wwPDB X-ray Structure Validation Summary Report (i)

Jan 20, 2024 – 09:48 pm GMT

PDB ID	:	7NZO
Title	:	D-lyxose isomerasefrom the hyperthermophilic archaeon Thermofilum sp
Authors	:	De Rose, S.A.; Isupov, M.N.; Littlechild, J.A.; Schoenheit, P.
Deposited on	:	2021-03-24
Resolution	:	1.67 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:

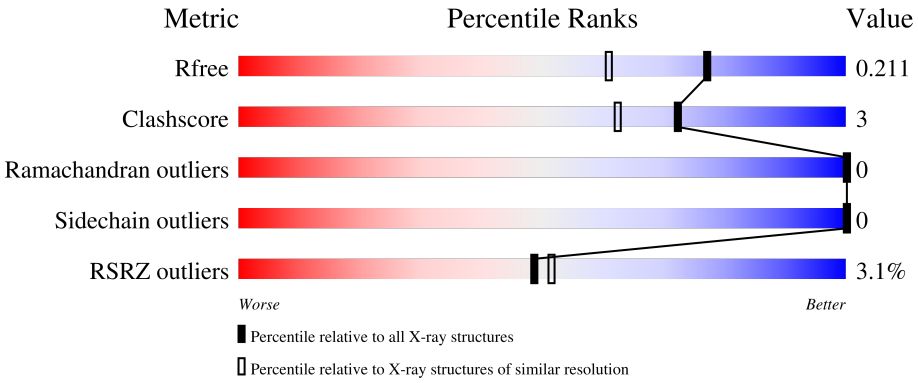
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.67 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}	130704	6780 (1.70-1.66)
Clashscore	141614	7310 (1.70-1.66)
Ramachandran outliers	138981	7173 (1.70-1.66)
Sidechain outliers	138945	7172 (1.70-1.66)
RSRZ outliers	127900	6661 (1.70-1.66)

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	AAA	204	80% 6% 14%
1	BBB	204	5% 82% 5% 13%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 3274 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 D-lyxose/D-mannose family sugar isomerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	176	Total	C	N	O	S	0	4	0
			1472	947	249	268	8			
1	BBB	178	Total	C	N	O	S	0	8	0
			1509	975	254	271	9			

There are 48 discrepancies between the modelled and reference sequences:

