



Full wwPDB X-ray Structure Validation Report ⓘ

Sep 23, 2024 – 02:38 PM JST

PDB ID : 8Z2N
Title : Crystal structure of 5-phosphomethyl-2'-deoxyuridine (5-PmdU) glycinyl-transferase gp46/PUGT from Pseudomonads phage PaMx11 in complex with dsDNA
Authors : Wen, Y.; Guo, W.T.; Wu, B.X.
Deposited on : 2024-04-13
Resolution : 2.30 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 ⓘ 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 : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.002 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.38.2

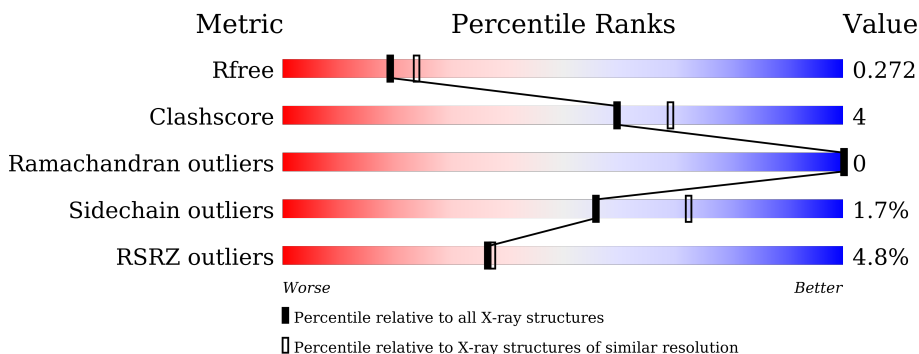
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

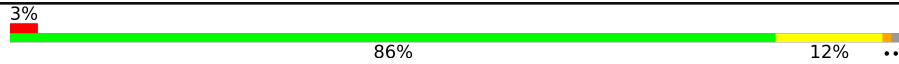
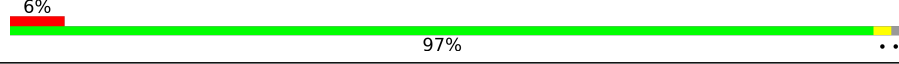
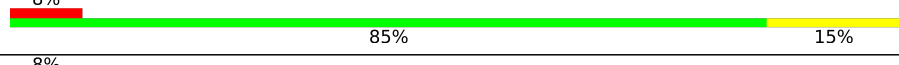
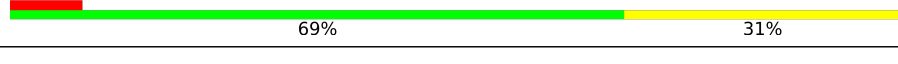
The reported resolution of this entry is 2.30 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	5963 (2.30-2.30)
Clashscore	180529	6698 (2.30-2.30)
Ramachandran outliers	177936	6640 (2.30-2.30)
Sidechain outliers	177891	6640 (2.30-2.30)
RSRZ outliers	164620	5963 (2.30-2.30)

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 $\leq 5\%$. 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	B	292	 3% 86% 12% ..
1	CA00	292	 6% 97% ..
2	C	13	 8% 85% 15%
3	D	13	 8% 69% 31%

2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 5247 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 Glycyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	288	2337	1495	411	421	10	0	1	0
1	CA00	289	2337	1495	411	421	10	0	0	0

- Molecule 2 is a DNA chain called DNA (5'-D(*TP*AP*GP*TP*CP*GP*TP*CP*GP*AP*CP*TP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	C	13	248	119	42	75	12	0	0	0

- Molecule 3 is a DNA chain called DNA (5'-D(*TP*AP*GP*TP*CP*GP*AP*CP*GP*AP*CP*TP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	D	13	254	122	45	75	12	0	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total O S 5 4 1	0	0
4	CA00	1	Total O S 5 4 1	0	0

