



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 5, 2026 – 03:32 AM UTC

PDB ID : 9NEH / pdb_00009neh
Title : The 1.48 Angstrom crystal structure of galactose oxidase variant with genetically incorporated Cl2-Tyr495
Authors : Liu, A.; Li, J.; Graciano, A.
Deposited on : 2025-02-19
Resolution : 1.48 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

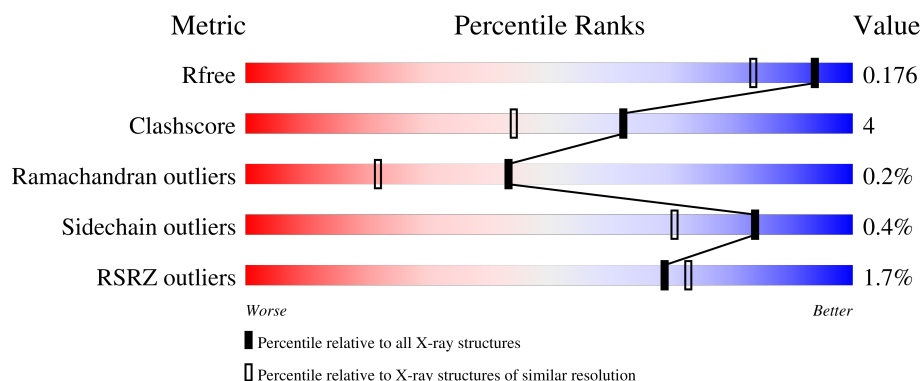
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

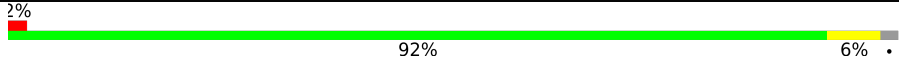
The reported resolution of this entry is 1.48 Å.

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}	180053	6779 (1.50-1.46)
Clashscore	190562	7025 (1.50-1.46)
Ramachandran outliers	187476	6917 (1.50-1.46)
Sidechain outliers	187428	6914 (1.50-1.46)
RSRZ outliers	180081	6781 (1.50-1.46)

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	648	

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
4	ACT	A	1004	-	-	X	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	ACT	A	1009	-	-	X	-
4	ACT	A	1013	-	-	X	-
5	GOL	A	1010	-	-	X	-

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 5511 atoms, of which 42 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 F5/8 type C domain-containing protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	638	Total	C	Cl	N	O	S	0	1	0
			4835	3021	2	839	955	18			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP A0AAN6HGK3
A	10	PRO	SER	conflict	UNP A0AAN6HGK3
A	70	VAL	MET	conflict	UNP A0AAN6HGK3
A	195	GLU	GLY	conflict	UNP A0AAN6HGK3
A	494	ALA	VAL	conflict	UNP A0AAN6HGK3
A	535	ASP	ASN	conflict	UNP A0AAN6HGK3
A	640	LEU	-	expression tag	UNP A0AAN6HGK3
A	641	GLU	-	expression tag	UNP A0AAN6HGK3
A	642	HIS	-	expression tag	UNP A0AAN6HGK3
A	643	HIS	-	expression tag	UNP A0AAN6HGK3
A	644	HIS	-	expression tag	UNP A0AAN6HGK3
A	645	HIS	-	expression tag	UNP A0AAN6HGK3
A	646	HIS	-	expression tag	UNP A0AAN6HGK3
A	647	HIS	-	expression tag	UNP A0AAN6HGK3

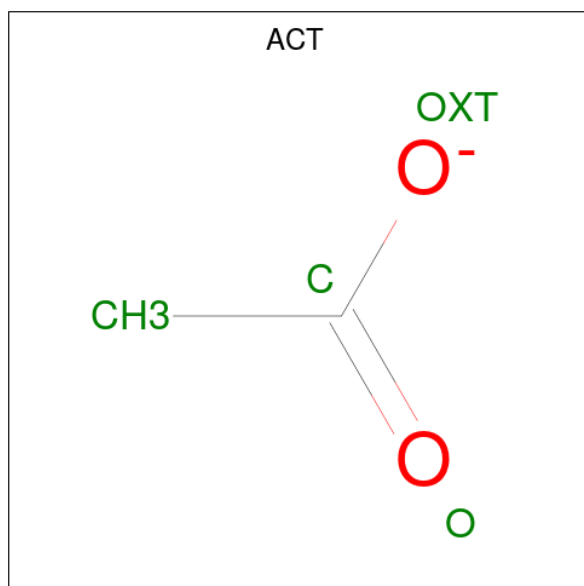
- Molecule 2 is CALCIUM ION (CCD ID: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

