



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 25, 2024 – 12:31 PM EDT

PDB ID : 6JTU  
Title : Crystal structure of MHETase from Ideonella sakaiensis  
Authors : Sagong, H.-Y.; Seo, H.; Kim, K.-J.  
Deposited on : 2019-04-12  
Resolution : 2.10 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	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
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)	:	2.37.1

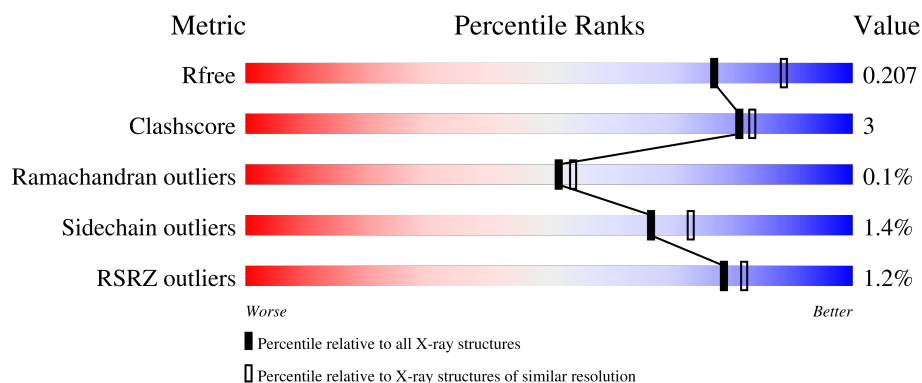
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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	613	<div> <div></div> <div>85% 6% 8%</div> </div>
1	B	613	<div> <div></div> <div>86% 5% 8%</div> </div>
1	C	613	<div> <div></div> <div>86% 6% 8%</div> </div>

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 13224 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 Mono(2-hydroxyethyl) terephthalate hydrolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	561	Total	C	N	O	S	0	0	0
			4140	2592	728	792	28			
1	B	561	Total	C	N	O	S	0	0	0
			4140	2592	728	792	28			
1	C	561	Total	C	N	O	S	0	0	0
			4140	2592	728	792	28			

There are 78 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-9	MET	-	initiating methionine	UNP A0A0K8P8E7
A	-8	MET	-	expression tag	UNP A0A0K8P8E7
A	-7	ILE	-	expression tag	UNP A0A0K8P8E7
A	-6	THR	-	expression tag	UNP A0A0K8P8E7
A	-5	LEU	-	expression tag	UNP A0A0K8P8E7
A	-4	ARG	-	expression tag	UNP A0A0K8P8E7
A	-3	LYS	-	expression tag	UNP A0A0K8P8E7
A	-2	LEU	-	expression tag	UNP A0A0K8P8E7
A	-1	PRO	-	expression tag	UNP A0A0K8P8E7
A	0	LEU	-	expression tag	UNP A0A0K8P8E7
A	1	ALA	-	expression tag	UNP A0A0K8P8E7
A	2	VAL	-	expression tag	UNP A0A0K8P8E7
A	3	ALA	-	expression tag	UNP A0A0K8P8E7
A	4	VAL	-	expression tag	UNP A0A0K8P8E7
A	5	ALA	-	expression tag	UNP A0A0K8P8E7
A	6	ALA	-	expression tag	UNP A0A0K8P8E7
A	7	GLY	-	expression tag	UNP A0A0K8P8E7
A	8	VAL	-	expression tag	UNP A0A0K8P8E7
A	9	MET	-	expression tag	UNP A0A0K8P8E7
A	10	SER	-	expression tag	UNP A0A0K8P8E7
A	11	ALA	-	expression tag	UNP A0A0K8P8E7
A	12	GLN	-	expression tag	UNP A0A0K8P8E7
A	13	ALA	-	expression tag	UNP A0A0K8P8E7

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	14	MET	-	expression tag	UNP A0A0K8P8E7
A	15	ALA	-	expression tag	UNP A0A0K8P8E7
A	16	MET	-	expression tag	UNP A0A0K8P8E7
B	-9	MET	-	initiating methionine	UNP A0A0K8P8E7
B	-8	MET	-	expression tag	UNP A0A0K8P8E7
B	-7	ILE	-	expression tag	UNP A0A0K8P8E7
B	-6	THR	-	expression tag	UNP A0A0K8P8E7
B	-5	LEU	-	expression tag	UNP A0A0K8P8E7
B	-4	ARG	-	expression tag	UNP A0A0K8P8E7
B	-3	LYS	-	expression tag	UNP A0A0K8P8E7
B	-2	LEU	-	expression tag	UNP A0A0K8P8E7
B	-1	PRO	-	expression tag	UNP A0A0K8P8E7
B	0	LEU	-	expression tag	UNP A0A0K8P8E7
B	1	ALA	-	expression tag	UNP A0A0K8P8E7
B	2	VAL	-	expression tag	UNP A0A0K8P8E7
B	3	ALA	-	expression tag	UNP A0A0K8P8E7
B	4	VAL	-	expression tag	UNP A0A0K8P8E7
B	5	ALA	-	expression tag	UNP A0A0K8P8E7
B	6	ALA	-	expression tag	UNP A0A0K8P8E7
B	7	GLY	-	expression tag	UNP A0A0K8P8E7
B	8	VAL	-	expression tag	UNP A0A0K8P8E7
B	9	MET	-	expression tag	UNP A0A0K8P8E7
B	10	SER	-	expression tag	UNP A0A0K8P8E7
B	11	ALA	-	expression tag	UNP A0A0K8P8E7
B	12	GLN	-	expression tag	UNP A0A0K8P8E7
B	13	ALA	-	expression tag	UNP A0A0K8P8E7
B	14	MET	-	expression tag	UNP A0A0K8P8E7
B	15	ALA	-	expression tag	UNP A0A0K8P8E7
B	16	MET	-	expression tag	UNP A0A0K8P8E7
C	-9	MET	-	initiating methionine	UNP A0A0K8P8E7
C	-8	MET	-	expression tag	UNP A0A0K8P8E7
C	-7	ILE	-	expression tag	UNP A0A0K8P8E7
C	-6	THR	-	expression tag	UNP A0A0K8P8E7
C	-5	LEU	-	expression tag	UNP A0A0K8P8E7
C	-4	ARG	-	expression tag	UNP A0A0K8P8E7
C	-3	LYS	-	expression tag	UNP A0A0K8P8E7
C	-2	LEU	-	expression tag	UNP A0A0K8P8E7
C	-1	PRO	-	expression tag	UNP A0A0K8P8E7
C	0	LEU	-	expression tag	UNP A0A0K8P8E7
C	1	ALA	-	expression tag	UNP A0A0K8P8E7
C	2	VAL	-	expression tag	UNP A0A0K8P8E7
C	3	ALA	-	expression tag	UNP A0A0K8P8E7