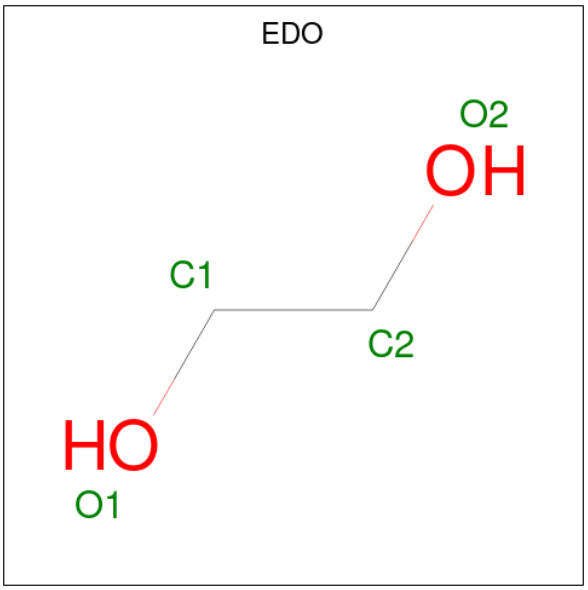
Chain	Residue	Modelled	Actual	Comment	Reference
AAA	-23	MET	-	initiating methionine	UNP A0A256XLS3
AAA	-22	GLY	-	expression tag	UNP A0A256XLS3
AAA	-21	HIS	-	expression tag	UNP A0A256XLS3
AAA	-20	HIS	-	expression tag	UNP A0A256XLS3
AAA	-19	HIS	-	expression tag	UNP A0A256XLS3
AAA	-18	HIS	-	expression tag	UNP A0A256XLS3
AAA	-17	HIS	-	expression tag	UNP A0A256XLS3
AAA	-16	HIS	-	expression tag	UNP A0A256XLS3
AAA	-15	HIS	-	expression tag	UNP A0A256XLS3
AAA	-14	HIS	-	expression tag	UNP A0A256XLS3
AAA	-13	HIS	-	expression tag	UNP A0A256XLS3
AAA	-12	HIS	-	expression tag	UNP A0A256XLS3
AAA	-11	SER	-	expression tag	UNP A0A256XLS3
AAA	-10	SER	-	expression tag	UNP A0A256XLS3
AAA	-9	GLY	-	expression tag	UNP A0A256XLS3
AAA	-8	HIS	-	expression tag	UNP A0A256XLS3
AAA	-7	ILE	-	expression tag	UNP A0A256XLS3
AAA	-6	ASP	-	expression tag	UNP A0A256XLS3
AAA	-5	ASP	-	expression tag	UNP A0A256XLS3
AAA	-4	ASP	-	expression tag	UNP A0A256XLS3
AAA	-3	ASP	-	expression tag	UNP A0A256XLS3
AAA	-2	LYS	-	expression tag	UNP A0A256XLS3
AAA	-1	HIS	-	expression tag	UNP A0A256XLS3
AAA	0	MET	-	expression tag	UNP A0A256XLS3
BBB	-23	MET	-	initiating methionine	UNP A0A256XLS3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
BBB	-22	GLY	-	expression tag	UNP A0A256XLS3
BBB	-21	HIS	-	expression tag	UNP A0A256XLS3
BBB	-20	HIS	-	expression tag	UNP A0A256XLS3
BBB	-19	HIS	-	expression tag	UNP A0A256XLS3
BBB	-18	HIS	-	expression tag	UNP A0A256XLS3
BBB	-17	HIS	-	expression tag	UNP A0A256XLS3
BBB	-16	HIS	-	expression tag	UNP A0A256XLS3
BBB	-15	HIS	-	expression tag	UNP A0A256XLS3
BBB	-14	HIS	-	expression tag	UNP A0A256XLS3
BBB	-13	HIS	-	expression tag	UNP A0A256XLS3
BBB	-12	HIS	-	expression tag	UNP A0A256XLS3
BBB	-11	SER	-	expression tag	UNP A0A256XLS3
BBB	-10	SER	-	expression tag	UNP A0A256XLS3
BBB	-9	GLY	-	expression tag	UNP A0A256XLS3
BBB	-8	HIS	-	expression tag	UNP A0A256XLS3
BBB	-7	ILE	-	expression tag	UNP A0A256XLS3
BBB	-6	ASP	-	expression tag	UNP A0A256XLS3
BBB	-5	ASP	-	expression tag	UNP A0A256XLS3
BBB	-4	ASP	-	expression tag	UNP A0A256XLS3
BBB	-3	ASP	-	expression tag	UNP A0A256XLS3
BBB	-2	LYS	-	expression tag	UNP A0A256XLS3
BBB	-1	HIS	-	expression tag	UNP A0A256XLS3
BBB	0	MET	-	expression tag	UNP A0A256XLS3

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	AAA	1	Total C O 4 2 2	0	0
2	AAA	1	Total C O 4 2 2	0	0
2	AAA	1	Total C O 4 2 2	0	0
2	AAA	1	Total C O 4 2 2	0	0
2	AAA	1	Total C O 4 2 2	0	0
2	AAA	1	Total C O 4 2 2	0	0
2	BBB	1	Total C O 4 2 2	0	0
2	BBB	1	Total C O 4 2 2	0	0
2	BBB	1	Total C O 4 2 2	0	0
2	BBB	1	Total C O 4 2 2	0	0
2	BBB	1	Total C O 4 2 2	0	0

- Molecule 3 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	AAA	1	Total Mn 1 1	0	0
3	BBB	1	Total Mn 1 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	AAA	131	Total O 131 131	0	0
4	BBB	116	Total O 116 116	0	0

3 Residue-property plots (i)

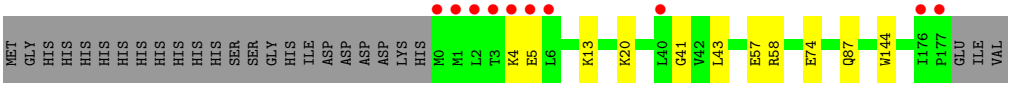
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: D-lyxose/D-mannose family sugar isomerase

Chain AAA:

- Molecule 1: D-lyxose/D-mannose family sugar isomerase

Chain BBB:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	47.77Å 86.91Å 54.03Å 90.00° 112.93° 90.00°	Depositor
Resolution (Å)	39.25 – 1.67 39.25 – 1.67	Depositor EDS
% Data completeness (in resolution range)	92.1 (39.25-1.67) 92.1 (39.25-1.67)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.16 (at 1.67Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.179 , 0.211 0.179 , 0.211	Depositor DCC
R_{free} test set	2056 reflections (4.73%)	wwPDB-VP
Wilson B-factor (Å^2)	32.8	Xtriage
Anisotropy	0.434	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 47.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.98	EDS
Total number of atoms	3274	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.62% 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: MN, EDO

