Full wwPDB X-ray Structure Validation Report (i)

May 13, 2024 – 04:59 pm BST

PDB ID	:	8RW3
Title	:	Crystal Structure of Agd31B, alpha-transglucosylase, complexed with a non-covalent 1,2- Cyclophellitol aziridine
Authors	:	Moran, E.; Davies, G.; Ofamn, T.; Heming, J.; Nin-Hill, A.; Kullmer, F.; Steneker, R.; Klein, A.; Bennett, M.; Ruijgrok, G.; Kok, K.; Aerts, J.; Van der Marel, G.; Rovira, C.; Artola, M.; Codee, J.; Overkleeft, H.
Deposited on	:	2024-02-02
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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

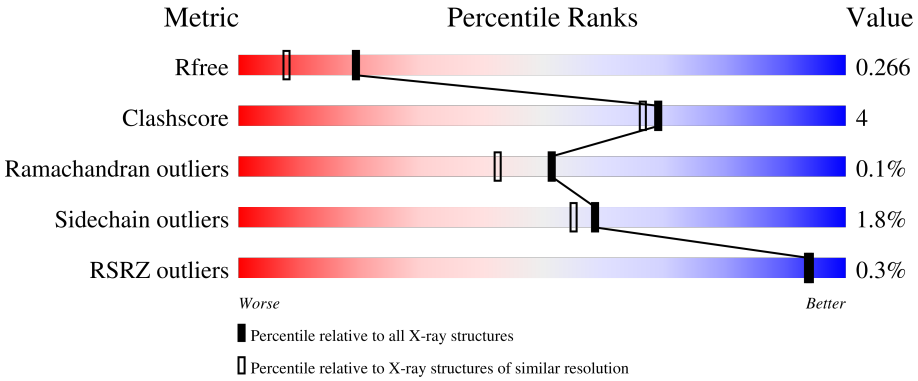
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 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 >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	817	86% 9% ·
1	B	817	85% 10% ·
1	C	817	86% 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	OXL	A	907	-	X	-	-

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 19733 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.

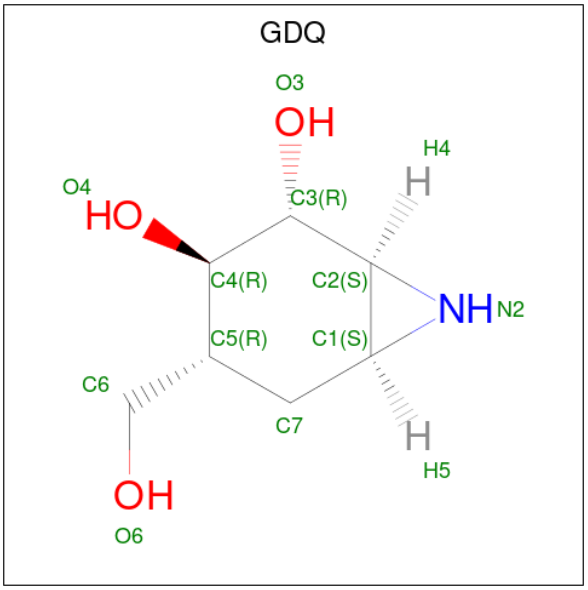
- Molecule 1 is a protein called Oligosaccharide 4-alpha-D-glucosyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	781	Total	C	N	O	S	0	0	0
			6172	3953	1045	1153	21			
1	A	781	Total	C	N	O	S	0	2	0
			6198	3969	1054	1154	21			
1	B	781	Total	C	N	O	S	0	1	0
			6207	3973	1053	1160	21			

There are 3 discrepancies between the modelled and reference sequences:

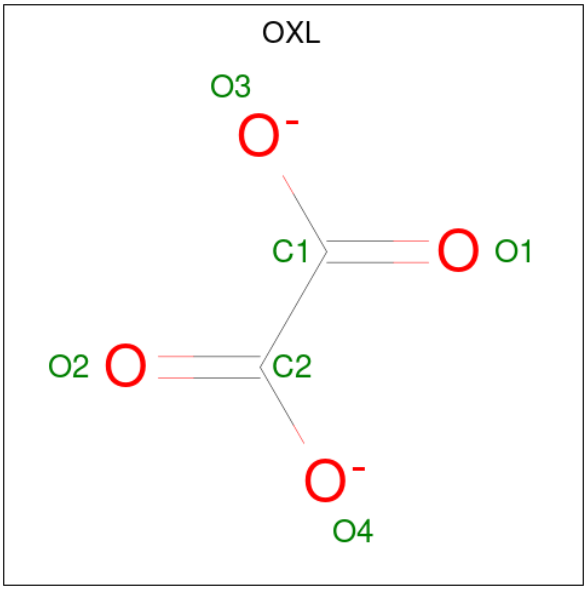
Chain	Residue	Modelled	Actual	Comment	Reference
C	817	LYS	-	expression tag	UNP B3PEE6
A	817	LYS	-	expression tag	UNP B3PEE6
B	817	LYS	-	expression tag	UNP B3PEE6

- Molecule 2 is (1 {S},2 {R},3 {R},4 {R},6 {S})-4-(hydroxymethyl)-7-azabicyclo[4.1.0]heptane-2,3-diol (three-letter code: GDQ) (formula: $C_7H_{13}NO_3$) (labeled as "Ligand of Interest" by depositor).



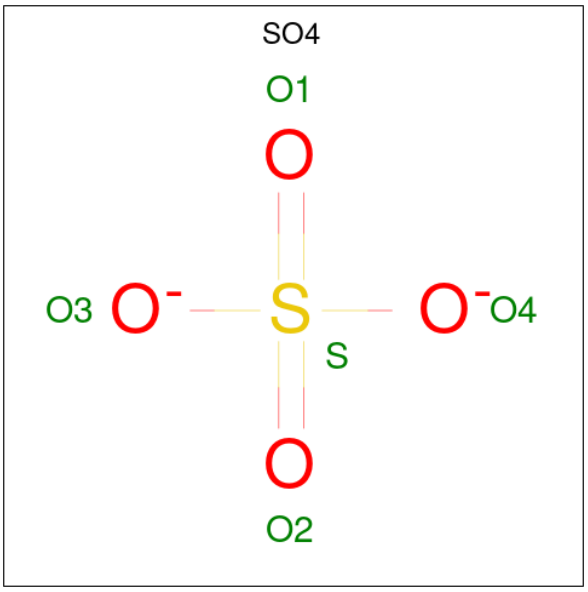
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	C	1	Total C N O 11 7 1 3	0	0
2	A	1	Total C N O 11 7 1 3	0	0
2	B	1	Total C N O 11 7 1 3	0	0

- Molecule 3 is OXALATE ION (three-letter code: OXL) (formula: C_2O_4).



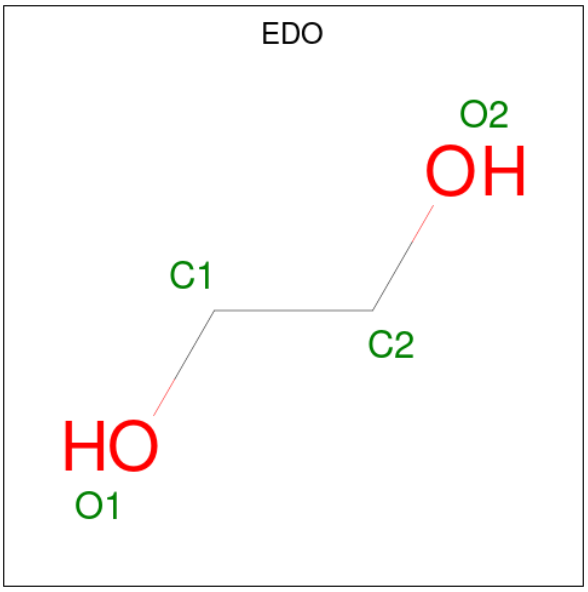
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	C	1	Total C O 6 2 4	0	0
3	A	1	Total C O 6 2 4	0	0
3	B	1	Total C O 6 2 4	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



