



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 19, 2021 – 11:47 am BST

PDB ID : 7OKP  
Title : Crystal structure of mouse CARM1 in complex with histone H3\_13-22 K18 acetylated  
Authors : Marechal, N.; Cura, V.; Troffer-Charlier, N.; Bonnefond, L.; Cavarelli, J.  
Deposited on : 2021-05-18  
Resolution : 2.20 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 : 2.23.2  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0267  
CCP4 : 7.1.010 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.23.2

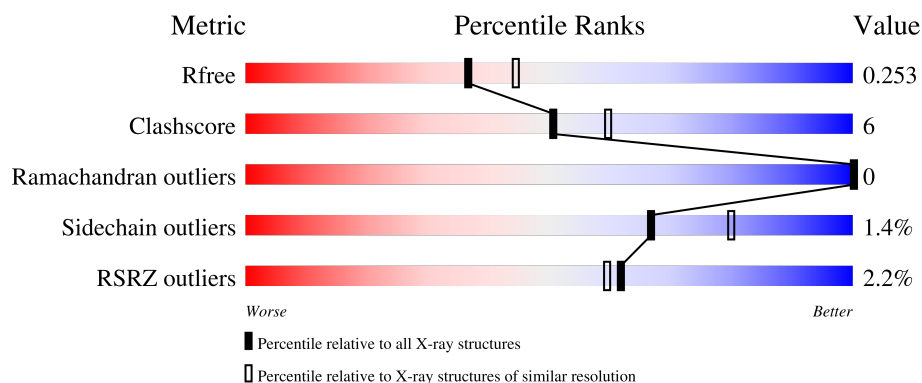
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	371	<div> <div>3%</div> <div>85%</div> <div>13%</div> <div>.</div> </div>
1	B	371	<div> <div>%</div> <div>78%</div> <div>15%</div> <div>7%</div> </div>
1	C	371	<div> <div>2%</div> <div>86%</div> <div>12%</div> <div>.</div> </div>
1	D	371	<div> <div>2%</div> <div>83%</div> <div>9%</div> <div>7%</div> </div>
2	E	11	<div> <div>18%</div> <div>100%</div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	11	 91% 9%
2	G	11	 9% 91% 9%
2	H	11	 91% 9%

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
3	MLI	B	501	-	-	X	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 23683 atoms, of which 11502 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 Histone-arginine methyltransferase CARM1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	363	Total	C	H	N	O	S	0	3	0
			5728	1864	2831	480	538	15			
1	B	344	Total	C	H	N	O	S	0	3	0
			5503	1792	2723	462	512	14			
1	C	362	Total	C	H	N	O	S	0	4	0
			5705	1858	2819	476	537	15			
1	D	344	Total	C	H	N	O	S	0	2	0
			5473	1782	2711	458	508	14			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	127	GLY	-	expression tag	UNP Q9WVG6
A	128	HIS	-	expression tag	UNP Q9WVG6
A	129	MET	-	expression tag	UNP Q9WVG6
B	127	GLY	-	expression tag	UNP Q9WVG6
B	128	HIS	-	expression tag	UNP Q9WVG6
B	129	MET	-	expression tag	UNP Q9WVG6
C	127	GLY	-	expression tag	UNP Q9WVG6
C	128	HIS	-	expression tag	UNP Q9WVG6
C	129	MET	-	expression tag	UNP Q9WVG6
D	127	GLY	-	expression tag	UNP Q9WVG6
D	128	HIS	-	expression tag	UNP Q9WVG6
D	129	MET	-	expression tag	UNP Q9WVG6

- Molecule 2 is a protein called Histone H3.3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	11	Total	C	H	N	O	0	0	0
			168	50	88	16	14			
2	F	11	Total	C	H	N	O	0	0	0
			168	50	88	16	14			

*Continued on next page...*

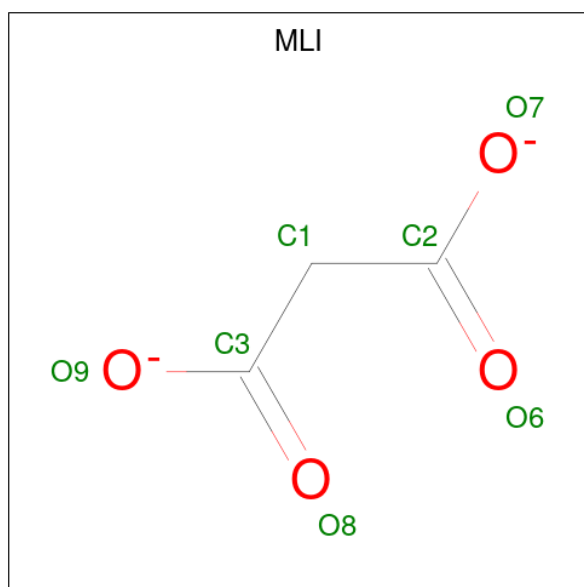
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	G	11	Total	C	H	N	O	0	0	0
			168	50	88	16	14			
2	H	11	Total	C	H	N	O	0	0	0
			168	50	88	16	14			

There are 4 discrepancies between the modelled and reference sequences:

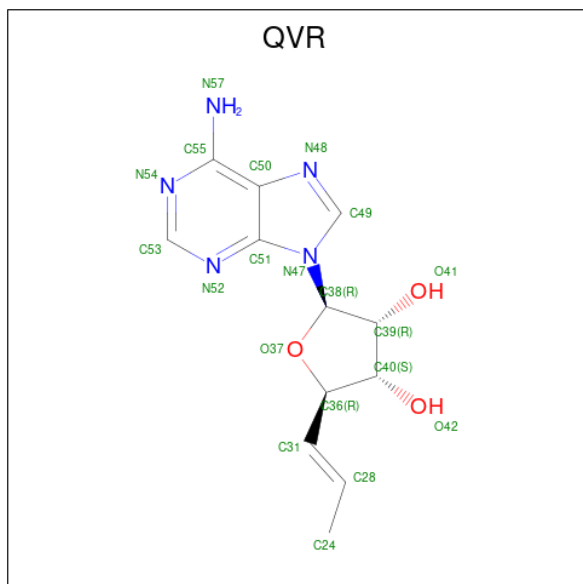
Chain	Residue	Modelled	Actual	Comment	Reference
E	12	ACE	-	acetylation	UNP P84243
F	12	ACE	-	acetylation	UNP P84243
G	12	ACE	-	acetylation	UNP P84243
H	12	ACE	-	acetylation	UNP P84243

- Molecule 3 is MALONATE ION (three-letter code: MLI) (formula:  $C_3H_2O_4$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	B	1	Total	C	H	O	0	0
			9	3	2	4		
3	C	1	Total	C	H	O	0	0
			9	3	2	4		
3	C	1	Total	C	H	O	0	0
			9	3	2	4		
3	C	1	Total	C	H	O	0	0
			9	3	2	4		
3	D	1	Total	C	H	O	0	0
			9	3	2	4		

