



Full wwPDB X-ray Structure Validation Report ⓘ

Sep 9, 2025 – 01:21 AM JST

PDB ID : 8GQ9 / pdb_00008gq9
Title : Crystal structure of lasso peptide epimerase MslH
Authors : Nakashima, Y.; Morita, H.
Deposited on : 2022-08-29
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 : 4-5-2 with Phenix2.0rc1
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 2.0rc1
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.006 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.45.1

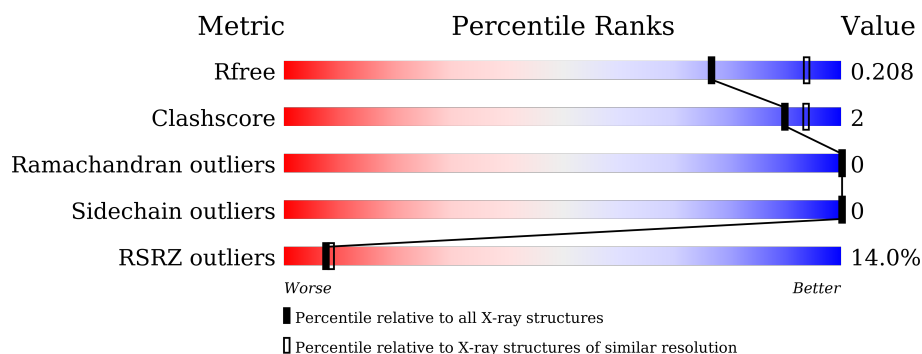
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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	A	440	14% 95% 5% .

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 3582 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.

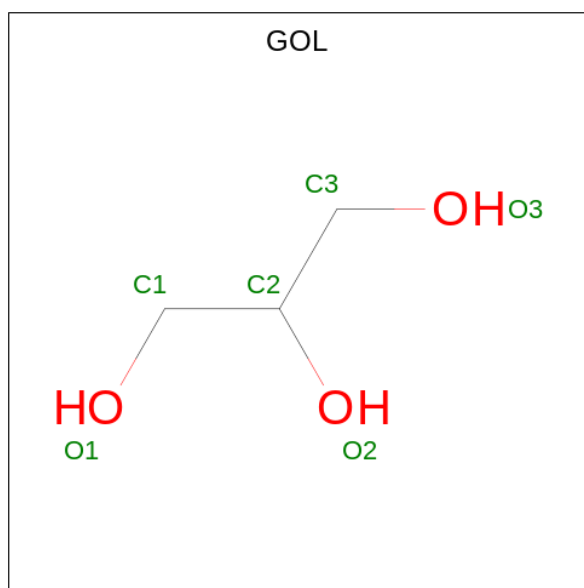
- Molecule 1 is a protein called Poly-gamma-glutamate synthesis protein (Capsule biosynthesis protein).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	437	Total	C	N	O	S	0	11	0
			3352	2085	612	642	13			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	34	ARG	HIS	conflict	UNP A0A4V2TW40
A	59	VAL	ALA	conflict	UNP A0A4V2TW40
A	101	MET	LEU	conflict	UNP A0A4V2TW40
A	287	HIS	ASP	conflict	UNP A0A4V2TW40
A	380	ALA	VAL	conflict	UNP A0A4V2TW40
A	425	ILE	VAL	conflict	UNP A0A4V2TW40

- Molecule 2 is GLYCEROL (CCD ID: GOL) (formula: $C_3H_8O_3$).

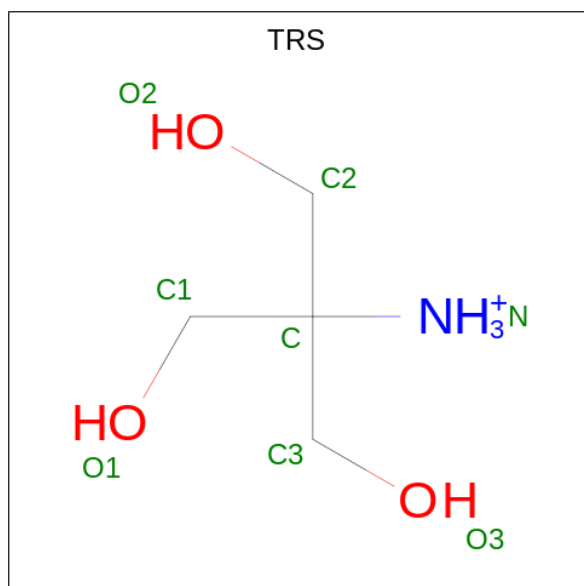


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 3 is CALCIUM ION (CCD ID: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Ca	0	0
			1	1		

