Full wwPDB X-ray Structure Validation Report (i)

Apr 28, 2025 – 10:30 AM EDT

PDB ID	:	9EK7 / pdb_00009ek7
Title	:	Crystal structure of MAIT TCR in complex with MR1-5FdU
Authors	:	Awad, W.; Rossjohn, J.
Deposited on	:	2024-12-01
Resolution	:	2.15 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

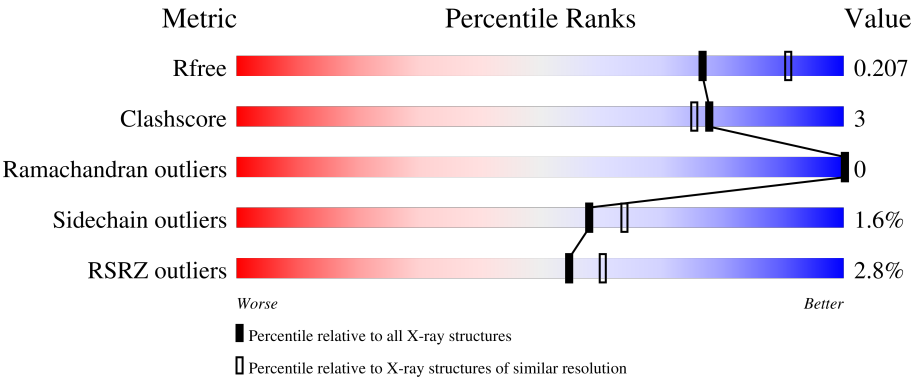
MolProbity	:	4-5-2 with Phenix2.0rc1
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.43.1

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.15 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	1881 (2.16-2.16)
Clashscore	180529	2047 (2.16-2.16)
Ramachandran outliers	177936	2027 (2.16-2.16)
Sidechain outliers	177891	2026 (2.16-2.16)
RSRZ outliers	164620	1882 (2.16-2.16)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	271	3% 92% 6%
1	C	271	5% 88% 10%
2	B	204	% 89% 7%
2	D	204	5% 90% 5%
3	E	246	4% 92% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	246	92% 7%
4	F	100	% 94%
4	H	100	% 94%

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
5	YC6	A	301[B]	X	-	-	-

2 Entry composition (i)

There are 11 unique types of molecules in this entry. The entry contains 13963 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
1	A	266	Total	C	N	O	S	0	6	0
			2196	1404	385	396	11			
1	C	268	Total	C	N	O	S	1	9	0
			2222	1427	384	399	12			

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 TCR alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	200	Total	C	N	O	S	1	8	0
			1576	1002	248	314	12			
2	D	193	Total	C	N	O	S	0	2	0
			1454	933	234	278	9			

- Molecule 3 is a protein called TCR beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	241	Total	C	N	O	S	0	6	0
			1877	1189	318	359	11			
3	G	244	Total	C	N	O	S	0	12	0
			1973	1243	342	375	13			

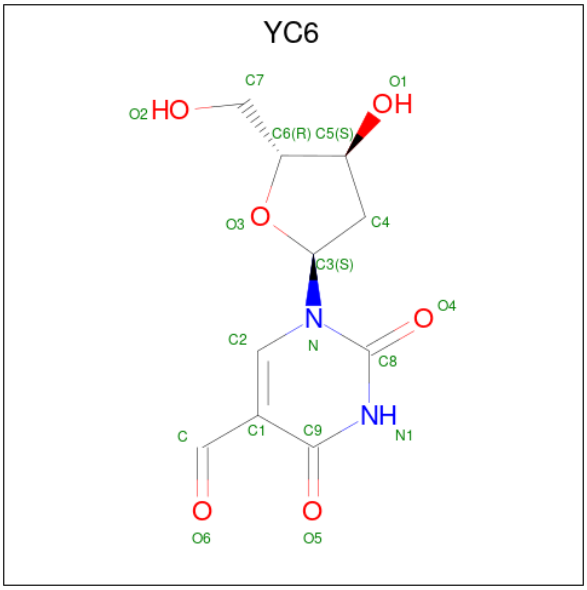
- Molecule 4 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	98	Total	C	N	O	S	0	2	0
			796	512	134	148	2			
4	H	98	Total	C	N	O	S	0	0	0
			787	505	134	145	3			

There are 2 discrepancies between the modelled and reference sequences:

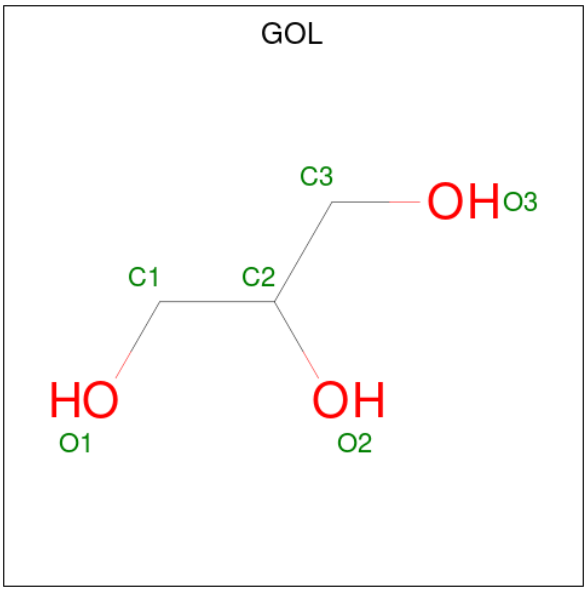
Chain	Residue	Modelled	Actual	Comment	Reference
F	0	MET	-	initiating methionine	UNP P61769
H	0	MET	-	initiating methionine	UNP P61769

- Molecule 5 is 1-(2-deoxy-alpha-D-erythro-pentofuranosyl)-5-methylpyrimidine-2,4(1H,3H)-dione (CCD ID: YC6) (formula: $C_{10}H_{12}N_2O_6$) (labeled as "Ligand of Interest" by depositor).