- Molecule 4 is (2 {R},3 {R},4 {S},5 {R})-2-(6-aminopurin-9-yl)-5-[( {E})-prop-1-enyl]oxolane-3,4-diol (three-letter code: QVR) (formula: C<sub>12</sub>H<sub>15</sub>N<sub>5</sub>O<sub>3</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	E	1	Total	C	H	N	O	0	0
			34	12	14	5	3		
4	F	1	Total	C	H	N	O	0	0
			34	12	14	5	3		
4	G	1	Total	C	H	N	O	0	0
			34	12	14	5	3		
4	H	1	Total	C	H	N	O	0	0
			34	12	14	5	3		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	126	Total	O	0	0
			126	126		
5	B	99	Total	O	0	0
			99	99		
5	C	101	Total	O	0	0
			101	101		
5	D	79	Total	O	0	0
			79	79		
5	E	4	Total	O	0	0
			4	4		
5	F	5	Total	O	0	0
			5	5		

Continued on next page...

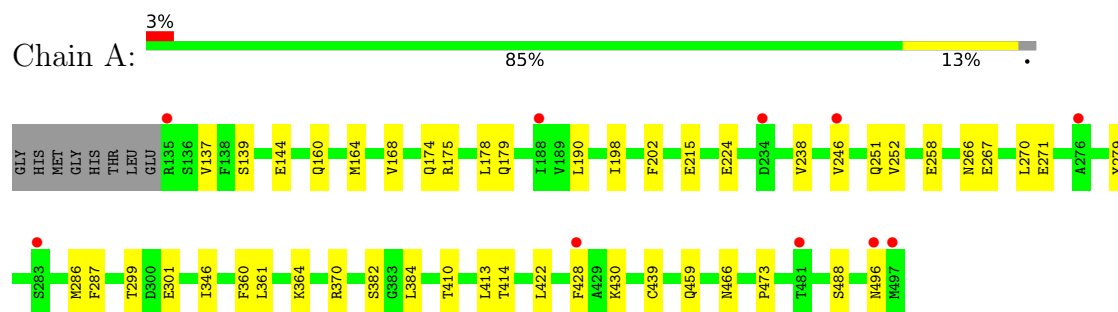
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	G	2	Total	O	0	0
			2	2		
5	H	5	Total	O	0	0
			5	5		

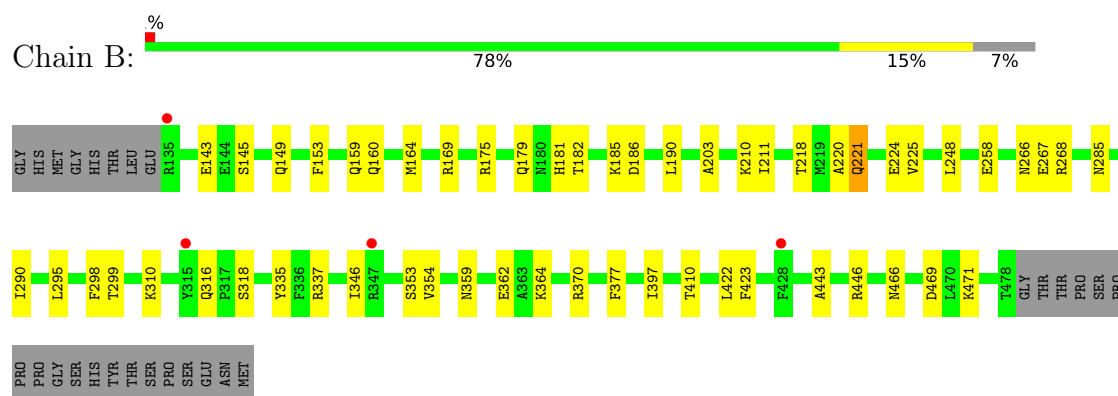
### 3 Residue-property plots

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

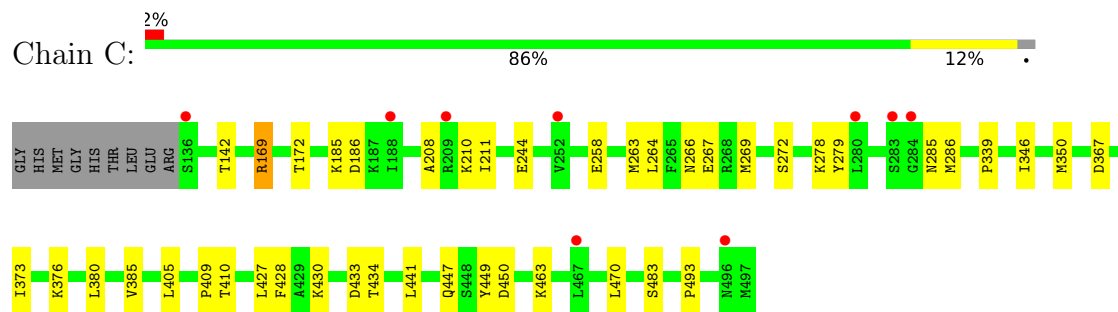
- Molecule 1: Histone-arginine methyltransferase CARM1



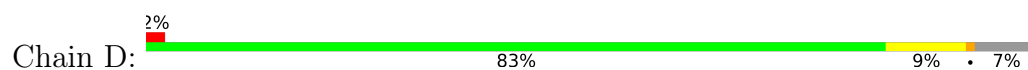
- Molecule 1: Histone-arginine methyltransferase CARM1



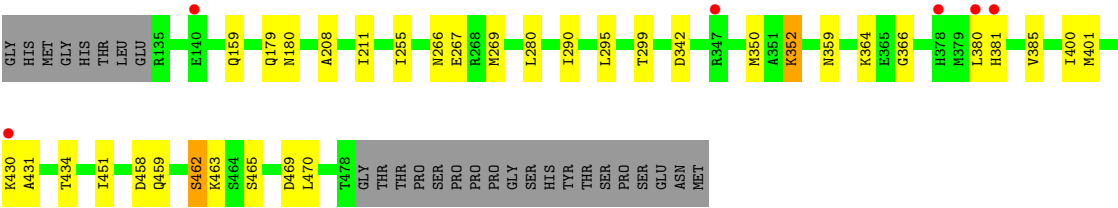
- Molecule 1: Histone-arginine methyltransferase CARM1



- Molecule 1: Histone-arginine methyltransferase CARM1



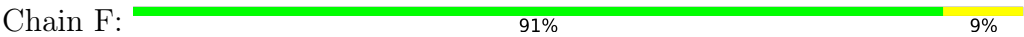




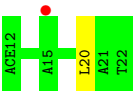
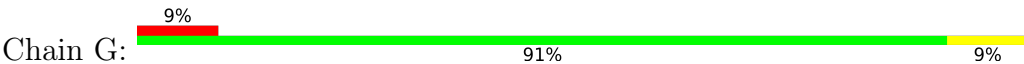
● Molecule 2: Histone H3.3



