W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

May 7, 2024 – 04:42 PM JST

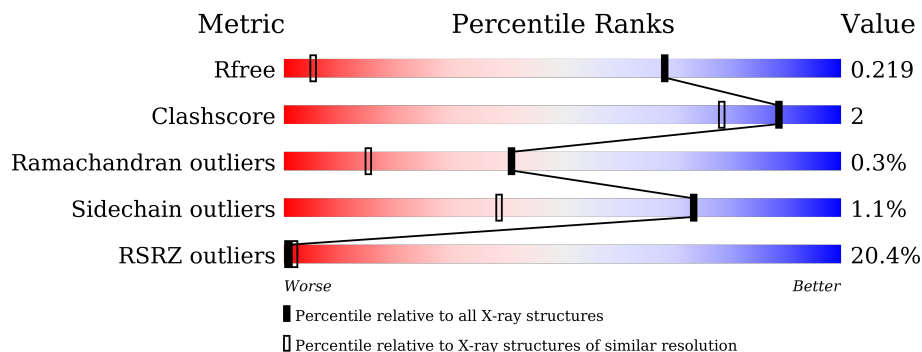|		
PDB ID	:	8K07
Title	:	Pseudouridine 5'-monophosphate glycosylase from Arabidopsis thaliana – citrate bound K185A mutant
Authors	:	Lee, J.Y.; Kim, S.H.; Rhee, S.K.
Deposited on	:	2023-07-07
Resolution	:	1.12 Å(reported)

---

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 1.12 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	Percentile Ranks	Value
Rfree		0.219
Clashscore		2
Ramachandran outliers		0.3%
Sidechain outliers		1.1%
RSRZ outliers		20.4%

Worse *Better*

▌Percentile relative to all X-ray structures

▯Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	1168 (1.14-1.10)
Clashscore	141614	1205 (1.14-1.10)
Ramachandran outliers	138981	1168 (1.14-1.10)
Sidechain outliers	138945	1165 (1.14-1.10)
RSRZ outliers	127900	1146 (1.14-1.10)

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	340	16% / 84% 5% 11%
1	B	340	15% / 85% · 11%
1	C	340	24% / 81% 5% 13%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 7474 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 Pseudouridine-5'-phosphate glycosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	301	Total 2218	C 1387	N 391	O 432	S 8	0	0	0
1	B	303	Total 2229	C 1393	N 393	O 435	S 8	0	0	0
1	C	295	Total 2176	C 1361	N 384	O 423	S 8	0	0	0

There are 33 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	185	ALA	LYS	engineered mutation	UNP Q84K35
A	331	LEU	-	expression tag	UNP Q84K35
A	332	GLU	-	expression tag	UNP Q84K35
A	333	HIS	-	expression tag	UNP Q84K35
A	334	HIS	-	expression tag	UNP Q84K35
A	335	HIS	-	expression tag	UNP Q84K35
A	336	HIS	-	expression tag	UNP Q84K35
A	337	HIS	-	expression tag	UNP Q84K35
A	338	HIS	-	expression tag	UNP Q84K35
A	339	HIS	-	expression tag	UNP Q84K35
A	340	HIS	-	expression tag	UNP Q84K35
B	186	ALA	LYS	engineered mutation	UNP Q84K35
B	332	LEU	-	expression tag	UNP Q84K35
B	333	GLU	-	expression tag	UNP Q84K35
B	334	HIS	-	expression tag	UNP Q84K35
B	335	HIS	-	expression tag	UNP Q84K35
B	336	HIS	-	expression tag	UNP Q84K35
B	337	HIS	-	expression tag	UNP Q84K35
B	338	HIS	-	expression tag	UNP Q84K35
B	339	HIS	-	expression tag	UNP Q84K35
B	340	HIS	-	expression tag	UNP Q84K35
B	341	HIS	-	expression tag	UNP Q84K35
C	185	ALA	LYS	engineered mutation	UNP Q84K35