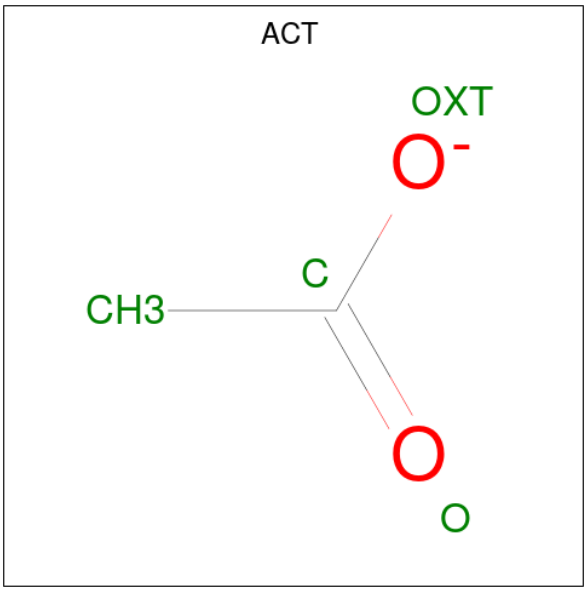
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	1
			34	20	4	10		

- Molecule 6 is GLYCEROL (CCD ID: GOL) (formula: $C_3H_8O_3$).



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	B	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	F	1	Total C O 6 3 3	0	0
6	F	1	Total C O 6 3 3	0	0

- Molecule 7 is ACETATE ION (CCD ID: ACT) (formula: $C_2H_3O_2$).

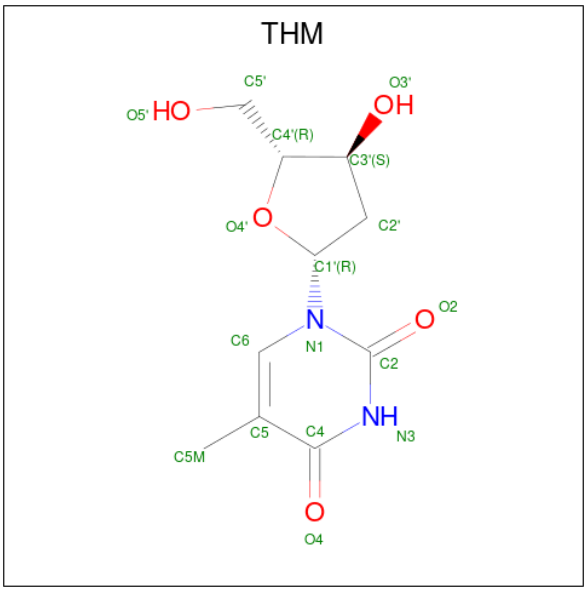


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C O 4 2 2	0	0
7	A	1	Total C O 4 2 2	0	0
7	B	1	Total C O 4 2 2	0	0
7	C	1	Total C O 4 2 2	0	0
7	E	1	Total C O 4 2 2	0	0

- Molecule 8 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

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

- Molecule 9 is THYMIDINE (CCD ID: THM) (formula: $C_{10}H_{14}N_2O_5$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	C	1	Total C N O 17 10 2 5	0	0

- Molecule 10 is SODIUM ION (CCD ID: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	E	1	Total Na 1 1	0	0
10	G	1	Total Na 1 1	0	0

- Molecule 11 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	A	166	Total O 166 166	0	0
11	B	150	Total O 150 150	0	0
11	C	179	Total O 179 179	0	0
11	D	79	Total O 79 79	0	0
11	E	82	Total O 82 82	0	0
11	F	66	Total O 66 66	0	0

Continued on next page...

Continued from previous page...

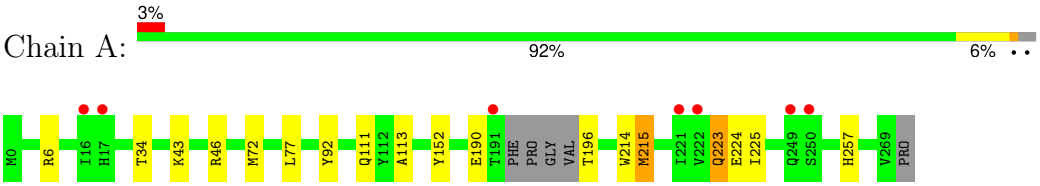
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	G	190	Total O 190 190	0	0
11	H	54	Total O 54 54	0	0

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.

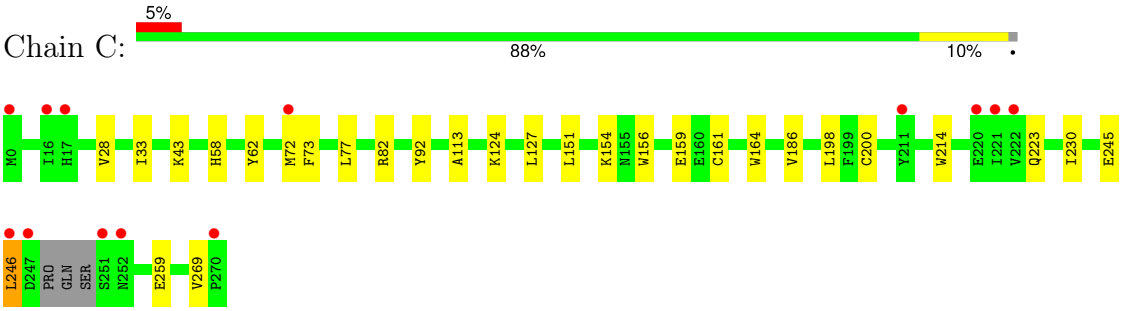
• Molecule 1: Major histocompatibility complex class I-related gene protein

Chain A:



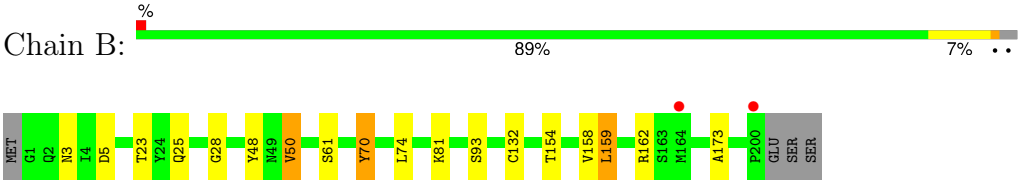
• Molecule 1: Major histocompatibility complex class I-related gene protein

Chain C:



