C	223	GLN
1	C	259	GLU
1	C	261	SER
1	C	268	GLN
3	D	68	LYS
3	D	150	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	D	159	LEU
3	D	176	ASN
3	D	179	ASP
4	E	79	GLU
4	E	97	THR
4	E	137	SER
4	E	194	ARG
4	E	222	ASP
2	F	70	PHE
3	G	93[A]	SER
3	G	93[B]	SER
3	G	200	PRO
4	H	19	MET
4	H	65[A]	ASN
4	H	65[B]	ASN
4	H	178[A]	LEU
4	H	178[B]	LEU
4	H	194	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (19) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	58	HIS
1	A	146	ASN
1	C	93	GLN
1	C	123	ASN
1	C	223	GLN
1	C	264	HIS
3	D	2	GLN
3	D	25	GLN
3	D	38	HIS
3	D	79	GLN
3	D	96	GLN
3	D	120	GLN
3	D	187	ASN
4	E	207	ASN
2	F	2	GLN
3	G	2	GLN
3	G	79	GLN
4	H	176	GLN
4	H	208	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 oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 20 ligands modelled in this entry, 3 are monoatomic - leaving 17 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	$\# Z  > 2$
6	GOL	F	102	-	5,5,5	0.33	0	5,5,5	0.36	0
6	GOL	G	501	-	5,5,5	0.36	0	5,5,5	0.45	0
5	PXL	C	301	1	11,11,12	0.90	1 (9%)	15,15,16	0.56	0
6	GOL	A	302	-	5,5,5	0.34	0	5,5,5	0.38	0
6	GOL	C	302	-	5,5,5	0.34	0	5,5,5	0.47	0
6	GOL	C	304	-	5,5,5	0.35	0	5,5,5	0.44	0
6	GOL	C	307	-	5,5,5	0.34	0	5,5,5	0.42	0
7	ACY	C	305	-	3,3,3	1.15	0	3,3,3	1.23	0
7	ACY	C	306	-	3,3,3	1.13	0	3,3,3	1.24	0
6	GOL	D	401	-	5,5,5	0.34	0	5,5,5	0.47	0
6	GOL	C	303	-	5,5,5	0.34	0	5,5,5	0.43	0
7	ACY	A	303	-	3,3,3	1.13	0	3,3,3	1.21	0
7	ACY	C	308	-	3,3,3	1.13	0	3,3,3	1.23	0
5	PXL	A	301	1	11,11,12	0.89	1 (9%)	15,15,16	0.58	0
6	GOL	A	305	-	5,5,5	0.34	0	5,5,5	0.45	0
7	ACY	A	304	-	3,3,3	1.15	0	3,3,3	1.17	0
6	GOL	F	101	-	5,5,5	0.36	0	5,5,5	0.63	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
6	GOL	F	102	-	-	0/4/4/4	-
6	GOL	G	501	-	-	0/4/4/4	-
5	PXL	C	301	1	-	0/2/2/4	0/1/1/1
6	GOL	A	302	-	-	1/4/4/4	-
6	GOL	C	302	-	-	2/4/4/4	-
6	GOL	C	304	-	-	0/4/4/4	-
6	GOL	C	307	-	-	4/4/4/4	-
6	GOL	D	401	-	-	1/4/4/4	-
6	GOL	C	303	-	-	4/4/4/4	-
5	PXL	A	301	1	-	0/2/2/4	0/1/1/1
6	GOL	A	305	-	-	2/4/4/4	-
6	GOL	F	101	-	-	4/4/4/4	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	301	PXL	C4A-C4	-2.76	1.46	1.51
5	A	301	PXL	C4A-C4	-2.67	1.46	1.51

There are no bond angle outliers.

There are no chirality outliers.

All (18) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	305	GOL	O1-C1-C2-C3
6	C	302	GOL	O1-C1-C2-C3
6	C	303	GOL	O1-C1-C2-O2
6	C	303	GOL	O1-C1-C2-C3
6	C	303	GOL	C1-C2-C3-O3
6	A	305	GOL	O1-C1-C2-O2
6	C	302	GOL	O1-C1-C2-O2
6	A	302	GOL	O1-C1-C2-C3
6	C	307	GOL	O1-C1-C2-C3
6	C	307	GOL	C1-C2-C3-O3
6	F	101	GOL	O1-C1-C2-C3
6	F	101	GOL	C1-C2-C3-O3