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	331	LEU	-	expression tag	UNP Q84K35
C	332	GLU	-	expression tag	UNP Q84K35
C	333	HIS	-	expression tag	UNP Q84K35
C	334	HIS	-	expression tag	UNP Q84K35
C	335	HIS	-	expression tag	UNP Q84K35
C	336	HIS	-	expression tag	UNP Q84K35
C	337	HIS	-	expression tag	UNP Q84K35
C	338	HIS	-	expression tag	UNP Q84K35
C	339	HIS	-	expression tag	UNP Q84K35
C	340	HIS	-	expression tag	UNP Q84K35

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Mn 1	0	0
2	B	1	Total 1	Mn 1	0	0
2	C	1	Total 1	Mn 1	0	0

- Molecule 3 is CITRIC ACID (three-letter code: CIT) (formula: $C_6H_8O_7$).
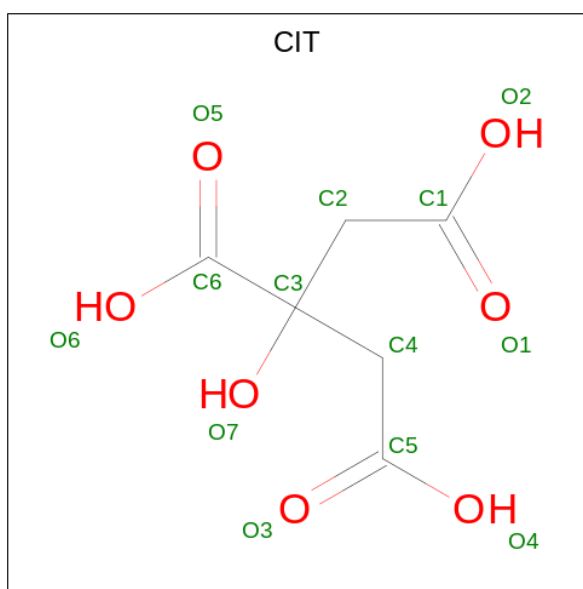
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 13	C 6	O 7	0	0
3	B	1	Total 13	C 6	O 7	0	0
3	C	1	Total 13	C 6	O 7	0	0

• Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	281	Total 281	O 281	0	0
4	B	289	Total 289	O 289	0	0
4	C	239	Total 239	O 239	0	0

