



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 6, 2026 – 10:14 PM UTC

PDB ID : 7YOH / pdb_00007yoh
Title : Crystal structure of I88L single mutant of O-acetylserine sulfhydrylase from Haemophilus influenzae in complex with high-affinity inhibitory peptide from serine acetyltransferase of Haemophilus influenzae at 2.5 Å
Authors : Kumar, N.; Rahisuddin, R.; Saini, N.; Singh, A.K.
Deposited on : 2022-08-01
Resolution : 2.50 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 Phenix2.0
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

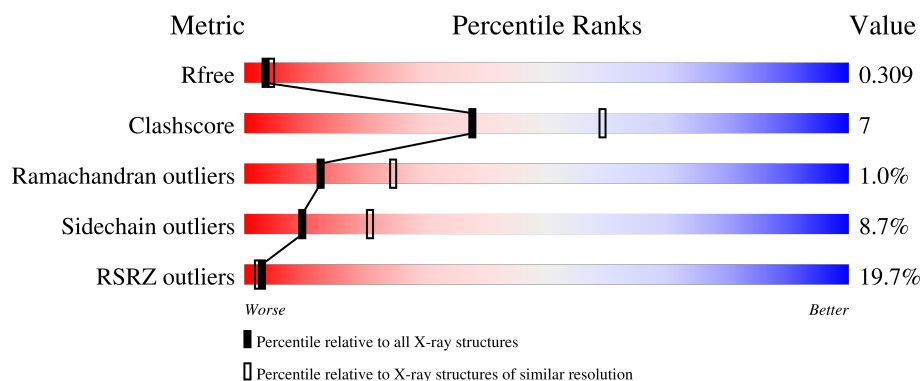
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	5829 (2.50-2.50)
Clashscore	190562	6492 (2.50-2.50)
Ramachandran outliers	187476	6378 (2.50-2.50)
Sidechain outliers	187428	6380 (2.50-2.50)
RSRZ outliers	180081	5833 (2.50-2.50)

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	316	18% 77% 16% ...
2	P	7	71% 86% 14%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 2304 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 Cysteine synthase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	308	Total	C	N	O	P	S	0	0	0
			2239	1401	391	438	1	8			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	88	LEU	ILE	engineered mutation	UNP P45040

- Molecule 2 is a protein called Peptide from Serine acetyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	P	7	Total	C	N	O	S	0	0	0
			47	27	9	10	1			

- Molecule 3 is CHLORIDE ION (CCD ID: CL) (formula: Cl) (labeled as "Ligand of Interest" by depositor).

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

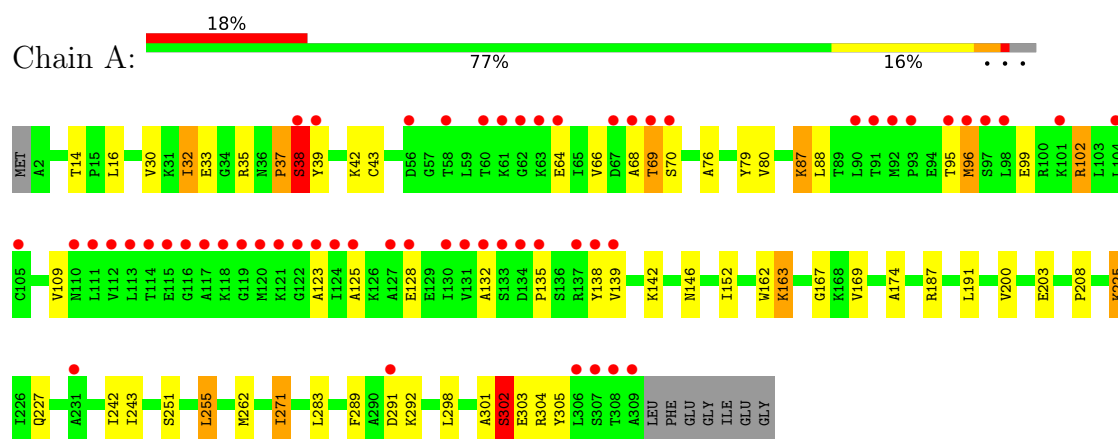
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	16	Total	O	0	0
			16	16		
4	P	1	Total	O	0	0
			1	1		

