wwPDB X-ray Structure Validation Summary Report (i)

Nov 12, 2026 – 12:09 pm GMT

PDB ID	:	28KT / pdb_000028kt
Title	:	Lipase from Geobacillus stearothermophilus T6 methanol and acid-stable variant H86Y/E251C/A269T/G332C/R374W
Authors	:	Yehishalom, N.; Arazi, Y.; Pazy, Y.; Fishman, A.
Deposited on	:	2026-02-05
Resolution	:	2.40 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 (i) 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 (i)) were used in the production of this report:

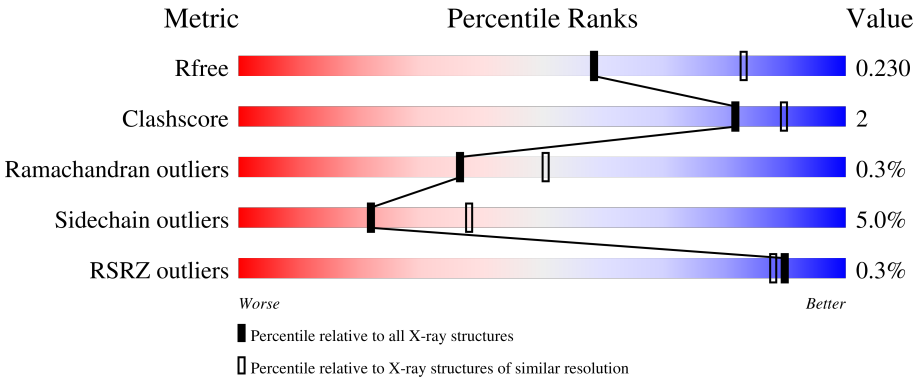
MolProbity	:	4-5-2 with Phenix2.2
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.2
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.015 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.51

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 2.40 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	4912 (2.40-2.40)
Clashscore	190562	5391 (2.40-2.40)
Ramachandran outliers	187476	5320 (2.40-2.40)
Sidechain outliers	187428	5321 (2.40-2.40)
RSRZ outliers	180081	4916 (2.40-2.40)

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 <=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	391	84% 14%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 6207 atoms, of which 2904 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 triacylglycerol lipase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	391	Total	C	H	N	O	S	0	5	0
			6034	1991	2904	557	572	10			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	86	TYR	HIS	engineered mutation	UNP Q93A71
A	251	CYS	GLU	engineered mutation	UNP Q93A71
A	269	THR	ALA	engineered mutation	UNP Q93A71
A	323	ALA	THR	conflict	UNP Q93A71
A	332	CYS	GLY	engineered mutation	UNP Q93A71
A	374	TRP	ARG	engineered mutation	UNP Q93A71
A	390	HIS	-	expression tag	UNP Q93A71
A	391	HIS	-	expression tag	UNP Q93A71
A	392	HIS	-	expression tag	UNP Q93A71
A	393	HIS	-	expression tag	UNP Q93A71
A	394	HIS	-	expression tag	UNP Q93A71
A	395	HIS	-	expression tag	UNP Q93A71

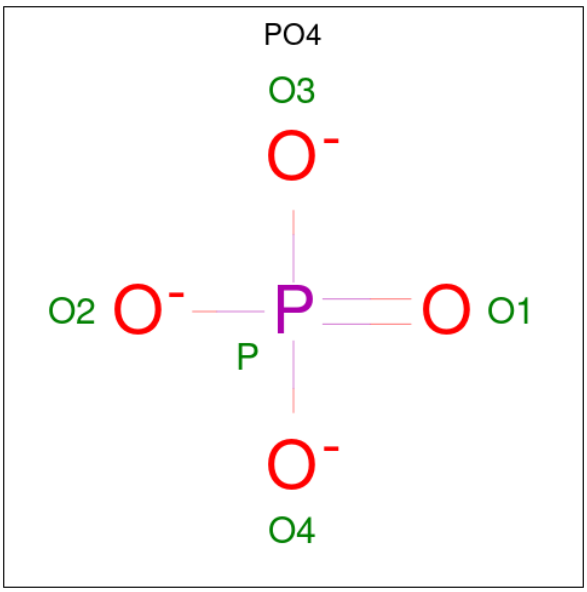
- Molecule 2 is ZINC ION (CCD ID: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