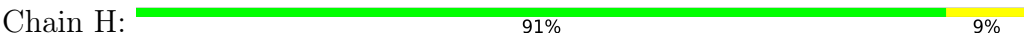
● Molecule 2: Histone H3.3



● Molecule 2: Histone H3.3



● Molecule 2: Histone H3.3



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.91Å 99.58Å 208.07Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.32 – 2.20 45.32 – 2.20	Depositor EDS
% Data completeness (in resolution range)	96.3 (45.32-2.20) 96.5 (45.32-2.20)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.12 (at 2.20Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, $R_{free}$	0.211 , 0.254 0.211 , 0.253	Depositor DCC
$R_{free}$ test set	3826 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	34.6	Xtriage
Anisotropy	0.689	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	23683	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 38.78 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 3.4997e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: ACE, QVR, ALY, MLI

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	0.28	0/2974	0.50	0/4035
1	B	0.29	0/2851	0.51	0/3863
1	C	0.28	0/2968	0.49	0/4028
1	D	0.27	0/2832	0.49	0/3837
2	E	0.35	0/65	0.49	0/85
2	F	0.35	0/65	0.55	0/85
2	G	0.38	0/65	0.53	0/85
2	H	0.35	0/65	0.57	0/85
All	All	0.28	0/11885	0.50	0/16103

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	2897	2831	2819	34	1
1	B	2780	2723	2710	39	1
1	C	2886	2819	2803	32	0
1	D	2762	2711	2696	24	0
2	E	80	88	88	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	F	80	88	88	1	0
2	G	80	88	88	1	0
2	H	80	88	88	1	0
3	B	7	2	2	3	0
3	C	21	6	6	0	0
3	D	7	2	2	0	0
4	E	20	14	0	1	0
4	F	20	14	0	0	0
4	G	20	14	0	1	0
4	H	20	14	0	2	0
5	A	126	0	0	12	1
5	B	99	0	0	12	1
5	C	101	0	0	7	0
5	D	79	0	0	3	0
5	E	4	0	0	0	0
5	F	5	0	0	0	0
5	G	2	0	0	1	0
5	H	5	0	0	2	0
All	All	12181	11502	11390	133	2

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 6.

All (133) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:269:MET:HE1	4:G:101:QVR:O37	1.72	0.88
1:A:488:SER:O	5:A:501:HOH:O	1.98	0.81
1:A:258:GLU:OE1	5:A:502:HOH:O	2.02	0.78
1:D:350:MET:HE2	1:D:385:VAL:HG22	1.65	0.78
1:D:159:GLN:OE1	5:D:601:HOH:O	2.03	0.76
1:A:246:VAL:O	5:A:503:HOH:O	2.04	0.76
1:A:301:GLU:OE2	5:A:504:HOH:O	2.04	0.75
2:G:20:LEU:O	5:G:201:HOH:O	2.05	0.75
1:B:179:GLN:O	5:B:603:HOH:O	2.04	0.74
1:C:450:ASP:OD2	5:C:602:HOH:O	2.06	0.73
1:D:342:ASP:OD1	5:D:602:HOH:O	2.05	0.73
1:A:267:GLU:OE2	5:A:505:HOH:O	2.05	0.73
4:H:101:QVR:O42	5:H:201:HOH:O	2.08	0.72
1:C:483:SER:OG	5:C:603:HOH:O	2.06	0.71
1:A:271:GLU:OE1	5:A:506:HOH:O	2.09	0.71

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:397:ILE:O	5:B:604:HOH:O	2.09	0.71
1:B:318:SER:O	1:D:462:SER:OG	2.06	0.69
1:A:384:LEU:HD21	1:A:428:PHE:CE2	2.28	0.69
1:A:422:LEU:O	1:A:466:ASN:ND2	2.22	0.68
1:D:208:ALA:HB3	1:D:211:ILE:HD11	1.75	0.68
1:A:370:ARG:NH1	5:A:508:HOH:O	2.16	0.68
1:C:449:TYR:O	5:C:604:HOH:O	2.11	0.68
1:B:335:TYR:O	5:B:605:HOH:O	2.11	0.67
1:C:267:GLU:O	5:C:605:HOH:O	2.15	0.65
1:B:362:GLU:OE1	5:B:606:HOH:O	2.14	0.64
1:A:160:GLN:O	1:A:164:MET:HG3	1.98	0.64
1:A:473:PRO:O	5:A:507:HOH:O	2.15	0.64
1:B:181:HIS:CD2	1:B:182:THR:HG23	2.34	0.62
1:C:169:ARG:NE	1:C:258:GLU:OE2	2.33	0.62
1:A:346:ILE:CD1	1:A:410:THR:HG22	2.29	0.61
1:C:376:LYS:NZ	1:C:434:THR:OG1	2.33	0.61
1:D:180:ASN:HA	1:D:400:ILE:HD12	1.81	0.61
1:C:185:LYS:HE3	1:C:186:ASP:OD2	2.02	0.59
1:D:350:MET:CE	1:D:385:VAL:HG22	2.31	0.59
1:C:208:ALA:HB3	1:C:211:ILE:HD11	1.85	0.58
1:B:290:ILE:HG22	1:B:359:ASN:HA	1.86	0.57
1:A:224:GLU:HG2	1:A:238:VAL:HB	1.86	0.57
5:A:525:HOH:O	3:B:501:MLI:H12	2.05	0.57
1:C:285:ASN:OD1	1:C:286:MET:N	2.38	0.57
1:B:143:GLU:OE1	1:B:446:ARG:CD	2.54	0.56
1:C:373:ILE:HG22	1:C:373:ILE:O	2.05	0.55
1:A:413:LEU:HD12	1:A:414:THR:N	2.23	0.54
1:C:376:LYS:HZ2	1:C:434:THR:CB	2.20	0.54
1:C:376:LYS:NZ	1:C:434:THR:CB	2.71	0.54
3:B:501:MLI:O7	5:B:607:HOH:O	2.19	0.54
1:C:428:PHE:O	5:C:606:HOH:O	2.18	0.52
1:A:168:VAL:HG11	5:A:556:HOH:O	2.10	0.51
1:A:346:ILE:HD11	1:A:410:THR:HG22	1.92	0.51
1:C:463:LYS:NZ	5:C:601:HOH:O	2.04	0.51
1:B:220:ALA:O	1:B:224:GLU:HG3	2.11	0.51
1:B:210:LYS:NZ	5:B:602:HOH:O	2.01	0.49
1:D:290:ILE:HG22	1:D:359:ASN:HA	1.94	0.49
1:B:169:ARG:NE	1:B:258:GLU:OE2	2.36	0.49
1:A:496:ASN:HA	5:A:522:HOH:O	2.12	0.48
1:B:354:VAL:HG23	1:B:377:PHE:CE1	2.48	0.48
1:C:367:ASP:O	5:C:607:HOH:O	2.20	0.48