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

- Molecule 3 is COPPER (II) ION (CCD ID: CU) (formula: Cu) (labeled as "Ligand of Interest" by depositor).

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

- Molecule 4 is ACETATE ION (CCD ID: ACT) (formula: $C_2H_3O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	H	O	0	0
			7	2	3	2		
4	A	1	Total	C	H	O	0	0
			7	2	3	2		
4	A	1	Total	C	H	O	0	0
			7	2	3	2		
4	A	1	Total	C	H	O	0	0
			7	2	3	2		
4	A	1	Total	C	O		0	0
			4	2	2			
4	A	1	Total	C	H	O	0	0
			7	2	3	2		

- Molecule 5 is GLYCEROL (CCD ID: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	H	O	0	0
			14	3	8	3		
5	A	1	Total	C	H	O	0	0
			14	3	8	3		
5	A	1	Total	C	H	O	0	0
			14	3	8	3		
5	A	1	Total	C	O		0	0
			6	3	3			

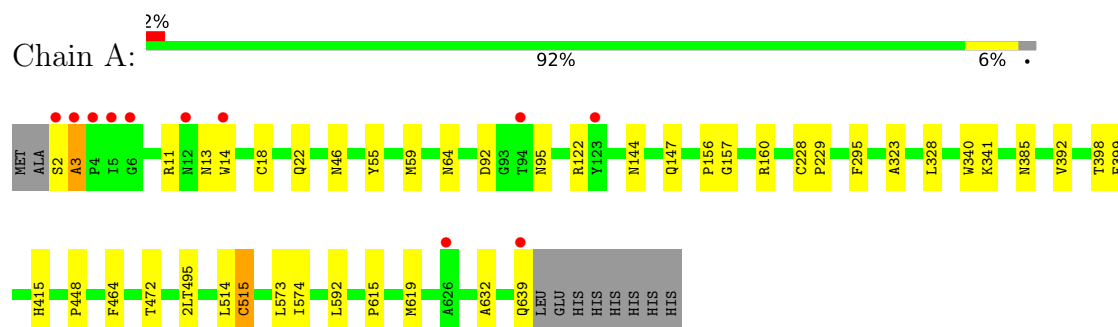
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	580	Total	O	0	0
			580	580		

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: F5/8 type C domain-containing protein



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	97.72Å 89.02Å 86.25Å 90.00° 117.96° 90.00°	Depositor
Resolution (Å)	48.56 – 1.48 48.56 – 1.48	Depositor EDS
% Data completeness (in resolution range)	99.4 (48.56-1.48) 93.5 (48.56-1.48)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.86 (at 1.48Å)	Xtriage
Refinement program	PHENIX (1.21.2-5419)	Depositor
R, R_{free}	0.161 , 0.177 0.161 , 0.176	Depositor DCC
R_{free} test set	2000 reflections (1.86%)	wwPDB-VP
Wilson B-factor (Å ²)	16.6	Xtriage
Anisotropy	0.110	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 38.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	5511	wwPDB-VP
Average B, all atoms (Å ²)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.71% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, ACT, CA, CU, 2LT

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	0.29	0/4952	0.57	0/6758

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

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	4835	0	4598	36	0
2	A	1	0	0	0	0
3	A	1	0	0	0	0
4	A	28	18	21	9	0
5	A	24	24	32	6	0
6	A	580	0	0	8	0
All	All	5469	42	4651	39	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.