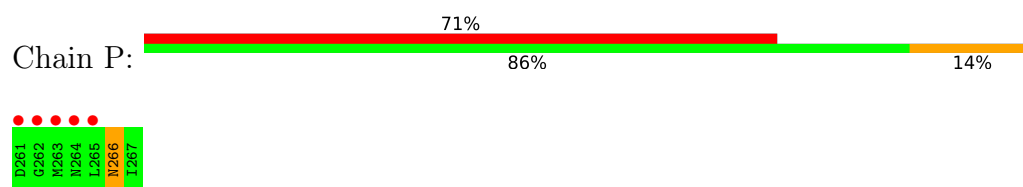
3 Residue-property plots [i](#)

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: Cysteine synthase



- Molecule 2: Peptide from Serine acetyltransferase



4 Data and refinement statistics

Property	Value	Source
Space group	I 41	Depositor
Cell constants a, b, c, α , β , γ	112.75Å 112.75Å 44.46Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.36 – 2.50 41.36 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.6 (41.36-2.50) 99.6 (41.36-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.18 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.243 , 0.295 0.255 , 0.309	Depositor DCC
R_{free} test set	515 reflections (5.22%)	wwPDB-VP
Wilson B-factor (Å ²)	35.7	Xtriage
Anisotropy	0.109	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 44.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.059 for -k,-h,-l	Xtriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	2304	wwPDB-VP
Average B, all atoms (Å ²)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.24% 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: LLP, CL

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	0.59	0/2246	1.24	7/3050 (0.2%)
2	P	0.55	0/46	1.50	1/59 (1.7%)
All	All	0.59	0/2292	1.25	8/3109 (0.3%)

There are no bond length outliers.

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	174	ALA	CA-C-N	6.07	125.86	120.10
1	A	174	ALA	C-N-CA	6.07	125.86	120.10
1	A	163	LYS	CB-CA-C	5.71	119.84	110.88
1	A	87	LYS	CB-CA-C	5.62	118.81	109.53
1	A	37	PRO	N-CA-C	-5.36	103.04	111.34
1	A	305	TYR	CA-C-N	5.33	129.69	121.53
1	A	305	TYR	C-N-CA	5.33	129.69	121.53
2	P	266	ASN	CA-CB-CG	-5.08	107.52	112.60

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	2239	0	2224	32	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	P	47	0	38	1	0
3	A	1	0	0	0	0
4	A	16	0	0	1	0
4	P	1	0	0	0	0
All	All	2304	0	2262	32	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 7.

All (32) 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:37:PRO:O	1:A:43:CYS:SG	2.24	0.96
1:A:255:LEU:HD13	1:A:271:ILE:HD12	1.48	0.95
1:A:69:THR:HA	2:P:266:ASN:OD1	1.99	0.63
1:A:291:ASP:C	1:A:292:LYS:HD3	2.30	0.56
1:A:200:VAL:HG21	1:A:283:LEU:HD11	1.89	0.55
1:A:38:SER:O	1:A:304:ARG:NH2	2.36	0.55
1:A:251:SER:O	1:A:255:LEU:HB2	2.07	0.54
1:A:76:ALA:O	1:A:80:VAL:HG23	2.08	0.53
1:A:302:SER:O	1:A:303:GLU:HB2	2.09	0.53
1:A:68:ALA:HB1	1:A:123:ALA:HB1	1.90	0.53
1:A:227:GLN:HA	1:A:227:GLN:OE1	2.08	0.52
1:A:33:GLU:HB2	1:A:298:LEU:O	2.08	0.52
1:A:301:ALA:C	1:A:302:SER:O	2.54	0.49
1:A:303:GLU:HB2	4:A:504:HOH:O	2.12	0.49
1:A:187:ARG:O	1:A:191:LEU:HB2	2.13	0.48
1:A:14:THR:HB	1:A:32:ILE:O	2.13	0.48
1:A:132:ALA:O	1:A:135:PRO:HD3	2.14	0.47
1:A:162:TRP:CE3	1:A:167:GLY:HA2	2.49	0.47
1:A:125:ALA:HA	1:A:128:GLU:OE1	2.15	0.47
1:A:37:PRO:O	1:A:39:TYR:N	2.42	0.46
1:A:96:MET:HE3	1:A:96:MET:HB3	1.63	0.44
1:A:289:PHE:HA	1:A:292:LYS:HG2	2.01	0.43
1:A:88:LEU:HD23	1:A:109:VAL:HG22	2.00	0.42
1:A:142:LYS:O	1:A:146:ASN:HB2	2.19	0.42
1:A:102:ARG:HE	1:A:102:ARG:HB2	1.34	0.42
1:A:225:LYS:O	1:A:271:ILE:HD13	2.20	0.41
1:A:64:GLU:O	1:A:138:TYR:HA	2.21	0.41
1:A:128:GLU:O	1:A:132:ALA:N	2.53	0.40
1:A:242:ILE:HD12	1:A:243:ILE:HD13	2.03	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:203:GLU:CD	1:A:208:PRO:HA	2.47	0.40
1:A:200:VAL:HG21	1:A:283:LEU:CD1	2.51	0.40
1:A:16:LEU:HA	1:A:30:VAL:O	2.21	0.40