3 Residue-property plots ⓘ

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.
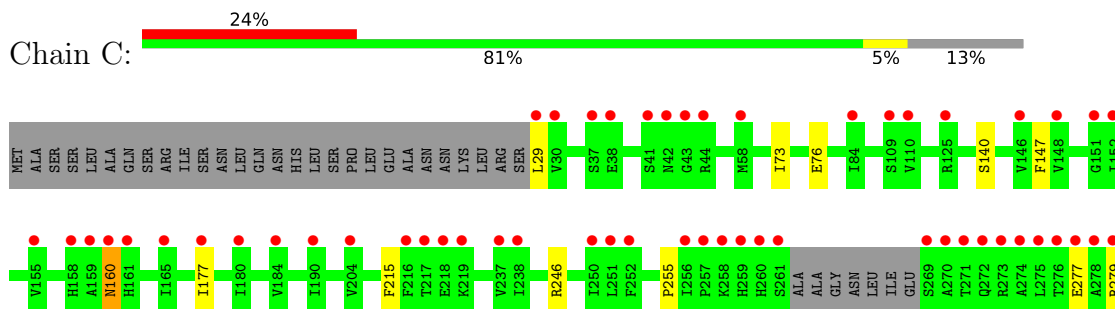
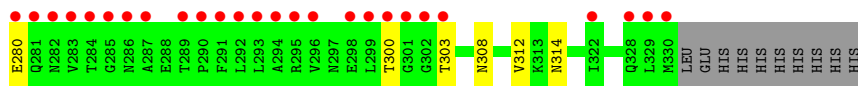• Molecule 1: Pseudouridine-5'-phosphate glycosidase
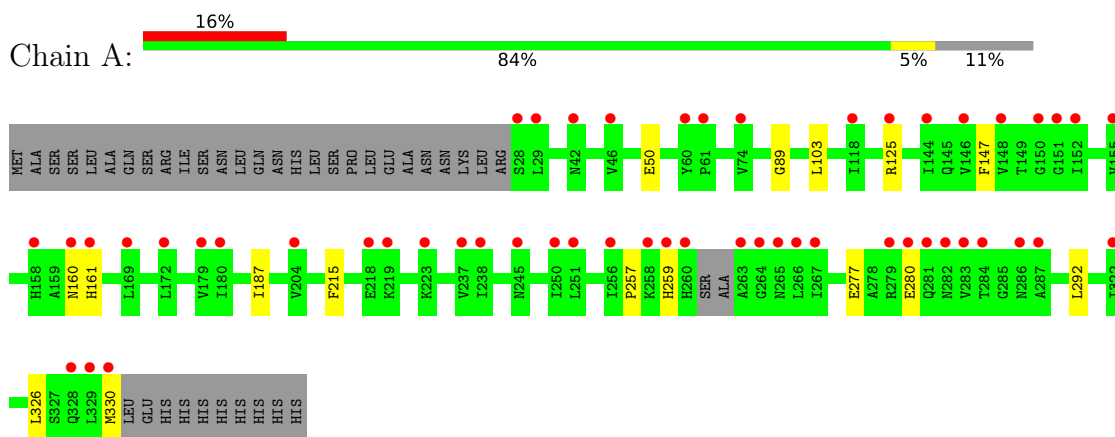
• Molecule 1: Pseudouridine-5'-phosphate glycosidase
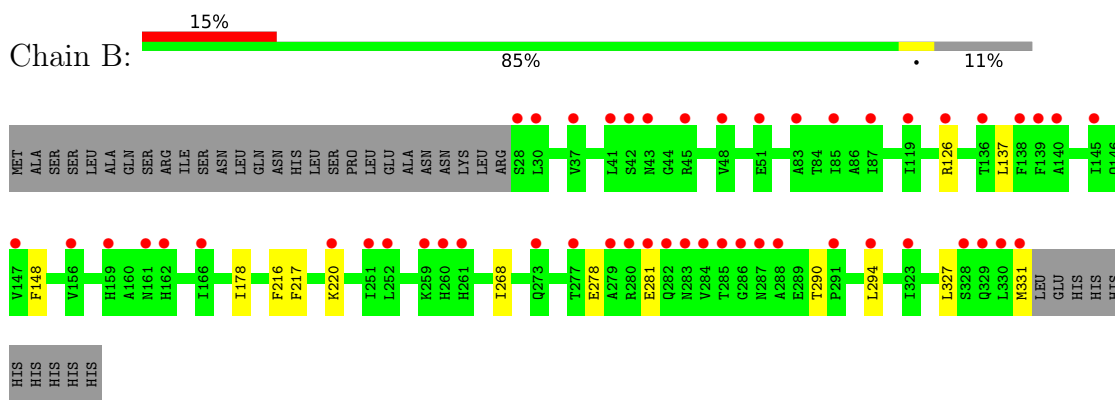
• Molecule 1: Pseudouridine-5'-phosphate glycosidase

4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	57.30Å 68.52Å 73.09Å 117.88° 94.31° 109.38°	Depositor
Resolution (Å)	24.80 – 1.12 24.80 – 1.12	Depositor EDS
% Data completeness (in resolution range)	92.3 (24.80-1.12) 92.3 (24.80-1.12)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.29 (at 1.12Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.205 , 0.219 0.206 , 0.219	Depositor DCC
R_{free} test set	1987 reflections (0.63%)	wwPDB-VP
Wilson B-factor (Å^2)	12.6	Xtriage
Anisotropy	0.421	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.47 , 56.5	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	7474	wwPDB-VP
Average B, all atoms (Å^2)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.38% 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: CIT, MN

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	0.34	0/2247	0.54	0/3052
1	B	0.34	0/2259	0.53	0/3070
1	C	0.32	0/2205	0.55	2/2995 (0.1%)
All	All	0.33	0/6711	0.54	2/9117 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	246	ARG	NE-CZ-NH1	6.99	123.79	120.30
1	C	246	ARG	NE-CZ-NH2	-5.95	117.32	120.30