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

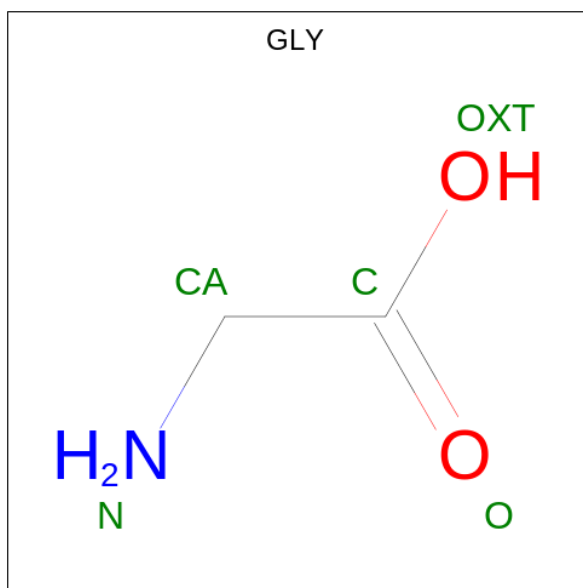
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total Ca 1 1	0	0
5	CA00	1	Total Ca 1 1	0	0

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



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

- Molecule 7 is GLYCINE (three-letter code: GLY) (formula: $C_2H_5NO_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	C	1	Total	C	N	O	0	0
			5	2	1	2		
7	CA00	1	Total	C	N	O	0	0
			5	2	1	2		

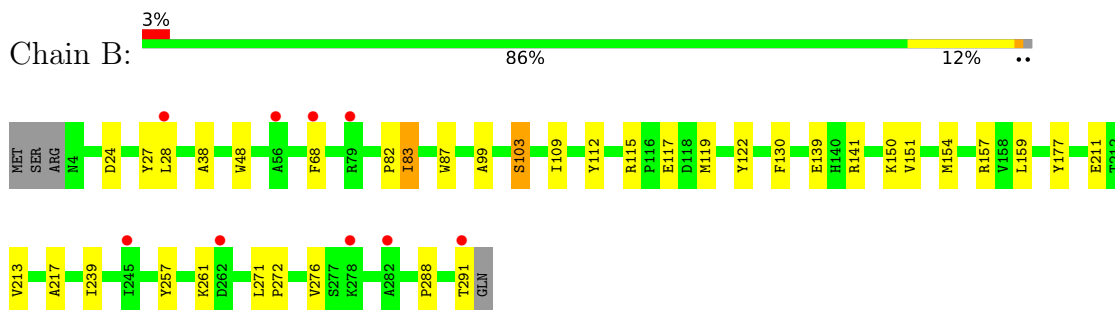
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	B	19	Total 19	O 19	0	0
8	C	3	Total 3	O 3	0	0
8	CA00	21	Total 21	O 21	0	0

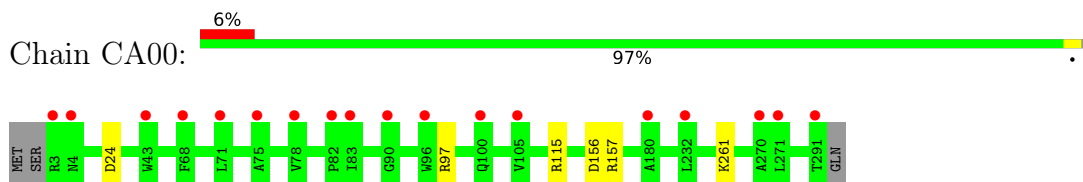
3 Residue-property plots [i](#)

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.