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	305/316 (96%)	276 (90%)	26 (8%)	3 (1%)	12	24
2	P	5/7 (71%)	5 (100%)	0	0	100	100
All	All	310/323 (96%)	281 (91%)	26 (8%)	3 (1%)	12	24

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	302	SER
1	A	38	SER
1	A	169	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.

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	225/247 (91%)	205 (91%)	20 (9%)	9	20
2	P	4/6 (67%)	4 (100%)	0	100	100
All	All	229/253 (90%)	209 (91%)	20 (9%)	9	21

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	32	ILE
1	A	35	ARG
1	A	38	SER
1	A	66	VAL
1	A	69	THR
1	A	70	SER
1	A	79	TYR
1	A	87	LYS
1	A	95	THR
1	A	96	MET
1	A	99	GLU
1	A	102	ARG
1	A	139	VAL
1	A	152	ILE
1	A	163	LYS
1	A	225	LYS
1	A	255	LEU
1	A	262	MET
1	A	271	ILE
1	A	302	SER

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

Mol	Chain	Res	Type
1	A	151	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	LLP	A	42	1	23,24,25	0.45	0	25,32,34	0.79	1 (4%)

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	LLP	A	42	1	-	4/16/17/19	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	42	LLP	OP4-C5'-C5	2.06	113.22	109.36

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	42	LLP	C4-C4'-NZ-CE
1	A	42	LLP	CA-CB-CG-CD
1	A	42	LLP	C5'-OP4-P-OP2
1	A	42	LLP	CE-CD-CG-CB

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 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

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	307/316 (97%)	1.11	57 (18%) 3 2	20, 32, 83, 102	0
2	P	7/7 (100%)	3.39	5 (71%) 0 0	49, 56, 84, 89	0
All	All	314/323 (97%)	1.16	62 (19%) 3 2	20, 32, 85, 102	0

All (62) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	122	GLY	6.9
1	A	123	ALA	6.1
1	A	124	ILE	6.1
1	A	119	GLY	5.8
2	P	264	ASN	5.7
2	P	263	MET	5.7
1	A	69	THR	5.0
1	A	70	SER	4.8
1	A	90	LEU	4.6
1	A	116	GLY	4.6
1	A	125	ALA	4.6
1	A	131	VAL	4.5
1	A	117	ALA	4.5
1	A	114	THR	4.4
1	A	96	MET	4.3
1	A	132	ALA	4.2
1	A	138	TYR	4.2
1	A	133	SER	4.0
2	P	261	ASP	4.0
1	A	306	LEU	4.0
1	A	137	ARG	3.9
1	A	121	LYS	3.7
1	A	130	ILE	3.7
1	A	63	LYS	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	91	THR	3.7
1	A	135	PRO	3.5
1	A	98	LEU	3.5
1	A	309	ALA	3.4
1	A	97	SER	3.4
1	A	115	GLU	3.4
1	A	38	SER	3.3
1	A	112	VAL	3.3
1	A	120	MET	3.3
1	A	95	THR	3.2
1	A	118	LYS	3.2
1	A	113	LEU	3.0
1	A	93	PRO	3.0
1	A	308	THR	3.0
1	A	127	ALA	3.0
1	A	307	SER	2.8
2	P	262	GLY	2.8
1	A	104	LEU	2.8
2	P	265	LEU	2.7
1	A	92	MET	2.7
1	A	231	ALA	2.7
1	A	68	ALA	2.5
1	A	64	GLU	2.5
1	A	139	VAL	2.5
1	A	67	ASP	2.4
1	A	60	THR	2.4
1	A	128	GLU	2.3
1	A	61	LYS	2.3
1	A	134	ASP	2.3
1	A	56	ASP	2.2
1	A	58	THR	2.2
1	A	111	LEU	2.2
1	A	101	LYS	2.2
1	A	62	GLY	2.1
1	A	39	TYR	2.1
1	A	105	CYS	2.1
1	A	291	ASP	2.1
1	A	110	ASN	2.0

