



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 6, 2023 – 02:19 pm GMT

PDB ID : 7QPH  
Title : Crystal structure of mouse CARM1 in complex with histone H3\_22-31 K27 acetylated  
Authors : Marechal, N.; Cura, V.; Troffer-Charlier, N.; Bonnefond, L.; Cavarelli, J.  
Deposited on : 2022-01-04  
Resolution : 1.90 Å(reported)

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

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

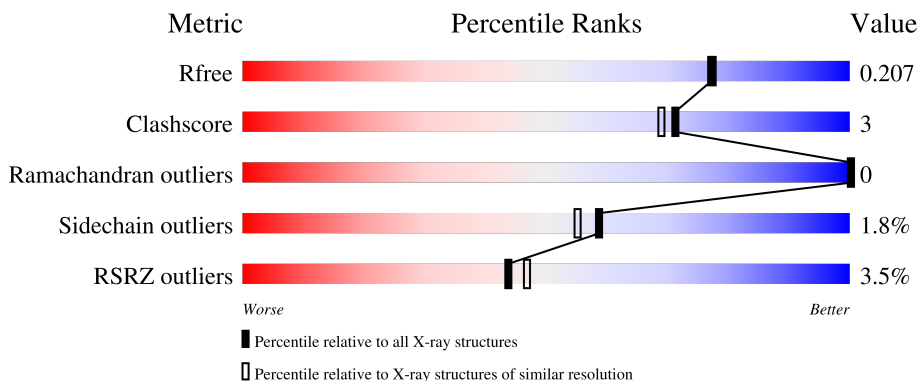
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	361	 3% 87% 7% 5%
1	B	361	 2% 88% 6% 5%
1	C	361	 4% 88% 7% 5%
1	D	361	 4% 84% 11% 5%
2	E	10	 10% 20% 10% 10% 60%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	10	
2	G	10	
2	H	10	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 22858 atoms, of which 10997 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	
			Total	C	H	N	O				S
1	A	343	5450	1776	2699	454	507	14	0	0	0
1	B	343	5450	1776	2699	454	507	14	0	0	0
1	C	343	5450	1776	2699	454	507	14	0	0	0
1	D	343	5450	1776	2699	454	507	14	0	0	0

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 22-31 K27 acetylated.

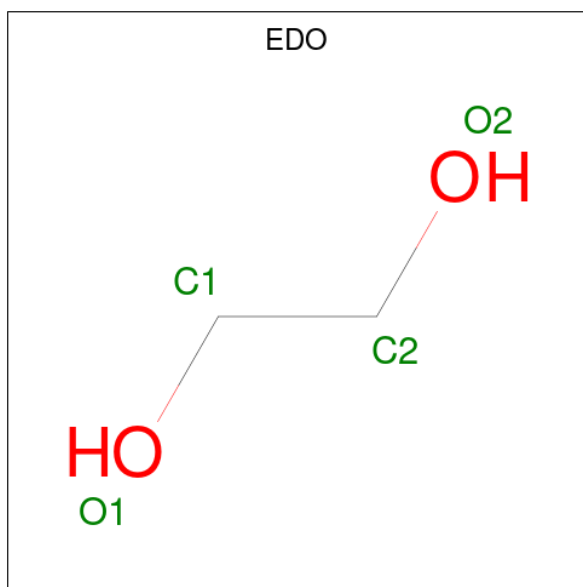
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
2	E	4	68	20	35	8	5	0	0	0
2	H	4	68	20	35	8	5	0	0	0

*Continued on next page...*

Continued from previous page...

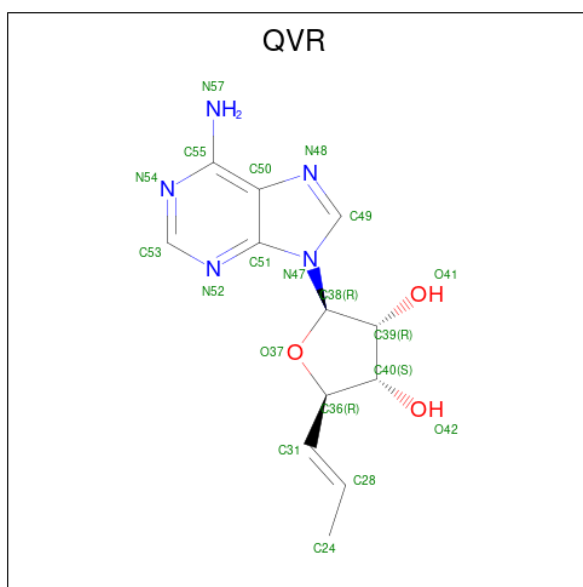
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
2	G	4	Total	C	H	N	O	0	0	0
			67	20	34	8	5			
2	F	4	Total	C	H	N	O	0	0	0
			68	20	35	8	5			

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	D	1	Total	C	H	O	0	0
			10	2	6	2		

- Molecule 4 is (2 {R},3 {R},4 {S},5 {R})-2-(6-aminopurin-9-yl)-5-[( {E})-prop-1-enyl]oxolan e-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
			Total	C	H	N	O		
4	E	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		
4	G	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		


- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	166	Total	O	0	0
			166	166		
5	B	134	Total	O	0	0
			134	134		
5	C	165	Total	O	0	0
			165	165		
5	D	166	Total	O	0	0
			166	166		
5	E	2	Total	O	0	0
			2	2		
5	H	4	Total	O	0	0
			4	4		
5	G	2	Total	O	0	0
			2	2		
5	F	2	Total	O	0	0
			2	2		

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

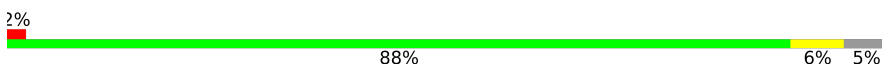
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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.