*Continued on next page...*

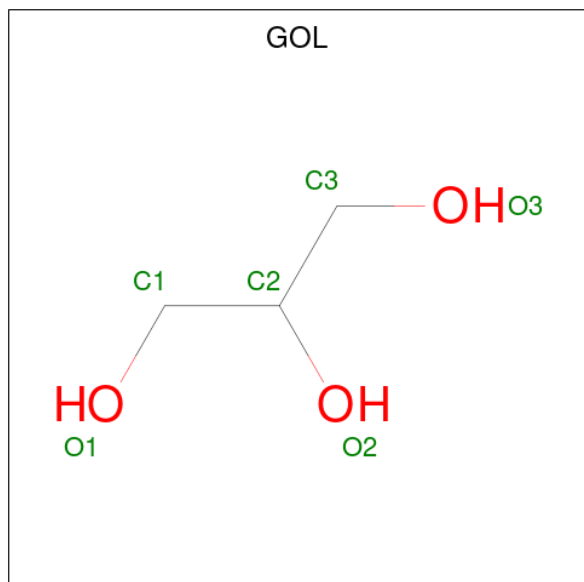
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	4	VAL	-	expression tag	UNP A0A0K8P8E7
C	5	ALA	-	expression tag	UNP A0A0K8P8E7
C	6	ALA	-	expression tag	UNP A0A0K8P8E7
C	7	GLY	-	expression tag	UNP A0A0K8P8E7
C	8	VAL	-	expression tag	UNP A0A0K8P8E7
C	9	MET	-	expression tag	UNP A0A0K8P8E7
C	10	SER	-	expression tag	UNP A0A0K8P8E7
C	11	ALA	-	expression tag	UNP A0A0K8P8E7
C	12	GLN	-	expression tag	UNP A0A0K8P8E7
C	13	ALA	-	expression tag	UNP A0A0K8P8E7
C	14	MET	-	expression tag	UNP A0A0K8P8E7
C	15	ALA	-	expression tag	UNP A0A0K8P8E7
C	16	MET	-	expression tag	UNP A0A0K8P8E7

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

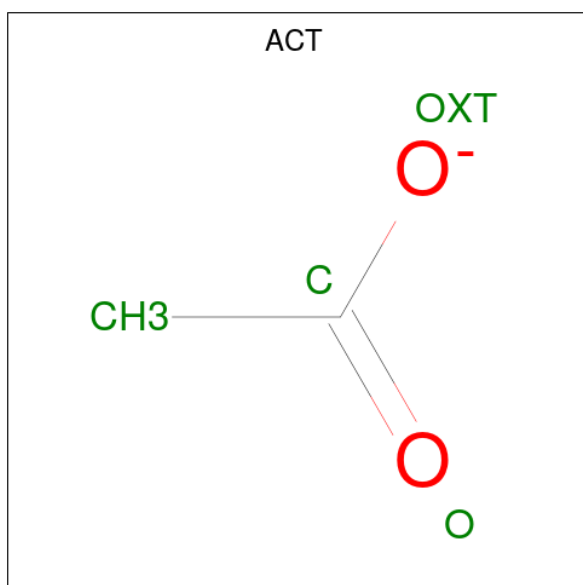
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Ca 1 1	0	0
2	B	1	Total Ca 1 1	0	0
2	C	1	Total Ca 1 1	0	0

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



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	C	1	Total C O 6 3 3	0	0
3	C	1	Total C O 6 3 3	0	0
3	C	1	Total C O 6 3 3	0	0

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



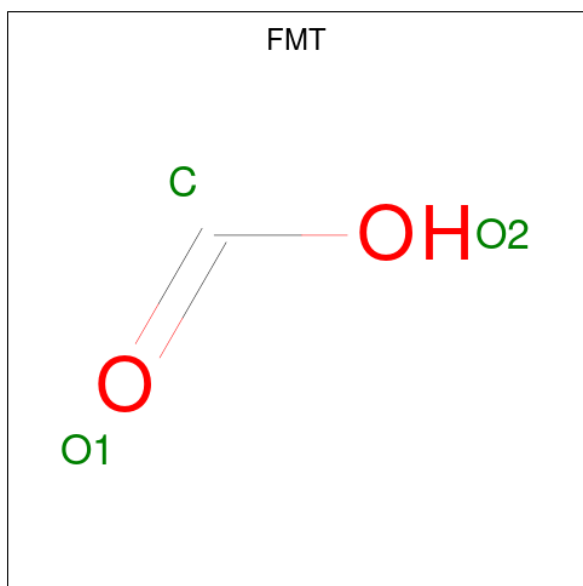
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 4 2 2	0	0
4	B	1	Total C O 4 2 2	0	0