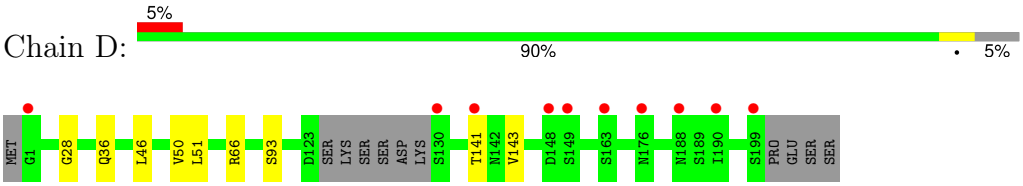
• Molecule 2: TCR alpha

Chain B:



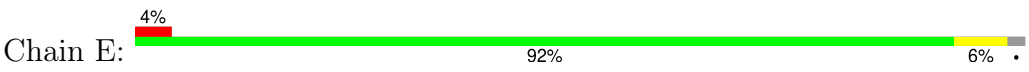
• Molecule 2: TCR alpha

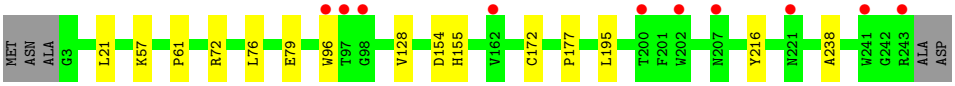
Chain D:



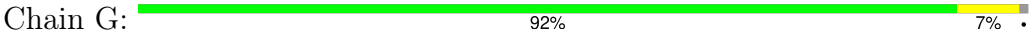
• Molecule 3: TCR beta

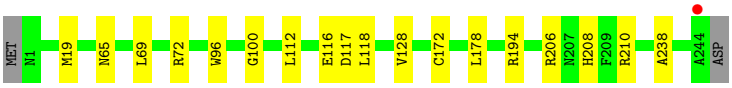
Chain E:



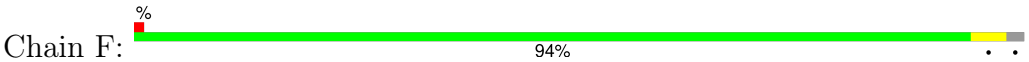


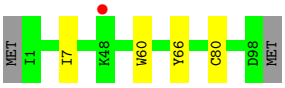
• Molecule 3: TCR beta





• Molecule 4: Beta-2-microglobulin





• Molecule 4: Beta-2-microglobulin

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	217.59Å 70.53Å 144.23Å 90.00° 104.68° 90.00°	Depositor
Resolution (Å)	49.74 – 2.15 49.74 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.8 (49.74-2.15) 99.8 (49.74-2.15)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.90 (at 2.14Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.178 , 0.206 0.179 , 0.207	Depositor DCC
R_{free} test set	5769 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	40.4	Xtriage
Anisotropy	0.390	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 43.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	13963	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.74% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CL, THM, YC6, GOL, ACT, NA

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	0.09	0/2279	0.26	0/3097
1	C	0.09	0/2314	0.28	0/3147
2	B	0.09	0/1635	0.32	0/2217
2	D	0.09	0/1493	0.28	0/2032
3	E	0.08	0/1943	0.28	0/2649
3	G	0.10	0/2054	0.30	0/2793
4	F	0.09	0/825	0.29	0/1124
4	H	0.08	0/810	0.26	0/1103
All	All	0.09	0/13353	0.29	0/18162

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

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	2196	0	2079	15	0
1	C	2222	0	2093	19	0
2	B	1576	0	1498	12	0
2	D	1454	0	1337	6	0
3	E	1877	0	1738	9	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	G	1973	0	1879	15	0
4	F	796	0	743	3	0
4	H	787	0	730	2	0
5	A	34	0	0	0	0
6	A	12	0	16	0	0
6	B	6	0	8	1	0
6	C	12	0	16	0	0
6	F	12	0	16	1	0
7	A	8	0	6	0	0
7	B	4	0	3	0	0
7	C	4	0	3	0	0
7	E	4	0	3	0	0
8	A	1	0	0	1	0
9	C	17	0	12	2	0
10	E	1	0	0	0	0
10	G	1	0	0	0	0
11	A	166	0	0	4	0
11	B	150	0	0	2	0
11	C	179	0	0	3	0
11	D	79	0	0	1	0
11	E	82	0	0	1	0
11	F	66	0	0	0	0
11	G	190	0	0	2	0
11	H	54	0	0	0	0
All	All	13963	0	12180	69	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 3.