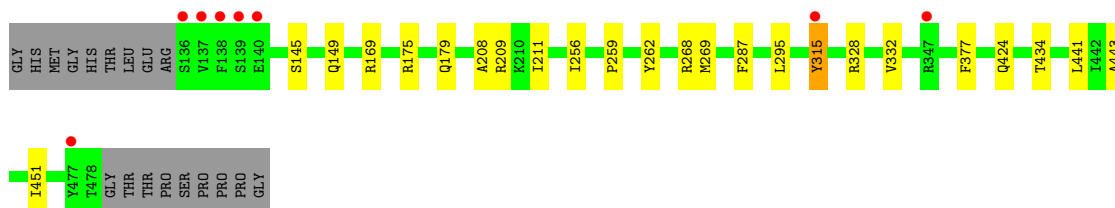
- Molecule 1: Histone-arginine methyltransferase CARM1

Chain A: 




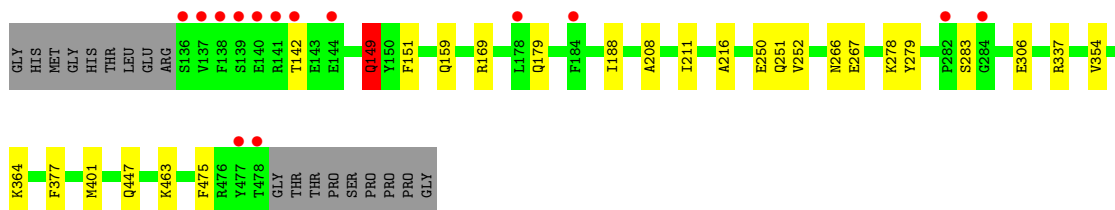
- Molecule 1: Histone-arginine methyltransferase CARM1

Chain B: 




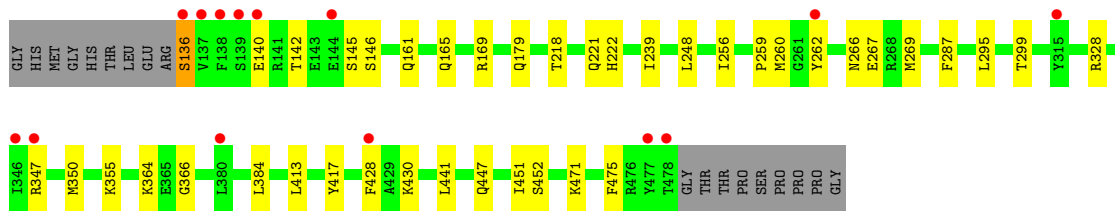
- Molecule 1: Histone-arginine methyltransferase CARM1

Chain C: 

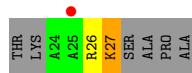


- Molecule 1: Histone-arginine methyltransferase CARM1

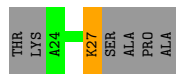
Chain D: 



● Molecule 2: Histone H3 22-31 K27 acetylated



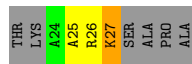
● Molecule 2: Histone H3 22-31 K27 acetylated



● Molecule 2: Histone H3 22-31 K27 acetylated



● Molecule 2: Histone H3 22-31 K27 acetylated





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	75.34Å 99.18Å 208.44Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.25 – 1.90 48.25 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.25-1.90) 99.9 (48.25-1.90)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.60 (at 1.90Å)	Xtrriage
Refinement program	PHENIX 1.20rc2_4400	Depositor
R, $R_{free}$	0.183 , 0.210 0.181 , 0.207	Depositor DCC
$R_{free}$ test set	6153 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.8	Xtrriage
Anisotropy	0.422	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 45.2	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	22858	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	50.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 50.14 % 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 6.7529e-05. 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: EDO, QVR, ALY

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	0.40	0/2821	0.54	0/3823
1	B	0.37	0/2821	0.52	0/3823
1	C	0.42	1/2821 (0.0%)	0.54	0/3823
1	D	0.39	0/2821	0.53	0/3823
2	E	1.32	0/20	1.50	0/25
2	F	1.47	1/20 (5.0%)	1.22	0/25
2	G	1.52	1/20 (5.0%)	1.09	0/25
2	H	1.01	0/20	1.28	0/25
All	All	0.41	3/11364 (0.0%)	0.54	0/15392

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	149	GLN	CB-CG	-6.81	1.34	1.52
2	G	26	ARG	N-CA	-5.43	1.35	1.46
2	F	26	ARG	N-CA	-5.31	1.35	1.46

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	2751	2699	2699	23	0
1	B	2751	2699	2699	13	0
1	C	2751	2699	2699	14	0
1	D	2751	2699	2699	23	0
2	E	33	35	34	3	0
2	F	33	35	34	3	0
2	G	33	34	34	1	0
2	H	33	35	34	1	0
3	D	4	6	6	1	0
4	E	20	14	0	0	0
4	F	20	14	0	0	0
4	G	20	14	0	0	0
4	H	20	14	0	0	0
5	A	166	0	0	3	3
5	B	134	0	0	1	2
5	C	165	0	0	3	0
5	D	166	0	0	3	0
5	E	2	0	0	0	0
5	F	2	0	0	0	0
5	G	2	0	0	0	0
5	H	4	0	0	0	0
All	All	11861	10997	10938	77	3

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 (77) 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:234:ASP:HB2	5:A:525:HOH:O	1.84	0.76
1:D:441:LEU:HD22	1:D:451:ILE:HG12	1.77	0.66
1:D:179:GLN:OE1	5:D:601:HOH:O	2.14	0.65
1:C:306:GLU:OE2	5:C:501:HOH:O	2.15	0.63
1:A:347:ARG:HG3	1:A:347:ARG:HH11	1.63	0.63
1:B:209:ARG:NH2	5:B:504:HOH:O	2.32	0.63
1:D:142:THR:HG23	1:D:447:GLN:HG3	1.80	0.63
1:D:364:LYS:HE3	1:D:366:GLY:H	1.63	0.63
2:E:27:ALY:HH31	2:E:27:ALY:HD3	1.82	0.62
1:D:259:PRO:HA	1:D:269:MET:HG2	1.84	0.59
1:B:259:PRO:HA	1:B:269:MET:HG2	1.85	0.58
1:B:262:TYR:OH	2:F:25:ALA:HB1	2.05	0.57
3:D:501:EDO:H11	5:D:742:HOH:O	2.02	0.57