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:144:GLU:HA	1:A:144:GLU:OE1	2.13	0.48
1:B:143:GLU:OE1	1:B:446:ARG:HD3	2.13	0.48
1:B:285:ASN:HB2	5:B:683:HOH:O	2.14	0.48
1:A:198:ILE:HD11	1:A:202:PHE:CZ	2.49	0.48
1:B:353:SER:N	5:B:608:HOH:O	2.20	0.48
1:B:316:GLN:HB3	5:B:601:HOH:O	2.13	0.48
1:D:179:GLN:HG2	1:D:401:MET:SD	2.54	0.48
1:A:346:ILE:HD13	1:A:410:THR:HG22	1.95	0.48
1:A:346:ILE:HD11	1:A:410:THR:HA	1.95	0.47
1:C:346:ILE:HD13	1:C:410:THR:HG23	1.95	0.47
1:A:266:ASN:O	1:A:267:GLU:HB2	2.13	0.47
1:B:310:LYS:NZ	5:B:610:HOH:O	2.28	0.47
1:B:143:GLU:OE1	1:B:446:ARG:HD2	2.15	0.47
1:B:346:ILE:HG21	1:B:410:THR:HG23	1.96	0.47
1:D:352:LYS:O	1:D:352:LYS:HD2	2.15	0.47
1:A:459:GLN:OE1	1:A:459:GLN:N	2.48	0.47
1:B:469[A]:ASP:OD2	1:B:471:LYS:HE2	2.15	0.47
1:B:469[A]:ASP:OD2	1:B:471:LYS:HG2	2.14	0.46
1:D:266:ASN:O	1:D:267:GLU:HB2	2.14	0.46
1:C:172:THR:HG21	1:C:405:LEU:HB2	1.97	0.46
1:D:299:THR:HB	1:D:350:MET:SD	2.56	0.46
1:D:430:LYS:HD2	1:D:431:ALA:O	2.16	0.46
1:A:251:GLN:HE22	1:A:279:TYR:HA	1.81	0.46
1:D:380:LEU:O	1:D:381:HIS:ND1	2.48	0.46
1:B:218:THR:HG23	5:B:635:HOH:O	2.15	0.46
1:B:218:THR:O	1:B:221:GLN:HG3	2.16	0.46
1:A:174:GLN:O	1:A:178:LEU:HG	2.15	0.46
1:B:266:ASN:O	1:B:267:GLU:HB2	2.16	0.46
1:A:175:ARG:HG2	1:A:179:GLN:NE2	2.31	0.45
1:A:286:MET:HG3	1:A:361:LEU:HD23	1.98	0.45
1:B:190:LEU:HD13	1:B:248:LEU:HD21	1.99	0.45
1:C:142:THR:HG23	1:C:447:GLN:HG3	1.99	0.45
1:C:346:ILE:CD1	1:C:409:PRO:HG2	2.47	0.45
1:B:298:PHE:HZ	1:B:422:LEU:HD23	1.82	0.45
1:D:451:ILE:HD11	1:D:470:LEU:HB2	1.99	0.45
1:B:153:PHE:CZ	1:B:159:GLN:HG3	2.53	0.44
1:B:185:LYS:HE3	1:B:186:ASP:OD2	2.17	0.44
1:C:385:VAL:HG23	1:C:427:LEU:O	2.18	0.44
1:A:175:ARG:O	1:A:179:GLN:HG3	2.18	0.44
1:B:225:VAL:HG21	3:B:501:MLI:O9	2.18	0.43
1:C:339:PRO:HG2	1:C:470:LEU:HA	2.00	0.43

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:299:THR:O	1:B:299:THR:HG23	2.19	0.43
1:B:145:SER:OG	1:B:446:ARG:NH1	2.50	0.43
1:C:266:ASN:O	1:C:267:GLU:HB2	2.19	0.43
1:A:137:VAL:HG23	5:A:511:HOH:O	2.19	0.43
1:B:337:ARG:HD3	1:B:469[B]:ASP:OD2	2.18	0.43
1:C:376:LYS:HZ2	1:C:434:THR:HB	1.83	0.42
1:C:278:LYS:HE3	1:C:279:TYR:CZ	2.55	0.42
1:A:270:LEU:HD22	1:A:360:PHE:CE1	2.54	0.42
1:D:459:GLN:OE1	1:D:459:GLN:N	2.45	0.42
1:A:190:LEU:HB2	1:A:252:VAL:HG11	2.01	0.42
1:B:203:ALA:HB3	1:B:211:ILE:HD13	2.01	0.42
1:D:469:ASP:OD1	5:D:603:HOH:O	2.21	0.42
1:B:221:GLN:HG2	5:B:690:HOH:O	2.18	0.42
1:A:413:LEU:HD12	1:A:414:THR:O	2.20	0.41
1:B:149:GLN:NE2	2:F:14:LYS:HA	2.34	0.41
1:C:441:LEU:HD12	1:C:441:LEU:N	2.35	0.41
1:A:299:THR:O	1:A:299:THR:HG23	2.19	0.41
1:B:268:ARG:CZ	1:B:443:ALA:HB1	2.50	0.41
1:D:255:ILE:HG13	1:D:280:LEU:HD13	2.02	0.41
1:C:244:GLU:OE2	1:C:272:SER:OG	2.33	0.41
1:D:364:LYS:HD3	1:D:366:GLY:H	1.86	0.41
1:B:160:GLN:O	1:B:164:MET:HG3	2.21	0.41
1:D:269:MET:HE1	4:H:101:QVR:O37	2.21	0.41
1:B:290:ILE:O	1:B:290:ILE:HD12	2.21	0.41
1:C:430:LYS:NZ	1:C:433:ASP:OD2	2.49	0.41
2:H:19:GLN:NE2	5:H:202:HOH:O	2.12	0.41
1:A:215:GLU:OE2	4:E:101:QVR:O41	2.39	0.41
1:C:263:MET:O	1:C:264:LEU:HB3	2.21	0.41
1:B:423:PHE:HA	1:B:466:ASN:OD1	2.22	0.40
1:D:266:ASN:O	1:D:267:GLU:CB	2.69	0.40
1:C:350:MET:CE	1:C:385:VAL:HG22	2.52	0.40
1:D:299:THR:O	1:D:299:THR:HG23	2.22	0.40
1:D:434:THR:OG1	1:D:458:ASP:OD2	2.33	0.40
1:C:266:ASN:O	1:C:267:GLU:CB	2.70	0.40

All (2) 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)
5:A:573:HOH:O	5:B:693:HOH:O[4_556]	1.92	0.28
1:A:382:SER:O	1:B:175:ARG:NH2[2_655]	2.12	0.08