*Continued on next page...*

*Continued from previous page...*

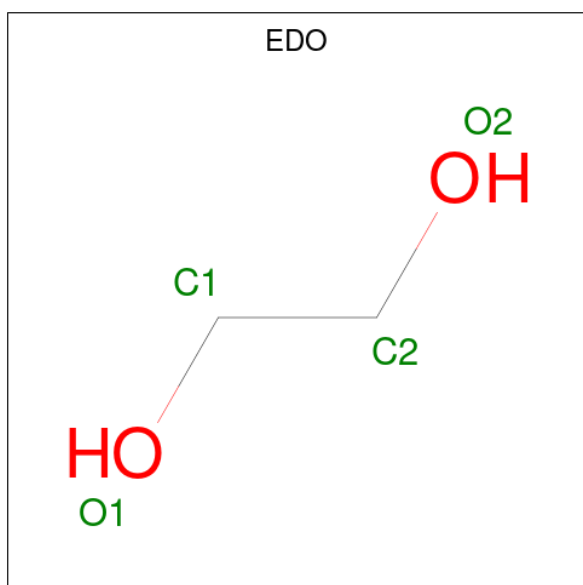
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	C	1	Total	C	O	0	0
			4	2	2		

- Molecule 5 is FORMIC ACID (three-letter code: FMT) (formula:  $\text{CH}_2\text{O}_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			3	1	2		

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $\text{C}_2\text{H}_6\text{O}_2$ ).



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

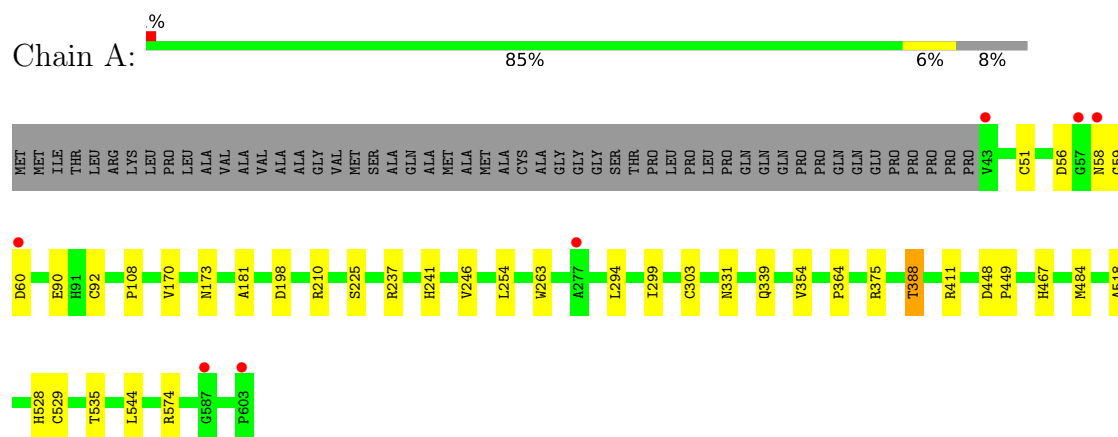
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	205	Total 205	O 205	0	0
7	B	260	Total 260	O 260	0	0
7	C	253	Total 253	O 253	0	0

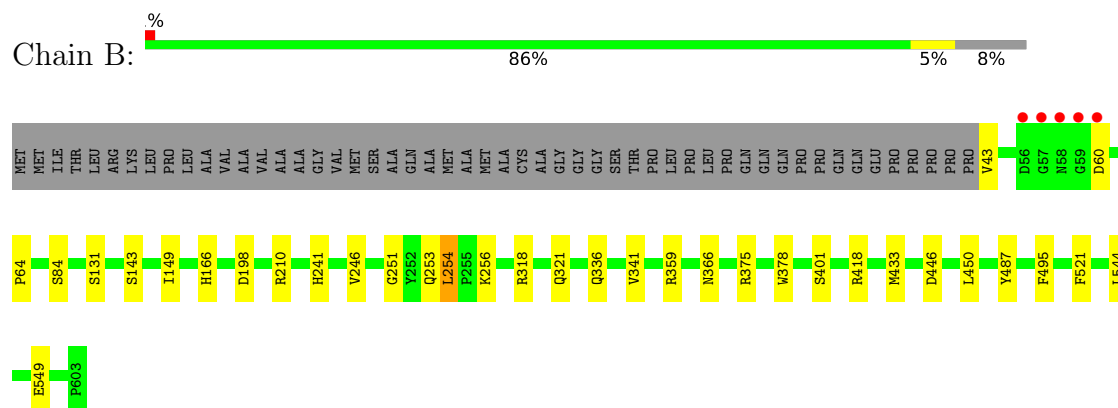
### 3 Residue-property plots [i](#)

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.