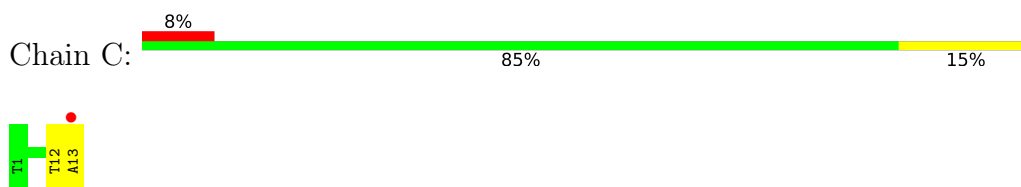
- Molecule 1: Glycyltransferase



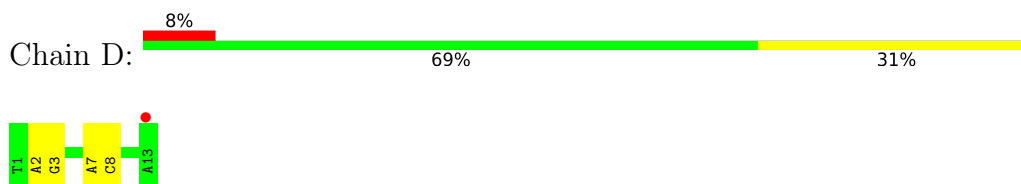
- Molecule 1: Glycyltransferase



- Molecule 2: DNA (5'-D(*TP*AP*GP*TP*CP*GP*TP*CP*GP*AP*CP*TP*A)-3')



- Molecule 3: DNA (5'-D(*TP*AP*GP*TP*CP*GP*AP*CP*GP*AP*CP*TP*A)-3')



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	57.79Å 92.03Å 167.64Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.00 – 2.30 29.00 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.8 (29.00-2.30) 99.7 (29.00-2.30)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.69 (at 2.31Å)	Xtrriage
Refinement program	PHENIX 1.21rc1_5156	Depositor
R, R_{free}	0.226 , 0.275 0.225 , 0.272	Depositor DCC
R_{free} test set	2041 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	45.7	Xtrriage
Anisotropy	0.821	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 38.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	5247	wwPDB-VP
Average B, all atoms (Å ²)	57.0	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 34.29 % 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 6.9872e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical 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: CA, GOL, SO4

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	B	0.51	0/2405	0.66	0/3266
1	CA00	0.51	0/2405	0.65	0/3266
2	C	0.96	0/276	1.02	0/424
3	D	0.90	0/284	0.96	0/436
All	All	0.57	0/5370	0.70	0/7392

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	B	0	2
1	CA00	0	3
All	All	0	5

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	115	ARG	Sidechain
1	B	157	ARG	Sidechain
1	CA00	115	ARG	Sidechain
1	CA00	157	ARG	Sidechain
1	CA00	97	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	B	2337	0	2253	19	0
1	CA00	2337	0	0	0	0
2	C	248	0	139	1	0
3	D	254	0	143	2	0
4	B	5	0	0	0	0
4	CA00	5	0	0	0	0
5	B	1	0	0	0	0
5	CA00	1	0	0	0	0
6	C	6	0	8	0	0
7	C	5	0	2	2	0
7	CA00	5	0	0	0	0
8	B	19	0	0	0	0
8	C	3	0	0	1	0
8	CA00	21	0	0	0	0
All	All	5247	0	2545	23	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 4.

All (23) 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)
7:C:102:GLY:N	8:C:201:HOH:O	2.13	0.81
1:B:38:ALA:HB1	1:B:117:GLU:HG2	1.74	0.67
1:B:150:LYS:HZ3	7:C:102:GLY:N	1.92	0.67
1:B:82:PRO:HB2	1:B:276:VAL:HG11	1.85	0.57
1:B:103:SER:OG	1:B:141:ARG:NE	2.40	0.54
1:B:83:ILE:HD11	1:B:272:PRO:O	2.09	0.52
1:B:99:ALA:O	1:B:103:SER:HB3	2.09	0.52
1:B:122:TYR:OH	1:B:139:GLU:OE2	2.22	0.51
1:B:257:TYR:CD2	1:B:261:LYS:HD3	2.47	0.49
1:B:271:LEU:HB3	1:B:272:PRO:HD3	1.96	0.48
1:B:211:GLU:H	1:B:211:GLU:CD	2.16	0.48
1:B:217:ALA:HB1	1:B:239:ILE:HD13	1.98	0.46
1:B:28:LEU:HD23	1:B:288:PRO:HG3	1.98	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:68:PHE:HB3	1:B:109:ILE:HD11	2.00	0.44
1:B:112:TYR:CZ	1:B:119:MET:HA	2.53	0.44
2:C:12:DT:H5''	2:C:13:DA:H5'	1.98	0.43
1:B:48:TRP:CE2	1:B:151:VAL:HG22	2.53	0.43
1:B:177:TYR:CZ	1:B:213:VAL:HG21	2.55	0.42
3:D:7:DA:H1'	3:D:8:DC:H5'	2.02	0.41
1:B:27:TYR:CE1	1:B:154:MET:HG2	2.55	0.41
1:B:159:LEU:HD23	1:B:159:LEU:HA	1.84	0.41
1:B:82:PRO:HD3	1:B:87:TRP:CH2	2.56	0.41
3:D:2:DA:C2	3:D:3:DG:C4	3.09	0.41