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	361/371 (97%)	346 (96%)	15 (4%)	0	100	100
1	B	344/371 (93%)	328 (95%)	16 (5%)	0	100	100
1	C	361/371 (97%)	346 (96%)	15 (4%)	0	100	100
1	D	342/371 (92%)	326 (95%)	16 (5%)	0	100	100
2	E	8/11 (73%)	8 (100%)	0	0	100	100
2	F	8/11 (73%)	8 (100%)	0	0	100	100
2	G	8/11 (73%)	8 (100%)	0	0	100	100
2	H	8/11 (73%)	8 (100%)	0	0	100	100
All	All	1440/1528 (94%)	1378 (96%)	62 (4%)	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	317/323 (98%)	312 (98%)	5 (2%)	62	76
1	B	302/323 (94%)	298 (99%)	4 (1%)	69	81
1	C	316/323 (98%)	313 (99%)	3 (1%)	78	88
1	D	300/323 (93%)	295 (98%)	5 (2%)	60	74
2	E	6/6 (100%)	6 (100%)	0	100	100
2	F	6/6 (100%)	6 (100%)	0	100	100

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	G	6/6 (100%)	6 (100%)	0	100	100
2	H	6/6 (100%)	6 (100%)	0	100	100
All	All	1259/1316 (96%)	1242 (99%)	17 (1%)	67	80

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	139	SER
1	A	287	PHE
1	A	364	LYS
1	A	430	LYS
1	A	439	CYS
1	B	221	GLN
1	B	295	LEU
1	B	364	LYS
1	B	370	ARG
1	C	169	ARG
1	C	210	LYS
1	C	493	PRO
1	D	295	LEU
1	D	352	LYS
1	D	462	SER
1	D	463	LYS
1	D	465	SER

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

Mol	Chain	Res	Type
1	A	149	GLN
1	A	179	GLN
1	A	251	GLN
1	A	378	HIS
1	B	149	GLN
1	B	152	GLN
1	B	159	GLN
1	B	165	GLN
1	B	181	HIS
1	C	251	GLN
1	D	312	ASN
2	E	19	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

4 non-standard protein/DNA/RNA residues are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	ALY	H	18	2	10,11,12	0.81	0	7,12,14	0.55	0
2	ALY	F	18	2	10,11,12	0.81	0	7,12,14	0.64	0
2	ALY	G	18	2	10,11,12	0.72	0	7,12,14	0.81	0
2	ALY	E	18	2	10,11,12	0.77	0	7,12,14	0.67	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	ALY	H	18	2	-	0/9/10/12	-
2	ALY	F	18	2	-	1/9/10/12	-
2	ALY	G	18	2	-	0/9/10/12	-
2	ALY	E	18	2	-	0/9/10/12	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	F	18	ALY	CA-CB-CG-CD

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

9 ligands are modelled in this entry.

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	QVR	H	101	2	19,22,22	0.63	0	19,32,32	0.72	1 (5%)
3	MLI	B	501	-	0,6,6	-	-	0,7,7	-	-
3	MLI	C	502	-	0,6,6	-	-	0,7,7	-	-
4	QVR	F	101	2	19,22,22	0.60	0	19,32,32	0.71	1 (5%)
3	MLI	C	501	-	0,6,6	-	-	0,7,7	-	-
4	QVR	G	101	2	19,22,22	0.62	0	19,32,32	0.76	1 (5%)
3	MLI	C	503	-	0,6,6	-	-	0,7,7	-	-
3	MLI	D	501	-	0,6,6	-	-	0,7,7	-	-
4	QVR	E	101	2	19,22,22	0.62	0	19,32,32	0.74	1 (5%)

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	QVR	H	101	2	-	0/3/23/23	0/3/3/3
3	MLI	B	501	-	-	0/0/4/4	-
3	MLI	C	502	-	-	0/0/4/4	-
4	QVR	F	101	2	-	0/3/23/23	0/3/3/3
3	MLI	C	501	-	-	0/0/4/4	-
4	QVR	G	101	2	-	0/3/23/23	0/3/3/3
3	MLI	C	503	-	-	0/0/4/4	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	MLI	D	501	-	-	0/0/4/4	-
4	QVR	E	101	2	-	0/3/23/23	0/3/3/3

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	E	101	QVR	C50-C55-N57	2.35	123.92	120.35
4	G	101	QVR	C50-C55-N57	2.28	123.82	120.35
4	H	101	QVR	C50-C55-N57	2.27	123.80	120.35
4	F	101	QVR	C50-C55-N57	2.24	123.76	120.35

There are no chirality outliers.