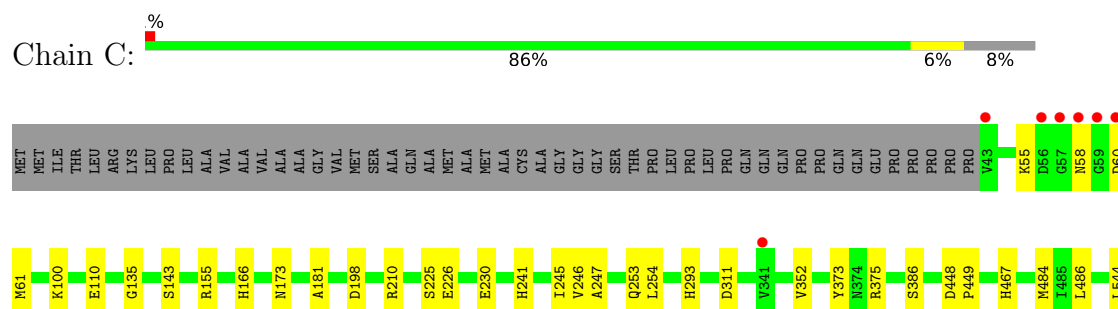
- Molecule 1: Mono(2-hydroxyethyl) terephthalate hydrolase



- Molecule 1: Mono(2-hydroxyethyl) terephthalate hydrolase



- Molecule 1: Mono(2-hydroxyethyl) terephthalate hydrolase





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	214.93Å 173.16Å 107.22Å 90.00° 119.86° 90.00°	Depositor
Resolution (Å)	33.21 – 2.10 33.19 – 2.10	Depositor EDS
% Data completeness (in resolution range)	97.6 (33.21-2.10) 97.7 (33.19-2.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.79 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, $R_{free}$	0.171 , 0.198 0.182 , 0.207	Depositor DCC
$R_{free}$ test set	9652 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	27.9	Xtriage
Anisotropy	0.114	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 40.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.014 for -h-2*k,l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	13224	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.00% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, ACT, CA, EDO, FMT

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	0.81	1/4248 (0.0%)	0.92	2/5785 (0.0%)
1	B	0.79	1/4248 (0.0%)	0.93	2/5785 (0.0%)
1	C	0.80	0/4248	0.91	0/5785
All	All	0.80	2/12744 (0.0%)	0.92	4/17355 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	90	GLU	CD-OE2	-5.70	1.19	1.25
1	B	549	GLU	CD-OE2	5.18	1.31	1.25

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	359	ARG	NE-CZ-NH2	6.92	123.76	120.30
1	A	388	THR	CA-CB-CG2	5.53	120.15	112.40
1	A	237	ARG	NE-CZ-NH1	5.41	123.00	120.30
1	B	318	ARG	NE-CZ-NH2	-5.04	117.78	120.30

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	4140	0	3931	21	0
1	B	4140	0	3929	18	0
1	C	4140	0	3929	22	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
3	A	6	0	8	0	0
3	B	36	0	48	2	0
3	C	18	0	24	4	0
4	A	4	0	3	0	0
4	B	4	0	3	0	0
4	C	4	0	3	0	0
5	A	3	0	1	1	0
6	A	8	0	12	0	0
7	A	205	0	0	0	0
7	B	260	0	0	2	0
7	C	253	0	0	5	0
All	All	13224	0	11891	60	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 (60) 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:253:GLN:HE22	3:C:704:GOL:H2	1.50	0.77
1:B:321:GLN:HE22	1:B:378:TRP:H	1.33	0.76
1:A:51:CYS:HG	1:A:92:CYS:HG	0.83	0.75
1:C:373:TYR:O	3:C:704:GOL:H11	1.90	0.71
1:A:60:ASP:HB2	1:A:210:ARG:HH12	1.55	0.70
1:C:166:HIS:HE1	7:C:863:HOH:O	1.81	0.63
1:A:56:ASP:HB2	1:A:59:GLY:H	1.64	0.63
1:C:60:ASP:OD1	1:C:210:ARG:NH2	2.32	0.62
1:B:166:HIS:HD2	7:B:925:HOH:O	1.85	0.58
1:A:246:VAL:HG23	1:A:544:LEU:HD22	1.87	0.57
1:C:247:ALA:HB3	1:C:486:LEU:HD23	1.88	0.55
1:A:60:ASP:HB3	1:A:210:ARG:HH22	1.71	0.55
1:B:166:HIS:HE1	7:B:851:HOH:O	1.88	0.55
1:A:198:ASP:OD1	1:A:241:HIS:HE1	1.90	0.54
1:C:135:GLY:HA2	1:C:166:HIS:O	2.07	0.54
1:C:155:ARG:NH2	7:C:801:HOH:O	2.35	0.54
1:B:84:SER:O	1:B:149:ILE:HG13	2.08	0.54