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:222:HIS:HD2	5:D:744:HOH:O	1.87	0.57
1:A:445:LYS:N	1:A:445:LYS:HD3	2.21	0.55
1:A:301:GLU:OE2	5:A:501:HOH:O	2.19	0.54
1:C:149:GLN:OE1	1:C:149:GLN:O	2.26	0.53
1:D:239:ILE:HD12	1:D:248:LEU:CD2	2.39	0.52
1:B:256:ILE:HG22	1:B:287:PHE:HB2	1.93	0.51
1:D:364:LYS:HE3	1:D:366:GLY:N	2.25	0.51
1:A:445:LYS:N	1:A:445:LYS:CD	2.74	0.51
1:D:136:SER:O	1:D:140:GLU:HG2	2.11	0.51
1:D:347:ARG:HB3	1:D:347:ARG:NH1	2.26	0.51
1:D:266:ASN:O	1:D:267:GLU:HB2	2.12	0.50
1:A:378:HIS:ND1	1:A:434:THR:HG22	2.27	0.50
1:A:445:LYS:H	1:A:445:LYS:HE2	1.78	0.49
1:B:441:LEU:HD22	1:B:451:ILE:HG12	1.95	0.49
1:C:208:ALA:HB3	1:C:211:ILE:HD11	1.95	0.49
1:A:143:GLU:HG2	1:A:145:SER:OG	2.13	0.48
1:C:142:THR:HG23	1:C:447:GLN:HG3	1.96	0.48
1:D:218:THR:O	1:D:221:GLN:HG2	2.13	0.48
1:C:250:GLU:HG3	1:C:251:GLN:O	2.13	0.48
1:C:159:GLN:NE2	5:C:502:HOH:O	2.17	0.48
1:A:224:GLU:OE2	1:A:228:LYS:HE3	2.14	0.47
2:F:27:ALY:HE3	2:F:27:ALY:HH31	1.80	0.46
1:A:262:TYR:HH	1:A:475:PHE:HE1	1.63	0.46
1:A:445:LYS:H	1:A:445:LYS:CE	2.30	0.46
1:D:364:LYS:HE3	1:D:366:GLY:HA3	1.97	0.46
1:D:142:THR:CG2	1:D:447:GLN:HG3	2.47	0.45
1:A:144:GLU:HA	1:A:144:GLU:OE1	2.17	0.45
1:C:337:ARG:HD3	5:C:644:HOH:O	2.16	0.45
1:B:268:ARG:CZ	1:B:443:ALA:HB1	2.46	0.45
1:B:315:TYR:HD2	1:B:328:ARG:HE	1.65	0.45
1:A:256:ILE:HG22	1:A:287:PHE:HB2	1.99	0.45
1:A:242:LYS:O	1:A:246:VAL:HG13	2.17	0.44
1:B:208:ALA:HB3	1:B:211:ILE:HD11	1.99	0.44
1:B:328:ARG:NH1	1:B:332:VAL:HG21	2.32	0.44
1:A:463:LYS:HB3	1:A:463:LYS:HE2	1.78	0.44
1:C:188:ILE:HG22	1:C:252:VAL:HG12	2.00	0.43
1:C:266:ASN:O	1:C:267:GLU:HB2	2.18	0.43
1:A:242:LYS:HE2	5:A:524:HOH:O	2.18	0.43
1:C:278:LYS:HE2	1:C:279:TYR:CZ	2.54	0.43
1:B:175:ARG:HG3	1:B:179:GLN:NE2	2.34	0.43
1:D:256:ILE:HG22	1:D:287:PHE:HB2	2.01	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:149:GLN:HG2	2:F:27:ALY:HH33	2.02	0.42
1:D:299:THR:OG1	1:D:350:MET:SD	2.78	0.42
1:C:151:PHE:CE1	1:C:216:ALA:HB3	2.55	0.42
1:C:179:GLN:OE1	1:C:401:MET:SD	2.77	0.42
1:D:259:PRO:HD2	1:D:260:MET:SD	2.60	0.42
1:D:347:ARG:HB3	1:D:347:ARG:CZ	2.50	0.41
1:C:475:PHE:CE1	2:G:25:ALA:HB2	2.55	0.41
1:D:161:GLN:O	1:D:165:GLN:HG3	2.20	0.41
1:A:141:ARG:O	1:A:141:ARG:HG2	2.20	0.41
1:A:415:HIS:CD2	2:E:26:ARG:HB2	2.55	0.41
1:C:354:VAL:HG13	1:C:377:PHE:CE1	2.55	0.41
1:D:299:THR:HG23	1:D:384:LEU:O	2.21	0.41
2:E:27:ALY:HH31	2:E:27:ALY:CD	2.45	0.41
1:A:377:PHE:O	1:A:434:THR:HA	2.20	0.41
1:D:428:PHE:CZ	1:D:430:LYS:HD2	2.56	0.41
1:B:377:PHE:O	1:B:434:THR:HA	2.21	0.41
2:H:27:ALY:HE3	2:H:27:ALY:HH31	1.83	0.41
1:A:160:GLN:O	1:A:164:MET:HG3	2.21	0.40
1:A:145:SER:HA	1:B:145:SER:HB3	2.03	0.40
1:A:175:ARG:O	1:A:179:GLN:HG3	2.21	0.40
1:D:262:TYR:CZ	1:D:475:PHE:HE1	2.39	0.40
1:D:413:LEU:HD11	1:D:417:TYR:CD2	2.56	0.40
1:A:208:ALA:HB3	1:A:211:ILE:HD11	2.02	0.40

All (3) 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:641:HOH:O	5:A:648:HOH:O[2_655]	1.74	0.46
5:A:647:HOH:O	5:B:609:HOH:O[4_556]	1.96	0.24
5:A:575:HOH:O	5:B:619:HOH:O[2_655]	2.11	0.09