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($\text{\AA}^2$)	Q<0.9
1	LLP	A	42	24/25	0.92	0.10	25,31,34,35	0

6.3 Carbohydrates [i](#)

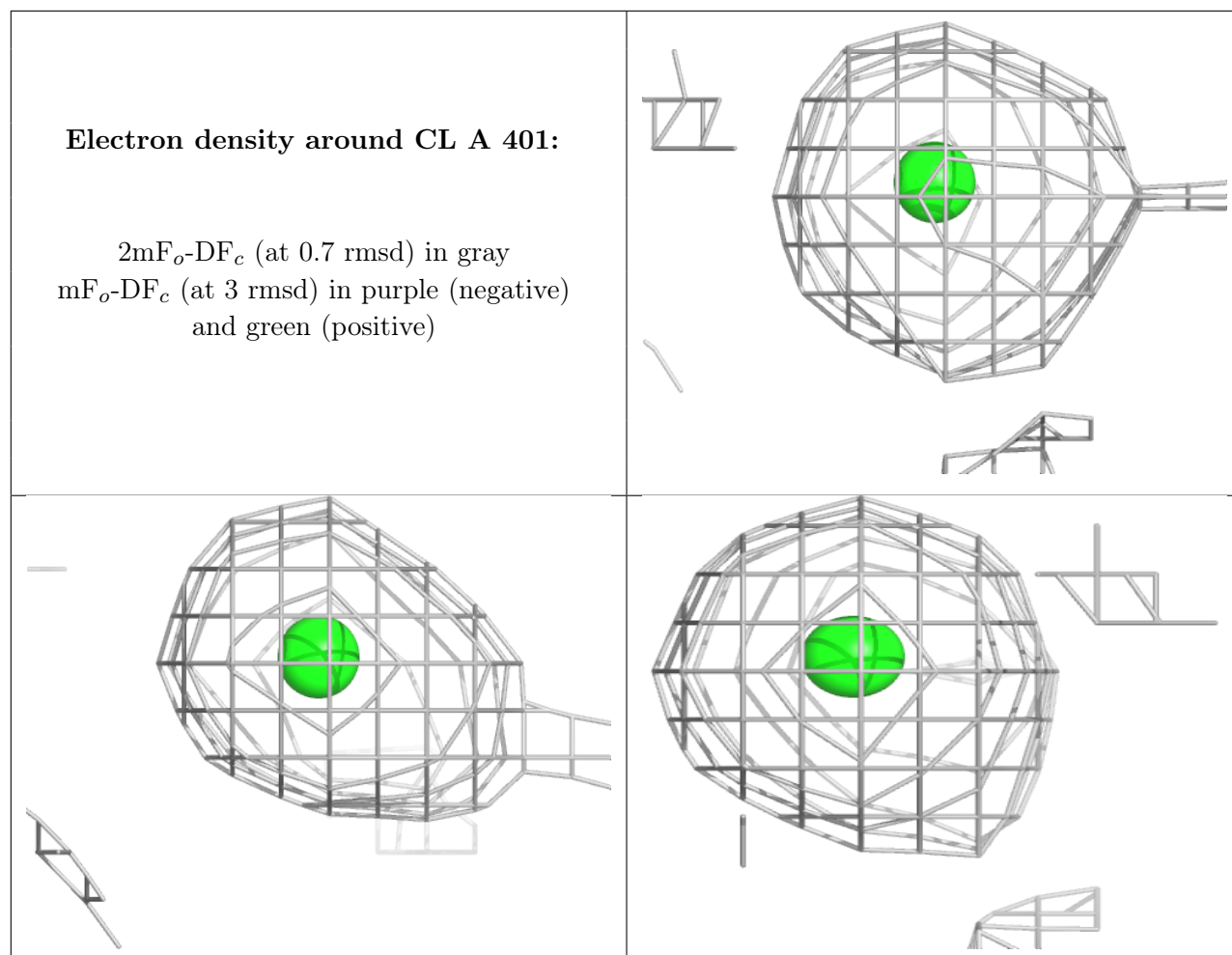
There are no oligosaccharides in this entry.

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($\text{\AA}^2$)	Q<0.9
3	CL	A	401	1/1	0.97	0.05	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.



6.5 Other polymers [i](#)

There are no such residues in this entry.