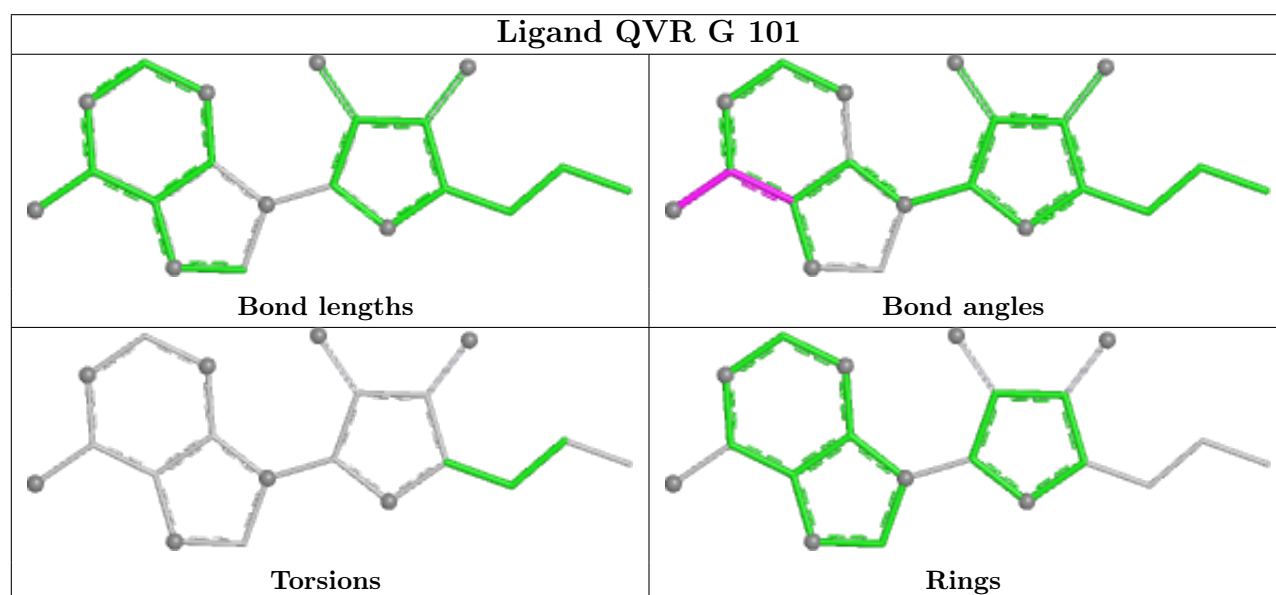
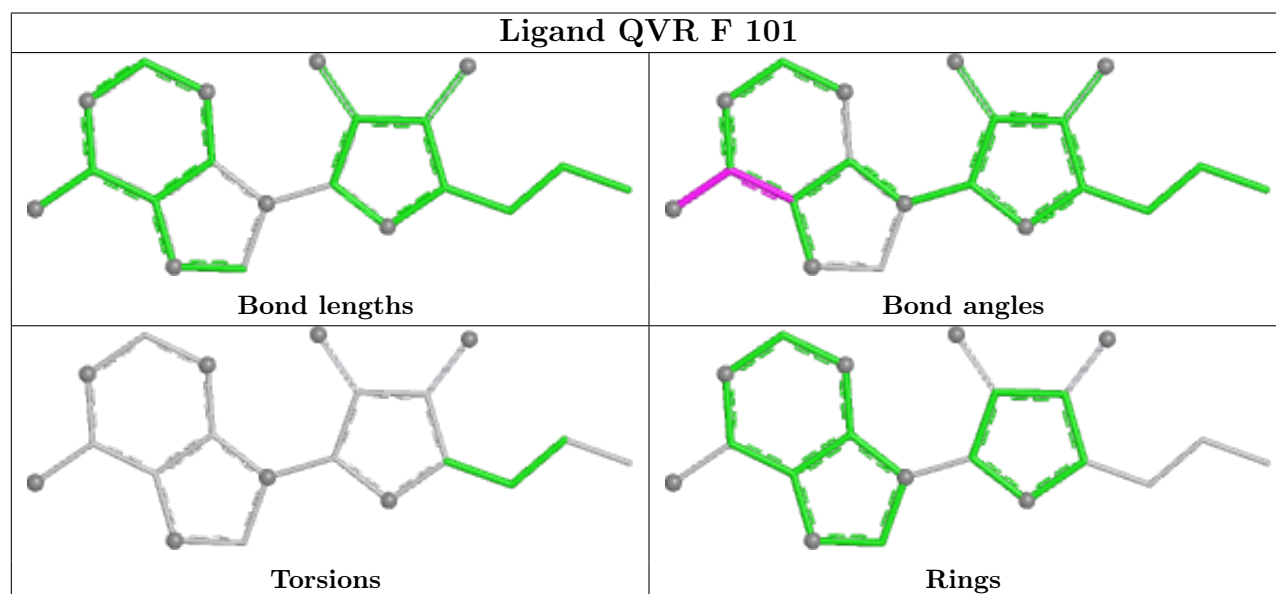
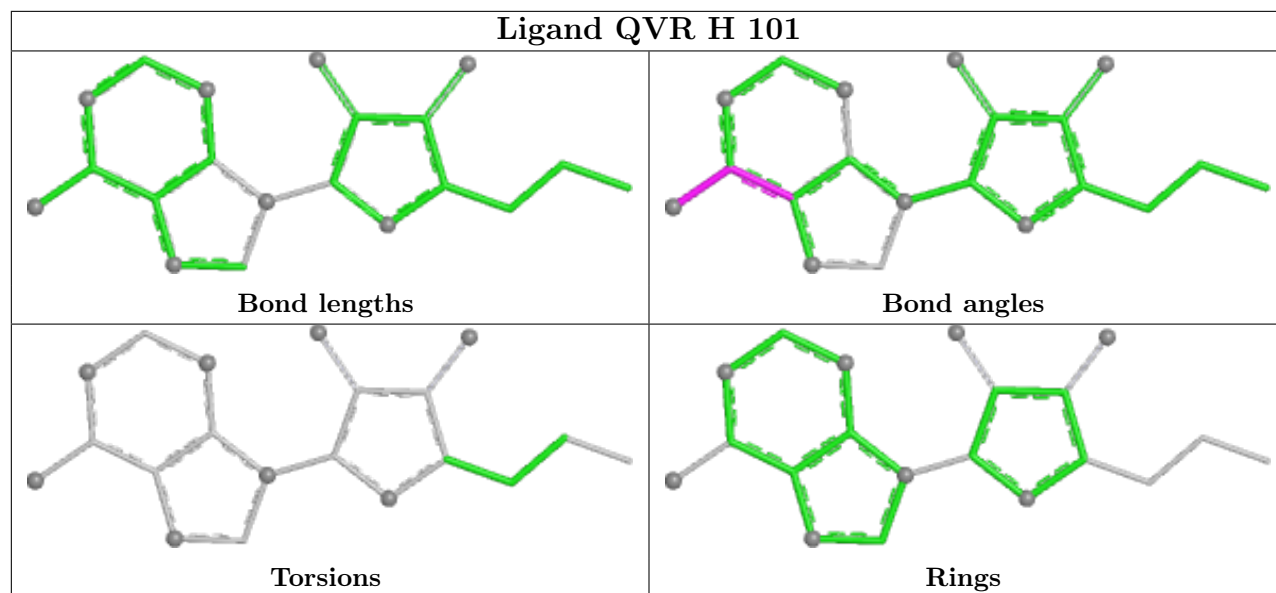
There are no torsion outliers.

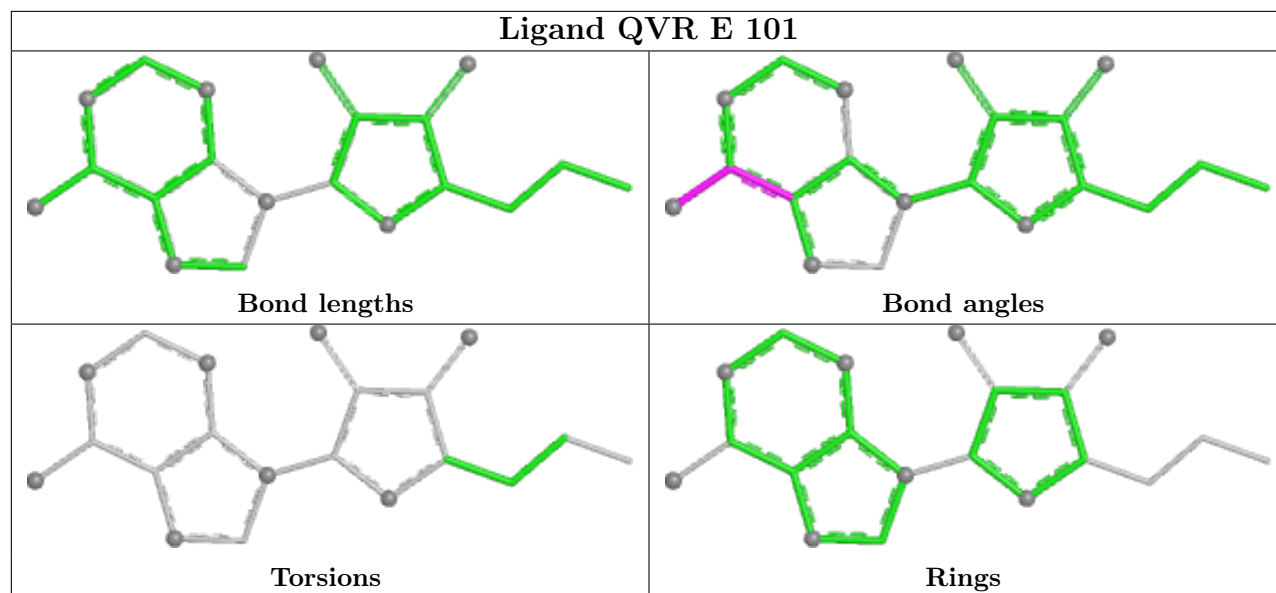
There are no ring outliers.

