



wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 18, 2022 – 04:05 pm BST

PDB ID : 7BJQ
Title : Crystal structure of RecJCdc45 from Methanothermobacter thermoautotrophicus in complex with ssDNA
Authors : De March, M.; Onesti, S.
Deposited on : 2021-01-14
Resolution : 2.70 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 : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : 2.29
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0267
CCP4 : 7.1.010 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.29

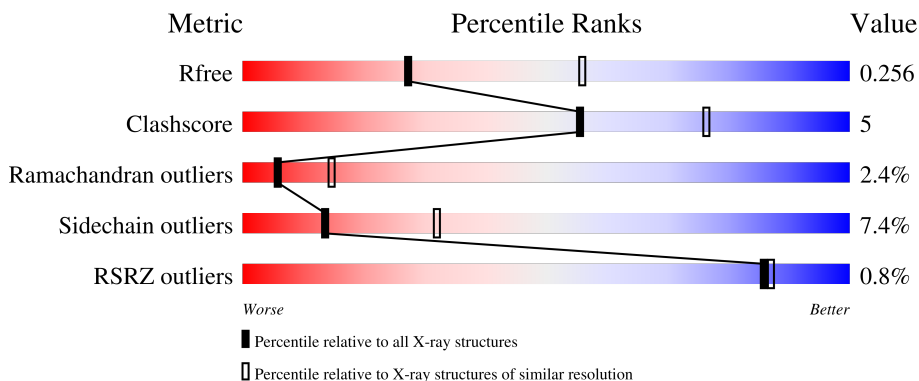
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	464	 81% 13% . .
1	B	464	 79% 14% . . .
2	D	6	 83% 17%
2	E	6	 50% 50%
2	F	6	 50% 50%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	6	 <p>A horizontal bar chart representing the quality of chain. The bar is divided into three segments: a red segment on the left labeled '17%', a large green segment in the middle labeled '83%', and a yellow segment on the right labeled '17%'.</p>

2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 7031 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 Conserved protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	445	Total 3266	C 2055	N 570	O 628	S 13	0	0	0
1	B	445	Total 3230	C 2036	N 566	O 616	S 12	0	0	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	458	ALA	-	expression tag	UNP O27473
A	459	GLU	-	expression tag	UNP O27473
A	460	ASN	-	expression tag	UNP O27473
A	461	LEU	-	expression tag	UNP O27473
A	462	TYR	-	expression tag	UNP O27473
A	463	PHE	-	expression tag	UNP O27473
A	464	GLN	-	expression tag	UNP O27473
B	458	ALA	-	expression tag	UNP O27473
B	459	GLU	-	expression tag	UNP O27473
B	460	ASN	-	expression tag	UNP O27473
B	461	LEU	-	expression tag	UNP O27473
B	462	TYR	-	expression tag	UNP O27473
B	463	PHE	-	expression tag	UNP O27473
B	464	GLN	-	expression tag	UNP O27473

- Molecule 2 is a DNA chain called DNA (5'-D(*5P**P*-P*CP*AP*TP*GP*GP*CP*-P*3P*')-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	D	6	Total 123	C 58	N 23	O 36	P 6	0	0	0
2	E	6	Total 123	C 58	N 23	O 36	P 6	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	6	Total	C	N	O	P	0	0	0
			123	58	23	36	6			
2	G	6	Total	C	N	O	P	0	0	0
			123	58	23	36	6			

- 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	A	2	Total	Mn	0	0
			2	2		
3	B	2	Total	Mn	0	0
			2	2		

