



Full wwPDB X-ray Structure Validation Report ⓘ

Jul 28, 2022 – 01:14 pm BST

PDB ID : 7NMC
Title : Crystal structure of beta-2-microglobulin D76E mutant
Authors : Guthertz, N.
Deposited on : 2021-02-23
Resolution : 1.20 Å(reported)

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

We welcome your comments at validation@mail.wwpdb.org

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 : 2.29
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0267
CCP4 : 7.1.010 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.29

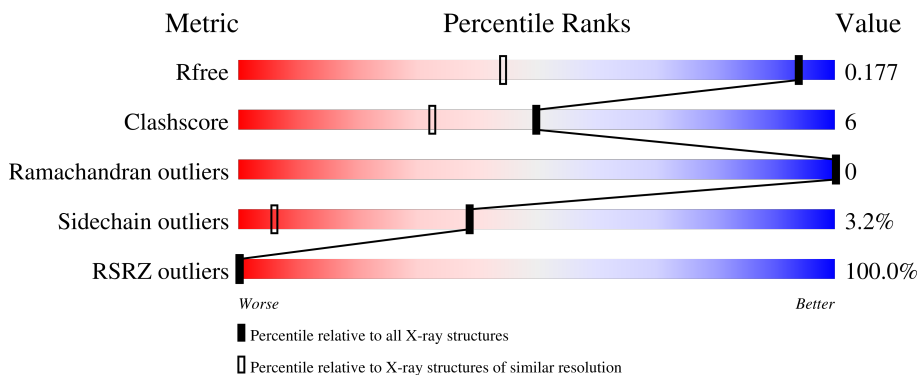
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.20 Å.

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



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1223 (1.22-1.18)
Clashscore	141614	1286 (1.22-1.18)
Ramachandran outliers	138981	1240 (1.22-1.18)
Sidechain outliers	138945	1239 (1.22-1.18)
RSRZ outliers	127900	1200 (1.22-1.18)

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	AAA	100	

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
2	GOL	AAA	100	-	-	-	X

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	GOL	AAA	102	-	-	-	X

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 1898 atoms, of which 854 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 Beta-2-microglobulin.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	AAA	100	1679	544	824	143	164	4	47	3	0

There are 2 discrepancies between the modelled and reference sequences:

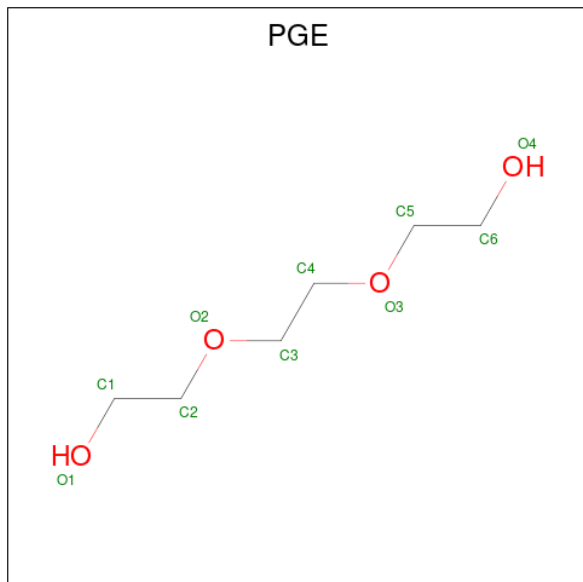
Chain	Residue	Modelled	Actual	Comment	Reference
AAA	0	MET	-	initiating methionine	UNP P61769
AAA	76	GLU	ASP	engineered mutation	UNP P61769

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
2	AAA	1	14	3	8	3	2	0
2	AAA	1	14	3	8	3	2	0

- Molecule 3 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: C₆H₁₄O₄) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
3	AAA	1	24	6	14	4	1	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	AAA	167	167	167	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: Beta-2-microglobulin



4 Data and refinement statistics

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	54.63Å 28.87Å 67.51Å 90.00° 102.02° 90.00°	Depositor
Resolution (Å)	46.60 – 1.20 37.86 – 1.20	Depositor EDS
% Data completeness (in resolution range)	96.3 (46.60-1.20) 96.3 (37.86-1.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.88 (at 1.20Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.158 , 0.179 0.156 , 0.177	Depositor DCC
R_{free} test set	1499 reflections (4.78%)	wwPDB-VP
Wilson B-factor (Å ²)	9.5	Xtriage
Anisotropy	0.367	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	1898	wwPDB-VP
Average B, all atoms (Å ²)	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.56% of the height of the origin peak. No significant pseudotranslation is detected.*