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

- 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 PHOSPHATE ION (CCD ID: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O P 5 4 1	0	0
4	A	1	Total O P 5 4 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	161	Total O 161 161	0	0

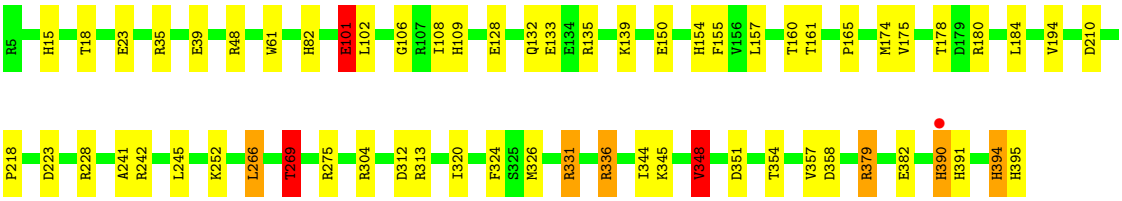
3 Residue-property plots (i)

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: triacylglycerol lipase

Chain A:





4 Data and refinement statistics (i)

Property	Value
Space group	P 21 21 21
Cell constants a, b, c, α, β, γ	48.71Å 72.77Å 105.95Å 90.00° 90.00° 90.00°
Resolution (Å)	9.41 – 2.40 9.41 – 2.40
% Data completeness (in resolution range)	97.6 (9.41-2.40) 97.6 (9.41-2.40)
R_{merge}	0.09
R_{sym}	(Not available)
$< I/\sigma(I) >$ [1]	5.92 (at 2.39Å)
Refinement program	REFMAC 5.8.0431 (refmacat 0.4.126), REFMAC 5.8.0431 (refmacat 0.4
R, R_{free}	0.156 , 0.228 0.161 , 0.230
R_{free} test set	785 reflections (5.11%)
Wilson B-factor (Å^2)	22.6
Anisotropy	0.291
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.56 , 72.4
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$
Estimated twinning fraction	No twinning to report.
F_o,F_c correlation	0.94
Total number of atoms	6207
Average B, all atoms (Å^2)	25.0

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.58% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: PO4, ZN, CA

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	0.88	0/3246	1.51	36/4416 (0.8%)

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	A	0	7

There are no bond length outliers.

The worst 5 of 36 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	155[A]	PHE	CA-CB-CG	10.84	124.64	113.80
1	A	155[B]	PHE	CA-CB-CG	10.84	124.64	113.80
1	A	382	GLU	CB-CA-C	9.82	130.41	110.38
1	A	379	ARG	CA-CB-CG	8.74	131.58	114.10
1	A	218	PRO	CB-CA-C	8.43	122.02	110.98

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	304	ARG	Sidechain
1	A	313	ARG	Sidechain
1	A	331	ARG	Sidechain
1	A	336[A]	ARG	Sidechain
1	A	35	ARG	Sidechain

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	3130	2904	2965	15	0
2	A	1	0	0	0	0
3	A	1	0	0	0	0
4	A	10	0	0	0	0
5	A	161	0	0	1	0
All	All	3303	2904	2965	15	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 2.

The worst 5 of 15 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:395:HIS:C	5:A:518:HOH:O	2.00	1.04
1:A:184:LEU:HD23	1:A:184:LEU:C	2.31	0.56
1:A:102:LEU:HD21	1:A:108:ILE:HG23	1.90	0.53
1:A:101:GLU:HG2	1:A:106:GLY:HA3	1.93	0.51
1:A:184:LEU:HD23	1:A:184:LEU:O	2.12	0.50

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	394/391 (101%)	379 (96%)	14 (4%)	1 (0%)	36 50

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	394	HIS

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	325/321 (101%)	308 (95%)	17 (5%)	21 36

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	348	VAL
1	A	390	HIS
1	A	180	ARG
1	A	245	LEU
1	A	252	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	185	GLN
1	A	388	GLN
1	A	392	HIS

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 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	PO4	A	404	-	4,4,4	0.80	0	6,6,6	0.51	0
4	PO4	A	403	-	4,4,4	0.76	0	6,6,6	0.43	0

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, 95^{th} 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	391/391 (100%)	-0.83	1 (0%)	90	88	10, 21, 45, 86	3 (0%)

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	390	HIS	2.2

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

There are no oligosaccharides in this entry.

