



Full wwPDB X-ray Structure Validation Report ⓘ

Nov 22, 2021 – 01:06 pm GMT

PDB ID : 6SNS
Title : DNA mismatch repair proteins MLH1 and MLH3
Authors : Dai, J.; Chervy, P.; Legrand, P.; Ropars, V.; Charbonnier, J.B.
Deposited on : 2019-08-27
Resolution : 2.60 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 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 : 2.23.2
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) : 2.23.2

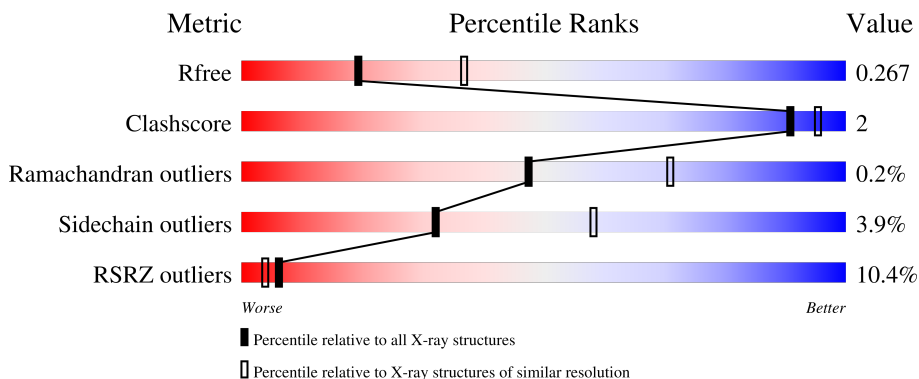
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	265	 15% 89% 8% ..
2	B	239	 4% 82% 8% 11%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 3885 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 DNA mismatch repair protein MLH1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	260	2125	1370	345	403	7	0	0	0

- Molecule 2 is a protein called DNA mismatch repair protein MLH3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	213	1724	1109	289	312	14	0	0	0

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
3	B	1	1	1	0	0

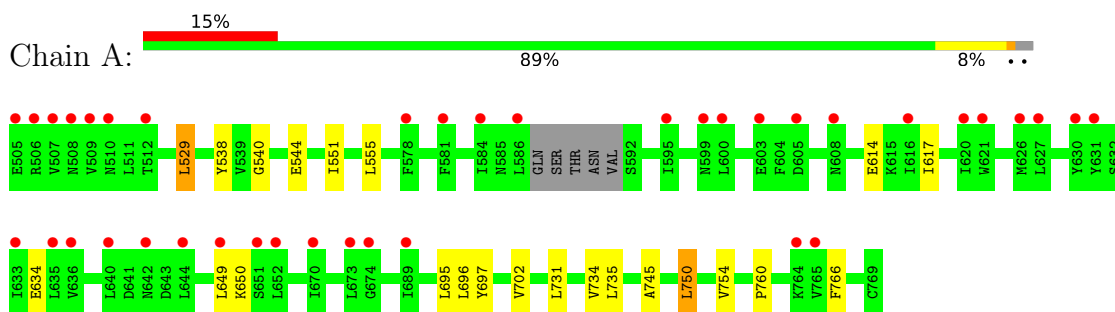
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	10	10	10	0	0
4	B	25	25	25	0	0

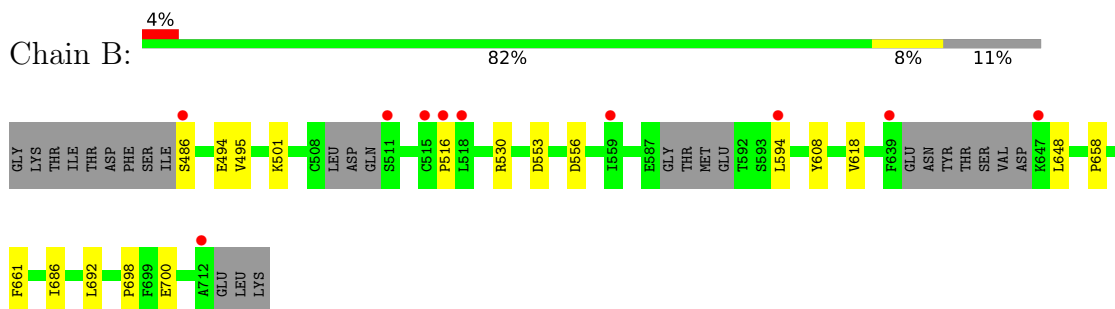
3 Residue-property plots [i](#)

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: DNA mismatch repair protein MLH1