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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, PGE

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	AAA	0.85	1/884 (0.1%)	0.85	2/1194 (0.2%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	AAA	98	ASP	C-N	17.80	1.75	1.34

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	AAA	98	ASP	O-C-N	-7.95	109.99	122.70
1	AAA	12	ARG	NE-CZ-NH1	5.37	122.99	120.30

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	AAA	855	824	820	8	1
2	AAA	12	16	16	0	0
3	AAA	10	14	14	2	0
4	AAA	167	0	0	4	1
All	All	1044	854	850	10	2

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

All (10) 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:AAA:98:ASP:C	1:AAA:99:MET:N	1.75	1.36
1:AAA:3:ARG:HD2	4:AAA:237:HOH:O	1.50	1.09
1:AAA:99:MET:HG3	4:AAA:321:HOH:O	1.73	0.87
3:AAA:101:PGE:H5	3:AAA:101:PGE:O2	1.89	0.72
1:AAA:99:MET:CG	4:AAA:321:HOH:O	2.40	0.64
3:AAA:101:PGE:O2	3:AAA:101:PGE:C5	2.56	0.51
1:AAA:97:ARG:NH2	1:AAA:97:ARG:HB3	2.27	0.48
1:AAA:73:THR:O	1:AAA:97:ARG:NH1	2.51	0.44
1:AAA:99:MET:C	4:AAA:204:HOH:O	2.57	0.43
1:AAA:97:ARG:CZ	1:AAA:97:ARG:CB	2.99	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)
4:AAA:338:HOH:O	4:AAA:338:HOH:O[2_555]	1.68	0.52
1:AAA:44:GLU:OE1	1:AAA:44:GLU:OE1[2_455]	1.75	0.45

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	AAA	101/100 (101%)	98 (97%)	3 (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	AAA	98/95 (103%)	95 (97%)	3 (3%)	40 6

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	17	ASN
1	AAA	70	PHE
1	AAA	97	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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](#)

3 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	PGE	AAA	101	-	9,9,9	0.15	0	8,8,8	0.10	0
2	GOL	AAA	100	-	5,5,5	0.15	0	5,5,5	0.68	0
2	GOL	AAA	102	-	5,5,5	0.15	0	5,5,5	0.38	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	PGE	AAA	101	-	-	2/7/7/7	-
2	GOL	AAA	100	-	-	3/4/4/4	-
2	GOL	AAA	102	-	-	1/4/4/4	-

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
2	AAA	100	GOL	C1-C2-C3-O3
2	AAA	100	GOL	O2-C2-C3-O3
3	AAA	101	PGE	O3-C5-C6-O4
2	AAA	102	GOL	O1-C1-C2-C3
3	AAA	101	PGE	O2-C3-C4-O3
2	AAA	100	GOL	O1-C1-C2-C3