*Continued on next page...*

*Continued from previous page...*

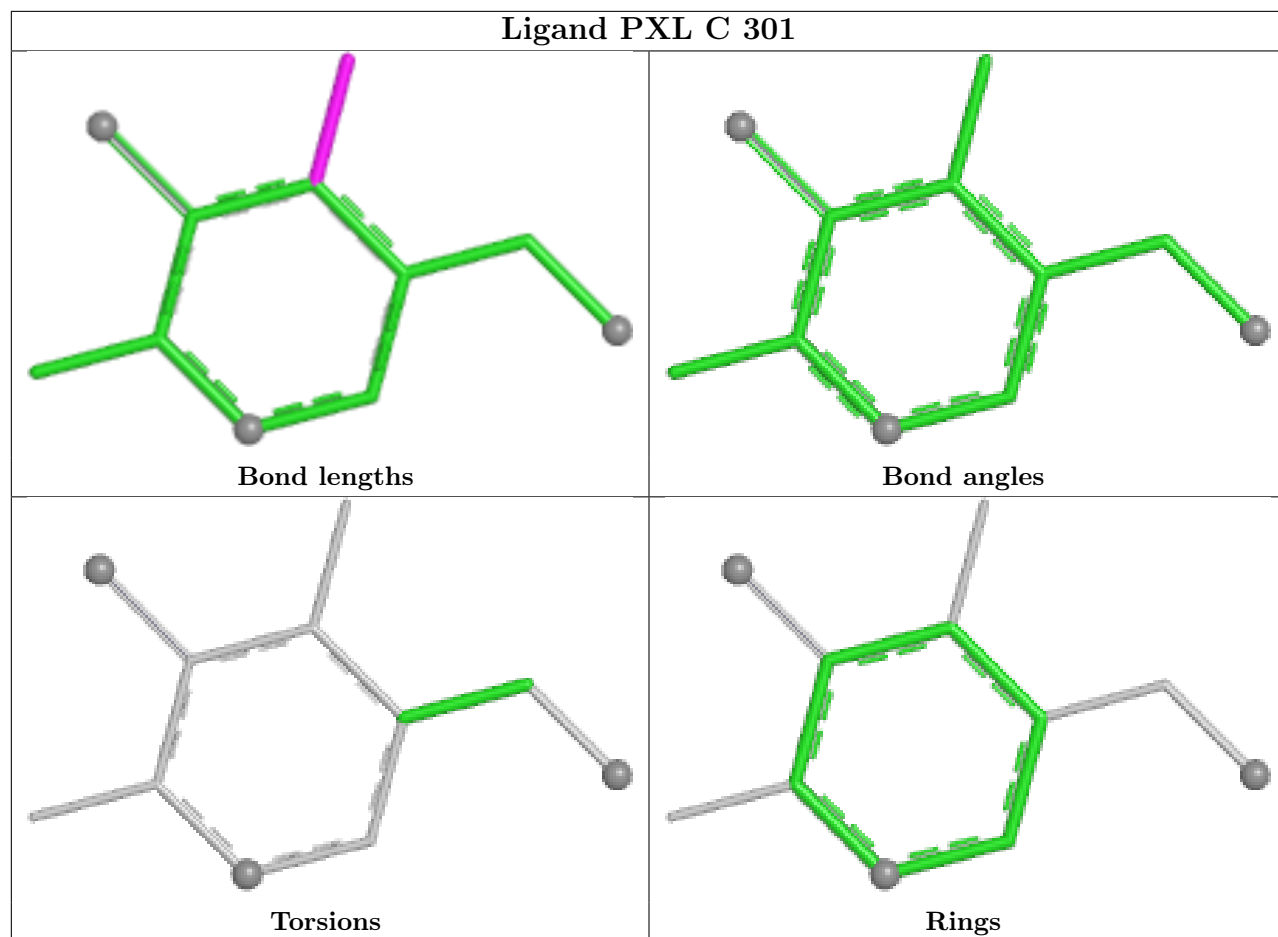
Mol	Chain	Res	Type	Atoms
6	C	303	GOL	O2-C2-C3-O3
6	D	401	GOL	O2-C2-C3-O3
6	C	307	GOL	O1-C1-C2-O2
6	F	101	GOL	O2-C2-C3-O3
6	C	307	GOL	O2-C2-C3-O3
6	F	101	GOL	O1-C1-C2-O2