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	2218	0	2264	7	0
1	B	2229	0	2275	8	0
1	C	2176	0	2222	11	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	13	0	4	0	0
3	B	13	0	4	0	0
3	C	13	0	4	0	0
4	A	281	0	0	1	0
4	B	289	0	0	2	0
4	C	239	0	0	2	0
All	All	7474	0	6773	26	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 26 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:C:160:ASN:H	1:C:160:ASN:ND2	1.59	1.00
1:C:160:ASN:HD22	1:C:160:ASN:N	1.66	0.91
1:C:314:ASN:OD1	4:C:501:HOH:O	2.01	0.78
1:C:160:ASN:H	1:C:160:ASN:HD22	0.83	0.77
1:A:103:LEU:O	4:A:501:HOH:O	2.04	0.74

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	297/340 (87%)	292 (98%)	4 (1%)	1 (0%)	41	15
1	B	301/340 (88%)	299 (99%)	1 (0%)	1 (0%)	41	15
1	C	291/340 (86%)	287 (99%)	3 (1%)	1 (0%)	41	15
All	All	889/1020 (87%)	878 (99%)	8 (1%)	3 (0%)	41	15

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	216	PHE
1	A	215	PHE
1	C	215	PHE

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	239/274 (87%)	237 (99%)	2 (1%)	81	51
1	B	240/274 (88%)	238 (99%)	2 (1%)	81	51
1	C	235/274 (86%)	231 (98%)	4 (2%)	60	22
All	All	714/822 (87%)	706 (99%)	8 (1%)	73	38

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	279	ARG
1	C	160	ASN
1	C	29	LEU
1	B	148	PHE
1	C	147	PHE

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	160	ASN
1	C	56	HIS
1	C	160	ASN

5.3.3   RNA (i)

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 6 ligands modelled in this entry, 3 are monoatomic - leaving 3 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	$\#|Z| > 2$
3	CIT	B	402	2	12,12,12	1.00	0	17,17,17	1.55	1 (5%)
3	CIT	A	402	2	12,12,12	1.06	0	17,17,17	1.48	4 (23%)
3	CIT	C	402	2	12,12,12	1.08	0	17,17,17	1.79	4 (23%)

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	CIT	B	402	2	-	4/16/16/16	-
3	CIT	A	402	2	-	2/16/16/16	-
3	CIT	C	402	2	-	4/16/16/16	-

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	402	CIT	O6-C6-C3	4.42	120.72	113.05
3	B	402	CIT	O6-C6-C3	4.29	120.49	113.05

Continued on next page...

WORLDWIDE wwPDB PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	402	CIT	O6-C6-C3	2.95	118.17	113.05
3	C	402	CIT	O4-C5-C4	2.76	123.21	114.35
3	A	402	CIT	O4-C5-C4	2.66	122.88	114.35

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	402	CIT	C4-C3-C6-O5
3	B	402	CIT	C4-C3-C6-O6
3	C	402	CIT	C4-C3-C6-O6
3	A	402	CIT	C3-C4-C5-O3
3	C	402	CIT	C3-C4-C5-O3

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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	301/340 (88%)	1.40	53 (17%)	1	2	10, 15, 30, 38	0
1	B	303/340 (89%)	1.31	50 (16%)	1	3	11, 15, 29, 39	0
1	C	295/340 (86%)	1.80	80 (27%)	0	1	10, 18, 41, 50	0
All	All	899/1020 (88%)	1.50	183 (20%)	1	2	10, 16, 35, 50	0

The worst 5 of 183 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	282	ASN	8.7
1	C	283	VAL	7.8
1	A	265	ASN	7.7
1	C	160	ASN	7.7
1	C	301	GLY	7.7