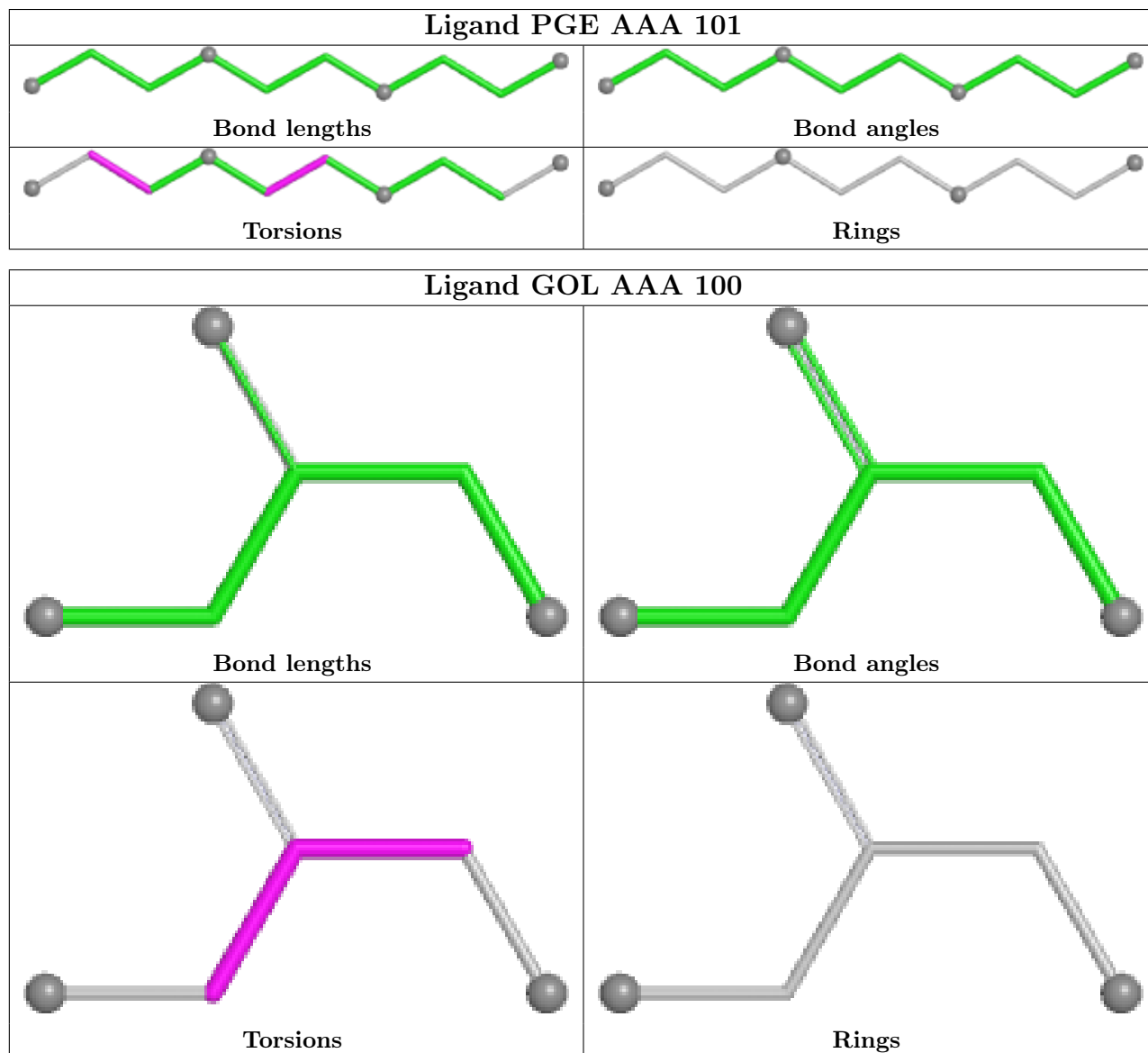
There are no ring outliers.