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	B	287/292 (98%)	282 (98%)	5 (2%)	0	100	100
1	CA00	287/292 (98%)	283 (99%)	4 (1%)	0	100	100
All	All	574/584 (98%)	565 (98%)	9 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	241/244 (99%)	236 (98%)	5 (2%)	48	66
1	CA00	241/244 (99%)	238 (99%)	3 (1%)	67	81
All	All	482/488 (99%)	474 (98%)	8 (2%)	56	72

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	24	ASP
1	B	83	ILE
1	B	103	SER
1	B	130	PHE
1	B	291	THR
1	CA00	24	ASP
1	CA00	156	ASP
1	CA00	261	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	72	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 2 are monoatomic - leaving 5 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	SO4	CA00	301	-	4,4,4	0.74	0	6,6,6	0.22	0
4	SO4	B	301	-	4,4,4	0.77	0	6,6,6	0.37	0
7	GLY	C	102	-	4,4,4	0.91	0	3,4,4	1.38	0
7	GLY	CA00	303	-	4,4,4	1.14	0	3,4,4	0.88	0
6	GOL	C	101	-	5,5,5	0.17	0	5,5,5	0.39	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
7	GLY	C	102	-	-	0/2/2/2	-
7	GLY	CA00	303	-	-	2/2/2/2	-
6	GOL	C	101	-	-	4/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	C	101	GOL	O1-C1-C2-C3
7	CA00	303	GLY	OXT-C-CA-N
6	C	101	GOL	C1-C2-C3-O3
6	C	101	GOL	O1-C1-C2-O2
6	C	101	GOL	O2-C2-C3-O3
7	CA00	303	GLY	O-C-CA-N

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	C	102	GLY	2	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, 95th 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	B	288/292 (98%)	0.60	9 (3%) 51 53	35, 55, 69, 83	1 (0%)
1	CA00	289/292 (98%)	0.75	18 (6%) 28 29	40, 55, 71, 88	0
2	C	13/13 (100%)	0.51	1 (7%) 21 22	48, 77, 81, 94	0
3	D	13/13 (100%)	0.62	1 (7%) 21 22	49, 78, 84, 95	0
All	All	603/610 (98%)	0.67	29 (4%) 36 37	35, 55, 74, 95	1 (0%)

All (29) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	13	DA	3.7
1	CA00	100	GLN	3.5
1	CA00	96	TRP	3.3
3	D	13	DA	3.3
1	B	79[A]	ARG	2.9
1	CA00	105	VAL	2.9
1	CA00	71	LEU	2.8
1	CA00	291	THR	2.8
1	CA00	68	PHE	2.7
1	CA00	82	PRO	2.7
1	CA00	4	ASN	2.5
1	B	56	ALA	2.5
1	B	262	ASP	2.4
1	CA00	78	VAL	2.3
1	CA00	232	LEU	2.3
1	CA00	83	ILE	2.3
1	CA00	43	TRP	2.2
1	CA00	75	ALA	2.2
1	B	291	THR	2.2
1	B	68	PHE	2.2
1	CA00	3	ARG	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	282	ALA	2.2
1	CA00	270	ALA	2.2
1	B	245	ILE	2.1
1	B	278	LYS	2.1
1	CA00	180	ALA	2.1
1	CA00	90	GLY	2.0
1	B	28	LEU	2.0
1	CA00	271	LEU	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](#)