4 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	H	101	QVR	2	0
3	B	501	MLI	3	0
4	G	101	QVR	1	0
4	E	101	QVR	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	363/371 (97%)	0.15	10 (2%) 53 51	29, 45, 65, 83	1 (0%)
1	B	344/371 (92%)	0.00	4 (1%) 79 77	29, 43, 58, 96	0
1	C	362/371 (97%)	0.15	9 (2%) 57 55	33, 49, 69, 83	1 (0%)
1	D	344/371 (92%)	0.09	6 (1%) 70 68	34, 48, 63, 86	1 (0%)
2	E	9/11 (81%)	0.71	2 (22%) 0 0	39, 44, 56, 76	0
2	F	9/11 (81%)	0.67	0 100 100	42, 46, 67, 69	0
2	G	9/11 (81%)	0.24	1 (11%) 5 4	43, 46, 62, 71	0
2	H	9/11 (81%)	0.59	0 100 100	45, 48, 64, 76	0
All	All	1449/1528 (94%)	0.11	32 (2%) 62 59	29, 46, 65, 96	3 (0%)

All (32) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	135	ARG	4.9
1	C	283	SER	4.3
1	D	347	ARG	3.9
1	D	380	LEU	3.8
1	A	283	SER	3.8
1	A	246	VAL	3.6
1	C	209	ARG	3.4
1	A	135	ARG	3.2
1	B	347	ARG	3.1
2	E	22	THR	2.8
1	A	428	PHE	2.8
1	A	496	ASN	2.7
1	D	381	HIS	2.7
1	A	481	THR	2.6
1	C	496	ASN	2.5
1	C	284	GLY	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	188	ILE	2.4
1	C	136	SER	2.4
1	C	188	ILE	2.4
2	G	15	ALA	2.3
2	E	15	ALA	2.3
1	C	252	VAL	2.3
1	D	140	GLU	2.2
1	B	315	TYR	2.2
1	A	234	ASP	2.2
1	D	378	HIS	2.2
1	A	497	MET	2.2
1	A	276	ALA	2.1
1	D	430	LYS	2.1
1	B	428	PHE	2.0
1	C	467	LEU	2.0
1	C	280	LEU	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [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	ALY	H	18	12/13	0.91	0.17	44,53,55,61	0
2	ALY	G	18	12/13	0.92	0.15	36,44,54,54	0
2	ALY	F	18	12/13	0.92	0.17	39,47,68,68	0
2	ALY	E	18	12/13	0.96	0.18	36,45,47,51	0

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

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

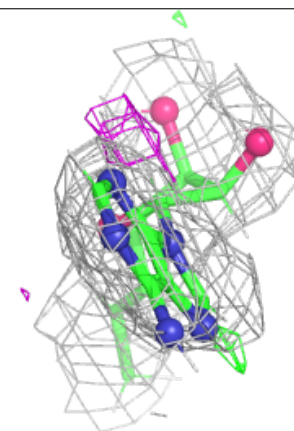
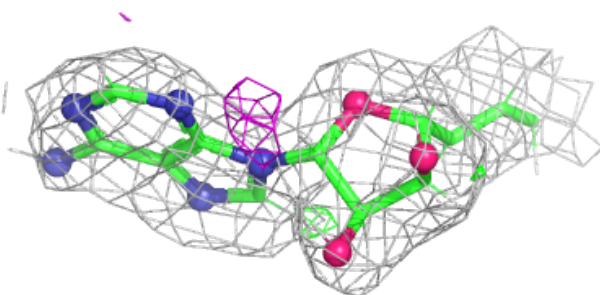
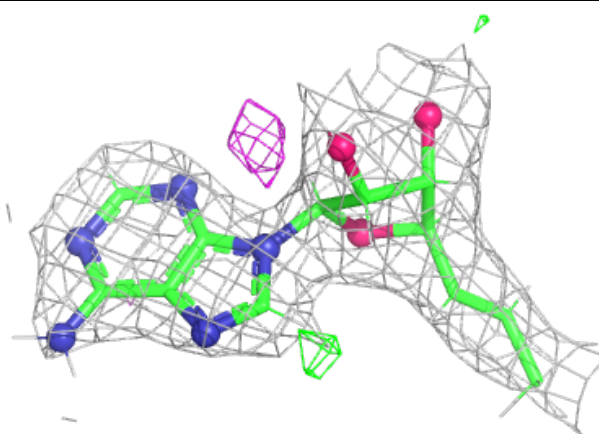


Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	MLI	C	501	7/7	0.70	0.22	51,66,80,82	0
3	MLI	C	502	7/7	0.70	0.28	62,72,83,83	0
3	MLI	C	503	7/7	0.72	0.20	54,61,72,72	0
3	MLI	B	501	7/7	0.75	0.41	42,51,56,59	0
3	MLI	D	501	7/7	0.84	0.18	54,62,70,70	0
4	QVR	G	101	20/20	0.91	0.13	38,48,59,61	0
4	QVR	H	101	20/20	0.94	0.11	41,49,61,63	0
4	QVR	E	101	20/20	0.95	0.11	34,41,54,59	0
4	QVR	F	101	20/20	0.95	0.12	35,40,48,50	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.