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

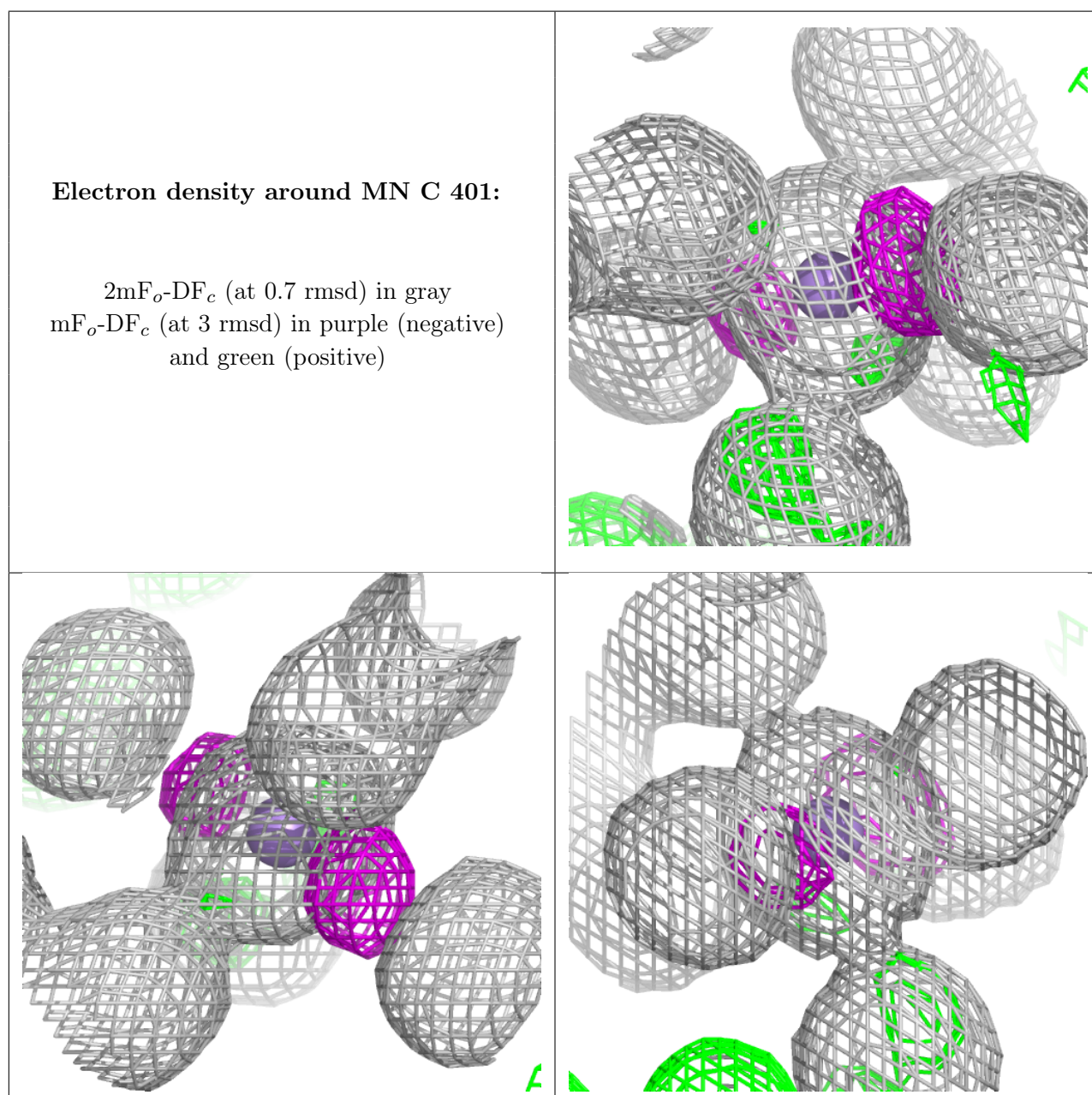There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides 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, 