## 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	341/361 (94%)	332 (97%)	9 (3%)	0	100	100
1	B	341/361 (94%)	327 (96%)	14 (4%)	0	100	100
1	C	341/361 (94%)	331 (97%)	10 (3%)	0	100	100
1	D	341/361 (94%)	330 (97%)	11 (3%)	0	100	100
2	E	2/10 (20%)	1 (50%)	1 (50%)	0	100	100
2	F	2/10 (20%)	2 (100%)	0	0	100	100
2	G	2/10 (20%)	2 (100%)	0	0	100	100
2	H	2/10 (20%)	0	2 (100%)	0	100	100
All	All	1372/1484 (92%)	1325 (97%)	47 (3%)	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	299/313 (96%)	296 (99%)	3 (1%)	76	76
1	B	299/313 (96%)	295 (99%)	4 (1%)	69	68
1	C	299/313 (96%)	294 (98%)	5 (2%)	60	57
1	D	299/313 (96%)	290 (97%)	9 (3%)	41	33
2	E	1/5 (20%)	1 (100%)	0	100	100
2	F	1/5 (20%)	1 (100%)	0	100	100
2	G	1/5 (20%)	1 (100%)	0	100	100
2	H	1/5 (20%)	1 (100%)	0	100	100
All	All	1200/1272 (94%)	1179 (98%)	21 (2%)	59	55

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

Mol	Chain	Res	Type
1	A	234	ASP
1	A	445	LYS
1	A	466	ASN
1	B	169	ARG
1	B	295	LEU
1	B	315	TYR
1	B	424	GLN
1	C	149	GLN
1	C	169	ARG
1	C	283	SER
1	C	364	LYS
1	C	463	LYS
1	D	136	SER
1	D	145	SER
1	D	146	SER
1	D	169	ARG
1	D	295	LEU
1	D	328	ARG
1	D	355	LYS
1	D	452	SER
1	D	471	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	159	GLN
1	A	466	ASN
1	B	179	GLN
1	C	159	GLN
1	C	316	GLN
1	D	159	GLN
1	D	179	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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	G	27	2	10,11,12	2.72	1 (10%)	7,12,14	4.23	3 (42%)
2	ALY	H	27	2	10,11,12	0.84	0	7,12,14	1.01	1 (14%)
2	ALY	E	27	2	10,11,12	0.75	0	7,12,14	1.95	3 (42%)
2	ALY	F	27	2	10,11,12	0.83	0	7,12,14	1.16	1 (14%)

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	G	27	2	-	4/9/10/12	-
2	ALY	H	27	2	-	5/9/10/12	-
2	ALY	E	27	2	-	7/9/10/12	-
2	ALY	F	27	2	-	4/9/10/12	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	G	27	ALY	CH3-CH	-7.96	1.34	1.50

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	G	27	ALY	OH-CH-CH3	-8.46	106.34	122.06
2	G	27	ALY	OH-CH-NZ	5.71	137.80	121.74
2	G	27	ALY	CE-NZ-CH	-3.89	116.57	122.56
2	E	27	ALY	CE-NZ-CH	3.05	127.24	122.56
2	E	27	ALY	CD-CE-NZ	2.98	120.73	112.21
2	F	27	ALY	CE-NZ-CH	2.43	126.30	122.56
2	E	27	ALY	CH3-CH-NZ	2.27	120.11	116.09
2	H	27	ALY	CE-NZ-CH	2.05	125.70	122.56

There are no chirality outliers.



All (20) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	E	27	ALY	N-CA-CB-CG
2	E	27	ALY	C-CA-CB-CG
2	E	27	ALY	O-C-CA-CB
2	H	27	ALY	N-CA-CB-CG
2	H	27	ALY	C-CA-CB-CG
2	G	27	ALY	N-CA-CB-CG
2	G	27	ALY	C-CA-CB-CG
2	F	27	ALY	O-C-CA-CB
2	G	27	ALY	CG-CD-CE-NZ
2	E	27	ALY	CD-CE-NZ-CH
2	E	27	ALY	OH-CH-NZ-CE
2	E	27	ALY	CH3-CH-NZ-CE
2	H	27	ALY	OH-CH-NZ-CE
2	H	27	ALY	CH3-CH-NZ-CE
2	F	27	ALY	OH-CH-NZ-CE
2	F	27	ALY	CH3-CH-NZ-CE
2	G	27	ALY	CA-CB-CG-CD
2	H	27	ALY	CG-CD-CE-NZ
2	F	27	ALY	CA-CB-CG-CD
2	E	27	ALY	CG-CD-CE-NZ

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	H	27	ALY	1	0
2	E	27	ALY	2	0
2	F	27	ALY	2	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

5 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
3	EDO	D	501	-	3,3,3	0.11	0	2,2,2	0.13	0
4	QVR	F	101	2	19,22,22	0.62	0	19,32,32	0.80	1 (5%)
4	QVR	G	101	2	19,22,22	0.61	0	19,32,32	0.84	1 (5%)
4	QVR	E	101	2	19,22,22	0.60	0	19,32,32	0.77	1 (5%)
4	QVR	H	101	2	19,22,22	0.58	0	19,32,32	0.72	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
3	EDO	D	501	-	-	0/1/1/1	-
4	QVR	F	101	2	-	0/3/23/23	0/3/3/3
4	QVR	G	101	2	-	2/3/23/23	0/3/3/3
4	QVR	E	101	2	-	2/3/23/23	0/3/3/3
4	QVR	H	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	F	101	QVR	C50-C55-N57	2.80	124.61	120.35
4	G	101	QVR	C50-C55-N57	2.44	124.06	120.35
4	H	101	QVR	C50-C55-N57	2.40	124.00	120.35
4	E	101	QVR	C50-C55-N57	2.39	123.98	120.35

There are no chirality outliers.