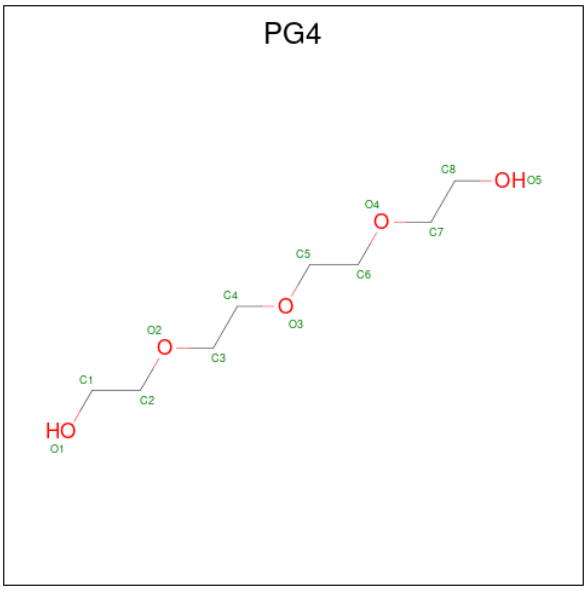
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	C	1	Total O S 5 4 1	0	0
4	C	1	Total O S 5 4 1	0	0
4	C	1	Total O S 5 4 1	0	0
4	C	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	B	1	Total O S 5 4 1	0	0
4	B	1	Total O S 5 4 1	0	0
4	B	1	Total O S 5 4 1	0	0
4	B	1	Total O S 5 4 1	0	0

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



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

- Molecule 6 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	C	1	Total	C	O	0	0
			13	8	5		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C O 13 8 5	0	0
6	B	1	Total C O 13 8 5	0	0

- Molecule 7 is water.

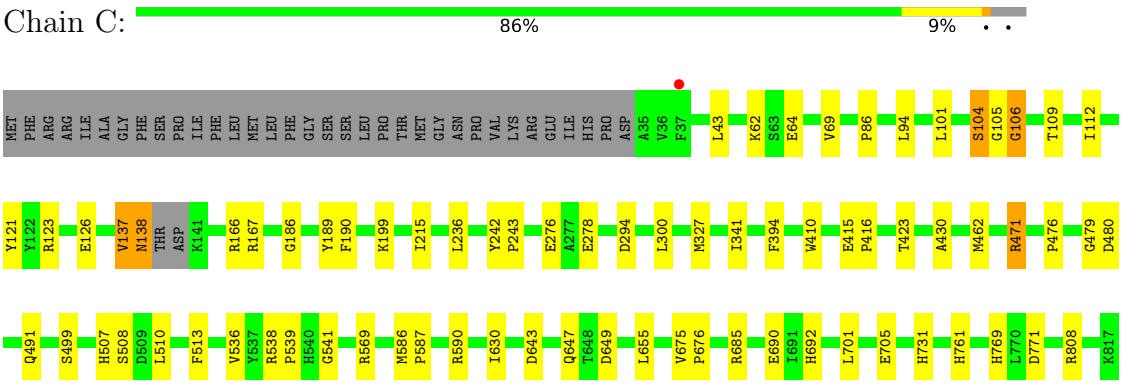
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	C	332	Total O 332 332	0	0
7	A	338	Total O 338 338	0	0
7	B	324	Total O 324 324	0	0

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.

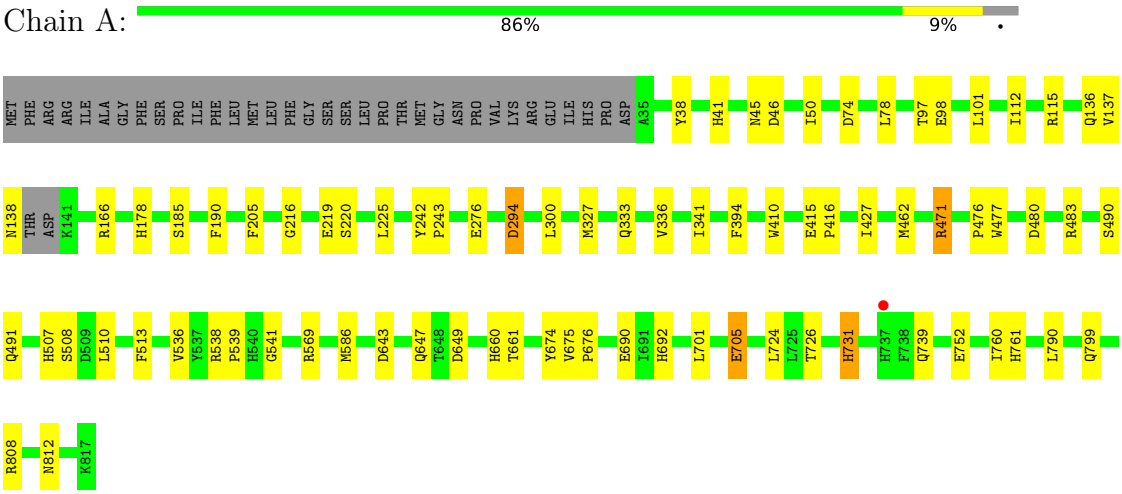
• Molecule 1: Oligosaccharide 4-alpha-D-glucosyltransferase

Chain C:



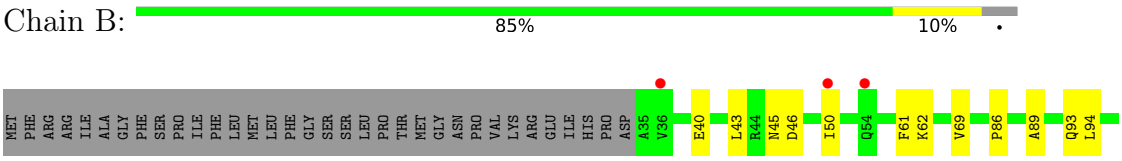
• Molecule 1: Oligosaccharide 4-alpha-D-glucosyltransferase

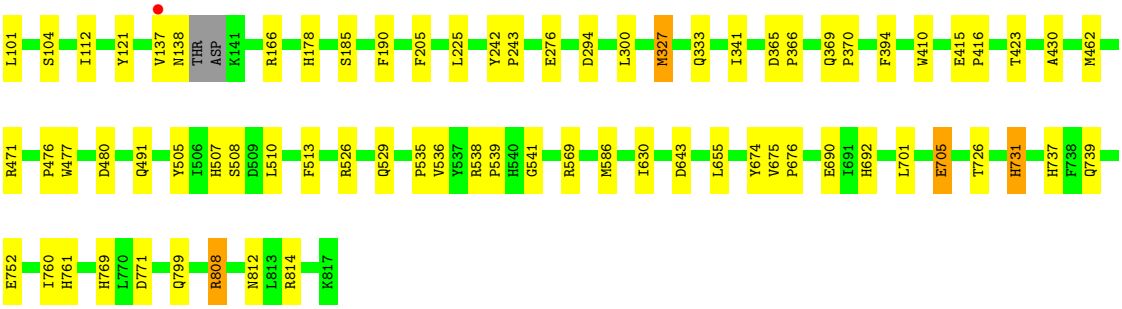
Chain A:



• Molecule 1: Oligosaccharide 4-alpha-D-glucosyltransferase

Chain B:




L101 S104 I112 Y121 V137 N138 THR ASP K141 R166 H178 S185 F190 F205 L225 Y242 P243 E276 D294 L300 M327 Q333 I341 D365 P366 Q369 P370 F394 W410 E415 P416 T423 A430 M462
R471 P476 W477 D460 Q491 Y505 I506 H507 S508 D509 L510 F513 R526 Q529 P535 V536 Y537 R538 P539 H540 G541 R569 M586 I630 D643 L655 Y674 V675 P676 E690 I691 H692 L701 E705 T726 H731 H737 F738 Q739
E752 I760 H761 H769 L770 D771 Q799 R808 N812 L813 R814 K817

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	197.02Å 341.25Å 102.81Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.82 – 1.90 49.77 – 1.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (49.82-1.90) 100.0 (49.77-1.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.23 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, R_{free}	0.217 , 0.261 0.224 , 0.266	Depositor DCC
R_{free} test set	13896 reflections (5.14%)	wwPDB-VP
Wilson B-factor (Å^2)	29.8	Xtriage
Anisotropy	0.510	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 30.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	0.477 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.470 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	19733	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.89% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

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

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: EDO, GDQ, OXL, SO4, PG4

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	0.44	0/6365	0.85	5/8638 (0.1%)
1	B	0.44	0/6373	0.86	6/8647 (0.1%)
1	C	0.44	0/6338	0.86	4/8605 (0.0%)
All	All	0.44	0/19076	0.86	15/25890 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	3
1	B	0	3
1	C	0	4
All	All	0	10

There are no bond length outliers.