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	AAA	0.51	0/1522	0.80	0/2059
1	BBB	0.46	0/1572	0.80	0/2128
All	All	0.49	0/3094	0.80	0/4187

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	AAA	1472	0	1475	12	0
1	BBB	1509	0	1533	8	0
2	AAA	24	0	36	5	0
2	BBB	20	0	30	4	0
3	AAA	1	0	0	0	0
3	BBB	1	0	0	0	0
4	AAA	131	0	0	4	0
4	BBB	116	0	0	2	0
All	All	3274	0	3074	21	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 3.

The worst 5 of 21 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:BBB:20:LYS:HE3	4:BBB:386:HOH:O	2.05	0.56
1:BBB:4[B]:LYS:HG2	1:BBB:5:GLU:OE2	2.07	0.54
1:AAA:13:LYS:HE2	4:AAA:417:HOH:O	2.10	0.52
1:AAA:94:TRP:NE1	2:AAA:204:EDO:H11	2.26	0.49
1:BBB:57[A]:GLU:HG2	1:BBB:58:ARG:HG3	1.93	0.49

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	AAA	178/204 (87%)	175 (98%)	3 (2%)	0	100	100
1	BBB	184/204 (90%)	181 (98%)	3 (2%)	0	100	100
All	All	362/408 (89%)	356 (98%)	6 (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	AAA	162/184 (88%)	162 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	BBB	168/184 (91%)	168 (100%)	0	100	100
All	All	330/368 (90%)	330 (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 (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 monosaccharides 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	$\#\|Z\| > 2$
2	EDO	BBB	204	-	3,3,3	0.16	0	2,2,2	0.29	0
2	EDO	BBB	201	-	3,3,3	0.15	0	2,2,2	0.29	0
2	EDO	BBB	202	-	3,3,3	0.24	0	2,2,2	0.08	0
2	EDO	AAA	201	-	3,3,3	0.20	0	2,2,2	0.39	0
2	EDO	AAA	206	-	3,3,3	0.19	0	2,2,2	0.25	0
2	EDO	AAA	203	-	3,3,3	0.06	0	2,2,2	0.48	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	EDO	BBB	203	-	3,3,3	0.12	0	2,2,2	0.40	0
2	EDO	AAA	204	-	3,3,3	0.12	0	2,2,2	0.16	0
2	EDO	AAA	202	-	3,3,3	0.29	0	2,2,2	0.15	0
2	EDO	AAA	205	-	3,3,3	0.10	0	2,2,2	0.32	0
2	EDO	BBB	205	-	3,3,3	0.13	0	2,2,2	0.48	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	EDO	BBB	204	-	-	1/1/1/1	-
2	EDO	BBB	201	-	-	0/1/1/1	-
2	EDO	BBB	202	-	-	0/1/1/1	-
2	EDO	AAA	201	-	-	1/1/1/1	-
2	EDO	AAA	206	-	-	1/1/1/1	-
2	EDO	AAA	203	-	-	1/1/1/1	-
2	EDO	BBB	203	-	-	0/1/1/1	-
2	EDO	AAA	204	-	-	0/1/1/1	-
2	EDO	AAA	202	-	-	0/1/1/1	-
2	EDO	AAA	205	-	-	1/1/1/1	-
2	EDO	BBB	205	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	AAA	201	EDO	O1-C1-C2-O2
2	AAA	206	EDO	O1-C1-C2-O2
2	BBB	204	EDO	O1-C1-C2-O2
2	BBB	205	EDO	O1-C1-C2-O2
2	AAA	203	EDO	O1-C1-C2-O2

There are no ring outliers.

5 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	BBB	201	EDO	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	AAA	201	EDO	1	0
2	BBB	203	EDO	2	0
2	AAA	204	EDO	3	0
2	AAA	202	EDO	1	0

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, 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	AAA	176/204 (86%)	-0.26	1 (0%) 89 91	26, 36, 55, 106	0
1	BBB	178/204 (87%)	0.21	10 (5%) 24 25	26, 37, 65, 115	0
All	All	354/408 (86%)	-0.02	11 (3%) 49 51	26, 37, 57, 115	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	0	MET	8.0
1	BBB	177	PRO	7.1
1	BBB	1[A]	MET	5.7
1	AAA	0	MET	4.2
1	BBB	2	LEU	3.7