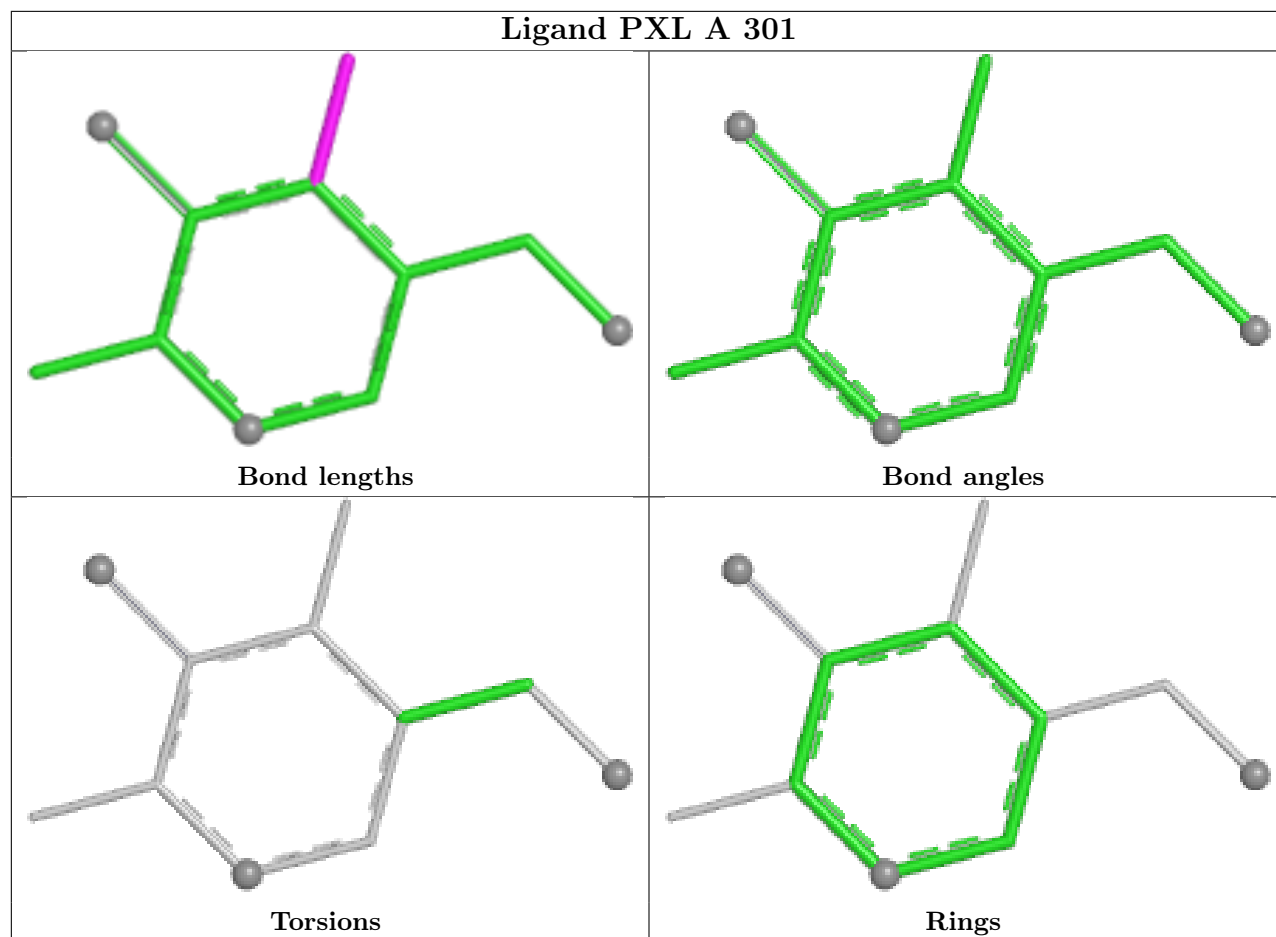
There are no ring outliers.

10 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	301	PXL	1	0
6	A	302	GOL	2	0
6	C	304	GOL	2	0
7	C	305	ACY	2	0
7	C	306	ACY	1	0
6	D	401	GOL	1	0
6	C	303	GOL	1	0
7	A	303	ACY	1	0
7	C	308	ACY	1	0
5	A	301	PXL	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