All (4) torsion outliers are listed below:

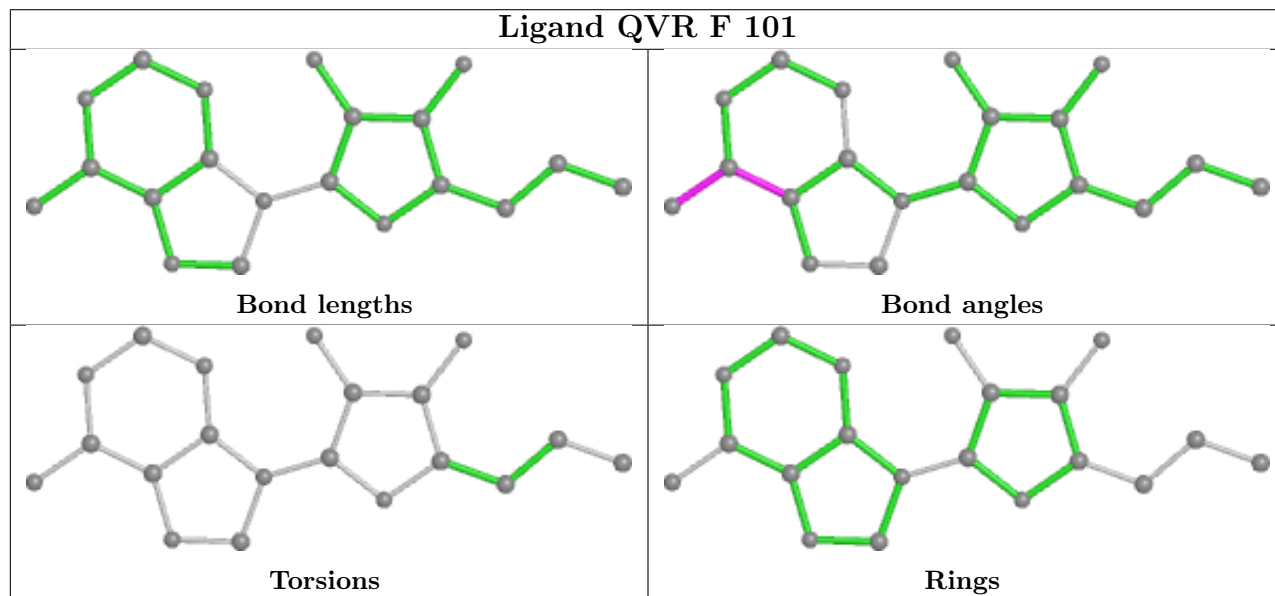
Mol	Chain	Res	Type	Atoms
4	G	101	QVR	C28-C31-C36-O37
4	E	101	QVR	C28-C31-C36-C40
4	G	101	QVR	C28-C31-C36-C40
4	E	101	QVR	C28-C31-C36-O37

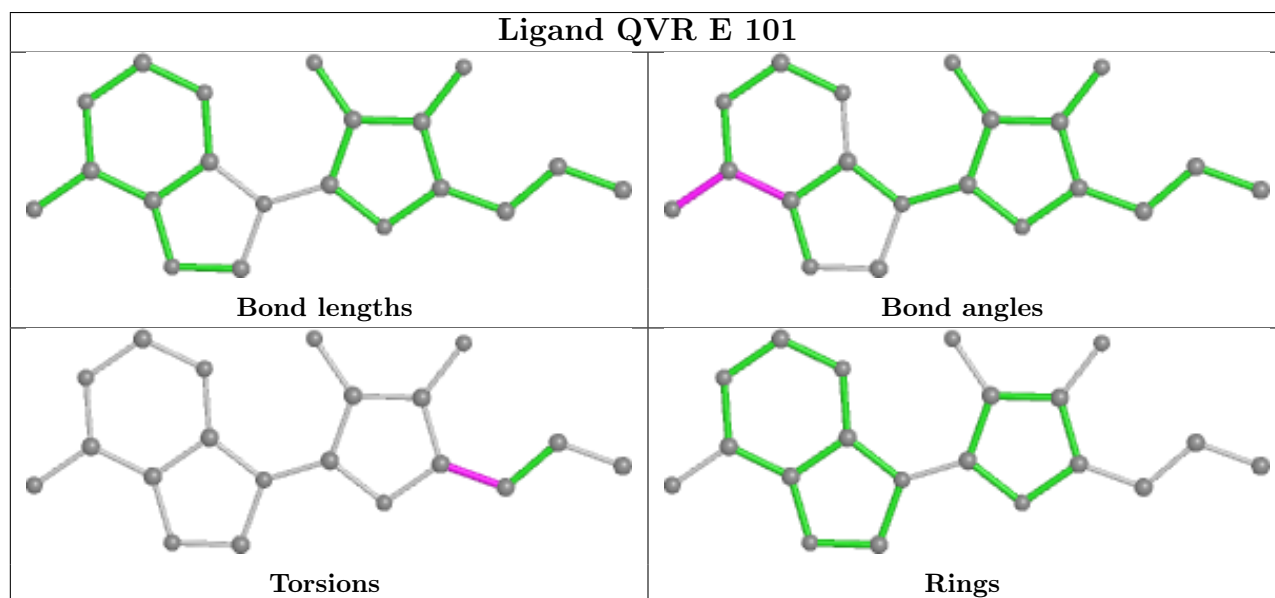
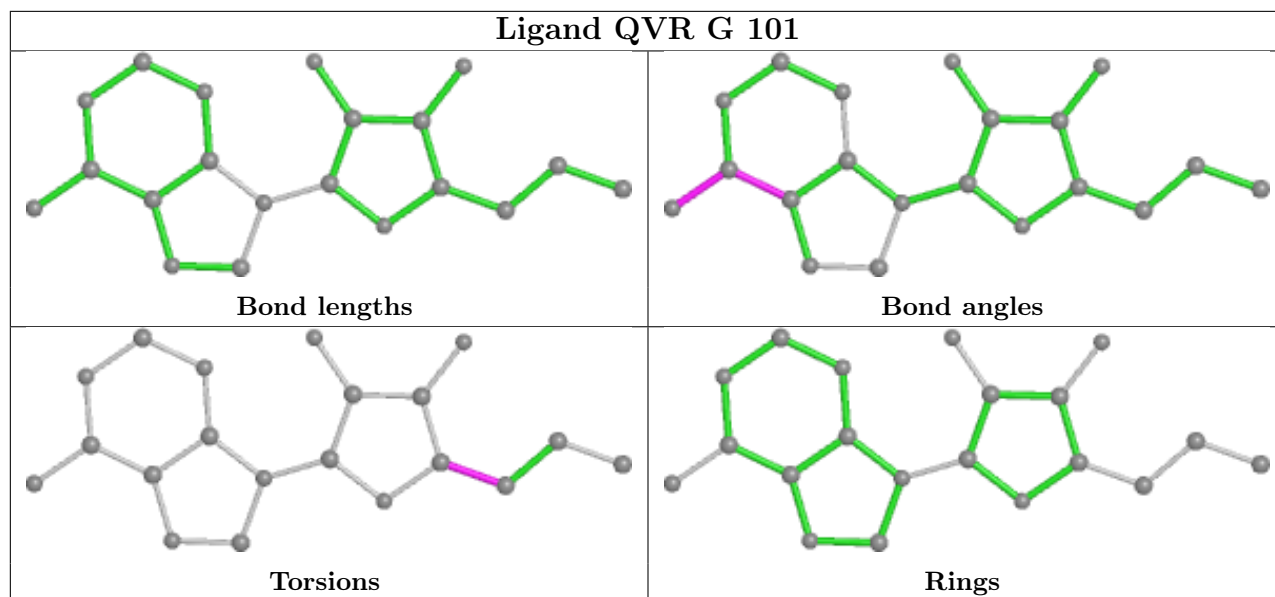
There are no ring outliers.