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.
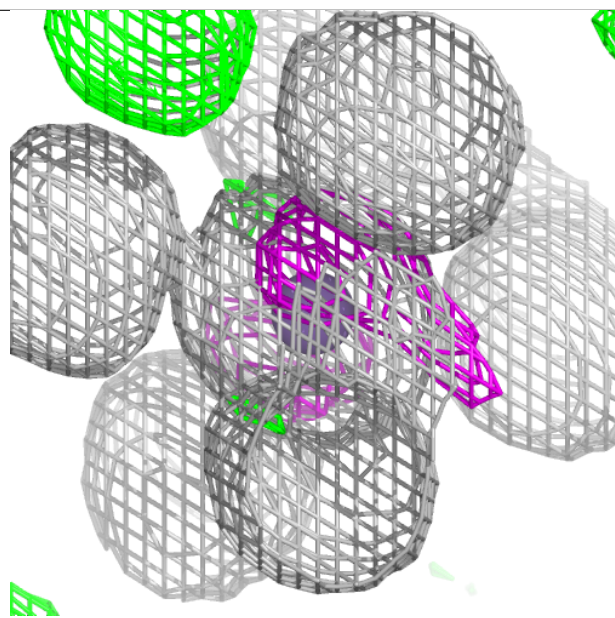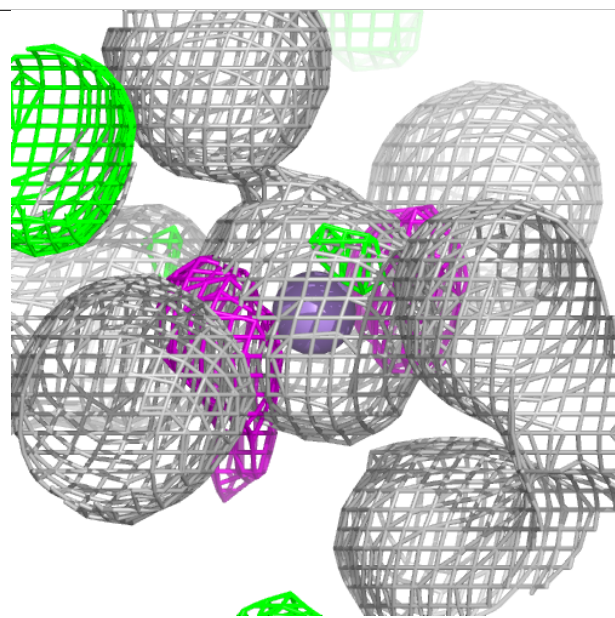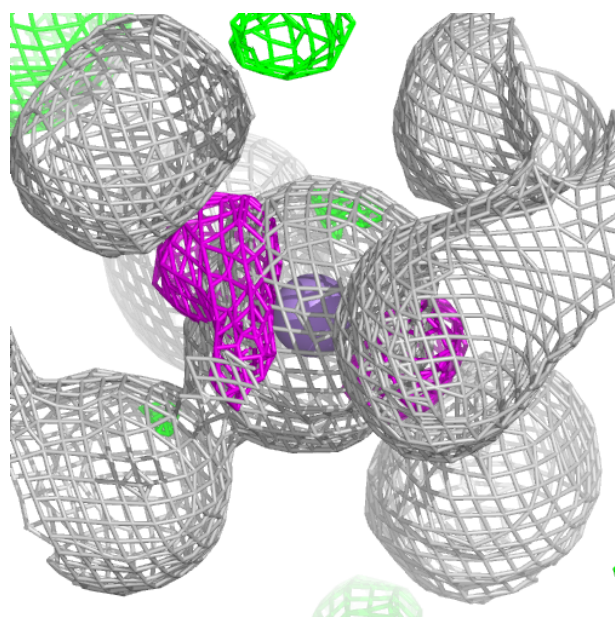
WORLDWIDE
wwPDB
PROTEIN DATA BANK