- Molecule 4 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (CCD ID: TRS) (formula: C₄H₁₂NO₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			8	4	1	3		

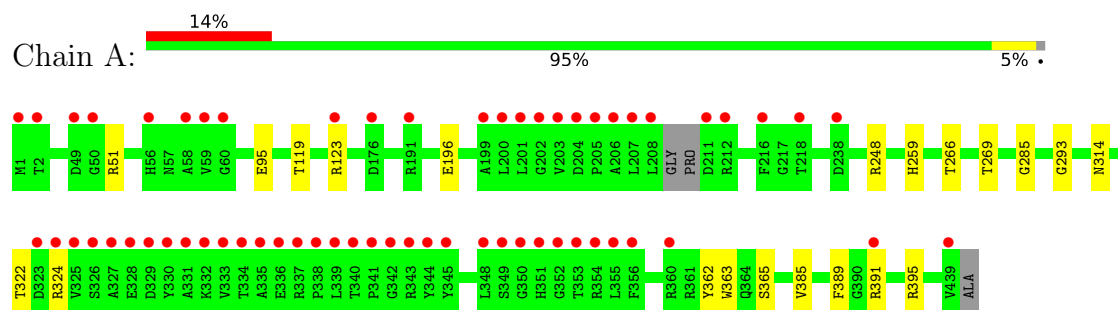
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	203	Total	O	0	0
			203	203		

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: Poly-gamma-glutamate synthesis protein (Capsule biosynthesis protein)



4 Data and refinement statistics

Property	Value	Source
Space group	I 4 2 2	Depositor
Cell constants a, b, c, α , β , γ	128.02Å 128.02Å 170.59Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.65 – 2.30 42.65 – 2.30	Depositor EDS
% Data completeness (in resolution range)	100.0 (42.65-2.30) 100.0 (42.65-2.30)	Depositor EDS
R_{merge}	0.26	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.38 (at 2.29Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.177 , 0.205 0.184 , 0.208	Depositor DCC
R_{free} test set	1568 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å ²)	37.3	Xtriage
Anisotropy	0.421	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 55.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	3582	wwPDB-VP
Average B, all atoms (Å ²)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.18% 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, CA, TRS

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	0.13	0/3420	0.33	0/4657

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3352	0	3260	11	1
2	A	18	0	24	0	0
3	A	1	0	0	0	0
4	A	8	0	10	0	0
5	A	203	0	0	2	0
All	All	3582	0	3294	11	1

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 2.

All (11) 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:391:ARG:HH21	1:A:395:ARG:HD2	1.54	0.72
1:A:322:THR:HG22	1:A:324:ARG:H	1.64	0.61
1:A:119:THR:O	1:A:123:ARG:HG3	2.02	0.59
1:A:391:ARG:NH2	5:A:807:HOH:O	2.38	0.55
1:A:314:ASN:ND2	5:A:811:HOH:O	2.43	0.52
1:A:266:THR:OG1	1:A:269:THR:OG1	2.20	0.48
1:A:248:ARG:HD3	1:A:285:GLY:HA3	1.94	0.48
1:A:363:TRP:HB2	1:A:389:PHE:HB2	1.97	0.47
1:A:365:SER:HB3	1:A:385:VAL:HB	1.99	0.44
1:A:51[B]:ARG:HG2	1:A:95:GLU:HB3	1.99	0.44
1:A:259:HIS:CE1	1:A:293:GLY:HA2	2.52	0.44