The worst 5 of 39 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:619:MET:HB3	1:A:632:ALA:HB1	1.51	0.92
1:A:22:GLN:HA	6:A:1105:HOH:O	1.70	0.90
1:A:157:GLY:H	5:A:1010:GOL:H31	1.43	0.83
1:A:157:GLY:H	5:A:1010:GOL:C3	1.95	0.79
1:A:472:THR:OG1	4:A:1008:ACT:H2	1.87	0.75

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	636/648 (98%)	615 (97%)	20 (3%)	1 (0%)	43	22

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	3	ALA

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	526/534 (98%)	524 (100%)	2 (0%)	84	70

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	514	LEU
1	A	515	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	133	ASN
1	A	355	ASN
1	A	413	ASN
1	A	489	ASN
1	A	552	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	$\# Z > 2$
1	2LT	A	495	1	13,14,15	0.87	0	14,19,21	1.46	3 (21%)

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
1	2LT	A	495	1	-	2/5/6/8	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	495	2LT	CZ-CE1-CL1	3.92	123.67	118.81
1	A	495	2LT	CE2-CD2-CG	-2.39	118.91	120.46
1	A	495	2LT	CD1-CE1-CZ	-2.03	119.06	121.77

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	495	2LT	CA-CB-CG-CD1
1	A	495	2LT	CA-CB-CG-CD2

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 13 ligands modelled in this entry, 2 are monoatomic - leaving 11 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	ACT	A	1003	-	3,3,3	0.77	0	3,3,3	1.70	2 (66%)
4	ACT	A	1013	-	3,3,3	0.94	0	3,3,3	1.47	1 (33%)
4	ACT	A	1011	-	3,3,3	1.47	1 (33%)	3,3,3	1.32	0
4	ACT	A	1009	-	3,3,3	0.73	0	3,3,3	1.68	1 (33%)
4	ACT	A	1008	-	3,3,3	1.27	0	3,3,3	1.55	0
5	GOL	A	1010	-	5,5,5	0.81	0	5,5,5	0.80	0
5	GOL	A	1012	-	5,5,5	0.93	0	5,5,5	1.13	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	ACT	A	1006	-	3,3,3	0.98	0	3,3,3	1.44	0
4	ACT	A	1004	-	3,3,3	0.82	0	3,3,3	1.59	1 (33%)
5	GOL	A	1007	-	5,5,5	0.73	0	5,5,5	0.88	0
5	GOL	A	1005	-	5,5,5	0.90	0	5,5,5	1.23	1 (20%)

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	GOL	A	1012	-	-	4/4/4/4	-
5	GOL	A	1010	-	-	0/4/4/4	-
5	GOL	A	1007	-	-	2/4/4/4	-
5	GOL	A	1005	-	-	2/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1011	ACT	CH3-C	2.21	1.57	1.49

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	1005	GOL	C3-C2-C1	-2.38	103.06	111.80
4	A	1009	ACT	OXT-C-O	2.27	130.44	122.03
4	A	1004	ACT	OXT-C-O	2.20	130.20	122.03
4	A	1003	ACT	OXT-C-O	2.11	129.86	122.03
4	A	1013	ACT	OXT-C-O	2.07	129.69	122.03

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	1005	GOL	O1-C1-C2-C3
5	A	1007	GOL	C1-C2-C3-O3
5	A	1012	GOL	O1-C1-C2-C3
5	A	1012	GOL	C1-C2-C3-O3
5	A	1005	GOL	O1-C1-C2-O2

There are no ring outliers.

6 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1003	ACT	1	0
4	A	1013	ACT	2	0
4	A	1009	ACT	3	0
4	A	1008	ACT	1	0
5	A	1010	GOL	6	0
4	A	1004	ACT	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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	637/648 (98%)	0.07	11 (1%) 69 73	14, 22, 41, 53	1 (0%)

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	2	SER	5.4
1	A	12	ASN	3.7
1	A	639	GLN	3.4
1	A	5	ILE	3.3
1	A	3	ALA	3.1

6.2 Non-standard residues in protein, DNA, RNA chains [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
1	2LT	A	495	14/15	0.98	0.06	14,16,21,22	0