All (69) 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)
3:G:65[A]:ASN:ND2	11:G:402:HOH:O	2.29	0.65
1:A:77:LEU:HD13	1:A:92:TYR:HB2	1.80	0.63
2:B:25:GLN:HA	6:B:802:GOL:H31	1.79	0.63
3:G:19[B]:MET:HE1	3:G:112:LEU:HD11	1.83	0.60
2:B:132[B]:CYS:SG	2:B:173:ALA:HB3	2.41	0.59
2:B:23:THR:HG22	2:B:70:TYR:HB2	1.84	0.59
2:B:162:ARG:NH1	11:B:904:HOH:O	2.36	0.59
3:E:72:ARG:NH1	11:E:404:HOH:O	2.35	0.58
1:C:77:LEU:HD13	1:C:92:TYR:HB2	1.87	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:190:GLU:HA	1:A:196:THR:HA	1.87	0.56
1:C:58:HIS:ND1	11:C:404:HOH:O	2.33	0.55
2:B:3:ASN:ND2	2:B:5:ASP:OD1	2.41	0.54
2:B:159:LEU:HB2	3:G:172[A]:CYS:HB2	1.89	0.53
3:G:128:VAL:HG23	3:G:238:ALA:HB3	1.90	0.53
4:F:7[B]:ILE:HD11	4:F:80:CYS:HB3	1.91	0.53
1:C:230:ILE:HD11	3:G:206[A]:ARG:HD2	1.91	0.53
1:C:198:LEU:HD13	1:C:269:VAL:HG21	1.91	0.52
3:G:210[A]:ARG:NH1	11:G:406:HOH:O	2.40	0.52
1:A:215:MET:HG3	1:A:257:HIS:CD2	2.44	0.52
2:B:28:GLY:HA3	2:B:93:SER:OG	2.11	0.51
1:A:34:THR:HB	1:A:43:LYS:HE2	1.92	0.51
1:A:113:ALA:HB2	4:H:60:TRP:CE2	2.46	0.50
1:A:152:TYR:CD1	3:G:100:GLY:HA3	2.47	0.50
1:A:214:TRP:H	1:A:225:ILE:HD13	1.77	0.50
1:C:186:VAL:HG11	1:C:269:VAL:HG22	1.94	0.50
1:C:72:MET:HG2	3:E:96:TRP:CE3	2.48	0.49
2:D:28:GLY:HA3	2:D:93:SER:OG	2.13	0.48
1:C:28:VAL:HG23	1:C:33:ILE:HD13	1.95	0.48
3:E:128:VAL:HG23	3:E:238:ALA:HB3	1.95	0.48
3:G:117[A]:ASP:OD1	3:G:118:LEU:N	2.46	0.48
2:B:81:LYS:NZ	11:B:906:HOH:O	2.45	0.47
3:E:155:HIS:HB3	3:E:216:TYR:HB2	1.97	0.47
3:E:21:LEU:HD12	3:E:76:LEU:HD23	1.95	0.47
1:A:6[B]:ARG:NH2	11:A:410:HOH:O	2.47	0.46
2:B:48:TYR:CE2	2:B:50:VAL:HG22	2.50	0.46
1:C:159:GLU:HG3	2:D:66[B]:ARG:HH21	1.80	0.46
1:A:111:GLN:NE2	11:A:411:HOH:O	2.49	0.46
2:B:159:LEU:HB2	3:G:172[B]:CYS:HB3	1.97	0.46
1:C:124:LYS:NZ	11:C:402:HOH:O	2.34	0.46
3:G:19[B]:MET:HB3	3:G:19[B]:MET:HE2	1.62	0.45
1:A:46:ARG:HD3	1:A:46:ARG:HA	1.87	0.45
2:B:154:THR:HG22	3:G:178[A]:LEU:HD23	1.99	0.45
1:C:72:MET:HG2	3:E:96:TRP:CZ3	2.52	0.45
3:E:57:LYS:HE3	3:E:61:PRO:HG3	1.98	0.45
4:F:66:TYR:OH	6:F:102:GOL:O2	2.32	0.44
1:C:113:ALA:HB2	4:F:60:TRP:CE2	2.52	0.44
3:G:69:LEU:H	3:G:69:LEU:HD22	1.82	0.44
3:E:154:ASP:CG	3:E:177:PRO:HG3	2.42	0.44
1:A:111:GLN:NE2	11:A:412:HOH:O	2.50	0.44
1:C:151:LEU:HD22	2:D:51:LEU:HD12	2.00	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:43:LYS:HE3	11:A:469:HOH:O	2.18	0.43
1:C:62:TYR:CE1	9:C:301:THM:H3'	2.53	0.43
1:A:72:MET:HE2	3:G:96:TRP:CG	2.53	0.42
3:G:208:HIS:HE1	3:G:210[A]:ARG:NH2	2.17	0.42
1:A:34:THR:HG21	8:A:306:CL:CL	2.57	0.42
3:E:79[B]:GLU:H	3:E:79[B]:GLU:HG3	1.58	0.42
1:C:164:TRP:CE2	9:C:301:THM:H5'2	2.55	0.42
2:D:50:VAL:O	2:D:66[B]:ARG:HD2	2.20	0.42
1:C:200:CYS:HB2	1:C:214:TRP:CZ2	2.55	0.41
1:A:223:GLN:H	1:A:223:GLN:HG2	1.61	0.41
2:D:66[A]:ARG:NH1	11:D:301:HOH:O	2.29	0.41
3:G:19[B]:MET:SD	3:G:112:LEU:HD21	2.60	0.41
2:B:61[B]:SER:OG	2:B:74:LEU:HB3	2.20	0.41
1:C:246:LEU:HA	1:C:246:LEU:HD12	1.82	0.41
2:D:36:GLN:HB2	2:D:46:LEU:HD11	2.03	0.41
1:C:127:LEU:HD21	1:C:154:LYS:HD2	2.02	0.40
1:C:156:TRP:CZ3	1:C:161:CYS:HB2	2.57	0.40
4:H:16:GLU:HB3	4:H:19:LYS:HD2	2.03	0.40
1:C:82:ARG:HD2	11:C:437:HOH:O	2.22	0.40

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	268/271 (99%)	265 (99%)	3 (1%)	0	100	100
1	C	273/271 (101%)	267 (98%)	6 (2%)	0	100	100
2	B	206/204 (101%)	204 (99%)	2 (1%)	0	100	100
2	D	191/204 (94%)	185 (97%)	6 (3%)	0	100	100
3	E	245/246 (100%)	239 (98%)	6 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	G	254/246 (103%)	252 (99%)	2 (1%)	0	100	100
4	F	98/100 (98%)	98 (100%)	0	0	100	100
4	H	96/100 (96%)	94 (98%)	2 (2%)	0	100	100
All	All	1631/1642 (99%)	1604 (98%)	27 (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	231/241 (96%)	228 (99%)	3 (1%)	65	71
1	C	232/241 (96%)	226 (97%)	6 (3%)	41	43
2	B	176/181 (97%)	172 (98%)	4 (2%)	45	49
2	D	147/181 (81%)	145 (99%)	2 (1%)	62	68
3	E	197/212 (93%)	195 (99%)	2 (1%)	73	78
3	G	216/212 (102%)	213 (99%)	3 (1%)	62	68
4	F	86/95 (90%)	86 (100%)	0	100	100
4	H	84/95 (88%)	83 (99%)	1 (1%)	67	73
All	All	1369/1458 (94%)	1348 (98%)	21 (2%)	58	66

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	215	MET
1	A	223	GLN
1	A	224	GLU
2	B	50	VAL
2	B	70	TYR
2	B	158	VAL
2	B	159	LEU
1	C	43	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	73	PHE
1	C	223	GLN
1	C	245	GLU
1	C	246	LEU
1	C	259	GLU
2	D	141	THR
2	D	143	VAL
3	E	172	CYS
3	E	195	LEU
3	G	72	ARG
3	G	116	GLU
3	G	194	ARG
4	H	36	GLU