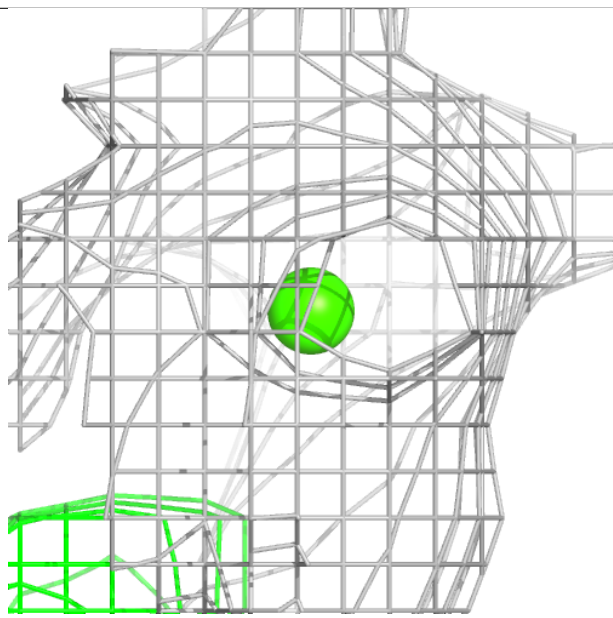
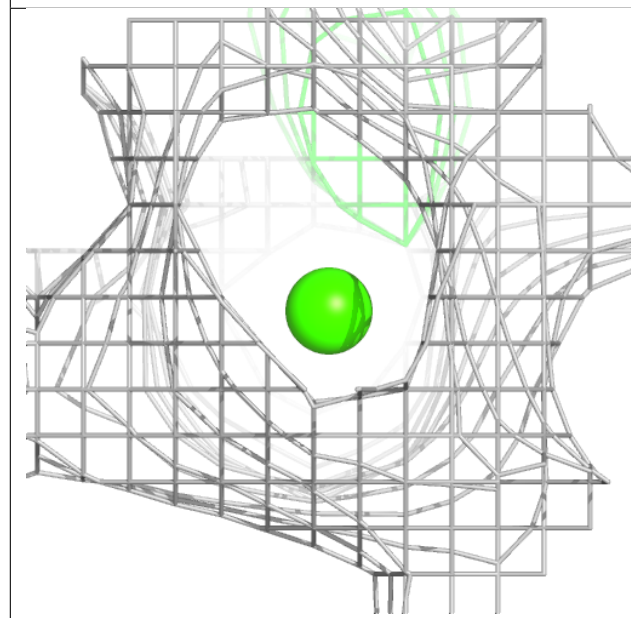
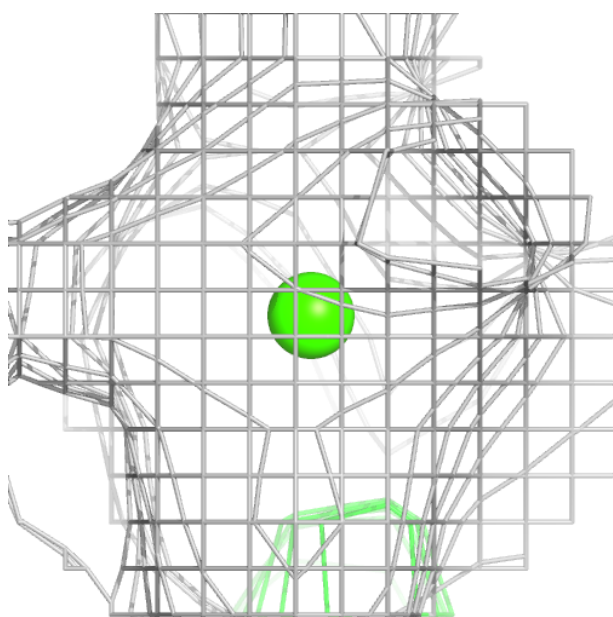
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