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,

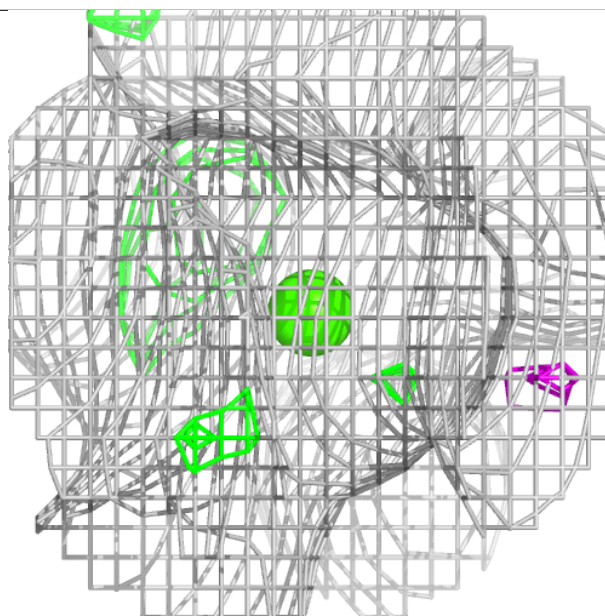
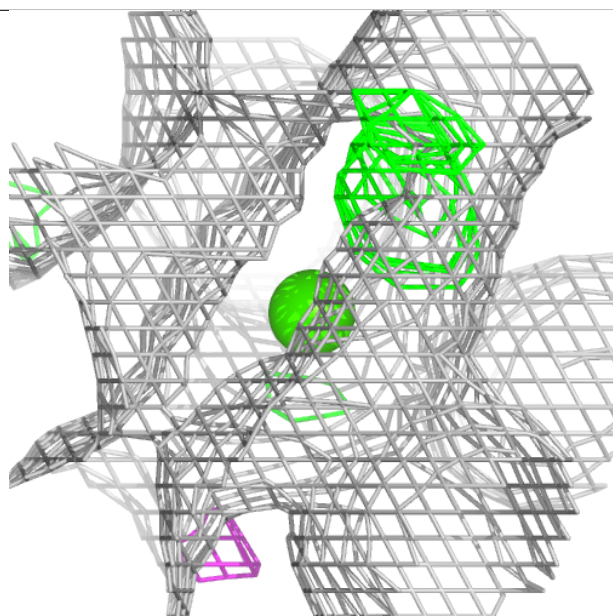
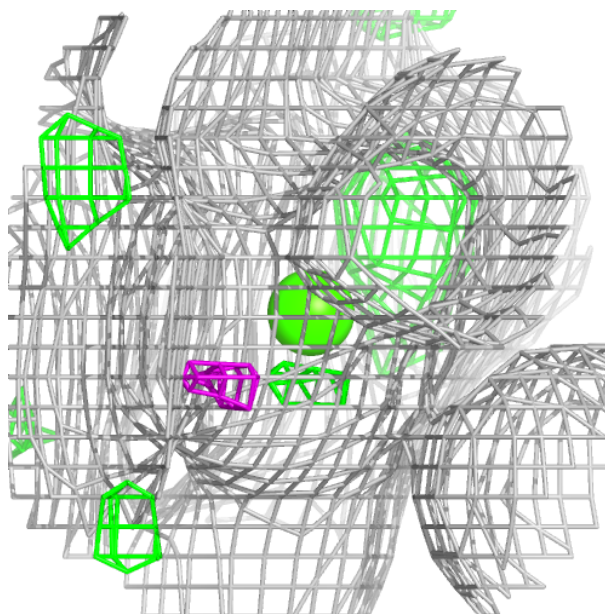
median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
4	ACT	A	1011	4/4	0.42	0.24	50,52,57,61	0
4	ACT	A	1006	4/4	0.71	0.16	28,35,44,44	0
5	GOL	A	1012	6/6	0.71	0.19	47,50,52,54	0
5	GOL	A	1005	6/6	0.75	0.15	60,72,78,82	0
4	ACT	A	1004	4/4	0.77	0.17	44,49,53,53	0
4	ACT	A	1003	4/4	0.78	0.15	49,53,59,59	0
4	ACT	A	1013	4/4	0.78	0.15	58,58,69,69	0
4	ACT	A	1008	4/4	0.85	0.21	24,25,29,36	0
5	GOL	A	1010	6/6	0.90	0.10	23,33,47,50	0
5	GOL	A	1007	6/6	0.90	0.10	24,33,40,45	0
4	ACT	A	1009	4/4	0.92	0.10	13,15,21,23	0
2	CA	A	1001	1/1	0.98	0.05	34,34,34,34	0
3	CU	A	1002	1/1	1.00	0.02	17,17,17,17	1

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.