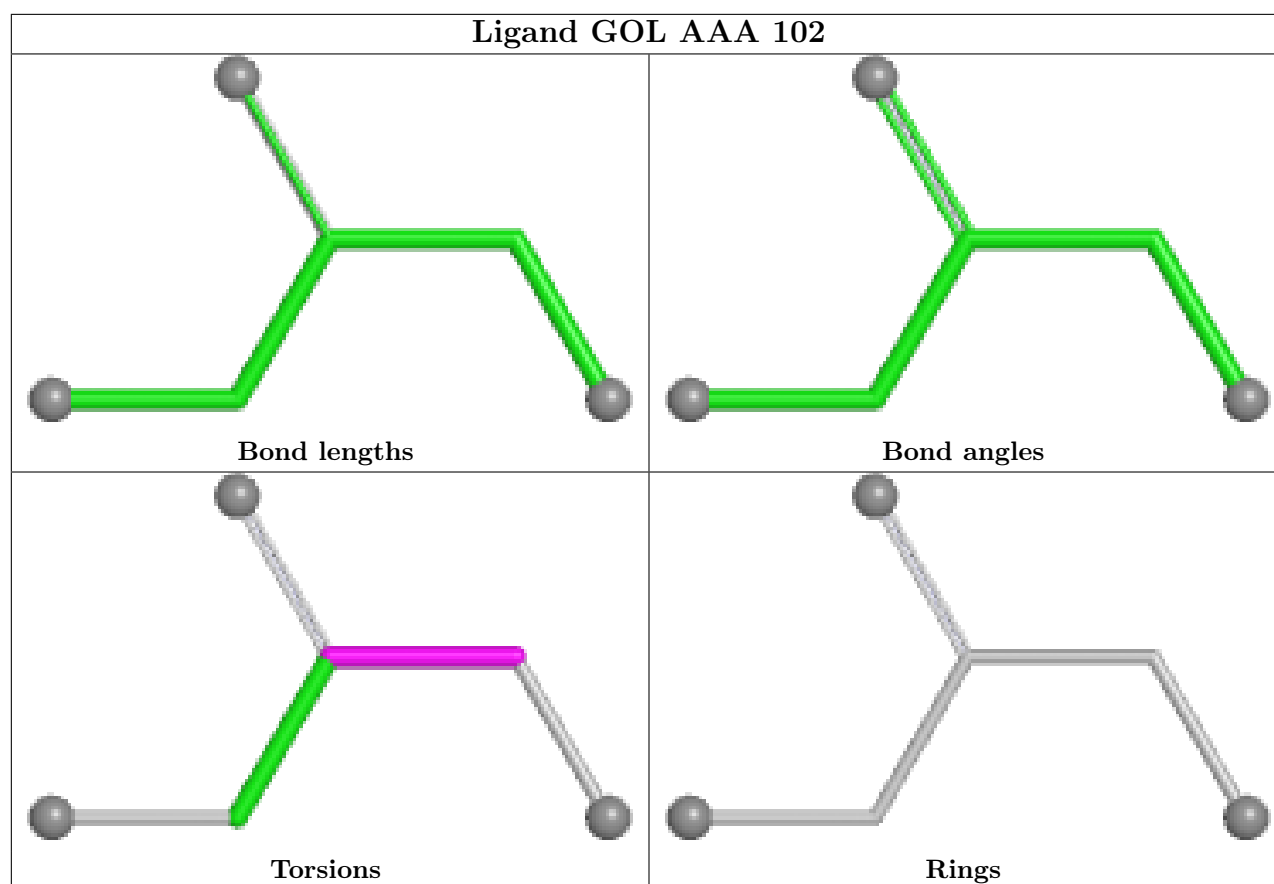
1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	AAA	101	PGE	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 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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	AAA	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	AAA	98:ASP	C	99:MET	N	1.75

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	AAA	100/100 (100%)	6.55	100 (100%) 0 0	5, 10, 30, 42	0