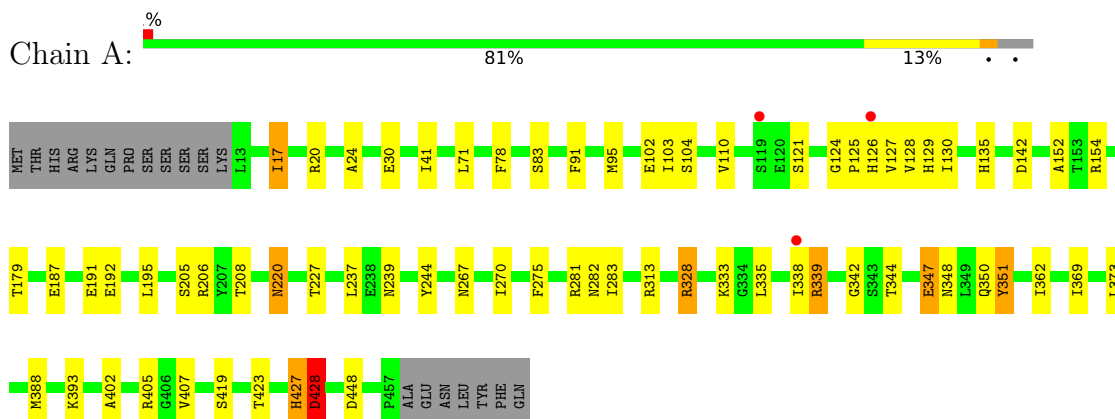
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	24	Total	O	0	0
			24	24		
4	B	14	Total	O	0	0
			14	14		
4	D	1	Total	O	0	0
			1	1		

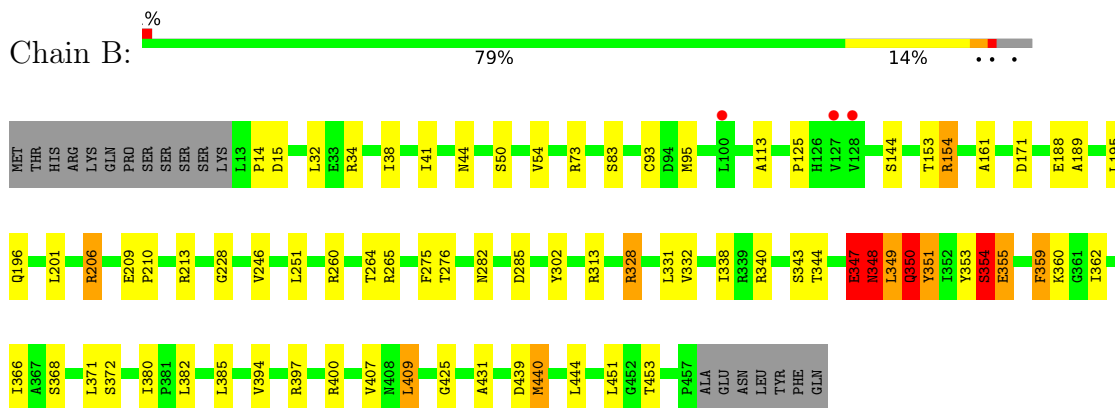
3 Residue-property plots [i](#)

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.

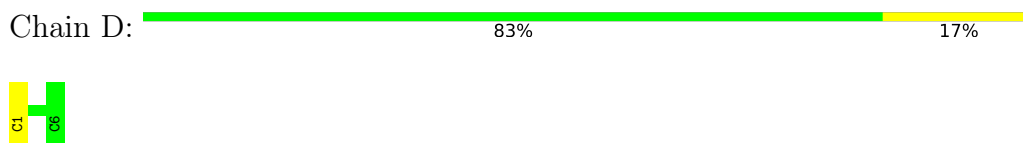
- Molecule 1: Conserved protein



- Molecule 1: Conserved protein



- Molecule 2: DNA (5'-D(*5P*P*-P*CP*AP*TP*GP*GP*CP*-P*3P*')-3')

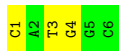


- Molecule 2: DNA (5'-D(*5P*P*-P*CP*AP*TP*GP*GP*CP*-P*3P*')-3')

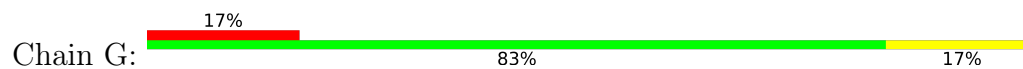




- Molecule 2: DNA (5'-D(*5P**P*-P*CP*AP*TP*GP*GP*CP*-P*3P*)-3')



- Molecule 2: DNA (5'-D(*5P**P*-P*CP*AP*TP*GP*GP*CP*-P*3P*)-3')