### 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	266/271 (98%)	-0.32	3 (1%) 77 75	30, 59, 102, 122	5 (1%)
1	C	266/271 (98%)	-0.38	2 (0%) 82 80	33, 57, 83, 110	8 (3%)
2	B	98/100 (98%)	-0.15	0 100 100	46, 76, 108, 118	0
2	F	100/100 (100%)	-0.44	0 100 100	44, 62, 86, 98	1 (1%)
3	D	197/204 (96%)	-0.08	1 (0%) 87 85	36, 69, 107, 136	5 (2%)
3	G	200/204 (98%)	-0.34	1 (0%) 87 85	33, 57, 89, 111	7 (3%)
4	E	242/246 (98%)	-0.08	3 (1%) 76 73	38, 71, 108, 123	7 (2%)
4	H	244/246 (99%)	-0.27	3 (1%) 76 73	37, 61, 88, 104	5 (2%)
All	All	1613/1642 (98%)	-0.26	13 (0%) 82 80	30, 62, 101, 136	38 (2%)

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	195	VAL	3.5
4	E	244	ALA	3.3
4	E	96	TRP	2.7
4	E	185	ASN	2.5
1	A	189	LYS	2.4
3	G	201	GLU	2.4
4	H	79	GLU	2.3
4	H	1	ASN	2.3
4	H	244	ALA	2.2
1	C	253	LEU	2.1
1	C	17	HIS	2.1
3	D	199	SER	2.1
1	A	223	GLN	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 [i](#)

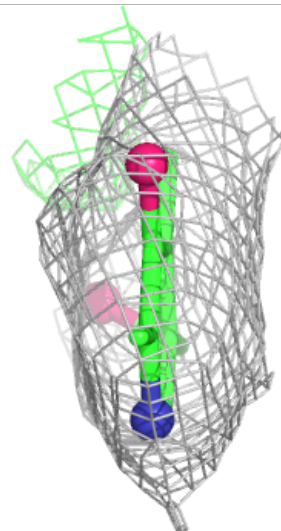
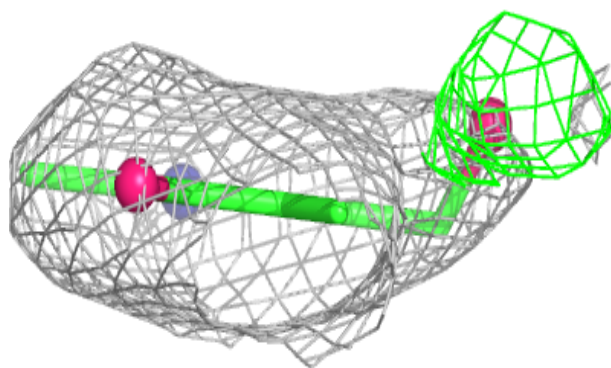
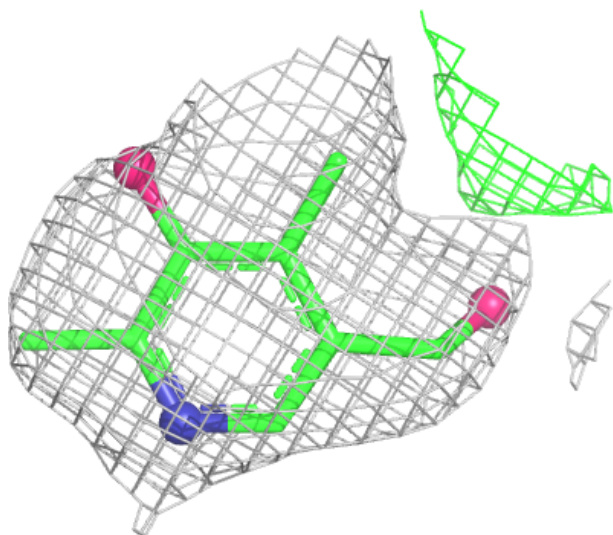
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