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	471	ARG	NE-CZ-NH2	-6.51	117.05	120.30
1	A	471	ARG	NE-CZ-NH2	-5.87	117.37	120.30
1	A	333	GLN	CB-CA-C	-5.72	98.95	110.40
1	C	471	ARG	NE-CZ-NH2	-5.65	117.48	120.30
1	A	705	GLU	CB-CA-C	5.54	121.48	110.40
1	C	278	GLU	CB-CA-C	5.50	121.41	110.40
1	B	726	THR	CA-CB-OG1	-5.46	97.52	109.00
1	C	199	LYS	CB-CA-C	-5.42	99.55	110.40
1	B	471	ARG	NE-CZ-NH1	5.38	122.99	120.30
1	B	333	GLN	CB-CA-C	-5.33	99.74	110.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	705	GLU	CB-CA-C	5.27	120.94	110.40
1	A	483	ARG	NE-CZ-NH1	5.18	122.89	120.30
1	A	726	THR	CA-CB-OG1	-5.14	98.21	109.00
1	C	705	GLU	CB-CA-C	5.13	120.66	110.40
1	B	327	MET	CG-SD-CE	-5.04	92.14	100.20

There are no chirality outliers.

All (10) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	471	ARG	Sidechain
1	A	538	ARG	Sidechain
1	A	569	ARG	Sidechain
1	B	538	ARG	Sidechain
1	B	569	ARG	Sidechain
1	B	814	ARG	Sidechain
1	C	471	ARG	Sidechain
1	C	538	ARG	Sidechain
1	C	569	ARG	Sidechain
1	C	685	ARG	Sidechain

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	6198	0	5967	43	1
1	B	6207	0	5984	51	1
1	C	6172	0	5929	39	0
2	A	11	0	0	1	0
2	B	11	0	0	3	0
2	C	11	0	0	3	0
3	A	6	0	0	0	0
3	B	6	0	0	0	0
3	C	6	0	0	0	0
4	A	20	0	0	0	0
4	B	20	0	0	0	0
4	C	20	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	4	0	6	0	0
5	B	4	0	6	0	0
5	C	4	0	6	2	0
6	A	13	0	18	1	0
6	B	13	0	18	1	0
6	C	13	0	18	1	0
7	A	338	0	0	3	0
7	B	324	0	0	2	0
7	C	332	0	0	1	1
All	All	19733	0	17952	134	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 4.

All (134) 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:692:HIS:HD2	1:A:761:HIS:HE1	1.19	0.88
1:C:692:HIS:CD2	1:C:761:HIS:HE1	1.92	0.87
1:A:692:HIS:CD2	1:A:761:HIS:HE1	1.94	0.86
1:C:692:HIS:HD2	1:C:761:HIS:HE1	1.26	0.82
1:B:692:HIS:CD2	1:B:761:HIS:HE1	1.99	0.81
1:C:692:HIS:HD2	1:C:761:HIS:CE1	1.99	0.79
1:A:692:HIS:HD2	1:A:761:HIS:CE1	2.00	0.79
1:B:692:HIS:HD2	1:B:761:HIS:HE1	1.34	0.76
1:C:480:ASP:OD1	2:C:901:GDQ:N2	2.20	0.73
1:B:692:HIS:HD2	1:B:761:HIS:CE1	2.06	0.73
1:B:480:ASP:OD2	2:B:901:GDQ:N2	2.23	0.71
1:B:480:ASP:OD1	2:B:901:GDQ:N2	2.25	0.70
1:A:480:ASP:OD1	2:A:901:GDQ:N2	2.28	0.66
1:B:526:ARG:HH11	1:B:529:GLN:HE22	1.44	0.65
1:A:643:ASP:OD1	1:A:761:HIS:HD2	1.80	0.65
1:C:692:HIS:CD2	1:C:761:HIS:CE1	2.78	0.64
1:A:647:GLN:NE2	1:A:649:ASP:OD1	2.29	0.63
1:B:737:HIS:HE1	1:B:812:ASN:OD1	1.81	0.63
1:B:752:GLU:HG2	1:B:808:ARG:NH1	2.12	0.63
1:A:115:ARG:HG3	7:A:1186:HOH:O	1.97	0.63
1:B:643:ASP:OD1	1:B:761:HIS:HD2	1.82	0.63
1:C:643:ASP:OD1	1:C:761:HIS:HD2	1.83	0.61
1:B:480:ASP:CG	2:B:901:GDQ:N2	2.54	0.61
1:C:647:GLN:NE2	1:C:649:ASP:OD1	2.34	0.58