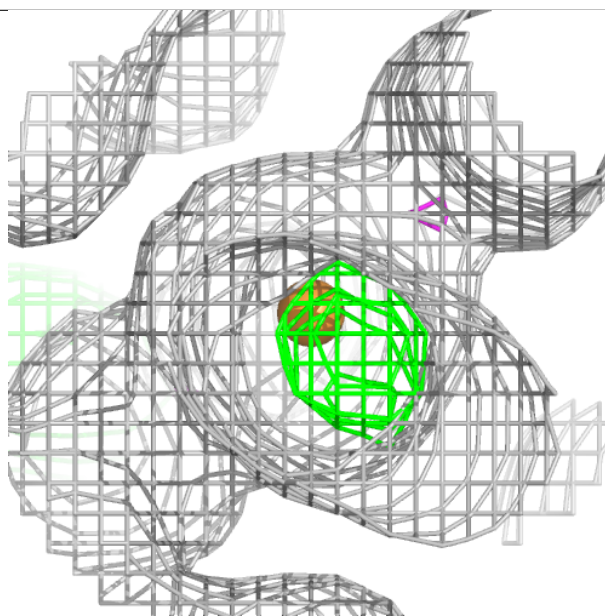
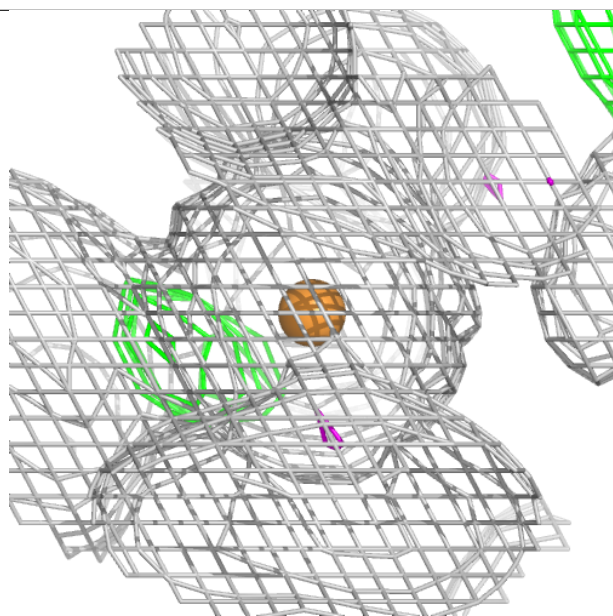
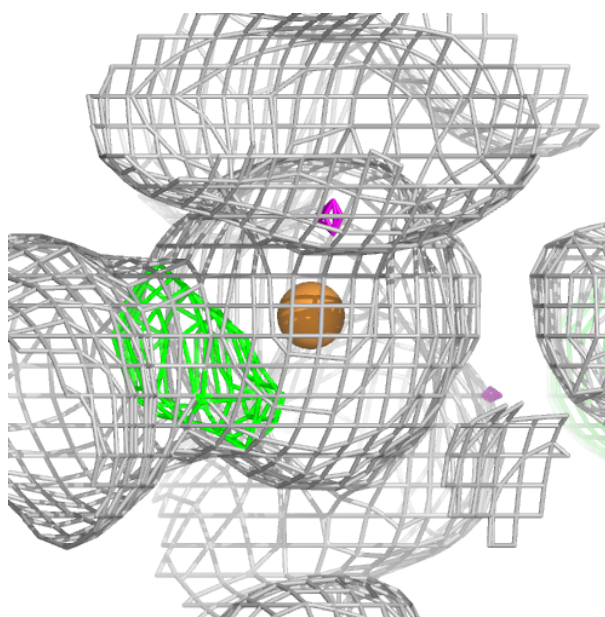
Electron density around CA A 1001:

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



Electron density around CU A 1002:

$2mF_o-DF_c$ (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.