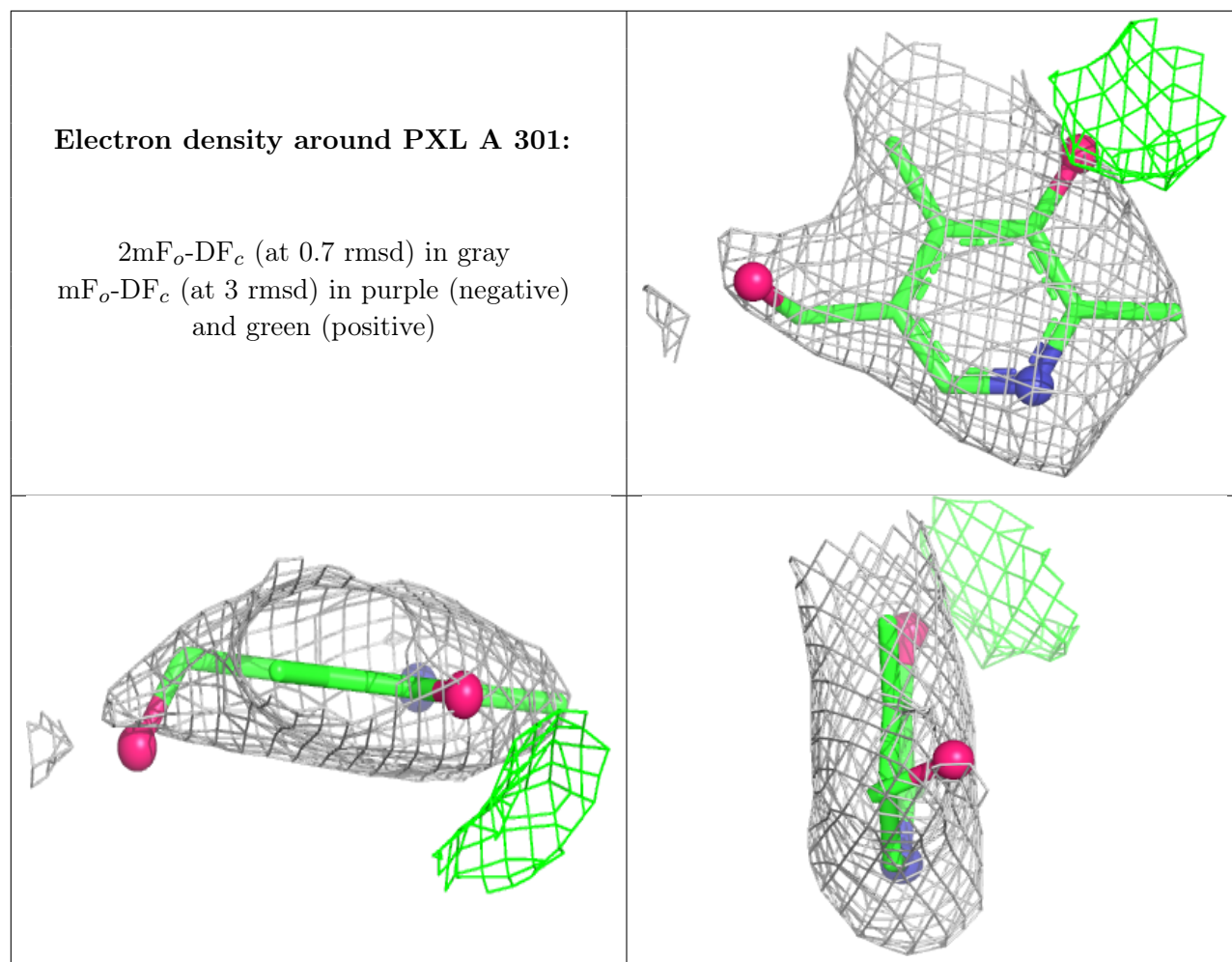
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
7	ACY	A	304	4/4	0.77	0.16	66,70,88,101	0
6	GOL	A	305	6/6	0.85	0.11	66,75,84,91	0
7	ACY	C	305	4/4	0.85	0.15	60,84,84,86	0
6	GOL	G	501	6/6	0.88	0.13	64,71,76,81	0
7	ACY	A	303	4/4	0.88	0.17	65,78,84,85	0
7	ACY	C	306	4/4	0.88	0.14	51,65,83,84	0
7	ACY	C	308	4/4	0.89	0.16	72,77,81,86	0
6	GOL	C	304	6/6	0.90	0.14	79,80,84,90	0
6	GOL	C	307	6/6	0.92	0.13	56,76,81,87	0
6	GOL	D	401	6/6	0.92	0.12	55,71,73,75	0
5	PXL	C	301	11/12	0.92	0.12	62,73,85,86	0
6	GOL	A	302	6/6	0.92	0.16	69,79,83,87	0
6	GOL	C	303	6/6	0.93	0.12	74,77,86,86	0
8	CL	C	309	1/1	0.94	0.09	75,75,75,75	0
9	CA	H	301	1/1	0.94	0.09	91,91,91,91	0
5	PXL	A	301	11/12	0.95	0.10	51,58,64,69	2
6	GOL	F	101	6/6	0.95	0.11	48,56,66,72	0
6	GOL	C	302	6/6	0.96	0.10	48,62,66,70	0
6	GOL	F	102	6/6	0.96	0.07	42,49,51,53	0
8	CL	A	306	1/1	0.97	0.07	74,74,74,74	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 PXL 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.