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:731:HIS:HD2	7:B:1067:HOH:O	1.84	0.58
1:A:136:GLN:HA	7:A:1130:HOH:O	2.04	0.57
1:B:137:VAL:O	1:B:138:ASN:C	2.42	0.57
1:B:50:ILE:HD11	1:B:112:ILE:HD13	1.86	0.56
1:B:692:HIS:CD2	1:B:761:HIS:CE1	2.84	0.56
1:B:369:GLN:HG3	1:B:370:PRO:HD2	1.87	0.56
1:C:480:ASP:OD2	2:C:901:GDQ:N2	2.39	0.55
1:C:462:MET:O	1:C:476:PRO:HA	2.06	0.55
1:B:93:GLN:HG2	1:B:104:SER:HB3	1.87	0.55
1:A:462:MET:O	1:A:476:PRO:HA	2.07	0.55
1:C:731:HIS:HD2	7:C:1119:HOH:O	1.88	0.55
1:C:480:ASP:CG	2:C:901:GDQ:N2	2.60	0.54
1:A:692:HIS:CD2	1:A:761:HIS:CE1	2.82	0.54
1:C:104:SER:OG	1:C:109:THR:OG1	2.18	0.54
1:C:137:VAL:O	1:C:138:ASN:C	2.46	0.54
1:A:675:VAL:HB	1:A:676:PRO:HD2	1.90	0.53
1:B:43:LEU:HD13	1:B:94:LEU:HD22	1.92	0.52
1:A:101:LEU:HB2	1:A:112:ILE:HB	1.91	0.52
1:B:507:HIS:CG	1:B:536:VAL:HB	2.45	0.52
1:C:43:LEU:HD13	1:C:94:LEU:HD22	1.91	0.52
1:C:630:ILE:HD12	1:C:655:LEU:HD13	1.92	0.52
1:A:415:GLU:N	1:A:416:PRO:HA	2.25	0.52
1:B:769:HIS:HE1	1:B:771:ASP:OD2	1.93	0.51
1:A:690:GLU:OE2	1:A:692:HIS:HE1	1.92	0.51
1:B:462:MET:O	1:B:476:PRO:HA	2.09	0.51
1:B:415:GLU:N	1:B:416:PRO:HA	2.25	0.51
1:C:415:GLU:N	1:C:416:PRO:HA	2.25	0.50
1:C:69:VAL:HG21	1:C:121:TYR:OH	2.12	0.50
1:C:690:GLU:OE2	1:C:692:HIS:HE1	1.95	0.49
1:A:341:ILE:HA	1:A:410:TRP:HB3	1.94	0.49
1:B:205:PHE:HB2	1:B:225:LEU:HD13	1.95	0.49
1:B:178:HIS:CE1	1:B:185:SER:HG	2.24	0.48
1:B:675:VAL:HB	1:B:676:PRO:HD2	1.96	0.48
1:B:690:GLU:OE2	1:B:692:HIS:HE1	1.95	0.48
1:C:112:ILE:N	1:C:112:ILE:HD12	2.28	0.48
1:A:276:GLU:HG3	1:A:327:MET:SD	2.55	0.47
1:C:513:PHE:HA	1:C:541:GLY:HA2	1.97	0.47
1:A:731:HIS:HD2	7:A:1061:HOH:O	1.97	0.47
1:A:205:PHE:HB2	1:A:225:LEU:HD13	1.96	0.46
1:B:341:ILE:HA	1:B:410:TRP:HB3	1.96	0.46
1:A:752:GLU:CG	1:A:808:ARG:NH1	2.78	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:526:ARG:HA	1:B:529:GLN:HE21	1.80	0.46
1:A:491:GLN:OE1	1:A:508:SER:HB3	2.16	0.46
1:A:243:PRO:HB3	1:A:586:MET:CE	2.45	0.46
1:C:64:GLU:HB3	1:C:106:GLY:HA3	1.98	0.45
1:C:186:GLY:H	5:C:906:EDO:H21	1.80	0.45
1:C:491:GLN:OE1	1:C:508:SER:HB3	2.16	0.45
1:C:675:VAL:HB	1:C:676:PRO:HD2	1.97	0.45
1:C:101:LEU:HB2	1:C:112:ILE:HB	1.98	0.45
1:A:510:LEU:HB2	1:A:539:PRO:HA	1.99	0.45
1:A:137:VAL:HG22	1:A:138:ASN:H	1.82	0.45
1:A:674:TYR:CZ	6:A:905:PG4:H71	2.51	0.44
1:A:243:PRO:HB3	1:A:586:MET:HE1	1.98	0.44
1:A:507:HIS:CG	1:A:536:VAL:HB	2.51	0.44
1:B:739:GLN:CD	1:B:812:ASN:HD21	2.21	0.44
1:B:243:PRO:HB3	1:B:586:MET:HE1	1.99	0.44
1:B:101:LEU:HB2	1:B:112:ILE:HB	1.98	0.44
1:B:341:ILE:HB	1:B:410:TRP:CE3	2.53	0.44
1:B:630:ILE:HD12	1:B:655:LEU:HD13	1.99	0.44
1:B:731:HIS:CD2	7:B:1067:HOH:O	2.66	0.44
6:C:907:PG4:C6	6:C:907:PG4:H32	2.44	0.43
1:A:705:GLU:HG3	1:A:724:LEU:HD11	1.99	0.43
1:B:45:ASN:O	1:B:46:ASP:HB2	2.17	0.43
1:C:189:TYR:CE2	1:C:479:GLY:HA3	2.53	0.43
1:A:513:PHE:HA	1:A:541:GLY:HA2	1.99	0.43
1:B:760:ILE:O	1:B:799:GLN:HA	2.18	0.43
1:C:507:HIS:CG	1:C:536:VAL:HB	2.54	0.43
1:B:769:HIS:CE1	1:B:771:ASP:OD2	2.69	0.43
1:A:50:ILE:HD11	1:A:112:ILE:HD13	2.01	0.43
1:A:341:ILE:HB	1:A:410:TRP:CE3	2.53	0.43
1:B:69:VAL:HG21	1:B:121:TYR:OH	2.19	0.43
1:B:674:TYR:CZ	6:B:904:PG4:H22	2.54	0.43
1:A:178:HIS:CE1	1:A:185:SER:HG	2.29	0.43
1:B:423:THR:HB	1:B:430:ALA:HB2	2.01	0.43
1:C:62:LYS:NZ	1:C:86:PRO:O	2.51	0.42
1:A:178:HIS:ND1	1:A:185:SER:OG	2.37	0.42
1:A:760:ILE:O	1:A:799:GLN:HA	2.19	0.42
1:B:477:TRP:HA	1:B:507:HIS:O	2.19	0.42
1:C:341:ILE:HB	1:C:410:TRP:CE3	2.54	0.42
1:C:423:THR:HB	1:C:430:ALA:HB2	2.01	0.42
1:B:276:GLU:HG3	1:B:327:MET:SD	2.59	0.42
1:B:491:GLN:OE1	1:B:508:SER:HB3	2.20	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:510:LEU:HB2	1:C:539:PRO:HA	2.02	0.42
1:A:78:LEU:HD22	1:A:427:ILE:HD12	2.01	0.42
1:B:752:GLU:HG2	1:B:808:ARG:HH12	1.84	0.42
1:A:294:ASP:O	1:A:336:VAL:HA	2.19	0.42
1:B:365:ASP:HB2	1:B:366:PRO:CD	2.49	0.42
1:C:105:GLY:O	1:C:106:GLY:C	2.58	0.42
1:C:186:GLY:H	5:C:906:EDO:C2	2.32	0.42
1:A:38:TYR:OH	1:A:41:HIS:ND1	2.41	0.42
1:B:62:LYS:NZ	1:B:86:PRO:O	2.51	0.41
1:A:660[A]:HIS:CD2	1:A:661:THR:HG23	2.56	0.41
1:B:510:LEU:HB2	1:B:539:PRO:HA	2.03	0.41
1:B:513:PHE:HA	1:B:541:GLY:HA2	2.01	0.41
1:A:45:ASN:O	1:A:46:ASP:HB2	2.18	0.41
1:A:97:THR:O	1:A:98:GLU:C	2.58	0.41
1:B:242:TYR:N	1:B:243:PRO:CD	2.84	0.41
1:B:243:PRO:HB3	1:B:586:MET:CE	2.50	0.41
1:C:243:PRO:HB3	1:C:586:MET:HE1	2.02	0.41
1:C:167:ARG:HA	1:C:215:ILE:HG22	2.02	0.41
1:A:739:GLN:CD	1:A:812:ASN:HD21	2.23	0.41
1:B:61:PHE:O	1:B:89:ALA:HA	2.21	0.41
1:C:242:TYR:N	1:C:243:PRO:CD	2.84	0.40
1:A:216:GLY:HA2	1:A:219:GLU:O	2.21	0.40
1:A:242:TYR:N	1:A:243:PRO:CD	2.84	0.40
1:B:505:TYR:HA	1:B:535:PRO:HB3	2.03	0.40
1:C:276:GLU:HG3	1:C:327:MET:SD	2.60	0.40
1:C:499:SER:HB3	1:C:587:PRO:CB	2.52	0.40
1:A:477:TRP:HA	1:A:507:HIS:O	2.21	0.40
1:C:769:HIS:HE1	1:C:771:ASP:OD2	2.04	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)
7:C:1332:HOH:O	7:C:1332:HOH:O[4_546]	1.95	0.25
1:A:812:ASN:ND2	1:B:369:GLN:OE1[8_545]	2.15	0.05

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	779/817 (95%)	747 (96%)	32 (4%)	0	100	100
1	B	778/817 (95%)	746 (96%)	32 (4%)	0	100	100
1	C	777/817 (95%)	746 (96%)	29 (4%)	2 (0%)	41	31
All	All	2334/2451 (95%)	2239 (96%)	93 (4%)	2 (0%)	51	42

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	106	GLY
1	C	137	VAL

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	642/690 (93%)	631 (98%)	11 (2%)	60	57
1	B	645/690 (94%)	635 (98%)	10 (2%)	62	60
1	C	637/690 (92%)	624 (98%)	13 (2%)	55	51
All	All	1924/2070 (93%)	1890 (98%)	34 (2%)	59	55

All (34) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	104	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	123	ARG
1	C	126	GLU
1	C	138	ASN
1	C	166	ARG
1	C	190	PHE
1	C	236	LEU
1	C	294	ASP
1	C	300	LEU
1	C	394	PHE
1	C	590	ARG
1	C	701	LEU
1	C	808	ARG
1	A	74	ASP
1	A	166	ARG
1	A	190	PHE
1	A	220	SER
1	A	294	ASP
1	A	300	LEU
1	A	394	PHE
1	A	490	SER
1	A	701	LEU
1	A	731	HIS
1	A	790	LEU
1	B	40	GLU
1	B	166	ARG
1	B	190	PHE
1	B	294	ASP
1	B	300	LEU
1	B	394	PHE
1	B	701	LEU
1	B	705	GLU
1	B	731	HIS
1	B	808	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (36) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	102	GLN
1	C	320	ASN
1	C	448	GLN
1	C	449	GLN
1	C	660	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	692	HIS
1	C	731	HIS
1	C	761	HIS
1	C	769	HIS
1	C	816	HIS
1	A	184	HIS
1	A	320	ASN
1	A	449	GLN
1	A	452	GLN
1	A	552	GLN
1	A	692	HIS
1	A	731	HIS
1	A	734	ASN
1	A	761	HIS
1	A	769	HIS
1	B	102	GLN
1	B	184	HIS
1	B	320	ASN
1	B	334	GLN
1	B	448	GLN
1	B	449	GLN
1	B	452	GLN
1	B	529	GLN
1	B	552	GLN
1	B	622	GLN
1	B	692	HIS
1	B	731	HIS
1	B	737	HIS
1	B	761	HIS
1	B	769	HIS
1	B	812	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