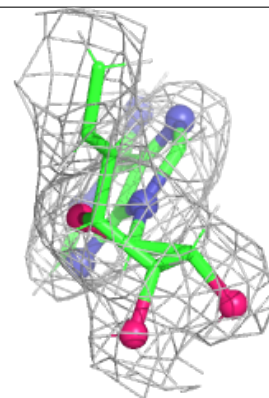
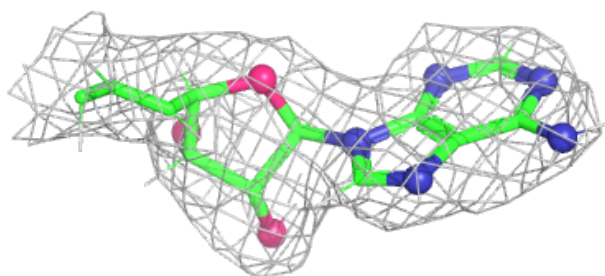
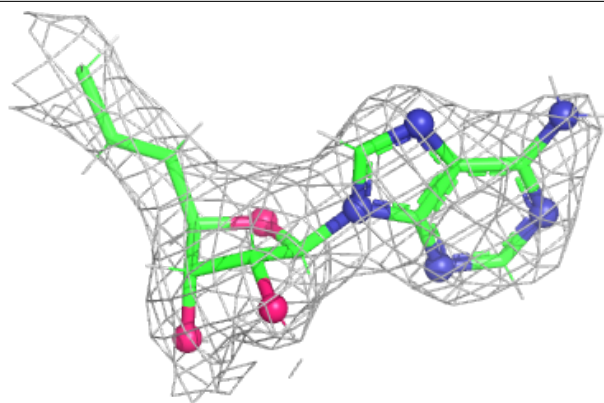
#### Electron density around QVR G 101:

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)

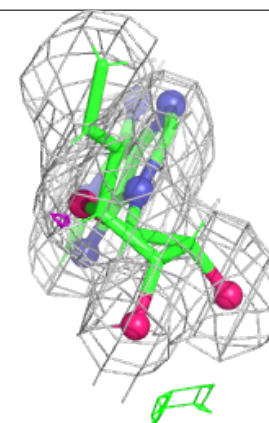
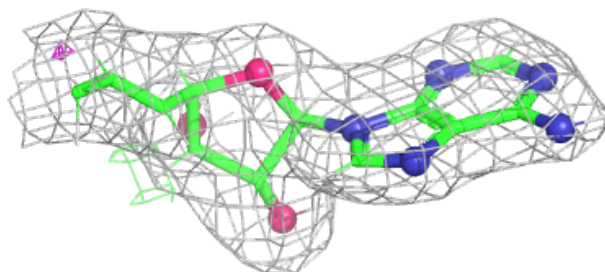
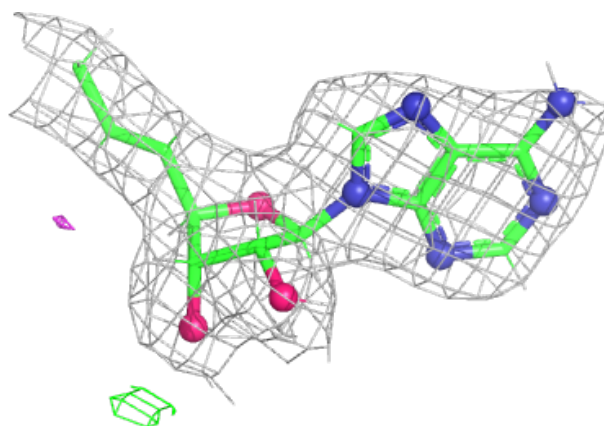


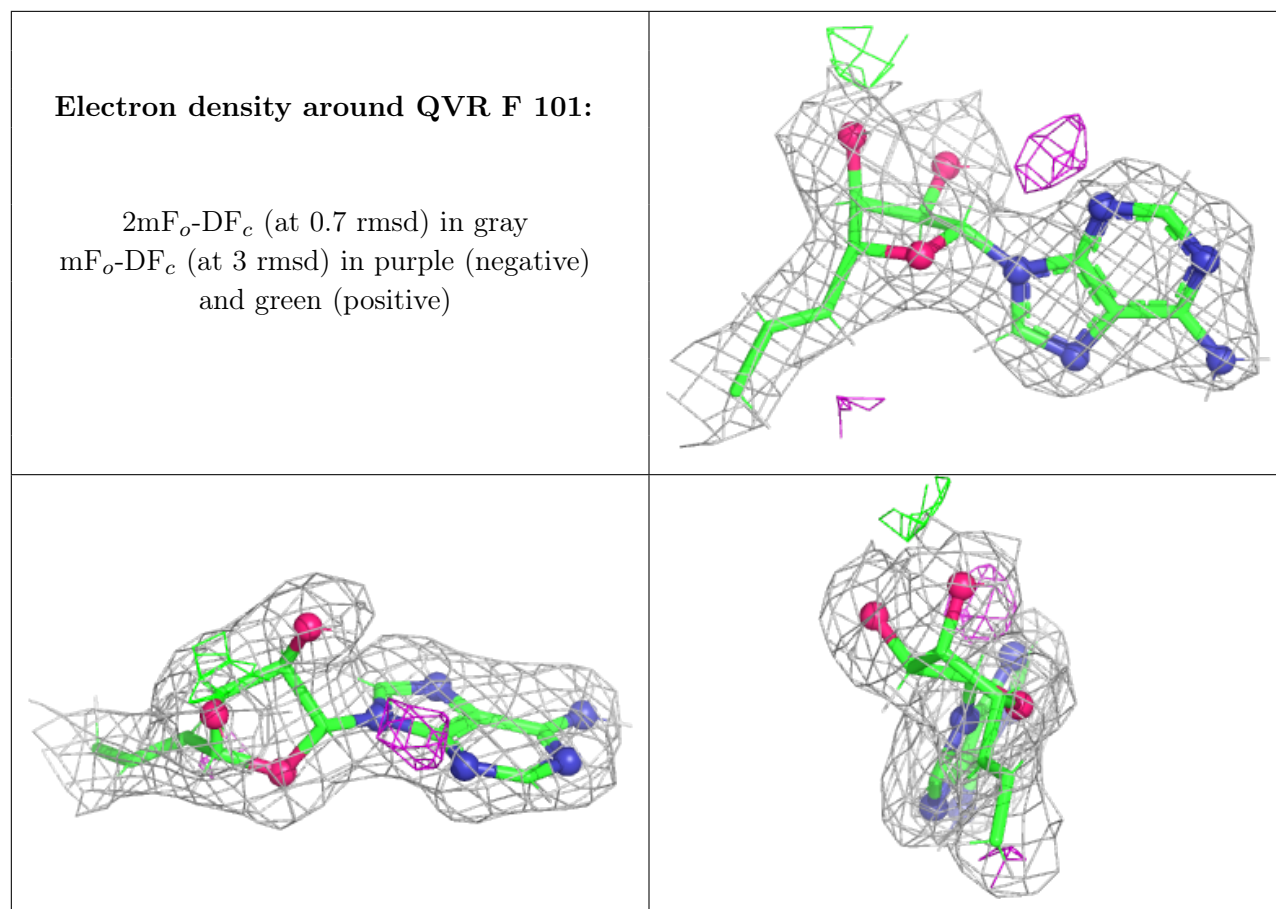
**Electron density around QVR H 101:**

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

**Electron density around QVR E 101:**

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





## 6.5 Other polymers ⓘ

There are no such residues in this entry.