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	141	GLN
1	A	268	GLN
2	B	19	GLN
2	B	21	ASN
2	B	169	ASN
1	C	42	GLN
1	C	70	GLN
1	C	83	HIS
1	C	93	GLN
2	D	19	GLN
2	D	21	ASN
2	D	36	GLN
2	D	38	HIS
2	D	120	GLN
3	E	204	ASN
3	E	214	GLN
4	F	13	HIS
4	H	13	HIS
4	H	31	HIS

5.3.3 RNA

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 18 ligands modelled in this entry, 3 are monoatomic - leaving 15 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
5	YC6	A	301[B]	1	18,18,19	0.55	0	26,26,27	0.36	0
7	ACT	B	801	-	3,3,3	1.39	1 (33%)	3,3,3	1.39	0
7	ACT	C	303	-	3,3,3	1.37	0	3,3,3	1.33	0
6	GOL	F	101	-	5,5,5	0.92	0	5,5,5	1.12	0
6	GOL	A	305	-	5,5,5	0.90	0	5,5,5	1.09	0
6	GOL	B	802	-	5,5,5	0.89	0	5,5,5	1.09	0
6	GOL	C	304	-	5,5,5	0.90	0	5,5,5	1.07	0
7	ACT	E	301	-	3,3,3	1.39	0	3,3,3	1.36	0
6	GOL	A	302	-	5,5,5	0.95	0	5,5,5	1.03	0
5	YC6	A	301[A]	1	18,18,19	0.55	0	26,26,27	0.49	0
7	ACT	A	303	-	3,3,3	1.36	0	3,3,3	1.38	0
6	GOL	F	102	-	5,5,5	0.93	0	5,5,5	1.07	0
7	ACT	A	304	-	3,3,3	1.39	0	3,3,3	1.34	0
9	THM	C	301	1	18,18,18	0.53	0	26,26,26	0.38	0
6	GOL	C	302	-	5,5,5	0.94	0	5,5,5	1.08	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	YC6	A	301[B]	1	1/1/3/4	0/6/18/20	0/2/2/2
6	GOL	F	101	-	-	2/4/4/4	-
6	GOL	A	305	-	-	0/4/4/4	-
6	GOL	B	802	-	-	2/4/4/4	-
6	GOL	C	304	-	-	0/4/4/4	-
6	GOL	A	302	-	-	1/4/4/4	-
5	YC6	A	301[A]	1	-	0/6/18/20	0/2/2/2
6	GOL	F	102	-	-	0/4/4/4	-
9	THM	C	301	1	-	0/6/18/18	0/2/2/2
6	GOL	C	302	-	-	0/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	B	801	ACT	CH3-C	2.00	1.57	1.49

There are no bond angle outliers.

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	A	301[B]	YC6	C3

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	802	GOL	C1-C2-C3-O3
6	F	101	GOL	O1-C1-C2-O2
6	F	101	GOL	O1-C1-C2-C3
6	B	802	GOL	O2-C2-C3-O3
6	A	302	GOL	C1-C2-C3-O3

There are no ring outliers.

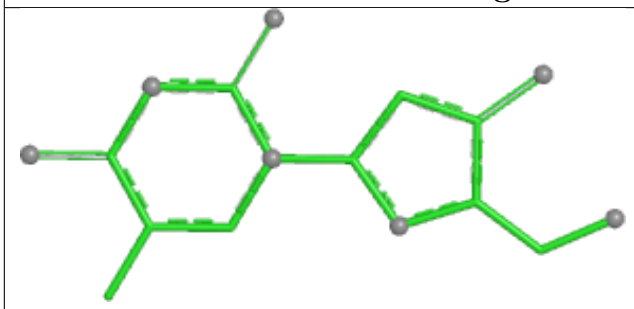
3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	802	GOL	1	0
6	F	102	GOL	1	0
9	C	301	THM	2	0

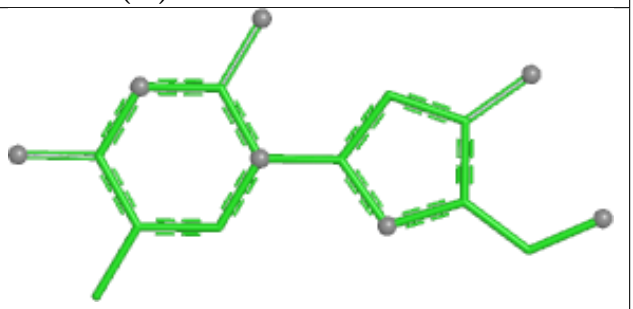
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In

addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

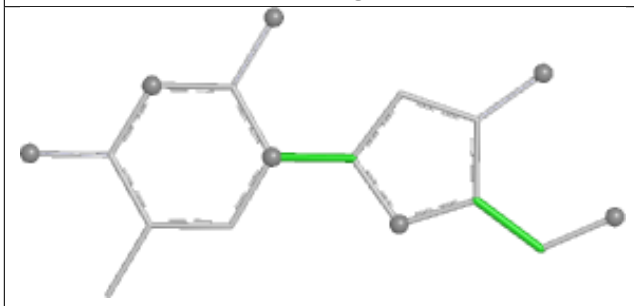
Ligand YC6 A 301 (B)