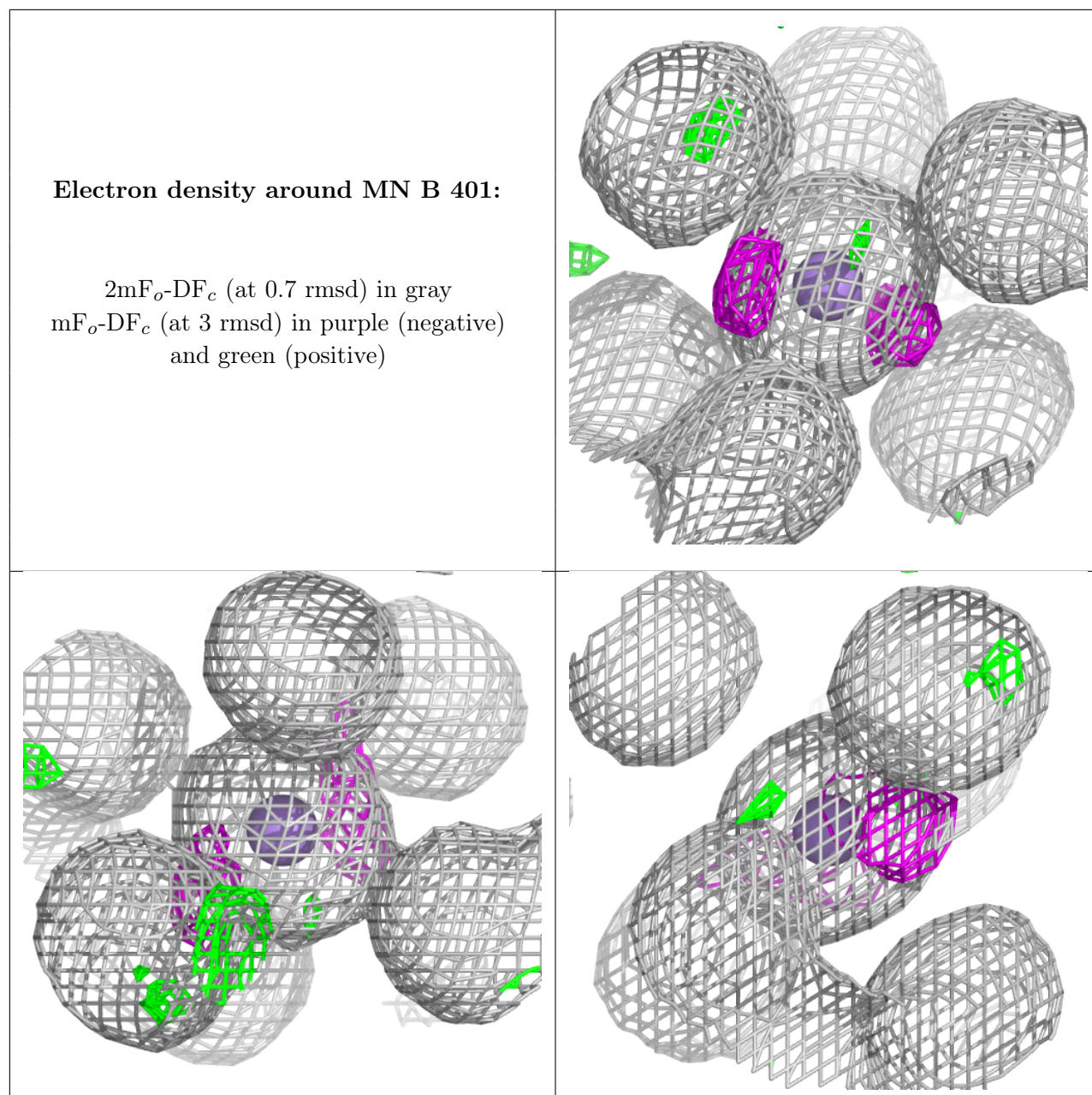| Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
3	CIT	A	402	13/13	0.95	0.10	12,13,21,22	0
3	CIT	B	402	13/13	0.95	0.12	11,13,22,24	0
3	CIT	C	402	13/13	0.96	0.10	13,15,23,26	0
2	MN	C	401	1/1	0.99	0.08	15,15,15,15	0
2	MN	A	401	1/1	1.00	0.07	14,14,14,14	0
2	MN	B	401	1/1	1.00	0.07	13,13,13,13	0

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 MN C 401:

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

Electron density around MN 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)

Electron density around MN B 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.