- Molecule 2: DNA mismatch repair protein MLH3



4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	91.86Å 103.09Å 135.14Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.13 – 2.60 48.13 – 2.60	Depositor EDS
% Data completeness (in resolution range)	66.4 (48.13-2.60) 66.4 (48.13-2.60)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.51 (at 2.61Å)	Xtriage
Refinement program	BUSTER 2.10.3	Depositor
R, R_{free}	0.215 , 0.246 0.228 , 0.267	Depositor DCC
R_{free} test set	683 reflections (5.13%)	wwPDB-VP
Wilson B-factor (Å ²)	84.2	Xtriage
Anisotropy	0.050	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	3885	wwPDB-VP
Average B, all atoms (Å ²)	114.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.99% 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:
ZN

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	0.41	0/2163	0.61	0/2921
2	B	0.43	0/1757	0.63	1/2365 (0.0%)
All	All	0.42	0/3920	0.62	1/5286 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	516	PRO	N-CA-CB	5.75	110.20	103.30

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	2125	0	2142	9	0
2	B	1724	0	1720	4	0
3	B	1	0	0	0	0
4	A	10	0	0	0	0
4	B	25	0	0	0	0
All	All	3885	0	3862	13	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.

All (13) 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)
2:B:658:PRO:HD2	2:B:661:PHE:HD2	1.68	0.58
1:A:544:GLU:HB3	1:A:702:VAL:HB	1.90	0.54
1:A:634:GLU:HB3	1:A:650:LYS:HB3	1.90	0.53
1:A:529:LEU:HD23	1:A:745:ALA:HB1	1.93	0.50
2:B:495:VAL:HG21	2:B:686:ILE:HA	1.92	0.49
1:A:538:TYR:HA	1:A:551:ILE:HG22	1.94	0.49
1:A:731:LEU:HA	1:A:735:LEU:HB2	2.01	0.42
1:A:614:GLU:HA	1:A:617:ILE:HD12	2.02	0.41
2:B:692:LEU:O	2:B:698:PRO:HB3	2.20	0.41
2:B:530:ARG:HD3	2:B:700:GLU:O	2.20	0.41
1:A:529:LEU:HD13	1:A:750:LEU:HD12	2.02	0.41
1:A:538:TYR:CZ	1:A:540:GLY:HA2	2.56	0.41
1:A:555:LEU:HD22	1:A:760:PRO:HG3	2.02	0.41

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	256/265 (97%)	246 (96%)	10 (4%)	0	100	100
2	B	203/239 (85%)	194 (96%)	8 (4%)	1 (0%)	29	52
All	All	459/504 (91%)	440 (96%)	18 (4%)	1 (0%)	47	71

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	556	ASP

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	241/247 (98%)	232 (96%)	9 (4%)	34	60
2	B	192/222 (86%)	184 (96%)	8 (4%)	30	55
All	All	433/469 (92%)	416 (96%)	17 (4%)	32	58

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	529	LEU
1	A	649	LEU
1	A	695	LEU
1	A	696	LEU
1	A	697	TYR
1	A	734	VAL
1	A	750	LEU
1	A	754	VAL
1	A	766	PHE
2	B	486	SER
2	B	494	GLU
2	B	501	LYS
2	B	553	ASP
2	B	594	LEU
2	B	608	TYR
2	B	618	VAL
2	B	648	LEU

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

Mol	Chain	Res	Type
2	B	638	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 monosaccharides 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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	516:PRO	C	518:LEU	N	3.62

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	260/265 (98%)	0.66	39 (15%) 2 1	61, 124, 214, 226	0
2	B	213/239 (89%)	0.16	10 (4%) 31 25	46, 83, 150, 195	0
All	All	473/504 (93%)	0.44	49 (10%) 6 4	46, 103, 209, 226	0

All (49) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	508	ASN	9.7
2	B	515	CYS	7.4
1	A	764	LYS	5.9
1	A	635	LEU	5.9
1	A	627	LEU	5.8
1	A	507	VAL	5.5
1	A	636	VAL	5.5
1	A	616	ILE	5.5
1	A	620	ILE	5.3
1	A	578	PHE	5.0
1	A	633	ILE	4.7
1	A	586	LEU	4.6
1	A	644	LEU	4.5
1	A	510	ASN	4.4
1	A	630	TYR	4.3
1	A	509	VAL	4.3
1	A	608	ASN	4.2
2	B	712	ALA	4.1
2	B	516	PRO	3.9
1	A	506	ARG	3.9
2	B	594	LEU	3.8
1	A	605	ASP	3.8
1	A	652	LEU	3.7
1	A	642	ASN	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	631	TYR	3.6
1	A	603	GLU	3.5
1	A	689	ILE	3.5
1	A	512	THR	3.5
2	B	486	SER	3.2
1	A	651	SER	3.0
1	A	595	ILE	2.9
1	A	626	MET	2.9
1	A	765	VAL	2.9
1	A	621	TRP	2.8
1	A	599	ASN	2.8
1	A	649	LEU	2.8
2	B	511	SER	2.8
1	A	600	LEU	2.7
2	B	639	PHE	2.7
2	B	559	ILE	2.5
2	B	647	LYS	2.4
1	A	640	LEU	2.4
1	A	584	ILE	2.4
1	A	670	ILE	2.3
1	A	673	LEU	2.3
2	B	518	LEU	2.2
1	A	674	GLY	2.2
1	A	581	PHE	2.2
1	A	505	GLU	2.0