All (1) 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)
1:A:196:GLU:OE1	1:A:362:TYR:OH[7_556]	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	444/440 (101%)	433 (98%)	11 (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	A	342/344 (99%)	342 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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

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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 5 ligands modelled in this entry, 1 is monoatomic - leaving 4 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	$\# Z > 2$
2	GOL	A	703	-	5,5,5	0.92	0	5,5,5	0.98	0
2	GOL	A	702	-	5,5,5	0.90	0	5,5,5	0.96	0
2	GOL	A	701	-	5,5,5	0.99	0	5,5,5	0.95	0
4	TRS	A	705	-	7,7,7	0.33	0	9,9,9	0.43	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
2	GOL	A	703	-	-	4/4/4/4	-
2	GOL	A	702	-	-	4/4/4/4	-
2	GOL	A	701	-	-	0/4/4/4	-
4	TRS	A	705	-	-	6/9/9/9	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (14) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	703	GOL	O1-C1-C2-C3
2	A	703	GOL	C1-C2-C3-O3
4	A	705	TRS	C2-C-C1-O1
4	A	705	TRS	C3-C-C1-O1
4	A	705	TRS	N-C-C1-O1
4	A	705	TRS	C1-C-C2-O2
4	A	705	TRS	C3-C-C2-O2
4	A	705	TRS	N-C-C2-O2
2	A	703	GOL	O1-C1-C2-O2
2	A	702	GOL	O1-C1-C2-C3
2	A	702	GOL	C1-C2-C3-O3
2	A	702	GOL	O1-C1-C2-O2
2	A	702	GOL	O2-C2-C3-O3
2	A	703	GOL	O2-C2-C3-O3

There are no ring outliers.

No monomer is involved in short contacts.

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	A	437/440 (99%)	0.30	61 (13%) 7 8	17, 36, 104, 166	11 (2%)

All (61) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	203	VAL	8.4
1	A	338	PRO	5.9
1	A	439	VAL	5.9
1	A	201	LEU	5.8
1	A	335	ALA	5.6
1	A	59	VAL	5.2
1	A	334	THR	5.1
1	A	200	LEU	5.0
1	A	1	MET	4.8
1	A	208	LEU	4.6
1	A	348	LEU	4.5
1	A	58	ALA	4.5
1	A	325	VAL	4.5
1	A	211	ASP	4.3
1	A	199	ALA	4.3
1	A	202	GLY	4.3
1	A	336	GLU	4.2
1	A	207	LEU	4.2
1	A	333	VAL	4.1
1	A	323	ASP	4.1
1	A	353	THR	4.1
1	A	351	HIS	4.1
1	A	360[A]	ARG	4.0
1	A	339	LEU	4.0
1	A	212	ARG	3.9
1	A	332	LYS	3.7
1	A	337	ARG	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	60	GLY	3.6
1	A	327	ALA	3.5
1	A	350	GLY	3.5
1	A	355	LEU	3.4
1	A	206	ALA	3.4
1	A	330	TYR	3.4
1	A	331	ALA	3.3
1	A	356	PHE	3.3
1	A	50[A]	GLY	3.3
1	A	328	GLU	3.2
1	A	352	GLY	3.1
1	A	205	PRO	3.0
1	A	2	THR	2.9
1	A	324	ARG	2.8
1	A	49[A]	ASP	2.8
1	A	345	TYR	2.7
1	A	326	SER	2.7
1	A	56	HIS	2.7
1	A	342	GLY	2.7
1	A	340	THR	2.7
1	A	354	ARG	2.7
1	A	344	TYR	2.5
1	A	218	THR	2.5
1	A	343	ARG	2.5
1	A	204	ASP	2.3
1	A	216	PHE	2.3
1	A	238[A]	ASP	2.3
1	A	349	SER	2.3
1	A	123	ARG	2.2
1	A	391	ARG	2.2
1	A	176	ASP	2.2
1	A	329	ASP	2.2
1	A	341	PRO	2.1
1	A	191	ARG	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 [i](#)

There are no oligosaccharides in this entry.