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:246:VAL:HG23	1:C:544:LEU:HD22	1.90	0.53
1:A:108:PRO:HG2	1:A:170:VAL:HG11	1.90	0.53
1:B:246:VAL:HG23	1:B:544:LEU:HD22	1.90	0.52
1:B:254:LEU:HD13	1:B:495:PHE:CE1	2.44	0.52
1:A:294:LEU:HD11	1:A:331:ASN:ND2	2.25	0.51
1:A:60:ASP:CB	1:A:210:ARG:HH12	2.22	0.51
1:A:411:ARG:HH21	5:A:704:FMT:H	1.74	0.51
1:C:352:VAL:HG23	7:C:1007:HOH:O	2.10	0.50
1:B:198:ASP:OD1	1:B:241:HIS:HE1	1.95	0.50
1:C:166:HIS:HD2	7:C:945:HOH:O	1.94	0.49
1:C:253:GLN:NE2	3:C:704:GOL:H2	2.24	0.48
1:C:143:SER:HA	3:C:702:GOL:H2	1.94	0.47
1:B:143:SER:HA	3:B:703:GOL:H2	1.97	0.47
1:C:226:GLU:OE2	1:C:230:GLU:OE1	2.32	0.47
1:A:331:ASN:C	1:A:331:ASN:OD1	2.53	0.47
1:A:528:HIS:O	1:A:529:CYS:HB2	2.15	0.46
1:C:293:HIS:HD2	7:C:1029:HOH:O	1.98	0.46
1:B:64:PRO:HA	3:B:705:GOL:H11	1.96	0.46
1:A:299:ILE:O	1:A:303:CYS:HB2	2.16	0.45
1:C:173:ASN:HB3	1:C:181:ALA:CB	2.46	0.45
1:B:254:LEU:HD13	1:B:495:PHE:HE1	1.80	0.45
1:A:173:ASN:HB3	1:A:181:ALA:CB	2.47	0.45
1:C:245:ILE:O	1:C:484:MET:HA	2.17	0.45
1:B:341:VAL:HG12	1:C:311:ASP:HA	2.00	0.44
1:B:418:ARG:HD3	1:B:433:MET:CE	2.48	0.44
1:A:108:PRO:HG2	1:A:170:VAL:CG1	2.48	0.44
1:B:366:ASN:HB2	1:B:446:ASP:OD1	2.18	0.44
1:B:253:GLN:HE22	1:B:256:LYS:NZ	2.15	0.44
1:B:60:ASP:OD1	1:B:210:ARG:NH2	2.49	0.43
1:A:263:TRP:CD2	1:A:364:PRO:HA	2.53	0.43
1:A:263:TRP:HB2	1:A:364:PRO:HB3	2.01	0.43
1:C:61:MET:HA	1:C:210:ARG:NH1	2.33	0.43
1:A:448:ASP:N	1:A:449:PRO:CD	2.82	0.43
1:C:100:LYS:HD2	1:C:110:GLU:OE1	2.19	0.42
1:A:535:THR:HG21	1:A:574:ARG:HB2	2.01	0.42
1:B:251:GLY:O	1:B:254:LEU:HB2	2.21	0.41
1:A:484:MET:O	1:A:518:ALA:HA	2.20	0.41
1:B:487:TYR:HA	1:B:521:PHE:O	2.20	0.41
1:C:198:ASP:OD1	1:C:241:HIS:HE1	2.03	0.41
1:C:448:ASP:N	1:C:449:PRO:CD	2.83	0.41
1:A:339:GLN:HE21	1:A:354:VAL:HB	1.86	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:55:LYS:O	1:C:58:ASN:HB2	2.21	0.41
1:B:198:ASP:OD1	1:B:241:HIS:CE1	2.74	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	559/613 (91%)	540 (97%)	18 (3%)	1 (0%)	47	49
1	B	559/613 (91%)	540 (97%)	19 (3%)	0	100	100
1	C	559/613 (91%)	540 (97%)	18 (3%)	1 (0%)	47	49
All	All	1677/1839 (91%)	1620 (97%)	55 (3%)	2 (0%)	51	54

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	225	SER
1	A	225	SER

### 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	407/446 (91%)	402 (99%)	5 (1%)	71	77

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	407/446 (91%)	400 (98%)	7 (2%)	60	67
1	C	407/446 (91%)	402 (99%)	5 (1%)	71	77
All	All	1221/1338 (91%)	1204 (99%)	17 (1%)	67	73

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

Mol	Chain	Res	Type
1	A	58	ASN
1	A	254	LEU
1	A	375	ARG
1	A	388	THR
1	A	467	HIS
1	B	43	VAL
1	B	131	SER
1	B	254	LEU
1	B	336	GLN
1	B	375	ARG
1	B	401	SER
1	B	450	LEU
1	C	254	LEU
1	C	375	ARG
1	C	386	SER
1	C	467	HIS
1	C	552	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (15) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	241	HIS
1	A	278	GLN
1	A	339	GLN
1	A	435	GLN
1	B	156	ASN
1	B	166	HIS
1	B	241	HIS
1	B	253	GLN
1	B	321	GLN
1	B	556	GLN
1	B	581	GLN
1	C	166	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	241	HIS
1	C	293	HIS
1	C	581	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