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	63.08Å 69.50Å 72.21Å 86.12° 69.26° 69.53°	Depositor
Resolution (Å)	67.38 – 2.70 67.39 – 2.70	Depositor EDS
% Data completeness (in resolution range)	97.7 (67.38-2.70) 97.7 (67.39-2.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.17 (at 2.69Å)	Xtrriage
Refinement program	REFMAC 5.8.0107	Depositor
R, R_{free}	0.196 , 0.258 0.199 , 0.256	Depositor DCC
R_{free} test set	1462 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	57.7	Xtrriage
Anisotropy	0.478	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.013 for -h,-l,-k	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	7031	wwPDB-VP
Average B, all atoms (Å ²)	66.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.26% 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: 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	0.59	0/3323	0.83	7/4510 (0.2%)
1	B	0.61	0/3287	0.88	7/4466 (0.2%)
2	D	0.40	0/137	0.82	0/209
2	E	0.53	0/137	1.00	0/209
2	F	0.50	0/137	0.91	1/209 (0.5%)
2	G	0.34	0/137	0.68	0/209
All	All	0.59	0/7158	0.85	15/9812 (0.2%)

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	1
1	B	0	2
All	All	0	3

There are no bond length outliers.

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	328	ARG	NE-CZ-NH1	6.23	123.42	120.30
1	B	350	GLN	N-CA-C	-5.96	94.92	111.00
1	A	339	ARG	NE-CZ-NH1	5.81	123.21	120.30
1	B	206	ARG	NE-CZ-NH2	-5.77	117.42	120.30
1	A	427	HIS	C-N-CA	5.77	136.12	121.70

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	333	LYS	Peptide
1	B	347	GLU	Peptide
1	B	354	SER	Peptide

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	3266	0	3124	27	0
1	B	3230	0	3057	39	0
2	D	123	0	68	1	0
2	E	123	0	68	2	0
2	F	123	0	68	1	0
2	G	123	0	68	1	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
4	A	24	0	0	0	0
4	B	14	0	0	0	0
4	D	1	0	0	0	0
All	All	7031	0	6453	69	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 5.

The worst 5 of 69 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:B:439:ASP:HA	1:B:440:MET:CB	2.08	0.83
1:A:427:HIS:HB3	1:A:428:ASP:HB3	1.66	0.78
1:B:397:ARG:NH2	2:E:2:DA:OP2	2.20	0.75
1:B:338:ILE:HD11	1:B:366:ILE:CG2	2.16	0.74
1:B:344:THR:CB	1:B:350:GLN:OE1	2.37	0.73

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	443/464 (96%)	407 (92%)	29 (6%)	7 (2%)	9	24
1	B	443/464 (96%)	405 (91%)	24 (5%)	14 (3%)	4	9
All	All	886/928 (96%)	812 (92%)	53 (6%)	21 (2%)	6	15

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	339	ARG
1	A	351	TYR
1	A	405	ARG
1	A	428	ASP
1	B	14	PRO

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	323/379 (85%)	302 (94%)	21 (6%)	17	38
1	B	311/379 (82%)	285 (92%)	26 (8%)	11	25
All	All	634/758 (84%)	587 (93%)	47 (7%)	13	32

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	196	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	313	ARG
1	B	209	GLU
1	B	265	ARG
1	B	348	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	282	ASN
1	B	182	ASN
1	A	445	GLN
1	A	390	GLN
1	B	36	ASN

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 4 ligands modelled in this entry, 4 are 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 [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	445/464 (95%)	-0.26	3 (0%) 87 89	36, 62, 95, 134	0
1	B	445/464 (95%)	-0.23	3 (0%) 87 89	36, 60, 90, 120	0
2	D	6/6 (100%)	-0.93	0 100 100	69, 80, 92, 101	0
2	E	6/6 (100%)	-0.29	0 100 100	65, 89, 104, 114	0
2	F	6/6 (100%)	-0.80	0 100 100	70, 78, 87, 117	0
2	G	6/6 (100%)	0.24	1 (16%) 1 1	107, 123, 157, 158	0
All	All	914/952 (96%)	-0.25	7 (0%) 86 87	36, 62, 94, 158	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	127	VAL	3.4
2	G	1	DC	3.1
1	A	119	SER	2.7
1	A	126	HIS	2.3
1	A	338	ILE	2.3