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

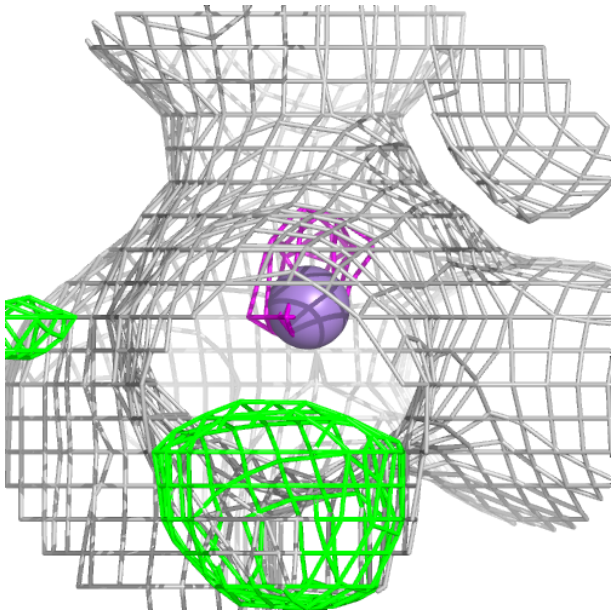
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

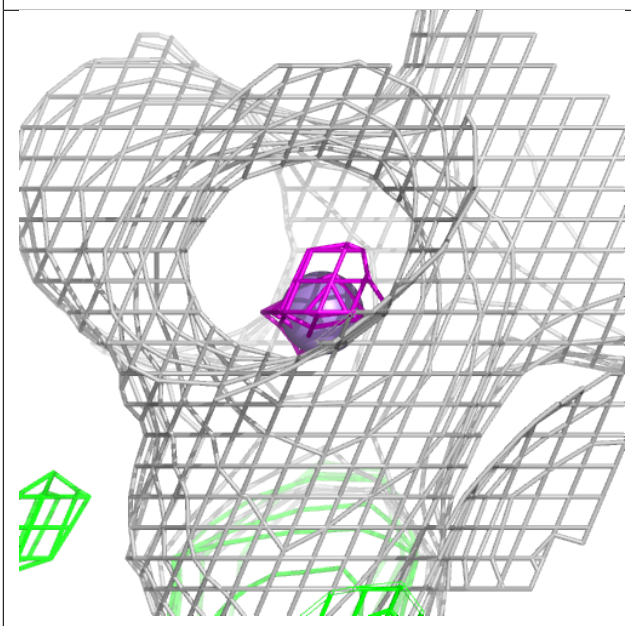
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	EDO	AAA	206	4/4	0.46	0.25	80,82,84,89	0
2	EDO	BBB	203	4/4	0.73	0.28	61,63,67,68	0
2	EDO	AAA	205	4/4	0.87	0.11	68,73,75,78	0
2	EDO	BBB	202	4/4	0.88	0.18	42,43,50,53	0
2	EDO	AAA	203	4/4	0.89	0.17	57,57,69,80	0
2	EDO	BBB	201	4/4	0.91	0.13	52,61,62,70	0
2	EDO	AAA	204	4/4	0.92	0.21	42,53,62,65	0
2	EDO	AAA	201	4/4	0.92	0.28	65,69,71,73	0
2	EDO	AAA	202	4/4	0.93	0.14	55,61,63,76	0
2	EDO	BBB	205	4/4	0.93	0.28	45,64,75,90	0
2	EDO	BBB	204	4/4	0.94	0.15	57,62,62,63	0
3	MN	BBB	206	1/1	0.99	0.05	36,36,36,36	0
3	MN	AAA	207	1/1	1.00	0.04	34,34,34,34	0

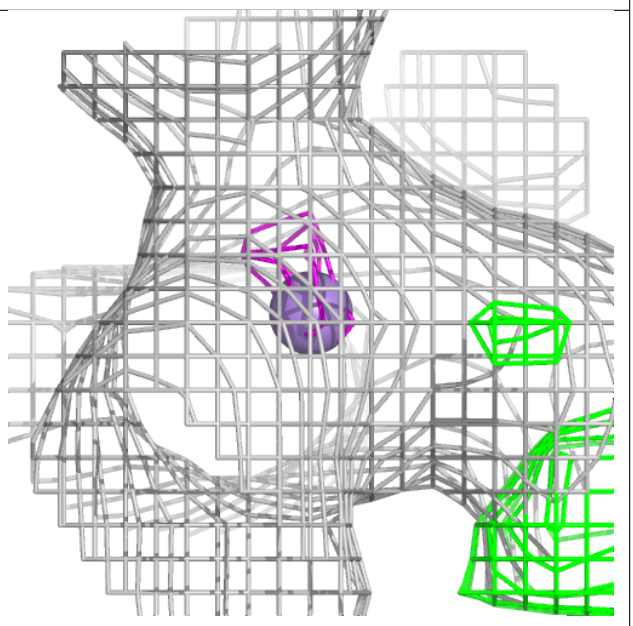
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around MN BBB 206:

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

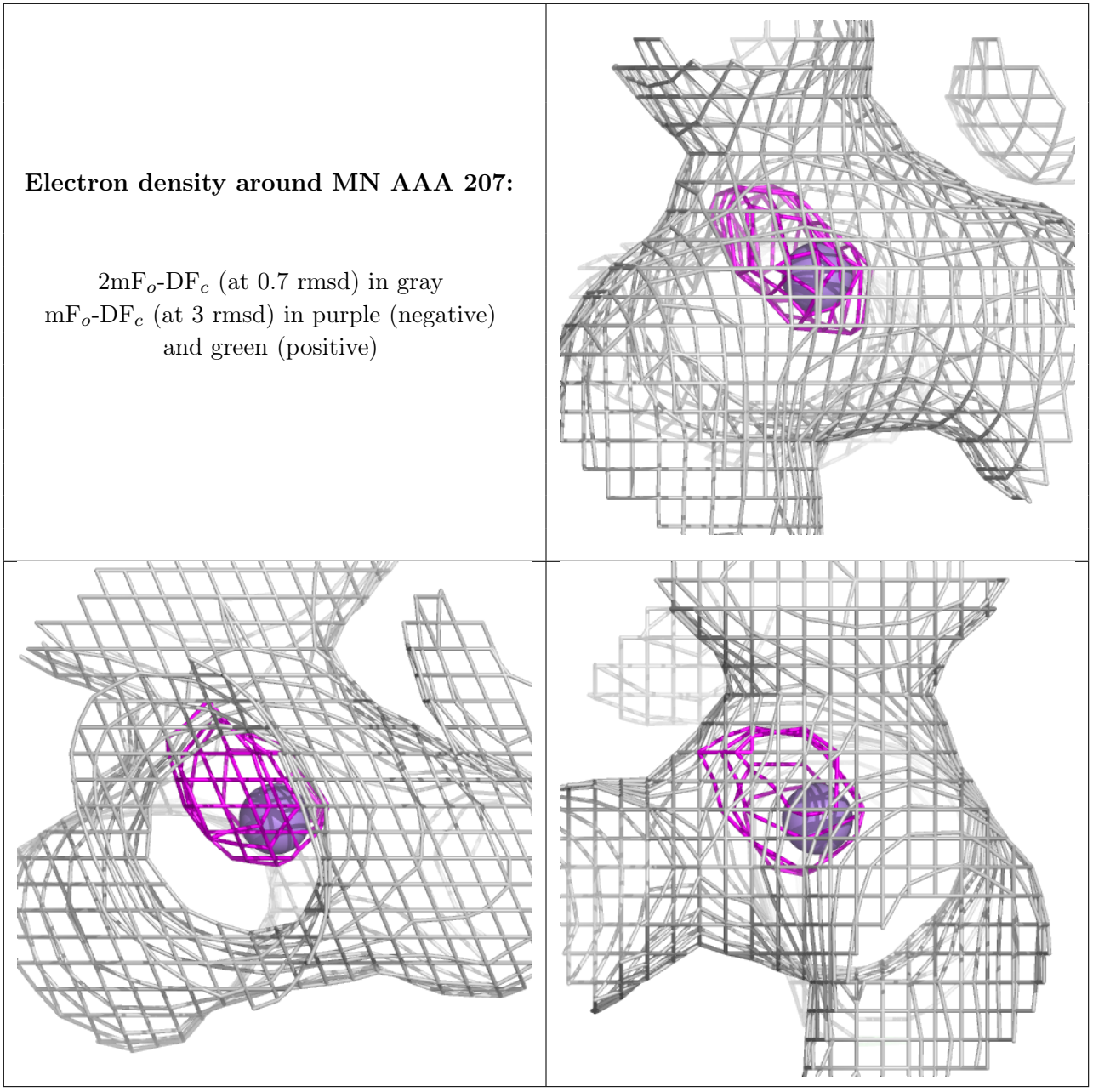






Electron density around MN AAA 207:

$2mF_o$-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.