24 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	SO4	A	908	-	4,4,4	0.30	0	6,6,6	0.08	0
3	OXL	C	902	-	5,5,5	2.04	2 (40%)	6,6,6	0.91	0
3	OXL	A	907	-	5,5,5	1.78	2 (40%)	6,6,6	1.12	0
4	SO4	C	905	-	4,4,4	0.37	0	6,6,6	0.16	0
4	SO4	B	906	-	4,4,4	0.35	0	6,6,6	0.16	0
4	SO4	C	903	-	4,4,4	0.37	0	6,6,6	0.10	0
5	EDO	A	902	-	3,3,3	0.51	0	2,2,2	0.65	0
4	SO4	A	904	-	4,4,4	0.35	0	6,6,6	0.15	0
4	SO4	B	902	-	4,4,4	0.39	0	6,6,6	0.11	0
4	SO4	A	906	-	4,4,4	0.39	0	6,6,6	0.08	0
2	GDQ	A	901	-	11,12,12	0.86	0	11,18,18	2.70	4 (36%)
5	EDO	B	907	-	3,3,3	0.42	0	2,2,2	0.10	0
4	SO4	C	904	-	4,4,4	0.39	0	6,6,6	0.08	0
4	SO4	C	908	-	4,4,4	0.32	0	6,6,6	0.12	0
4	SO4	A	903	-	4,4,4	0.40	0	6,6,6	0.11	0
4	SO4	B	905	-	4,4,4	0.40	0	6,6,6	0.12	0
6	PG4	C	907	-	12,12,12	0.40	0	11,11,11	0.32	0
6	PG4	A	905	-	12,12,12	0.34	0	11,11,11	0.23	0
6	PG4	B	904	-	12,12,12	0.32	0	11,11,11	0.22	0
4	SO4	B	908	-	4,4,4	0.32	0	6,6,6	0.13	0
2	GDQ	B	901	-	11,12,12	0.82	0	11,18,18	2.82	5 (45%)
2	GDQ	C	901	-	11,12,12	0.75	0	11,18,18	2.64	4 (36%)
5	EDO	C	906	-	3,3,3	0.35	0	2,2,2	0.59	0
3	OXL	B	903	-	5,5,5	2.45	2 (40%)	6,6,6	0.60	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	EDO	B	907	-	-	1/1/1/1	-
5	EDO	A	902	-	-	0/1/1/1	-
2	GDQ	B	901	-	-	2/2/23/23	0/2/2/2
6	PG4	C	907	-	-	8/10/10/10	-
2	GDQ	C	901	-	-	2/2/23/23	0/2/2/2
3	OXL	C	902	-	-	3/4/4/4	-
5	EDO	C	906	-	-	1/1/1/1	-
3	OXL	A	907	-	-	4/4/4/4	-
6	PG4	A	905	-	-	8/10/10/10	-
6	PG4	B	904	-	-	6/10/10/10	-
2	GDQ	A	901	-	-	2/2/23/23	0/2/2/2
3	OXL	B	903	-	-	3/4/4/4	-

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	903	OXL	O3-C1	-4.31	1.18	1.30
3	C	902	OXL	O4-C2	-3.34	1.20	1.30
3	A	907	OXL	O4-C2	-3.05	1.21	1.30
3	C	902	OXL	O3-C1	-2.85	1.22	1.30
3	B	903	OXL	O4-C2	-2.84	1.22	1.30
3	A	907	OXL	O3-C1	-2.50	1.23	1.30

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	901	GDQ	C5-C7-C1	6.41	116.52	108.49
2	C	901	GDQ	C5-C7-C1	6.40	116.52	108.49
2	A	901	GDQ	C5-C7-C1	6.17	116.22	108.49
2	A	901	GDQ	O4-C4-C3	-4.75	99.37	110.35
2	C	901	GDQ	O4-C4-C3	-3.95	101.21	110.35
2	B	901	GDQ	O4-C4-C3	-3.71	101.77	110.35
2	B	901	GDQ	C7-C5-C4	3.52	114.22	110.17
2	C	901	GDQ	C7-C5-C4	3.25	113.90	110.17
2	B	901	GDQ	O3-C3-C2	2.88	115.47	109.66
2	A	901	GDQ	C7-C5-C4	2.72	113.30	110.17
2	C	901	GDQ	O3-C3-C2	2.23	114.16	109.66

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	901	GDQ	O6-C6-C5	-2.05	106.66	111.36
2	A	901	GDQ	O6-C6-C5	-2.04	106.68	111.36

There are no chirality outliers.

All (40) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	901	GDQ	C4-C5-C6-O6
2	A	901	GDQ	C4-C5-C6-O6
3	A	907	OXL	O1-C1-C2-O2
3	A	907	OXL	O1-C1-C2-O4
3	A	907	OXL	O3-C1-C2-O2
3	A	907	OXL	O3-C1-C2-O4
6	C	907	PG4	C3-C4-O3-C5
6	A	905	PG4	O4-C7-C8-O5
6	C	907	PG4	O2-C3-C4-O3
6	A	905	PG4	O1-C1-C2-O2
6	C	907	PG4	O1-C1-C2-O2
6	B	904	PG4	O3-C5-C6-O4
2	C	901	GDQ	C7-C5-C6-O6
6	B	904	PG4	O1-C1-C2-O2
3	B	903	OXL	O3-C1-C2-O2
6	A	905	PG4	O2-C3-C4-O3
6	C	907	PG4	C8-C7-O4-C6
6	A	905	PG4	O3-C5-C6-O4
2	B	901	GDQ	C4-C5-C6-O6
2	A	901	GDQ	C7-C5-C6-O6
5	C	906	EDO	O1-C1-C2-O2
6	B	904	PG4	C5-C6-O4-C7
3	C	902	OXL	O1-C1-C2-O2
6	C	907	PG4	C4-C3-O2-C2
6	A	905	PG4	C5-C6-O4-C7
6	B	904	PG4	C6-C5-O3-C4
6	A	905	PG4	C1-C2-O2-C3
3	B	903	OXL	O1-C1-C2-O2
3	C	902	OXL	O1-C1-C2-O4
5	B	907	EDO	O1-C1-C2-O2
3	B	903	OXL	O3-C1-C2-O4
2	B	901	GDQ	C7-C5-C6-O6
6	A	905	PG4	C3-C4-O3-C5
3	C	902	OXL	O3-C1-C2-O4
6	C	907	PG4	C5-C6-O4-C7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
6	B	904	PG4	C8-C7-O4-C6
6	A	905	PG4	C4-C3-O2-C2
6	B	904	PG4	O2-C3-C4-O3
6	C	907	PG4	O3-C5-C6-O4
6	C	907	PG4	O4-C7-C8-O5

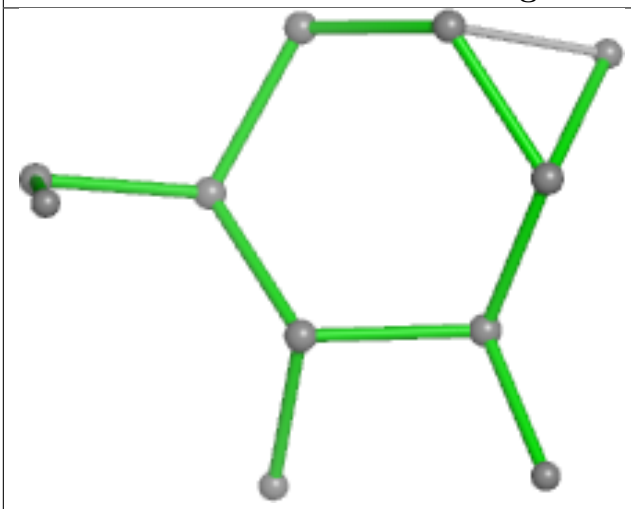
There are no ring outliers.