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

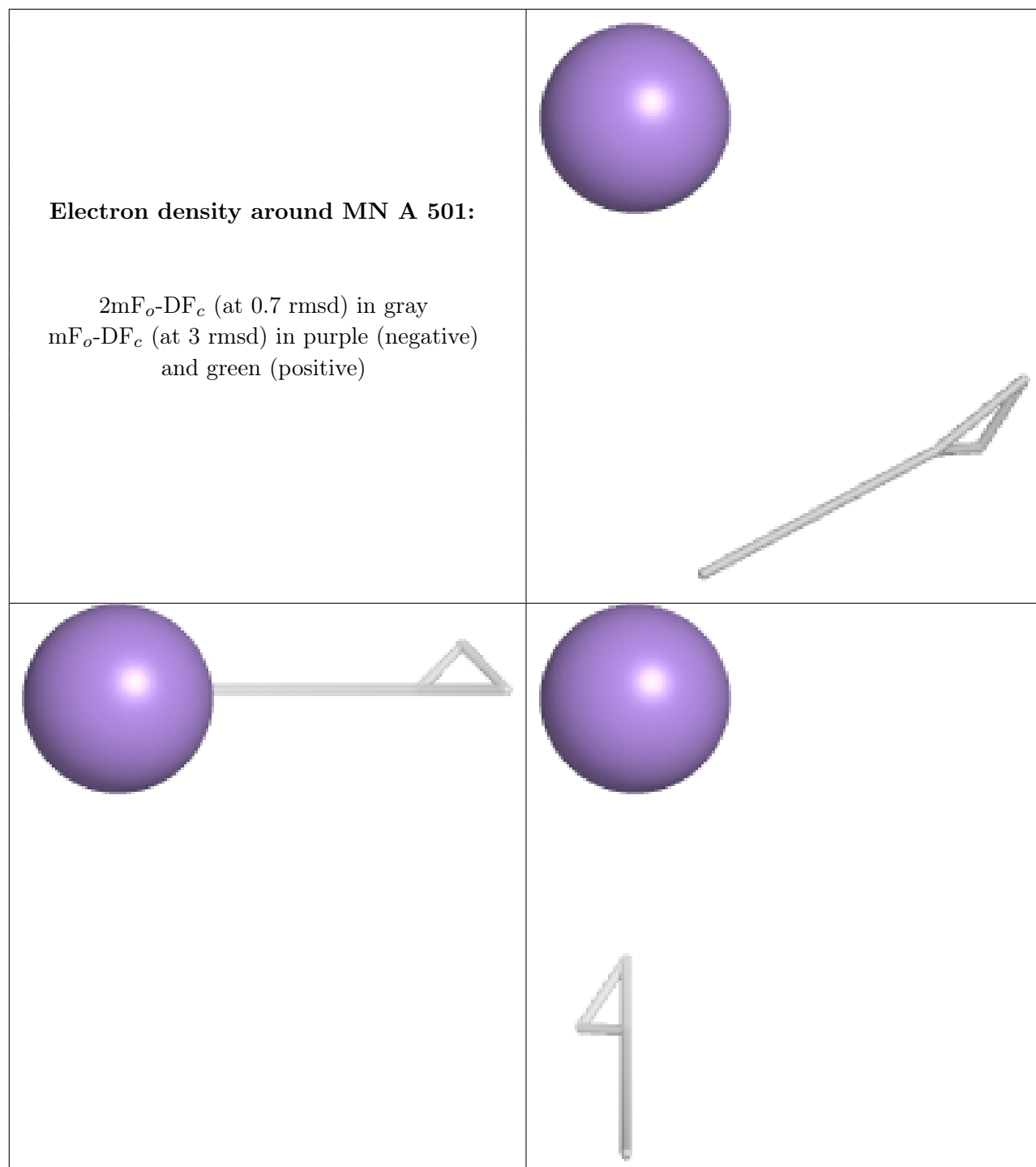
There are no monosaccharides in this entry.