All (100) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	60	TRP	13.2
1	AAA	99	MET	11.0
1	AAA	92	ILE	10.7
1	AAA	40	LEU	10.4
1	AAA	95	TRP	10.3
1	AAA	10	TYR	9.7
1	AAA	22	PHE	9.6
1	AAA	67	TYR	9.6
1	AAA	80	CYS	9.3
1	AAA	25	CYS	9.3
1	AAA	35	ILE	9.2
1	AAA	56	PHE	8.8
1	AAA	97	ARG	8.8
1	AAA	7	ILE	8.7
1	AAA	26	TYR	8.7
1	AAA	47	GLU	8.6
1	AAA	30	PHE	8.6
1	AAA	93	VAL	8.5
1	AAA	1	ILE	8.4
1	AAA	73	THR	8.4
1	AAA	82	VAL	8.4
1	AAA	65	LEU	8.4
1	AAA	62	PHE	8.3
1	AAA	70	PHE	8.3
1	AAA	37	VAL	8.3
1	AAA	49	VAL	8.3
1	AAA	48	LYS	8.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	AAA	27	VAL	8.1
1	AAA	85	VAL	8.1
1	AAA	39	LEU	8.1
1	AAA	9	VAL	8.1
1	AAA	66	TYR	8.1
1	AAA	64	LEU	7.6
1	AAA	63	TYR	7.4
1	AAA	75	LYS	7.4
1	AAA	23	LEU	7.4
1	AAA	43	GLY	7.4
1	AAA	78	TYR	7.2
1	AAA	86	THR	7.2
1	AAA	54	LEU	7.0
1	AAA	87	LEU	6.9
1	AAA	46	ILE	6.8
1	AAA	15	ALA	6.8
1	AAA	4	THR	6.7
1	AAA	58	LYS	6.5
1	AAA	74	GLU	6.4
1	AAA	83[A]	ASN	6.4
1	AAA	32	PRO	6.3
1	AAA	98	ASP	6.3
1	AAA	96	ASP	6.2
1	AAA	14	PRO	6.0
1	AAA	90	PRO	6.0
1	AAA	68	THR	5.9
1	AAA	5	PRO	5.9
1	AAA	42	ASN	5.6
1	AAA	45	ARG	5.6
1	AAA	79	ALA	5.6
1	AAA	94	LYS	5.6
1	AAA	3	ARG	5.6
1	AAA	77	GLU	5.6
1	AAA	20[A]	SER	5.5
1	AAA	17	ASN	5.5
1	AAA	0	MET	5.5
1	AAA	89	GLN	5.4
1	AAA	29	GLY	5.4
1	AAA	52	SER	5.3
1	AAA	6	LYS	5.3
1	AAA	11	SER	5.2
1	AAA	61	SER	5.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	AAA	31	HIS	5.1
1	AAA	16[A]	GLU	5.0
1	AAA	55	SER	4.9
1	AAA	84	HIS	4.9
1	AAA	21	ASN	4.9
1	AAA	33	SER	4.9
1	AAA	51	HIS	4.8
1	AAA	24	ASN	4.7
1	AAA	19	LYS	4.7
1	AAA	28	SER	4.7
1	AAA	71	THR	4.6
1	AAA	8	GLN	4.6
1	AAA	18	GLY	4.6
1	AAA	91	LYS	4.5
1	AAA	12	ARG	4.5
1	AAA	34	ASP	4.5
1	AAA	72	PRO	4.5
1	AAA	44	GLU	4.4
1	AAA	59	ASP	4.4
1	AAA	88	SER	4.4
1	AAA	50	GLU	4.3
1	AAA	2	GLN	4.2
1	AAA	81	ARG	4.2
1	AAA	76	GLU	4.1
1	AAA	57	SER	4.1
1	AAA	13	HIS	4.0
1	AAA	38	ASP	4.0
1	AAA	69	GLU	3.9
1	AAA	36	GLU	3.9
1	AAA	53	ASP	3.8
1	AAA	41	LYS	3.5

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](#)

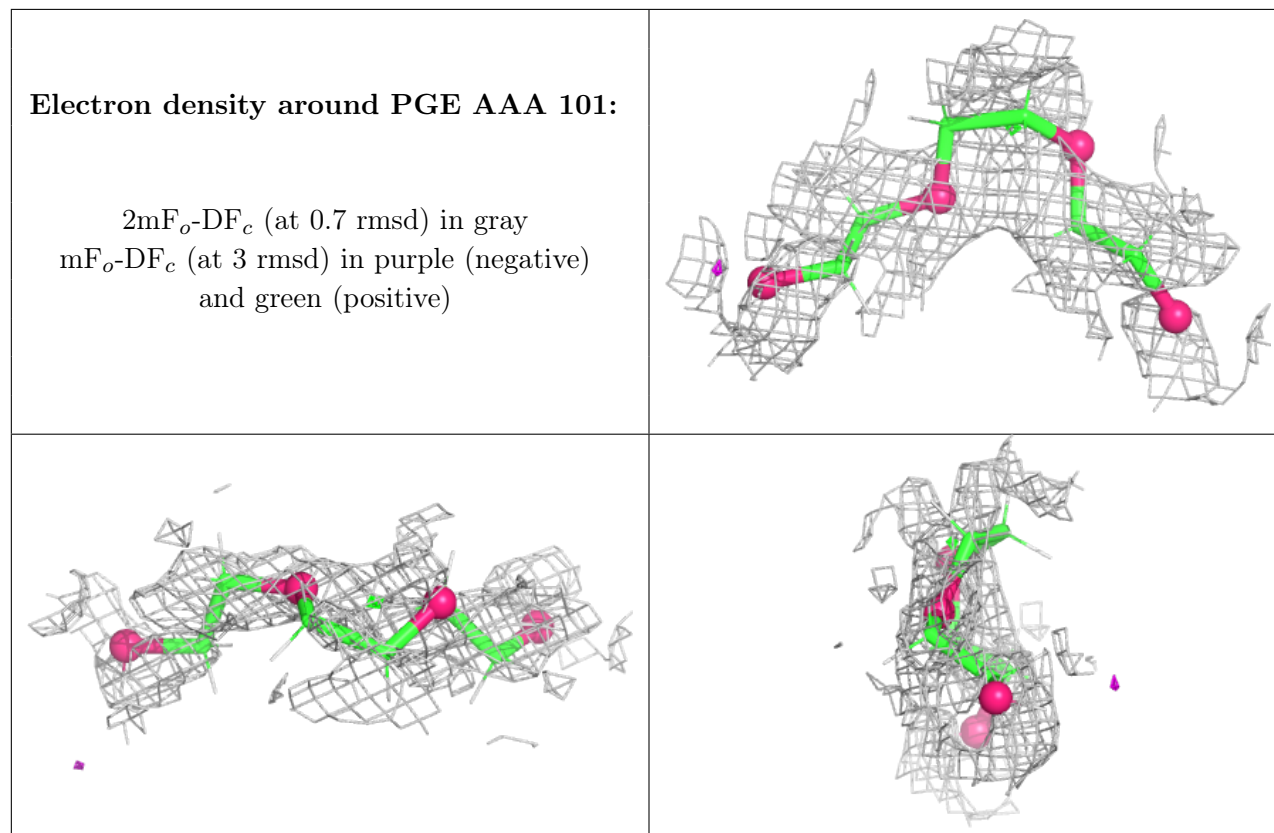
There are no monosaccharides in this entry.