6.4 Ligands

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, 95^{th} 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.

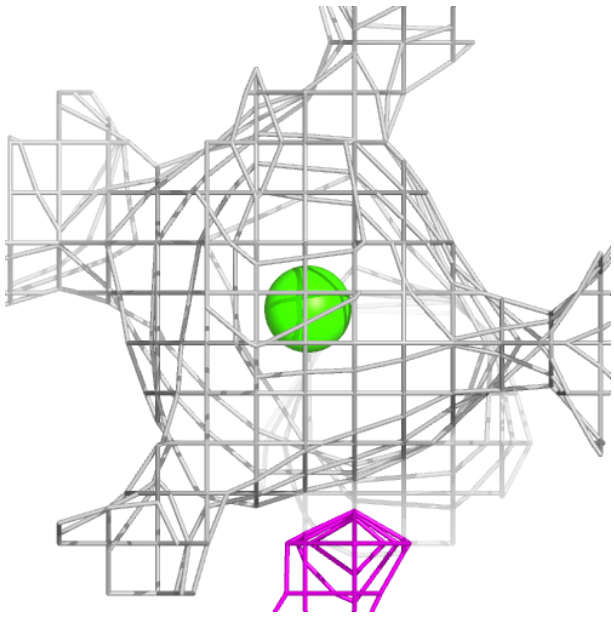
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	PO4	A	403	5/5	0.89	0.08	78,79,86,88	0
4	PO4	A	404	5/5	0.97	0.06	41,42,52,53	0
3	CA	A	402	1/1	0.98	0.02	32,32,32,32	0
2	ZN	A	401	1/1	0.99	0.01	20,20,20,20	0

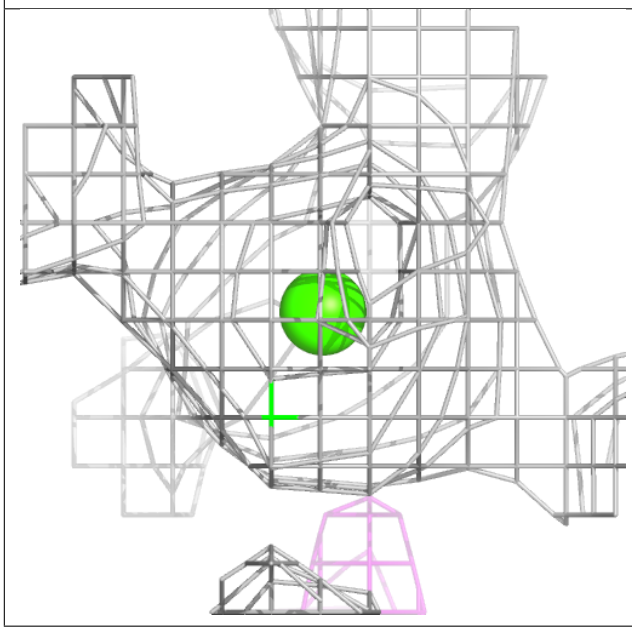
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers

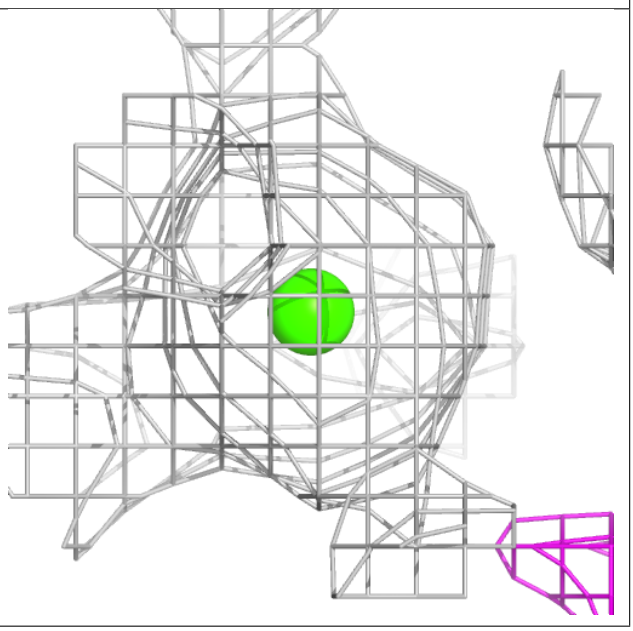
as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around CA A 402:

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

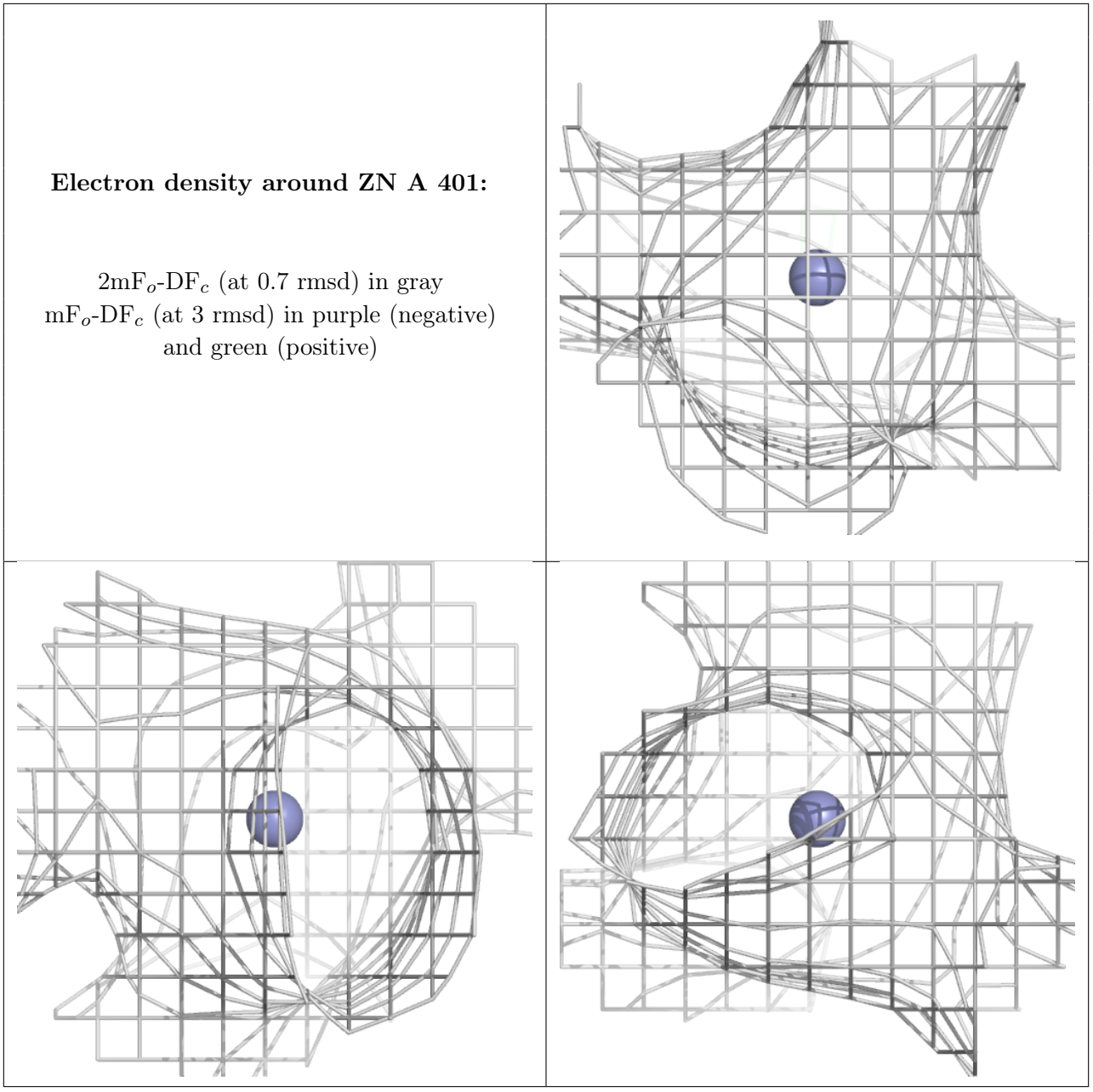






Electron density around ZN A 401:

$2mF_o$-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.