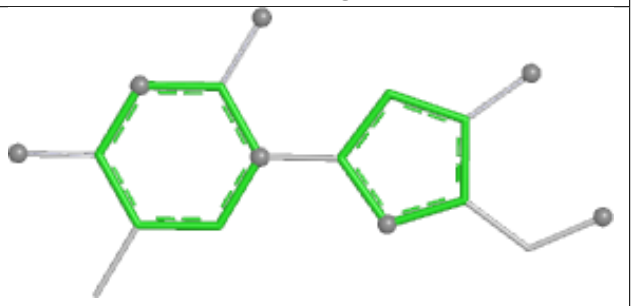
Bond lengths



Bond angles

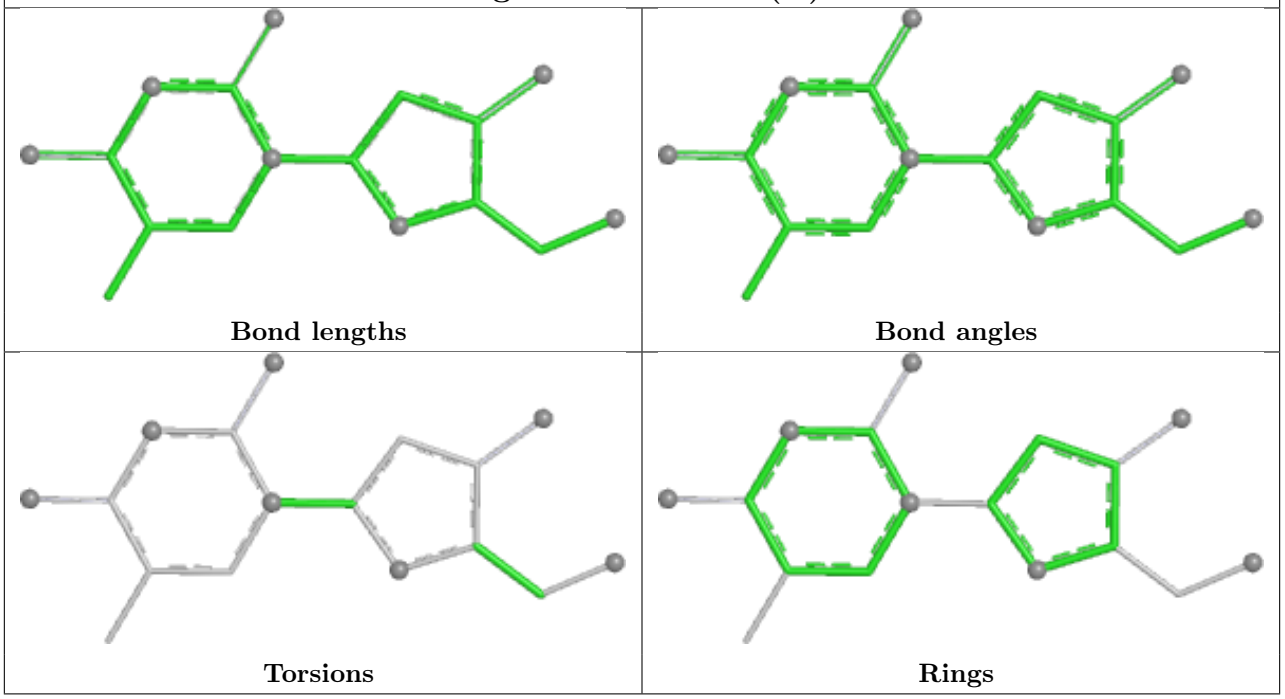


Torsions

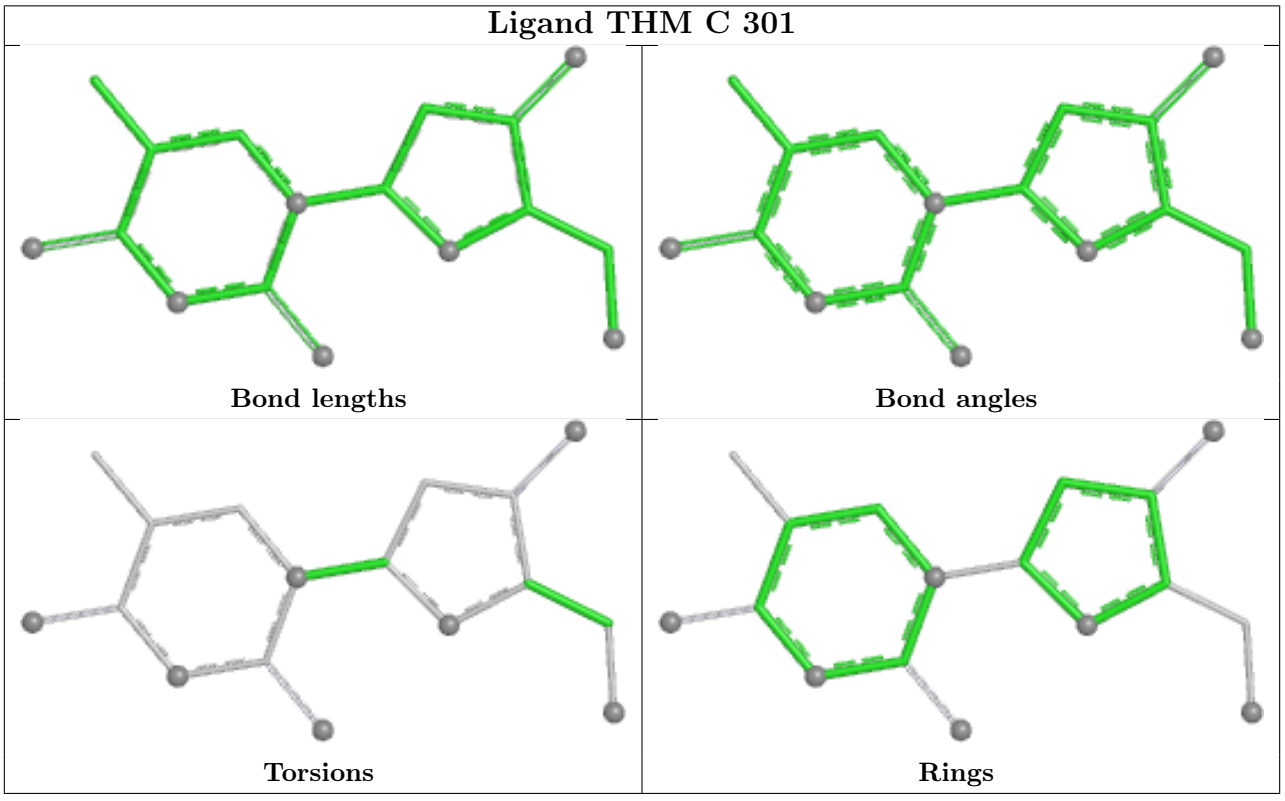


Rings

Ligand YC6 A 301 (A)