7 monomers are involved in 12 short contacts:

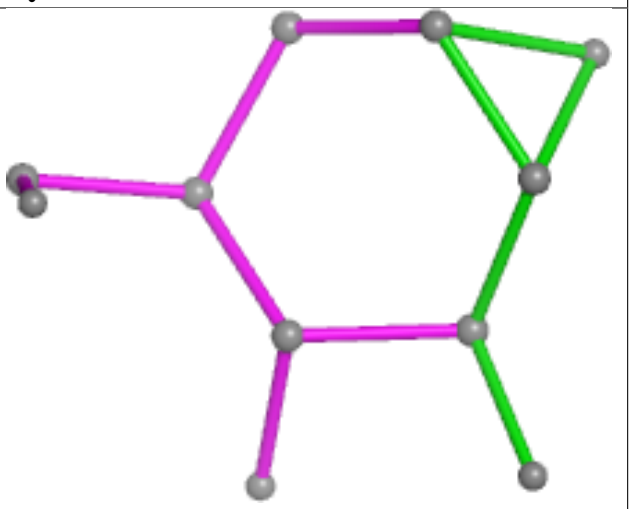
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	901	GDQ	1	0
6	C	907	PG4	1	0
6	A	905	PG4	1	0
6	B	904	PG4	1	0
2	B	901	GDQ	3	0
2	C	901	GDQ	3	0
5	C	906	EDO	2	0