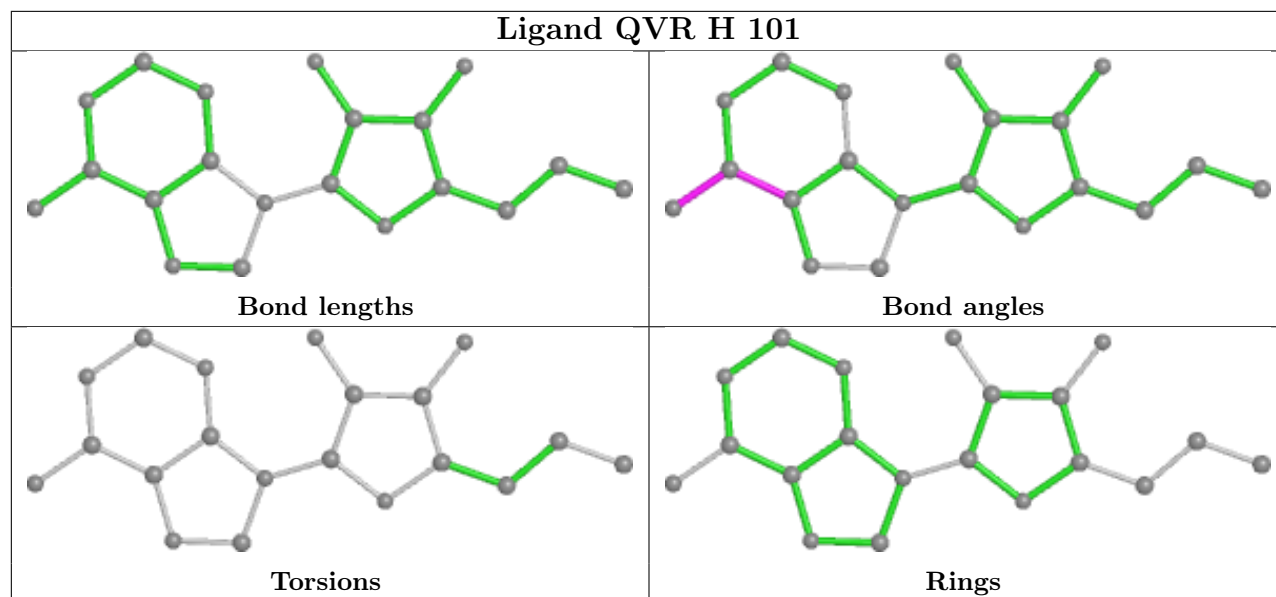
1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	501	EDO	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	343/361 (95%)	0.22	11 (3%) 47 50	26, 41, 61, 98	0
1	B	343/361 (95%)	0.20	8 (2%) 60 63	32, 45, 68, 86	0
1	C	343/361 (95%)	0.28	14 (4%) 37 40	26, 43, 73, 93	0
1	D	343/361 (95%)	0.17	14 (4%) 37 40	28, 42, 63, 93	0
2	E	3/10 (30%)	1.64	1 (33%) 0 0	67, 67, 82, 88	0
2	F	3/10 (30%)	1.20	0 100 100	68, 68, 77, 79	0
2	G	3/10 (30%)	1.31	1 (33%) 0 0	60, 60, 77, 89	0
2	H	3/10 (30%)	0.59	0 100 100	66, 66, 78, 80	0
All	All	1384/1484 (93%)	0.23	49 (3%) 44 47	26, 43, 69, 98	0

All (49) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	136	SER	8.1
1	C	137	VAL	7.9
1	D	137	VAL	7.7
1	A	137	VAL	7.3
1	C	138	PHE	6.9
1	A	138	PHE	6.8
1	C	140	GLU	6.0
1	B	137	VAL	5.9
1	A	139	SER	5.7
1	A	136	SER	5.6
1	D	477	TYR	5.6
1	B	139	SER	5.6
1	C	139	SER	5.2
1	A	140	GLU	5.1
1	B	136	SER	4.4
1	A	142	THR	4.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	141	ARG	4.3
1	C	477	TYR	4.3
1	B	138	PHE	4.0
1	C	142	THR	4.0
1	D	478	THR	4.0
1	D	140	GLU	3.9
1	D	136	SER	3.8
1	B	140	GLU	3.7
1	D	139	SER	3.5
1	C	478	THR	3.4
1	C	141	ARG	3.3
1	B	477	TYR	3.2
1	B	315	TYR	3.1
1	C	282	PRO	3.0
1	C	284	GLY	2.9
1	D	347	ARG	2.8
1	D	138	PHE	2.8
1	B	347	ARG	2.7
1	D	315	TYR	2.7
1	A	428	PHE	2.6
1	A	478	THR	2.5
1	D	428	PHE	2.5
1	D	346	ILE	2.5
1	C	178	LEU	2.4
1	A	477	TYR	2.4
1	D	262	TYR	2.4
2	E	25	ALA	2.3
1	D	144	GLU	2.3
1	C	184	PHE	2.2
1	C	144	GLU	2.2
1	A	144	GLU	2.1
2	G	24	ALA	2.1
1	D	380	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( $\text{\AA}^2$ )	Q<0.9
2	ALY	G	27	12/13	0.68	0.25	65,80,101,101	0
2	ALY	E	27	12/13	0.69	0.27	67,91,110,116	0
2	ALY	F	27	12/13	0.79	0.29	60,79,90,101	0
2	ALY	H	27	12/13	0.84	0.21	64,81,93,106	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.