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](#)

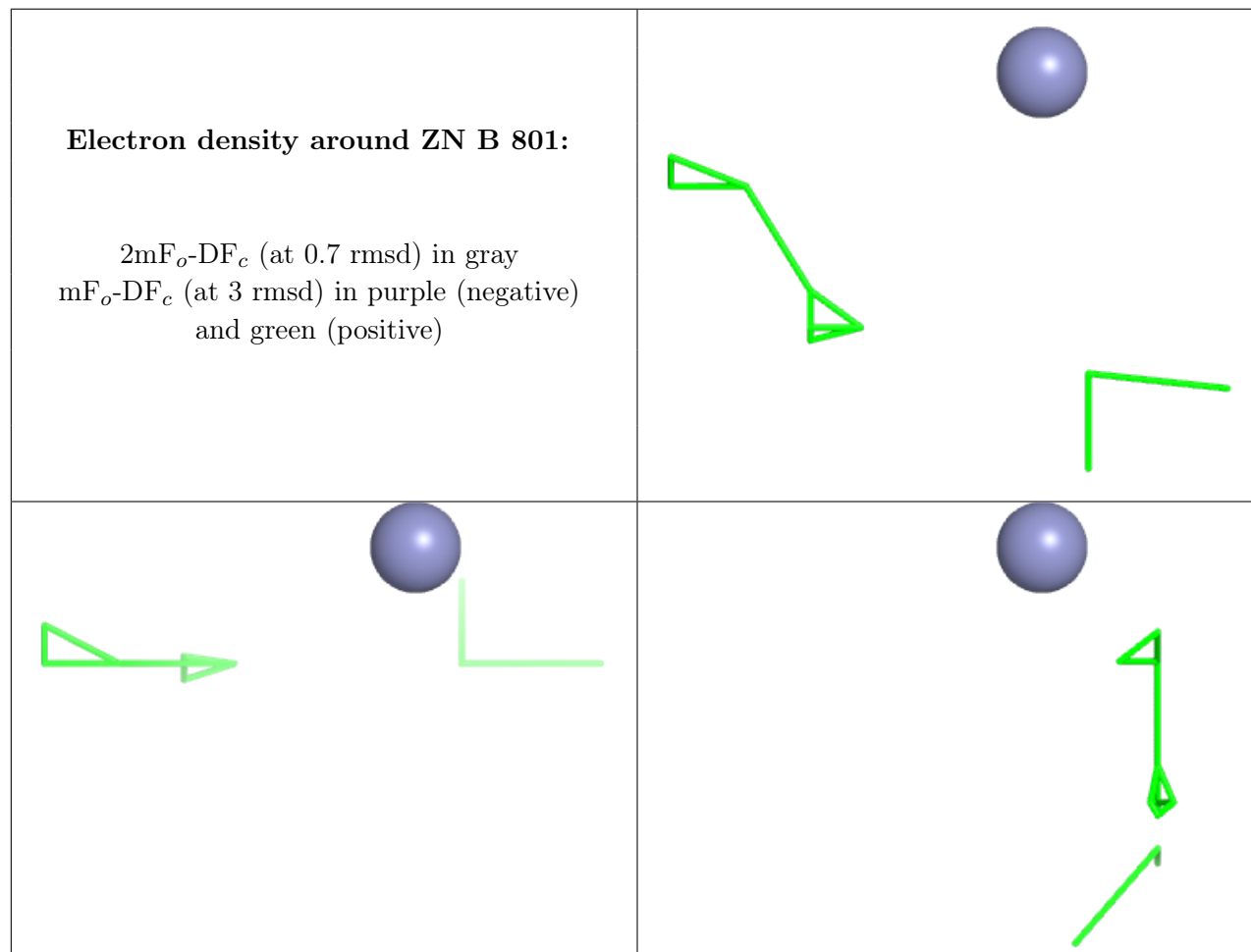
There are no monosaccharides in this entry.

6.4 Ligands [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(\AA^2)	Q<0.9
3	ZN	B	801	1/1	0.97	0.16	87,87,87,87	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.