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

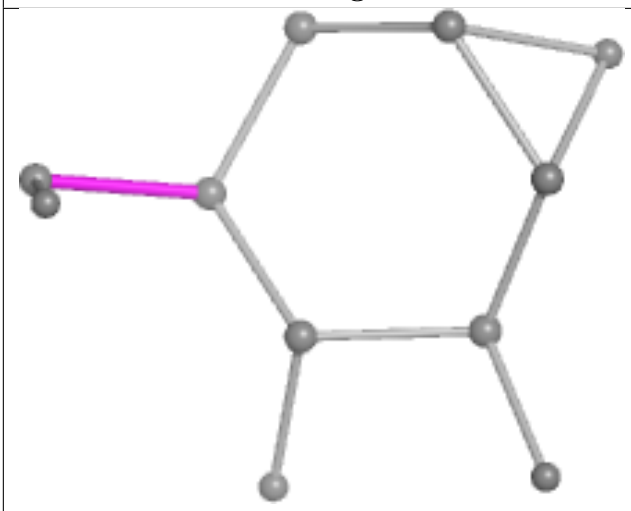
Ligand GDQ A 901



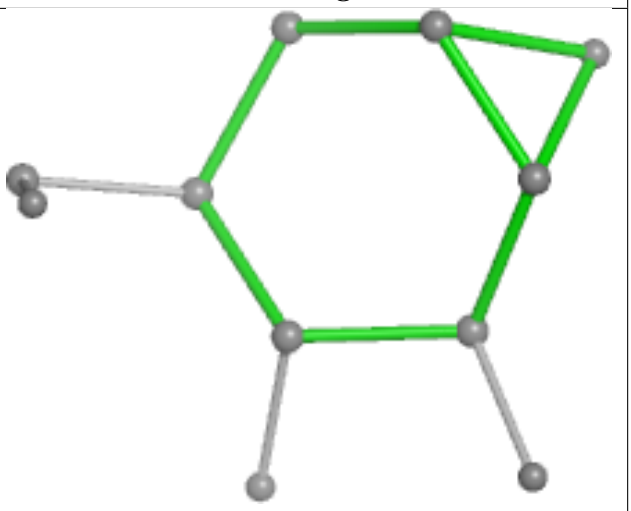
Bond lengths



Bond angles

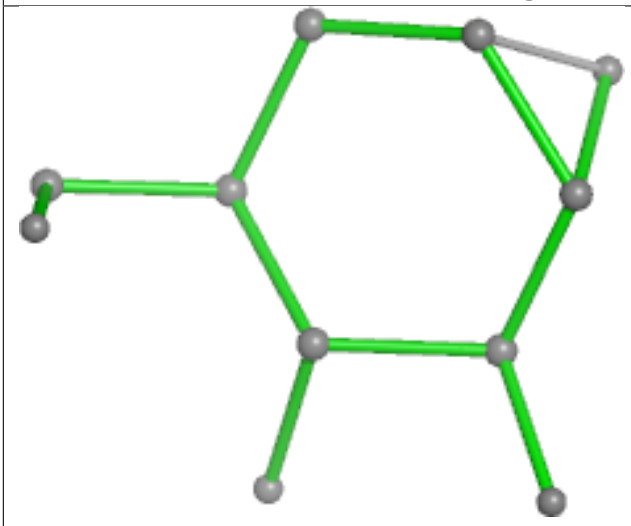


Torsions

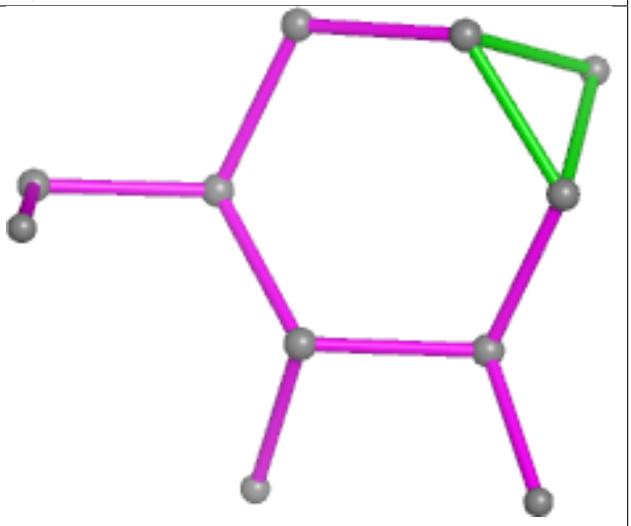


Rings

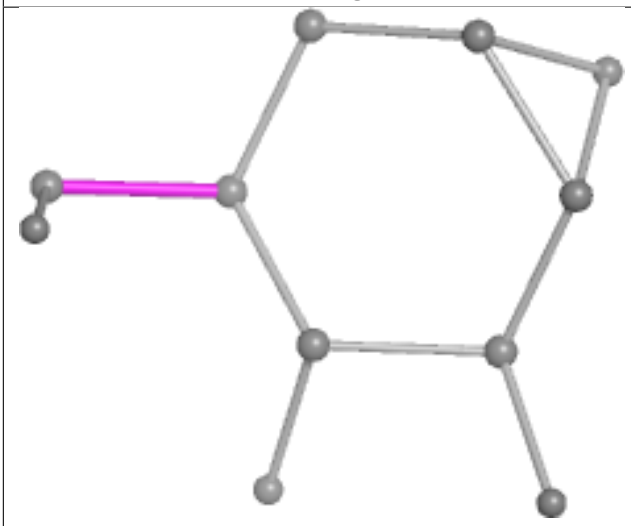
Ligand GDQ B 901



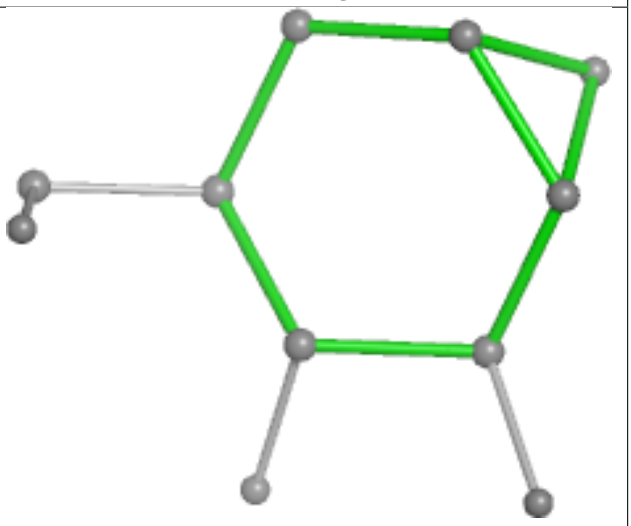
Bond lengths



Bond angles

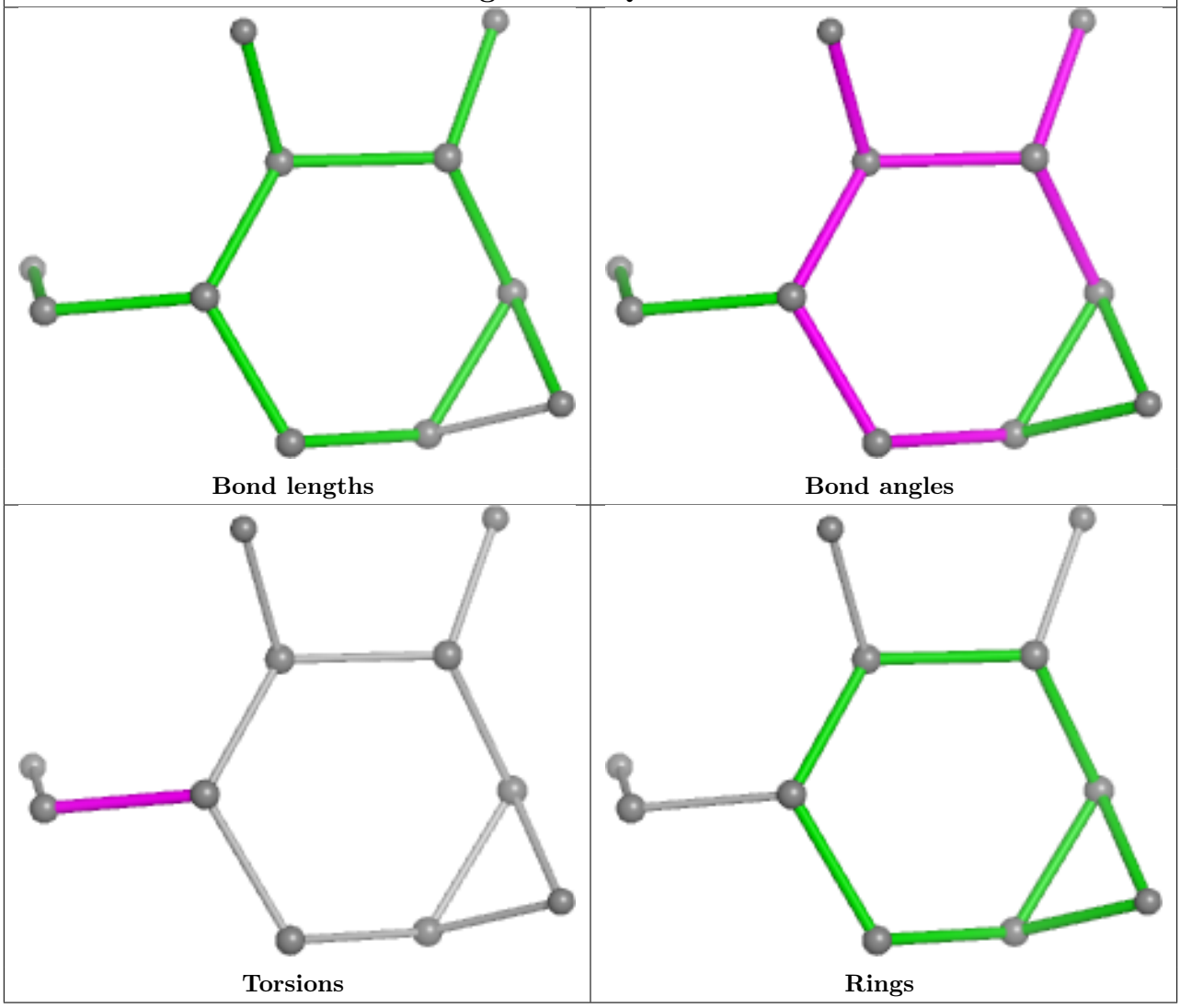


Torsions



Rings

Ligand GDQ C 901



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	781/817 (95%)	-0.12	1 (0%)	95	95	23, 33, 56, 79	8 (1%)
1	B	781/817 (95%)	-0.14	4 (0%)	91	92	23, 33, 55, 87	18 (2%)
1	C	781/817 (95%)	-0.14	1 (0%)	95	95	22, 33, 55, 78	17 (2%)
All	All	2343/2451 (95%)	-0.13	6 (0%)	94	94	22, 33, 56, 87	43 (1%)

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	737	HIS	3.3
1	B	54	GLN	3.1
1	B	137	VAL	2.6
1	C	37	PHE	2.2
1	B	36	VAL	2.1
1	B	50	ILE	2.0

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

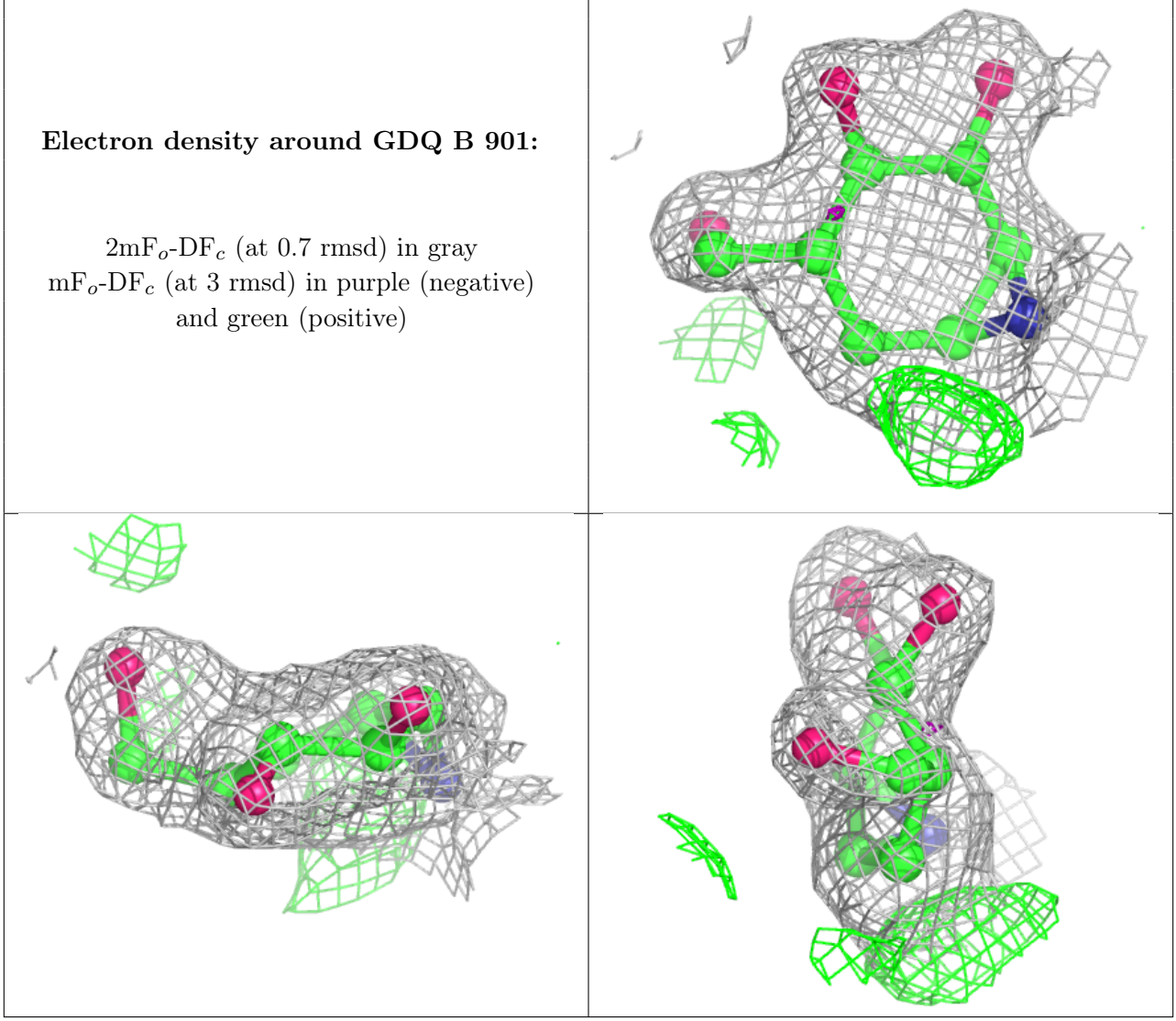
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	GDQ	B	901	11/11	0.80	0.18	40,48,55,63	0
2	GDQ	A	901	11/11	0.81	0.26	27,29,31,32	11
2	GDQ	C	901	11/11	0.82	0.22	41,49,55,57	0
3	OXL	A	907	6/6	0.84	0.15	40,48,57,64	0
6	PG4	A	905	13/13	0.85	0.18	42,53,63,66	0
6	PG4	C	907	13/13	0.86	0.15	41,47,62,64	0
6	PG4	B	904	13/13	0.88	0.15	41,49,61,62	0
5	EDO	B	907	4/4	0.90	0.19	42,46,54,60	0
5	EDO	C	906	4/4	0.91	0.11	41,44,48,53	0
4	SO4	A	908	5/5	0.92	0.15	62,65,82,82	0
4	SO4	C	908	5/5	0.92	0.13	62,65,78,79	0
3	OXL	B	903	6/6	0.93	0.09	33,41,52,59	0
3	OXL	C	902	6/6	0.93	0.15	39,49,63,63	0
4	SO4	B	908	5/5	0.94	0.11	63,67,84,86	0
5	EDO	A	902	4/4	0.94	0.11	41,43,45,46	0
4	SO4	B	905	5/5	0.96	0.10	66,67,72,73	0
4	SO4	A	903	5/5	0.96	0.08	66,66,70,71	0
4	SO4	C	904	5/5	0.97	0.10	65,66,71,73	0
4	SO4	C	905	5/5	0.97	0.08	52,52,58,66	0
4	SO4	A	904	5/5	0.97	0.09	49,53,57,64	0
4	SO4	B	906	5/5	0.98	0.10	50,53,57,64	0
4	SO4	C	903	5/5	0.98	0.07	40,42,45,45	0
4	SO4	B	902	5/5	0.98	0.09	38,41,44,47	0
4	SO4	A	906	5/5	0.98	0.09	38,41,42,47	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.