Of 19 ligands modelled in this entry, 3 are monoatomic - leaving 16 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
4	ACT	A	703	-	3,3,3	1.29	0	3,3,3	1.53	1 (33%)
3	GOL	B	704	-	5,5,5	0.27	0	5,5,5	0.52	0
3	GOL	C	703	-	5,5,5	0.47	0	5,5,5	1.27	1 (20%)
3	GOL	C	704	-	5,5,5	0.25	0	5,5,5	0.80	0
6	EDO	A	705	-	3,3,3	0.20	0	2,2,2	0.23	0
3	GOL	A	702	-	5,5,5	0.26	0	5,5,5	0.51	0
6	EDO	A	706	-	3,3,3	0.43	0	2,2,2	0.48	0
3	GOL	B	702	-	5,5,5	0.21	0	5,5,5	0.85	0
3	GOL	B	706	-	5,5,5	0.08	0	5,5,5	0.24	0
3	GOL	C	702	-	5,5,5	0.09	0	5,5,5	0.59	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	GOL	B	707	-	5,5,5	0.30	0	5,5,5	0.59	0
4	ACT	B	708	-	3,3,3	1.60	0	3,3,3	0.32	0
5	FMT	A	704	-	2,2,2	0.65	0	1,1,1	0.27	0
4	ACT	C	705	-	3,3,3	1.50	0	3,3,3	0.20	0
3	GOL	B	703	-	5,5,5	0.30	0	5,5,5	0.69	0
3	GOL	B	705	-	5,5,5	0.25	0	5,5,5	0.57	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
3	GOL	B	704	-	-	0/4/4/4	-
3	GOL	C	703	-	-	2/4/4/4	-
3	GOL	C	704	-	-	4/4/4/4	-
6	EDO	A	705	-	-	1/1/1/1	-
3	GOL	A	702	-	-	0/4/4/4	-
6	EDO	A	706	-	-	0/1/1/1	-
3	GOL	B	702	-	-	2/4/4/4	-
3	GOL	B	706	-	-	0/4/4/4	-
3	GOL	C	702	-	-	2/4/4/4	-
3	GOL	B	707	-	-	2/4/4/4	-
3	GOL	B	703	-	-	2/4/4/4	-
3	GOL	B	705	-	-	2/4/4/4	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	703	GOL	C3-C2-C1	2.32	120.71	111.70
4	A	703	ACT	O-C-CH3	-2.05	114.36	122.33

There are no chirality outliers.

All (17) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	702	GOL	O1-C1-C2-C3
3	B	707	GOL	O1-C1-C2-C3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
3	C	702	GOL	O1-C1-C2-C3
3	C	703	GOL	O1-C1-C2-C3
3	C	704	GOL	O1-C1-C2-C3
3	C	704	GOL	C1-C2-C3-O3
3	C	702	GOL	O1-C1-C2-O2
3	C	703	GOL	O1-C1-C2-O2
3	B	703	GOL	C1-C2-C3-O3
3	B	702	GOL	O1-C1-C2-O2
3	B	703	GOL	O2-C2-C3-O3
3	B	707	GOL	O1-C1-C2-O2
3	C	704	GOL	O2-C2-C3-O3
3	C	704	GOL	O1-C1-C2-O2
3	B	705	GOL	O1-C1-C2-C3
6	A	705	EDO	O1-C1-C2-O2
3	B	705	GOL	O2-C2-C3-O3

There are no ring outliers.

5 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	704	GOL	3	0
3	C	702	GOL	1	0
5	A	704	FMT	1	0
3	B	703	GOL	1	0
3	B	705	GOL	1	0

## 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	561/613 (91%)	-0.29	7 (1%) 79 82	22, 32, 47, 92	0
1	B	561/613 (91%)	-0.49	5 (0%) 84 86	22, 28, 43, 85	0
1	C	561/613 (91%)	-0.44	8 (1%) 75 78	23, 29, 42, 89	0
All	All	1683/1839 (91%)	-0.41	20 (1%) 79 82	22, 30, 46, 92	0

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	57	GLY	5.9
1	B	57	GLY	5.3
1	A	58	ASN	5.2
1	C	59	GLY	4.0
1	C	58	ASN	3.7
1	B	58	ASN	3.7
1	C	57	GLY	3.6
1	B	56	ASP	3.2
1	B	59	GLY	3.2
1	C	56	ASP	3.1
1	A	43	VAL	3.0
1	A	587	GLY	2.8
1	C	341	VAL	2.7
1	A	60	ASP	2.6
1	C	43	VAL	2.3
1	A	603	PRO	2.3
1	C	60	ASP	2.2
1	C	603	PRO	2.0
1	B	60	ASP	2.0
1	A	277	ALA	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