Ligand THM C 301



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^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	266/271 (98%)	-0.12	7 (2%)	57	62	19, 41, 83, 105	18 (6%)
1	C	268/271 (98%)	-0.12	13 (4%)	36	41	20, 39, 67, 90	27 (10%)
2	B	200/204 (98%)	-0.18	2 (1%)	79	82	20, 39, 64, 78	15 (7%)
2	D	193/204 (94%)	0.39	10 (5%)	34	39	25, 56, 99, 110	8 (4%)
3	E	241/246 (97%)	0.30	10 (4%)	42	47	25, 54, 91, 105	12 (4%)
3	G	244/246 (99%)	-0.26	1 (0%)	89	90	22, 39, 59, 92	21 (8%)
4	F	98/100 (98%)	-0.18	1 (1%)	79	82	21, 44, 69, 84	3 (3%)
4	H	98/100 (98%)	0.31	1 (1%)	79	82	35, 62, 88, 98	2 (2%)
All	All	1608/1642 (97%)	-0.00	45 (2%)	55	60	19, 44, 86, 110	106 (6%)

All (45) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	199	SER	5.3
1	A	191	THR	4.6
2	D	190	ILE	3.2
2	D	149	SER	3.2
3	E	243	ARG	3.1
3	E	96	TRP	3.1
2	D	176	ASN	3.1
1	A	221	ILE	3.0
3	E	207	ASN	3.0
1	A	222	VAL	3.0
1	C	16	ILE	2.8
2	D	148	ASP	2.8
1	C	251	SER	2.8
1	C	252	ASN	2.8
1	C	72	MET	2.8
2	B	200	PRO	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	247	ASP	2.7
3	E	241	TRP	2.7
4	H	0	MET	2.6
1	A	249	GLN	2.6
1	C	220	GLU	2.6
1	C	0	MET	2.5
2	D	130	SER	2.4
1	C	221	ILE	2.4
1	C	222	VAL	2.4
3	G	244	ALA	2.4
3	E	98	GLY	2.4
3	E	97	THR	2.3
3	E	221	ASN	2.3
1	C	246	LEU	2.3
2	D	141	THR	2.2
3	E	200	THR	2.2
1	C	270	PRO	2.1
1	A	17	HIS	2.1
1	C	17	HIS	2.1
2	B	164[A]	MET	2.1
3	E	162	VAL	2.1
1	A	16	ILE	2.1
1	A	250	SER	2.1
4	F	48	LYS	2.1
2	D	163	SER	2.1
2	D	188	ASN	2.0
2	D	1	GLY	2.0
1	C	211[A]	TYR	2.0
3	E	202	TRP	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)

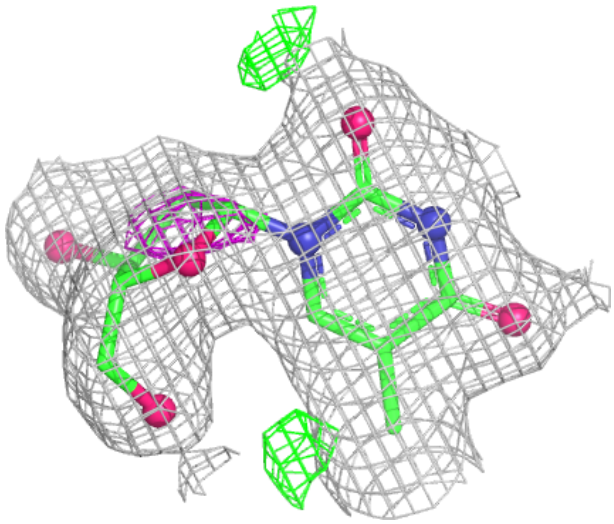
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

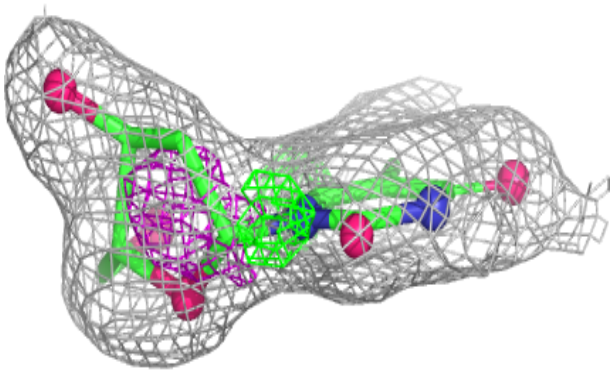
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	GOL	C	302	6/6	0.70	0.15	64,75,84,85	0
7	ACT	C	303	4/4	0.74	0.23	45,49,53,55	4
7	ACT	E	301	4/4	0.76	0.19	58,64,65,72	0
6	GOL	A	302	6/6	0.80	0.18	56,66,70,78	0
6	GOL	C	304	6/6	0.82	0.16	56,57,63,64	0
6	GOL	F	102	6/6	0.83	0.14	63,68,73,76	0
7	ACT	A	304	4/4	0.83	0.20	42,42,51,56	0
5	YC6	A	301[B]	17/18	0.85	0.13	35,40,44,45	17
5	YC6	A	301[A]	17/18	0.85	0.13	35,40,43,43	17
6	GOL	B	802	6/6	0.86	0.12	47,58,67,73	0
7	ACT	B	801	4/4	0.89	0.16	50,53,56,56	0
7	ACT	A	303	4/4	0.89	0.17	54,57,60,68	0
6	GOL	A	305	6/6	0.89	0.12	51,54,56,58	0
6	GOL	F	101	6/6	0.93	0.09	38,42,46,51	0
9	THM	C	301	17/17	0.94	0.08	35,40,44,44	11
8	CL	A	306	1/1	0.95	0.08	53,53,53,53	1
10	NA	E	302	1/1	0.96	0.05	54,54,54,54	0
10	NA	G	301	1/1	0.99	0.03	43,43,43,43	0

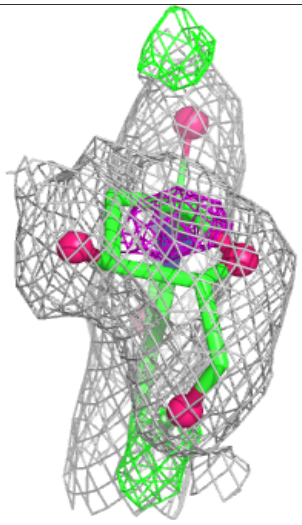
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around YC6 A 301 (B):