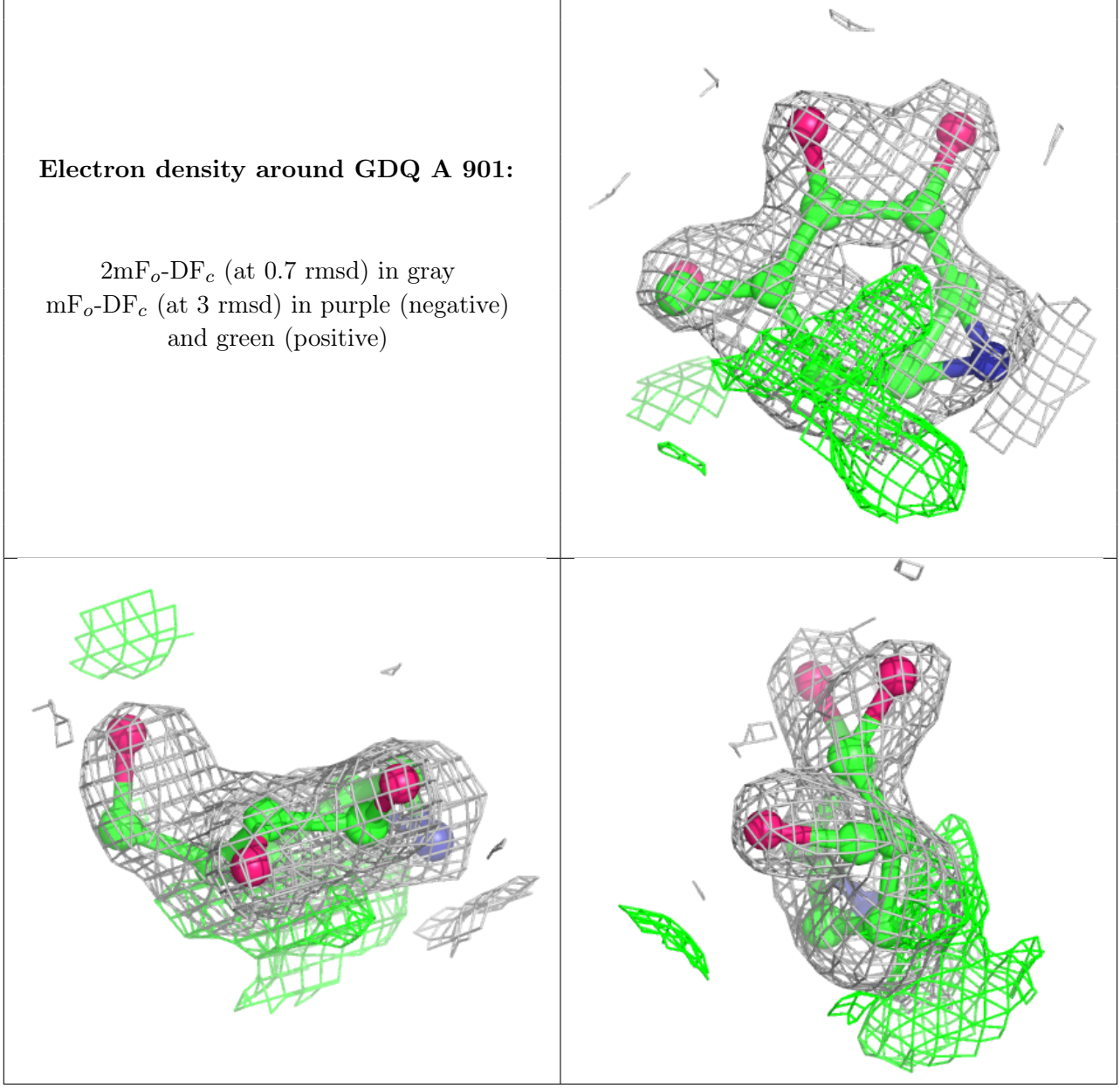
Electron density around GDQ B 901:

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



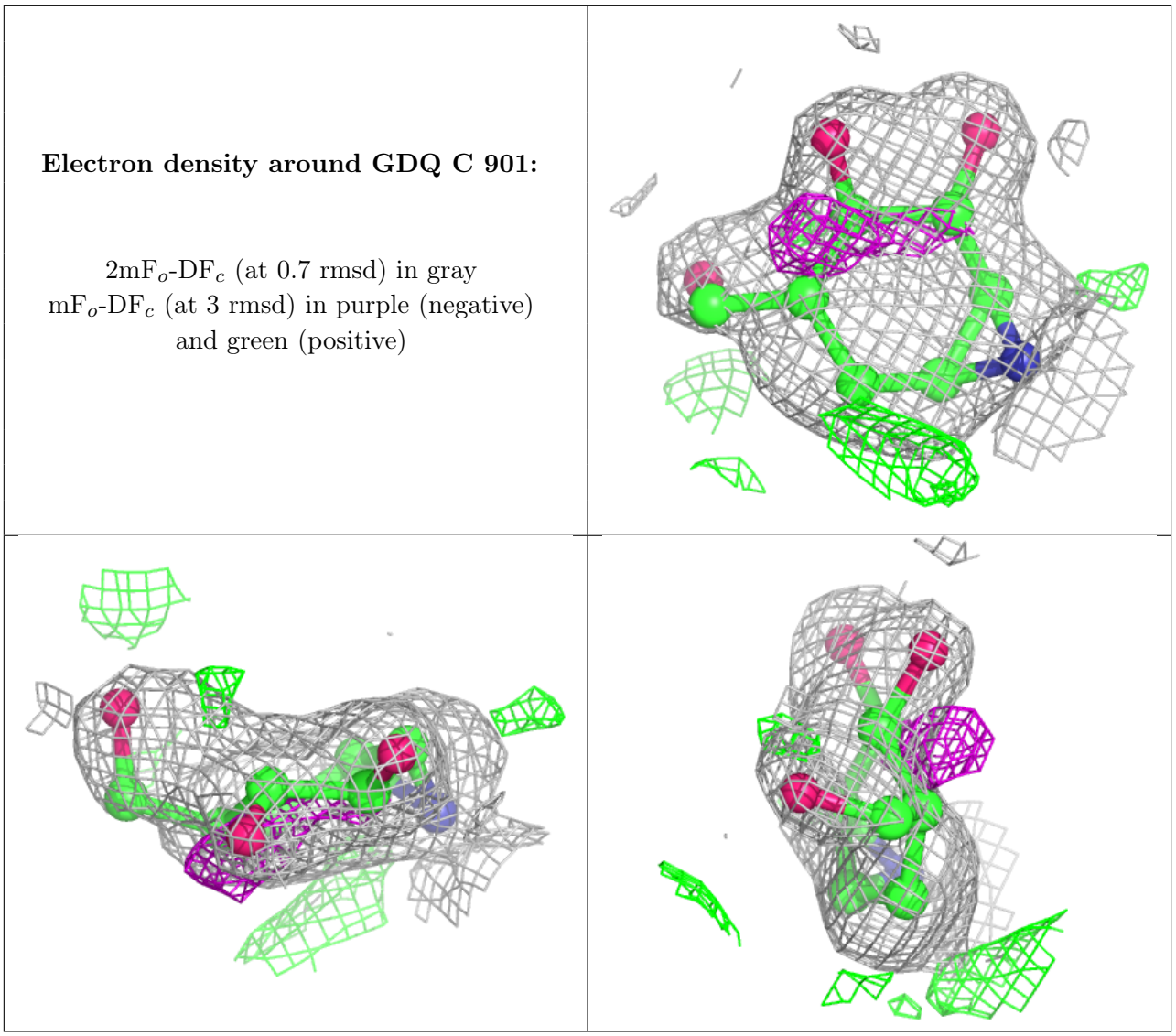
Electron density around GDQ A 901:

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



Electron density around GDQ C 901:

$2mF_o$-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.