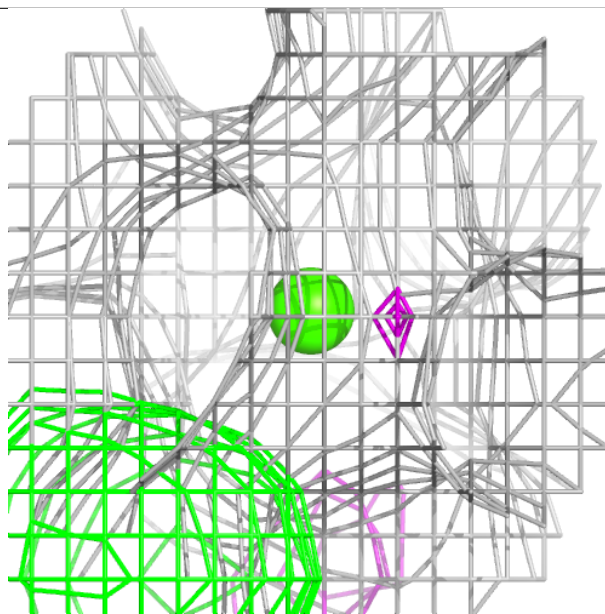
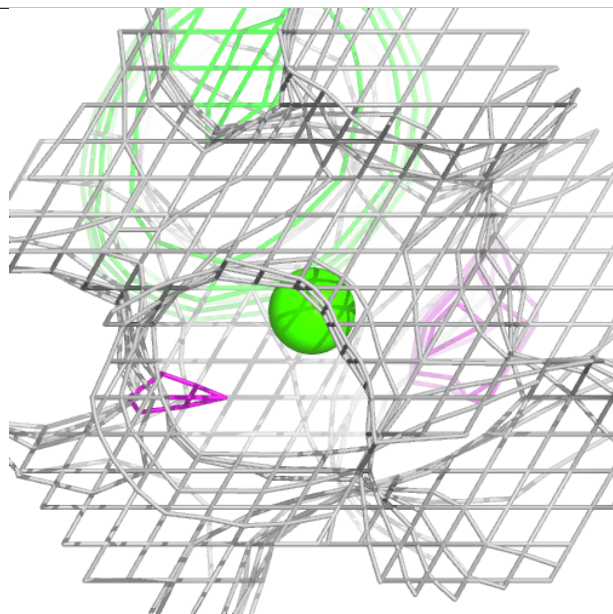
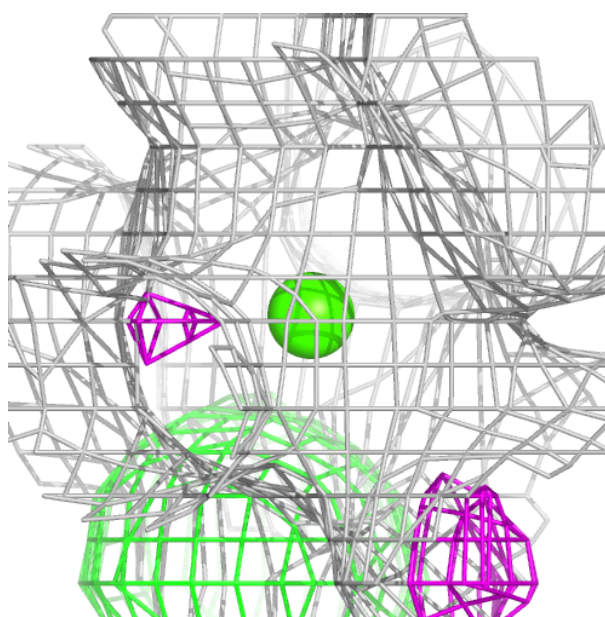
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	FMT	A	704	3/3	0.86	0.23	41,41,43,44	0
3	GOL	B	707	6/6	0.88	0.13	52,55,60,61	0
3	GOL	C	704	6/6	0.88	0.20	40,52,56,65	0
3	GOL	B	705	6/6	0.88	0.12	49,55,58,61	0
4	ACT	A	703	4/4	0.89	0.12	41,43,44,55	0
4	ACT	B	708	4/4	0.89	0.20	36,44,44,51	0
3	GOL	B	703	6/6	0.89	0.13	36,42,53,57	0
6	EDO	A	705	4/4	0.90	0.12	42,50,51,55	0
3	GOL	C	702	6/6	0.92	0.13	30,45,48,52	0
4	ACT	C	705	4/4	0.94	0.16	42,42,43,44	0
3	GOL	C	703	6/6	0.94	0.10	34,41,47,48	0
3	GOL	B	702	6/6	0.94	0.11	30,35,38,39	0
6	EDO	A	706	4/4	0.94	0.26	36,47,51,56	0
3	GOL	A	702	6/6	0.96	0.13	34,43,47,49	0
3	GOL	B	706	6/6	0.96	0.10	41,48,51,56	0
3	GOL	B	704	6/6	0.97	0.08	33,40,42,43	0
2	CA	B	701	1/1	0.99	0.05	32,32,32,32	0
2	CA	C	701	1/1	0.99	0.07	26,26,26,26	0
2	CA	A	701	1/1	0.99	0.06	36,36,36,36	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.

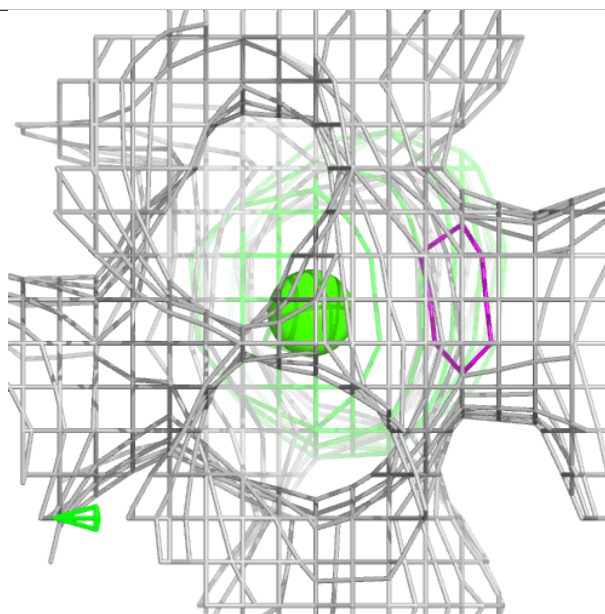
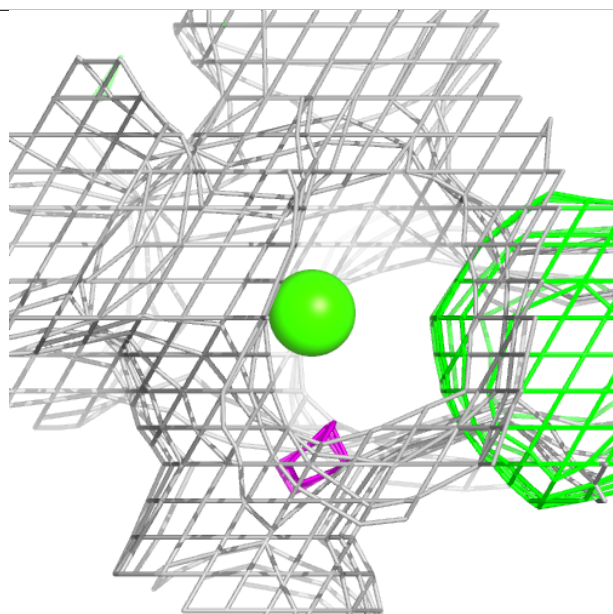
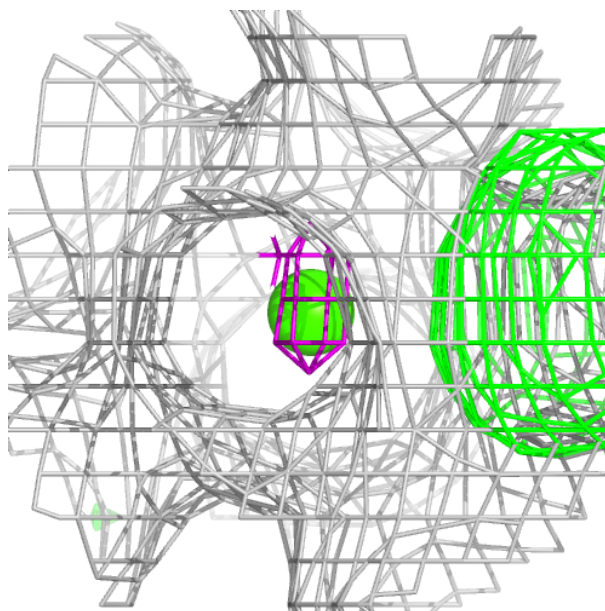
**Electron density around CA B 701:**