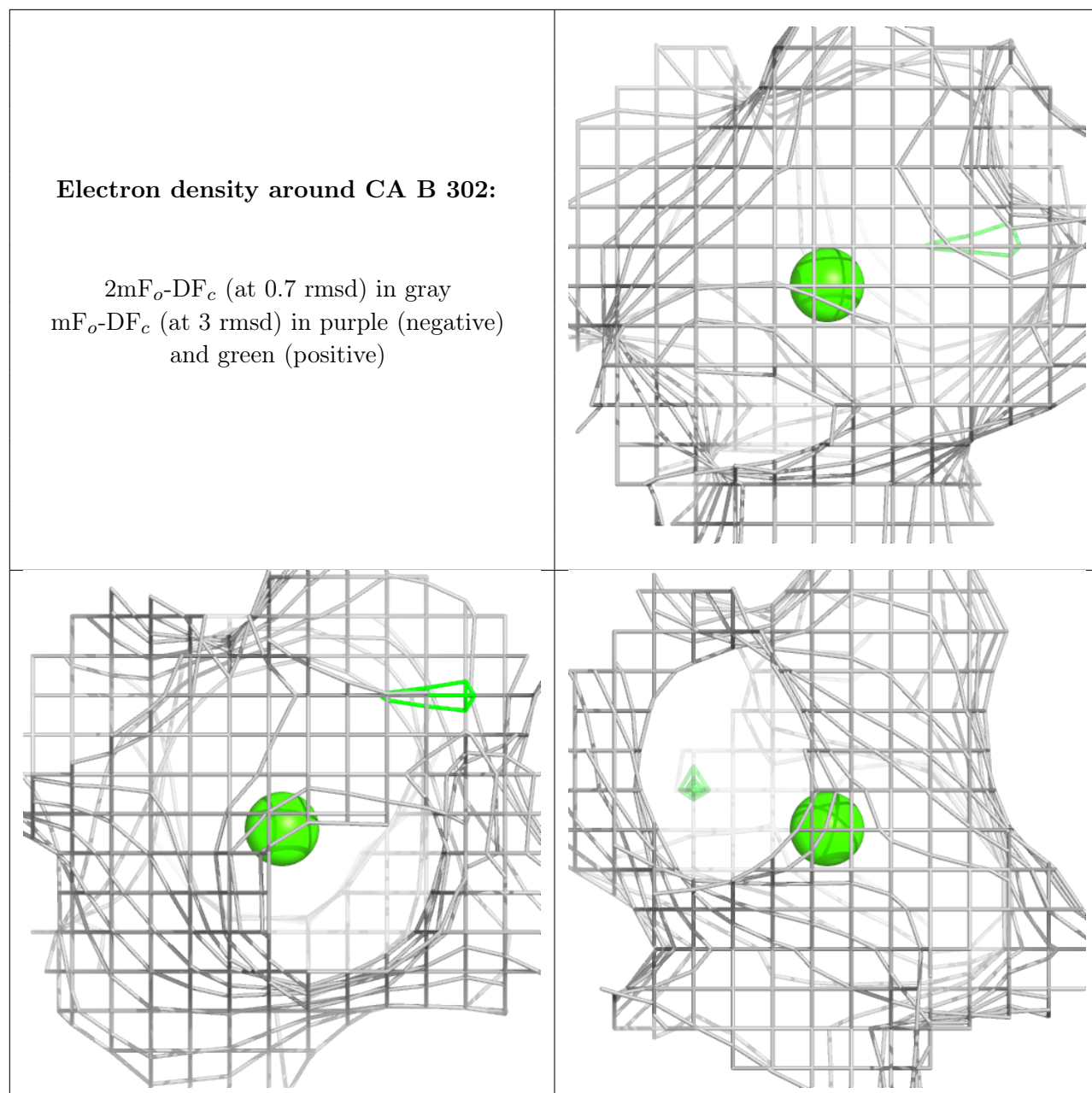
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
4	SO4	CA00	301	5/5	0.58	0.15	75,77,85,92	0
6	GOL	C	101	6/6	0.66	0.23	78,81,83,87	0
4	SO4	B	301	5/5	0.81	0.12	62,63,80,86	0
7	GLY	C	102	5/5	0.87	0.15	52,53,59,60	0
7	GLY	CA00	303	5/5	0.94	0.10	48,50,52,55	0
5	CA	CA00	302	1/1	0.99	0.04	49,49,49,49	0
5	CA	B	302	1/1	0.99	0.07	49,49,49,49	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 CA00 302:

$2mF_o-DF_c$ (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.