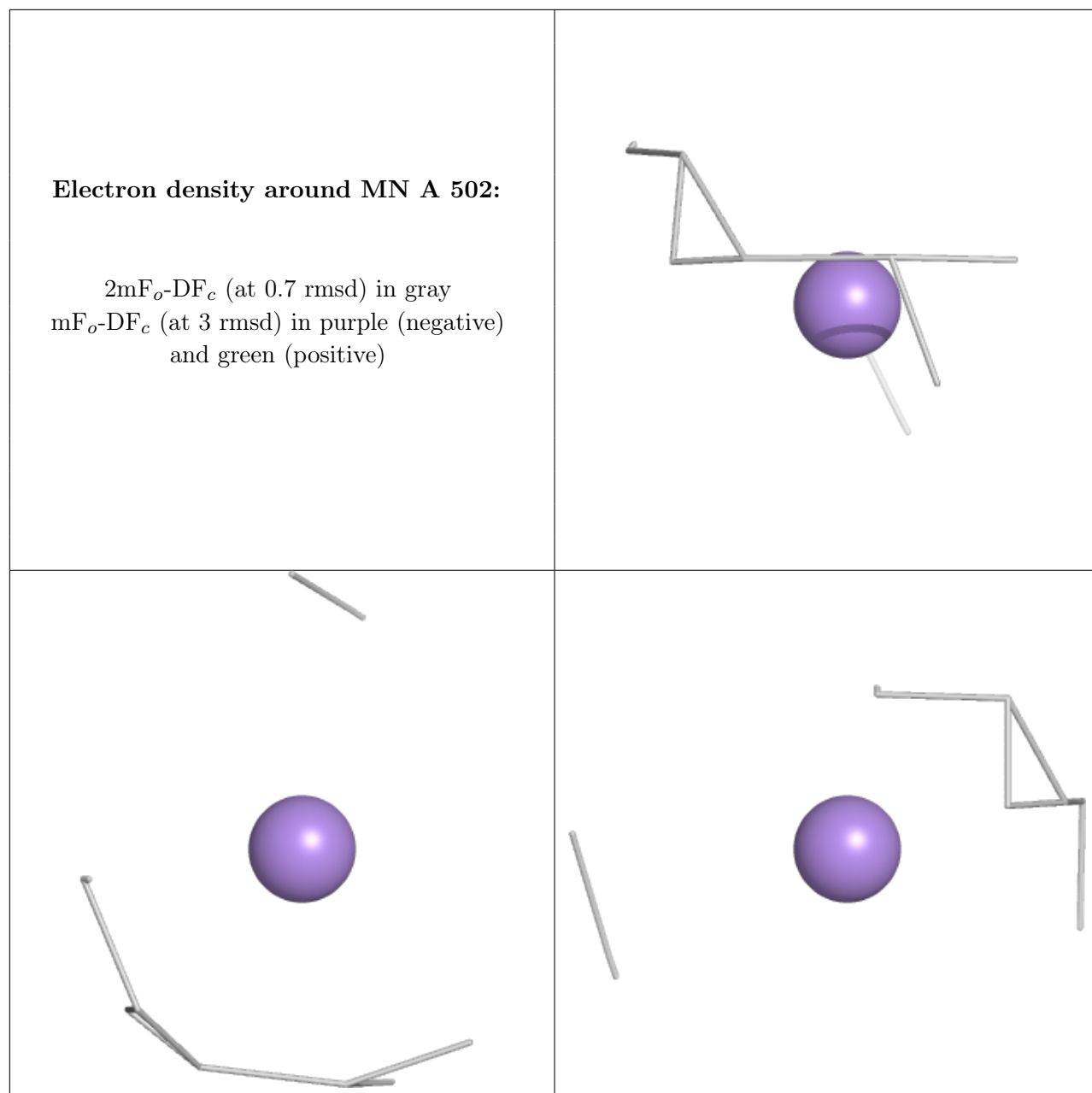
6.4 Ligands

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.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	MN	A	501	1/1	0.94	0.08	101,101,101,101	0
3	MN	A	502	1/1	0.98	0.08	98,98,98,98	0
3	MN	B	501	1/1	0.98	0.08	84,84,84,84	0
3	MN	B	502	1/1	0.99	0.07	89,89,89,89	0