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



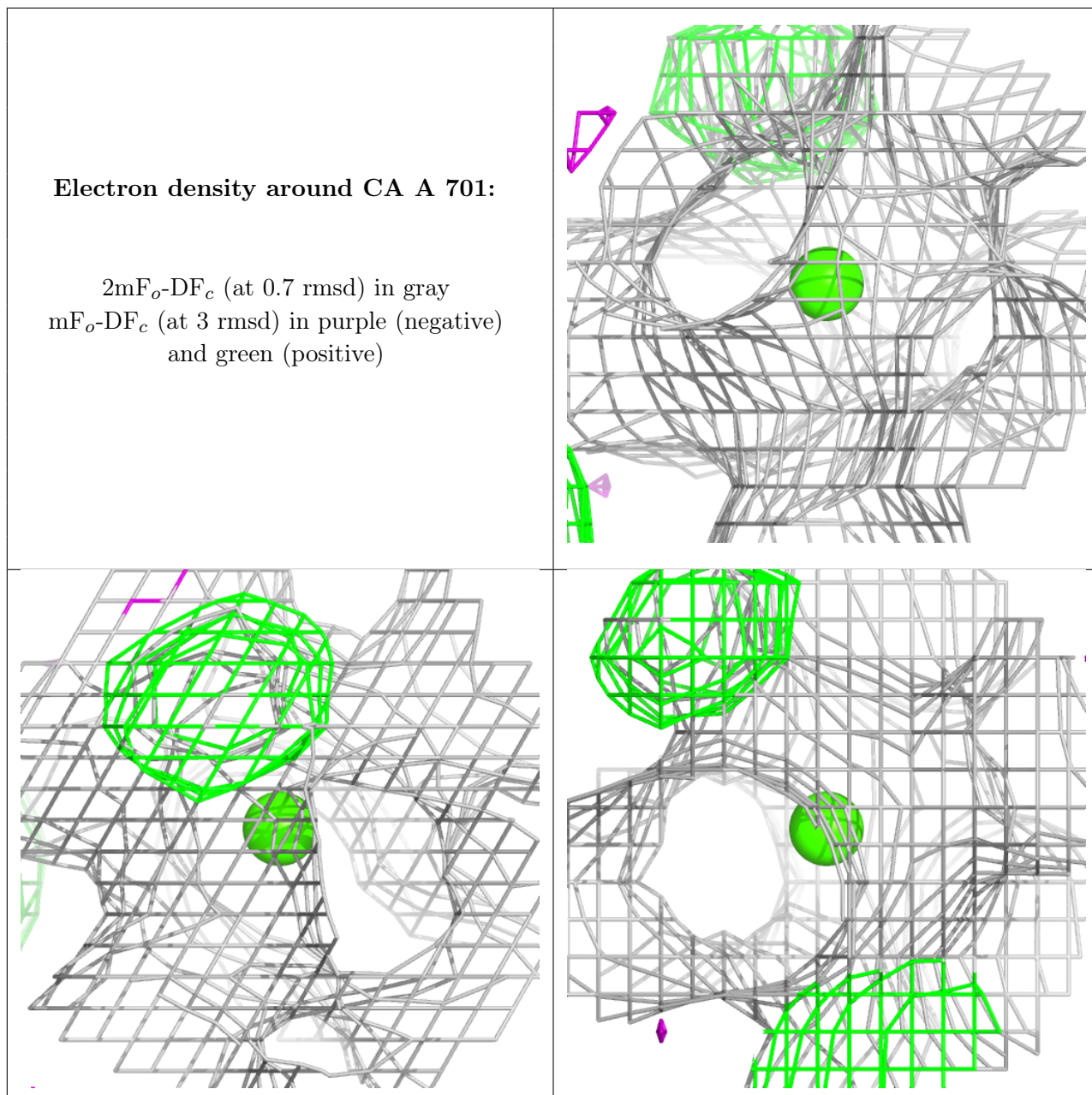
**Electron density around CA C 701:**

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



**Electron density around CA A 701:**

$2mF_o-DF_c$  (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.