6.4 Ligands [i](#)

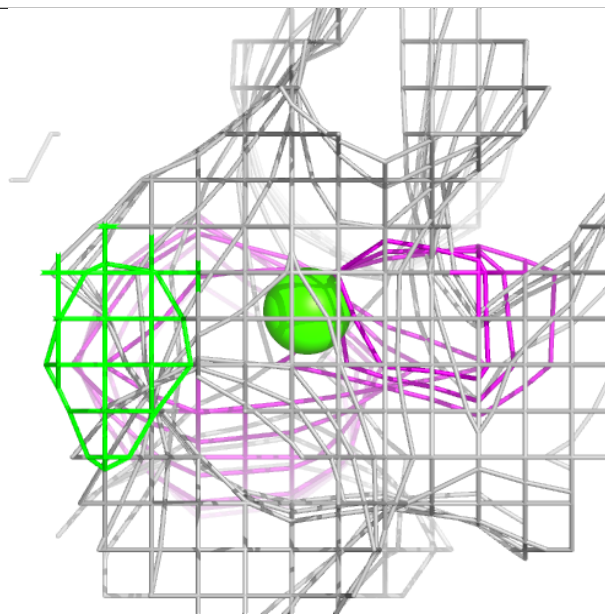
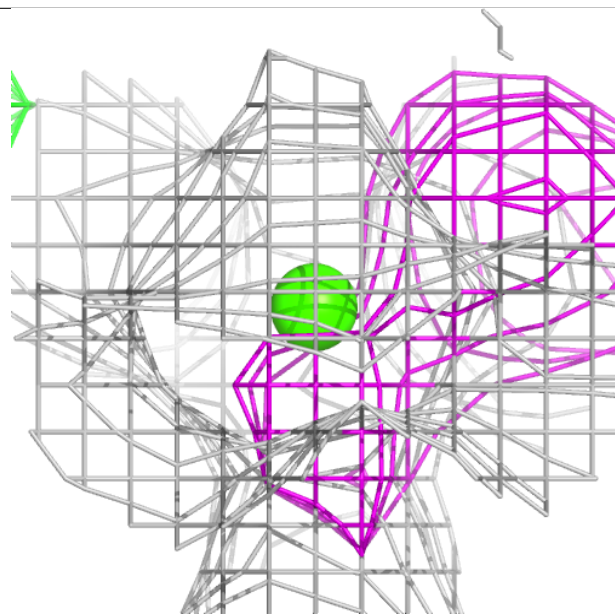
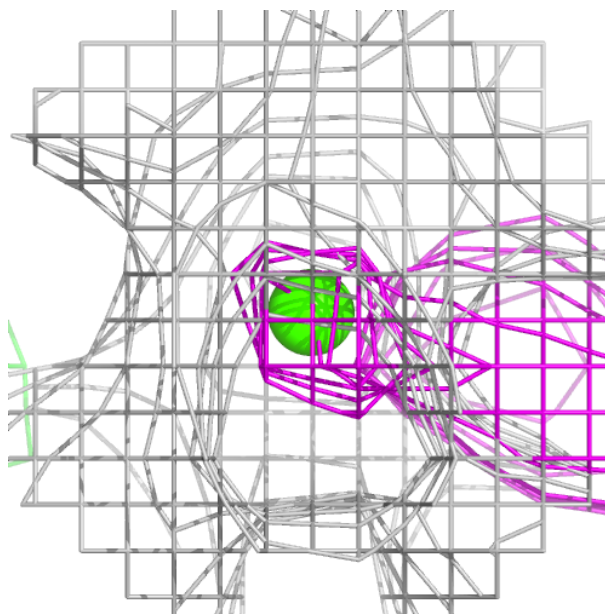
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
4	TRS	A	705	8/8	0.81	0.17	45,54,60,64	0
2	GOL	A	703	6/6	0.86	0.14	64,78,85,86	0
2	GOL	A	702	6/6	0.88	0.18	79,87,94,98	0
2	GOL	A	701	6/6	0.89	0.20	67,72,83,95	0
3	CA	A	704	1/1	0.95	0.18	68,68,68,68	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 A 704:

$2mF_o-DF_c$ (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.