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

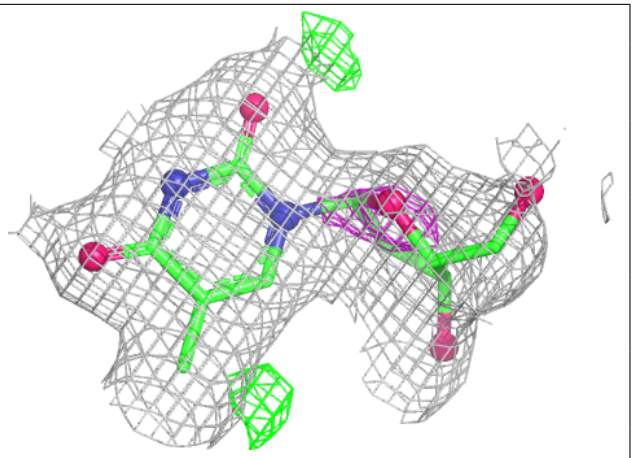


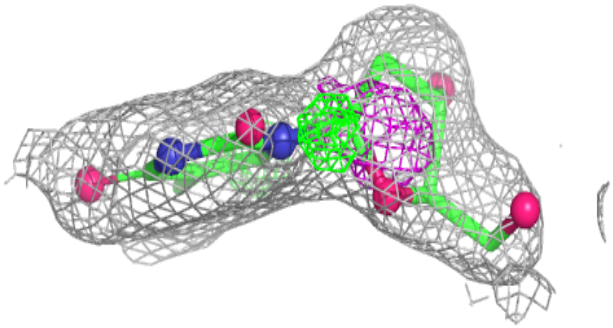


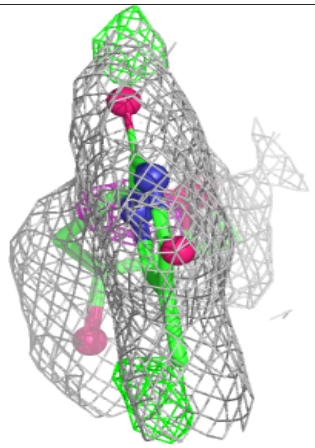


Electron density around YC6 A 301 (A):

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

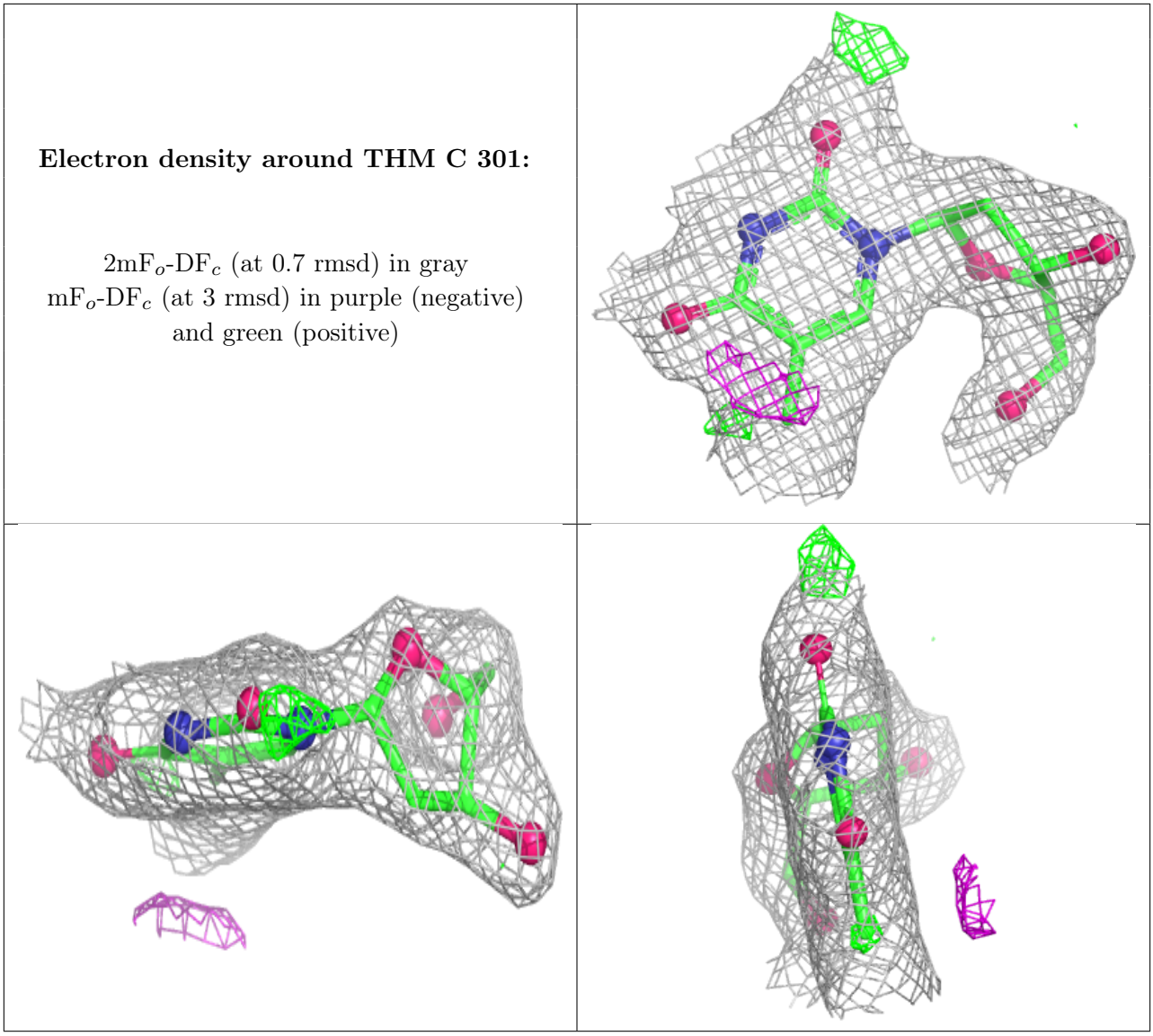






Electron density around THM C 301:

$2mF_o$-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.