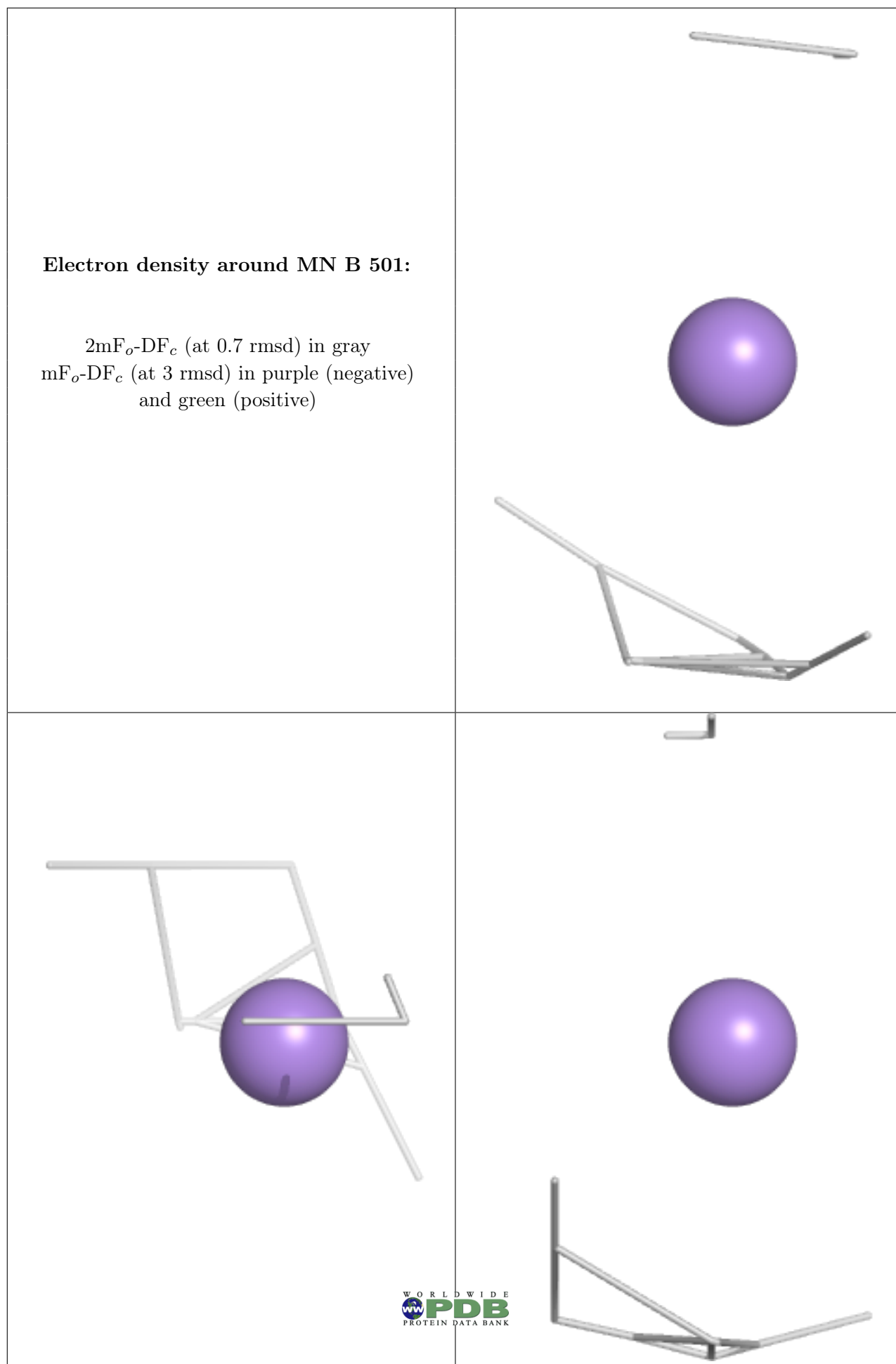
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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





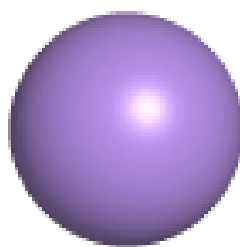
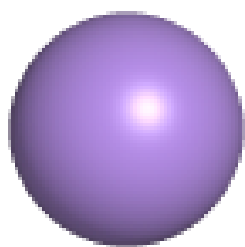
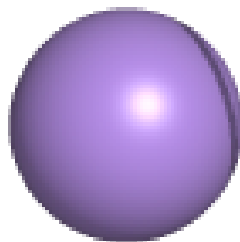
Electron density around MN B 501:

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



Electron density around MN B 502:

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



6.5 Other polymers [i](#)

There are no such residues in this entry.