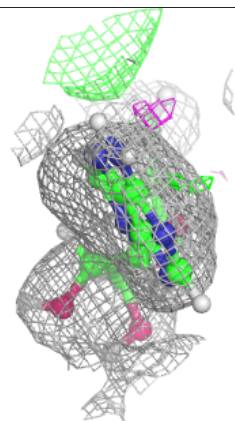
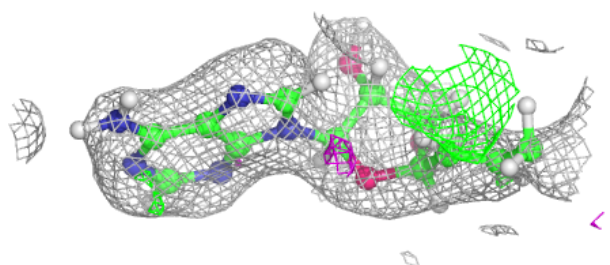
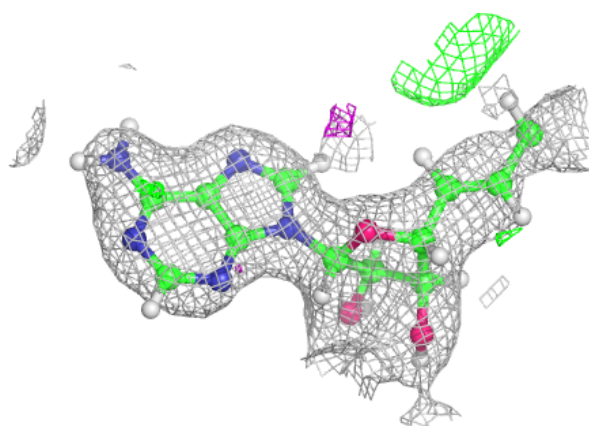
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	EDO	D	501	4/4	0.79	0.16	65,82,100,104	0
4	QVR	G	101	20/20	0.93	0.12	39,50,68,74	0
4	QVR	F	101	20/20	0.95	0.13	34,41,61,73	0
4	QVR	E	101	20/20	0.96	0.11	33,40,61,67	0
4	QVR	H	101	20/20	0.96	0.12	33,40,57,69	0

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.

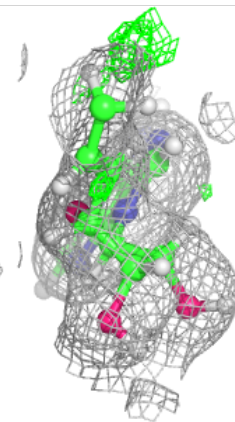
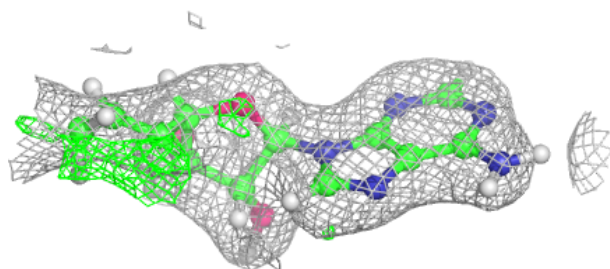
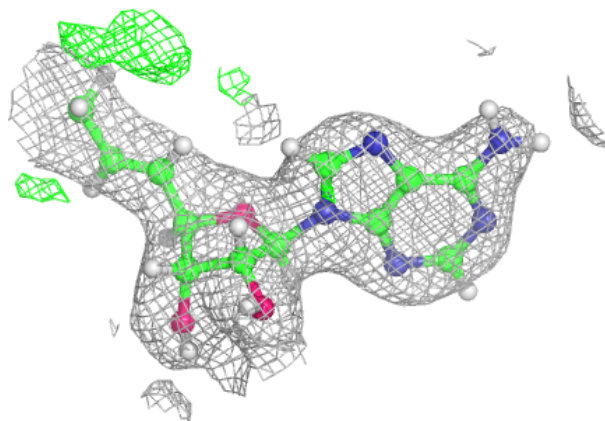