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.

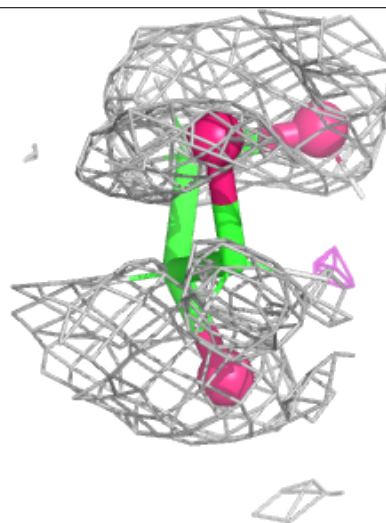
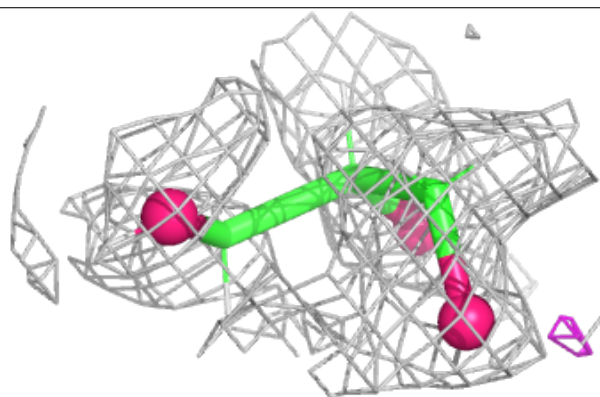
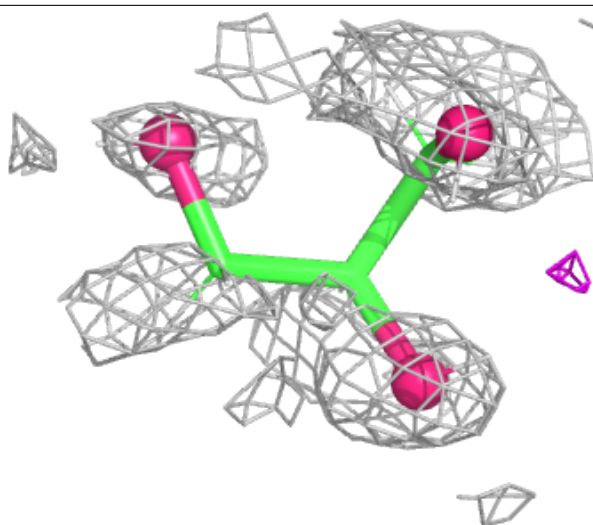
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	PGE	AAA	101	10/10	0.49	0.40	29,41,52,53	24
2	GOL	AAA	102	6/6	0.55	0.48	15,26,31,37	2
2	GOL	AAA	100	6/6	0.59	0.49	26,28,31,32	2