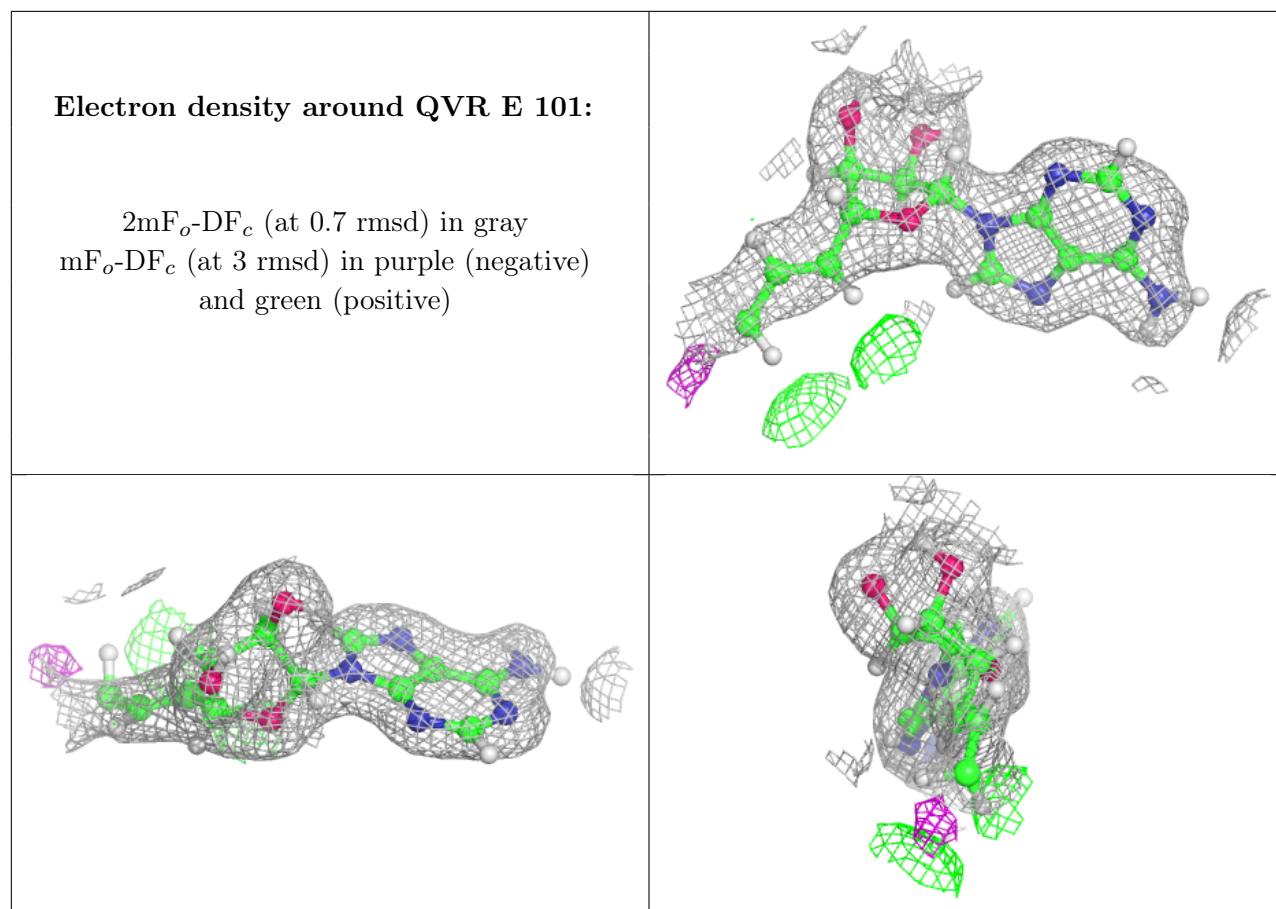
**Electron density around QVR G 101:**

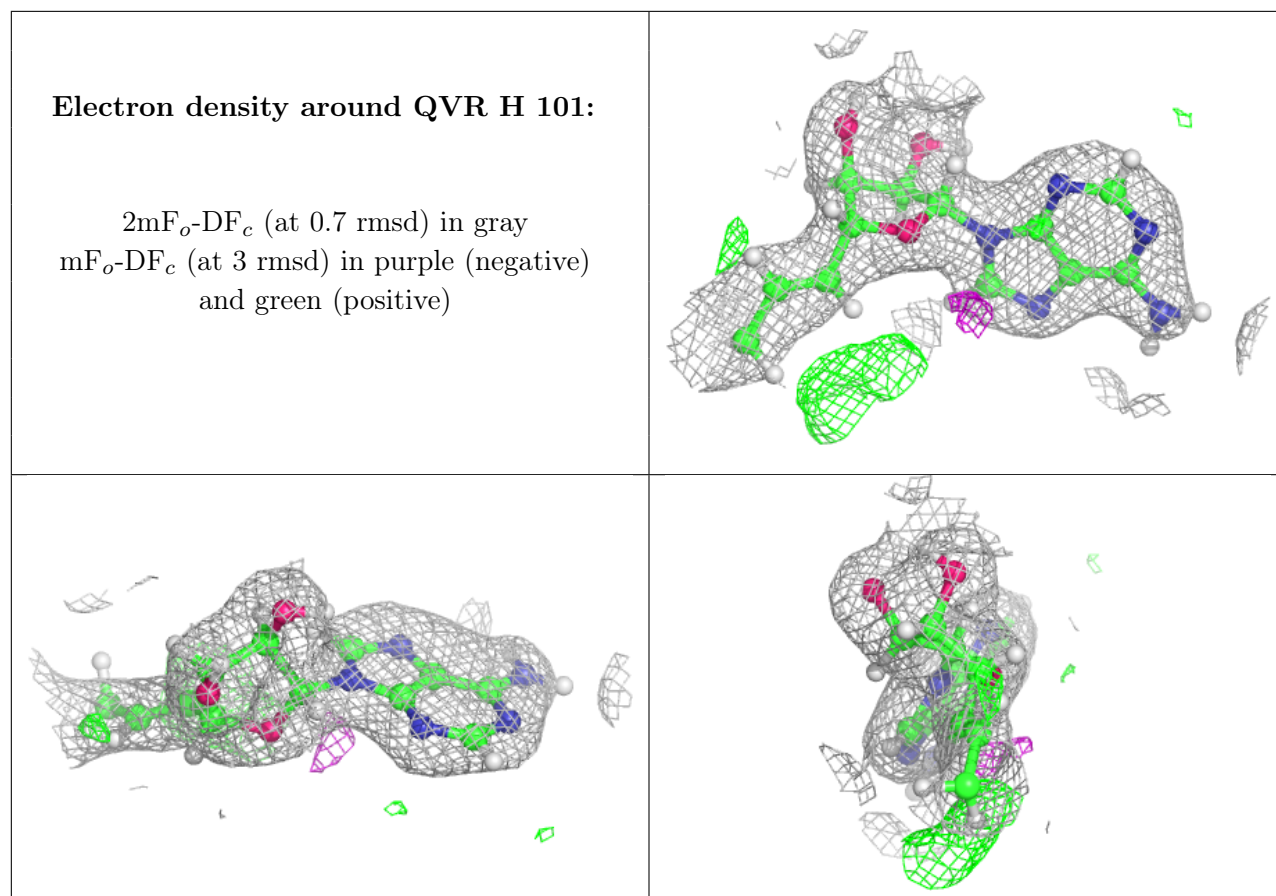
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

**Electron density around QVR F 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 [i](#)

There are no such residues in this entry.