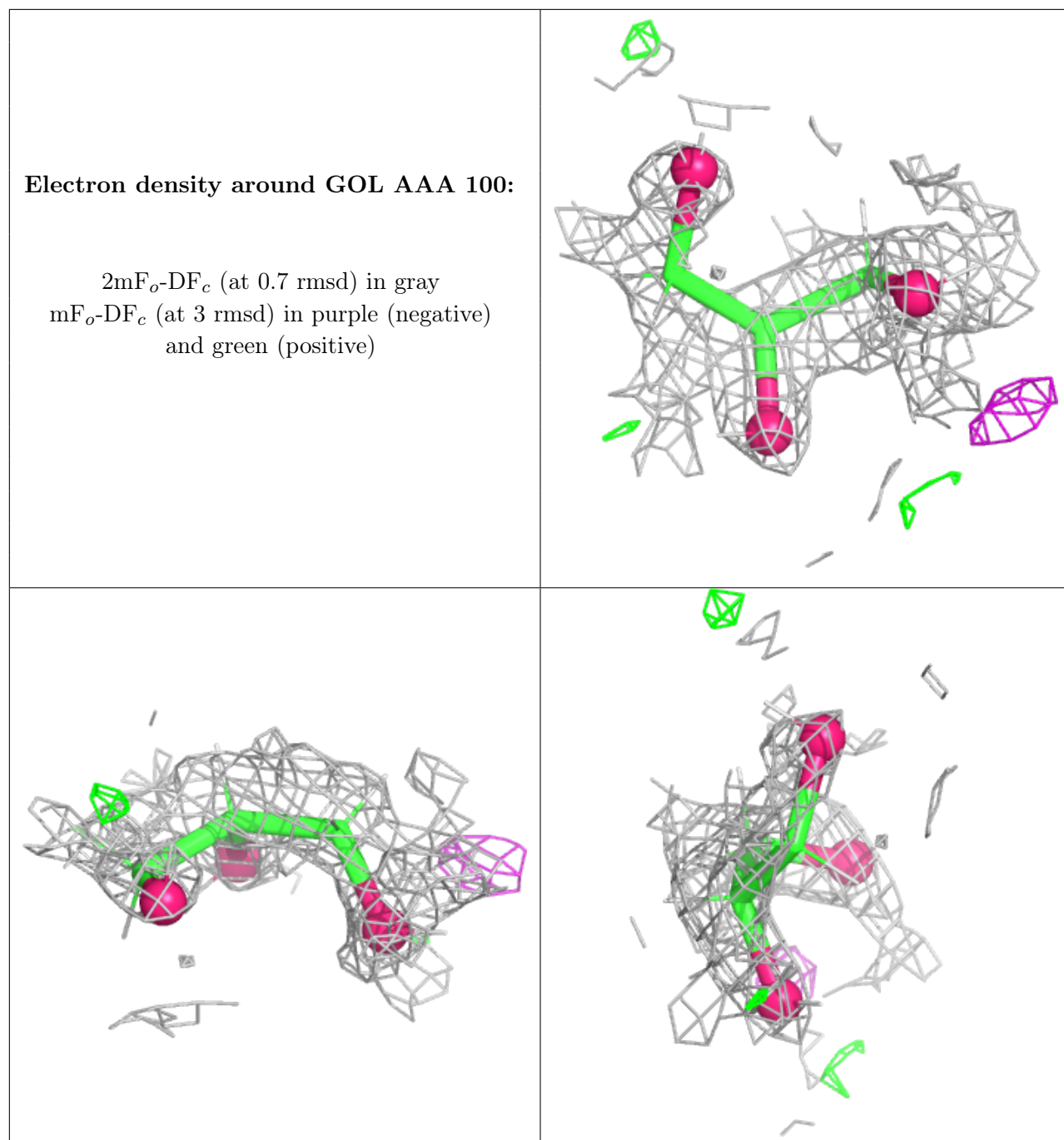
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



Electron density around GOL AAA 102:

$2mF_o-DF_c$ (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.