wwPDB X-ray Structure Validation Summary Report (i)

Dec 2, 2023 – 01:58 pm GMT

PDB ID	:	1ODG
Title	:	Very-short-patch DNA repair endonuclease bound to its reaction product site
Authors	:	Bunting, K.A.; Roe, S.M.; Headley, A.; Brown, T.; Savva, R.; Pearl, L.H.
Deposited on	:	2003-02-19
Resolution	:	2.80 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:

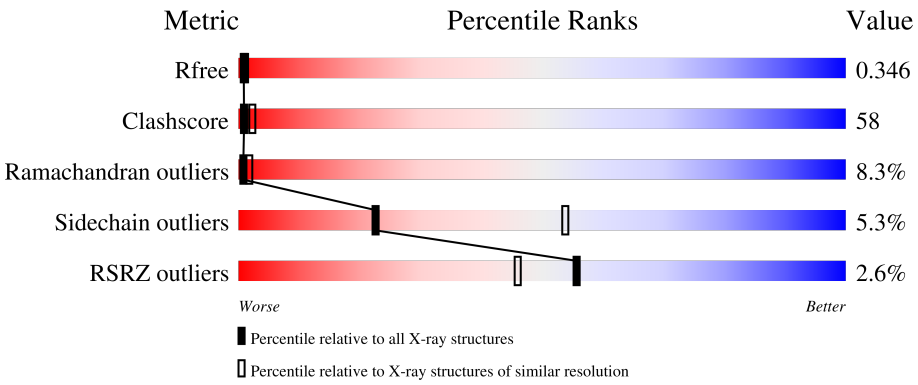
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	134	3%, 34%, 57%, 8%, •
2	F	12	75%, 25%
2	W	12	83%, 17%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 1672 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 ENDONUCLEASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	134	1092	694	200	193	5	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	F	12	249	118	45	74	12	0	0	0
			Total	C	N	O	P			
2	W	12	249	118	45	74	12	0	0	0

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	34	34	34	0	0
			Total	O		
4	F	27	27	27	0	0
			Total	O		
4	W	20	20	20	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 ENDONUCLEASE

Chain A: 3% 34% 57% 8%

A23 I24 E25 K26 R27 L28 A29 S30 L31 L32 T33 G34 Q35 G36 L37 A38 F39 R40 V41 Q42 D43 A44 S45 L46 P47 G48 R49 P50 D51 F52 D55 E56 Y57 R58 C59 V60 I61 F62 T63 H64 G65 C66 F67 W68 H69 H70 Y74 L75 F76 K77 V78 P79 A80 T81 R82 T83 E84 F85

W86 L87 E88 K89 I90 N93 I101 S102 R103 E106 W109 R110 V111 L112 I113 V114 W115 E116 C117 A118 L119 R120 G121 R122 E123 K124 L125 T126 D127 E128 A129 L130 T131 E132 R133 L134 E135 E136 W137 I138 C139 G140 E141 G142 A145 Q146 I147 D148 T149 Q150 G151 I152 H153 L154 L155

A156

• Molecule 2: 5'-D(*TP*AP*GP*GP*CP*5CM*TP*GP*GP*AP*TP*CP)-3'

Chain F: 75% 25%

T1 A2 G3 G4 C5 C6 T7 G8 G9 A10 T11 C12

• Molecule 2: 5'-D(*TP*AP*GP*GP*CP*5CM*TP*GP*GP*AP*TP*CP)-3'

Chain W: 83% 17%

T1 A2 G3 G4 C5 C6 T7 G8 G9 A10 T11 C12

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α, β, γ	102.81Å 102.81Å 64.29Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	24.70 – 2.80 52.12 – 2.75	Depositor EDS
% Data completeness (in resolution range)	98.7 (24.70-2.80) 98.7 (52.12-2.75)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.22 (at 2.77Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.287 , 0.356 0.278 , 0.346	Depositor DCC
R_{free} test set	960 reflections (9.42%)	wwPDB-VP
Wilson B-factor (Å^2)	81.1	Xtriage
Anisotropy	0.278	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 70.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.065 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	1672	wwPDB-VP
Average B, all atoms (Å^2)	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.68% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: 5CM, 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	0.51	0/1118	0.65	0/1513
2	F	0.71	1/255 (0.4%)	0.86	0/388
2	W	0.79	1/255 (0.4%)	0.82	0/388
All	All	0.59	2/1628 (0.1%)	0.72	0/2289

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
2	F	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	W	1	DT	OP3-P	-7.55	1.52	1.61
2	F	1	DT	OP3-P	-7.07	1.52	1.61

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	F	8	DG	Sidechain
2	F	9	DG	Sidechain

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	1092	0	1066	110	0
2	F	249	0	138	36	0
2	W	249	0	138	27	0
3	A	1	0	0	0	0
4	A	34	0	0	11	0
4	F	27	0	0	8	0
4	W	20	0	0	7	0
All	All	1672	0	1342	169	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 58.

The worst 5 of 169 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:F:2:DA:H2"	2:F:3:DG:H5"	1.38	1.03
1:A:24:ILE:HD11	1:A:119:LEU:HB3	1.43	1.01
1:A:129:ALA:HB1	1:A:133:ARG:HH21	1.32	0.94
2:F:2:DA:H62	2:W:7:DT:H3	1.16	0.94
2:W:7:DT:H1'	4:W:2008:HOH:O	1.68	0.91

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	132/134 (98%)	100 (76%)	21 (16%)	11 (8%)	1 2

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	138	ILE
1	A	141	GLU
1	A	37	LEU
1	A	42	GLN
1	A	123	GLU

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	113/114 (99%)	107 (95%)	6 (5%)	22 54

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	119	LEU
1	A	131	THR
1	A	153	HIS
1	A	31	LEU
1	A	24	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	42	GLN
1	A	64	HIS
1	A	69	HIS
1	A	70	HIS
1	A	150	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

2 non-standard protein/DNA/RNA residues are modelled in this entry.

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	5CM	W	6	2	17,21,22	0.63	0	24,30,33	1.55	4 (16%)
2	5CM	F	6	2	17,21,22	0.71	1 (5%)	24,30,33	1.58	3 (12%)

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	5CM	W	6	2	-	5/7/21/22	0/2/2/2
2	5CM	F	6	2	-	3/7/21/22	0/2/2/2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	6	5CM	C4-N3	2.24	1.38	1.34

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	6	5CM	C5-C4-N3	-5.07	116.21	121.67
2	W	6	5CM	C5-C4-N3	-4.80	116.50	121.67
2	W	6	5CM	C4-N3-C2	4.38	126.61	120.69
2	F	6	5CM	C4-N3-C2	4.32	126.53	120.69
2	W	6	5CM	N1-C2-N3	-2.39	114.45	118.81

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	F	6	5CM	O4'-C4'-C5'-O5'
2	F	6	5CM	C3'-C4'-C5'-O5'
2	W	6	5CM	C2'-C1'-N1-C6
2	W	6	5CM	O4'-C1'-N1-C6
2	W	6	5CM	O4'-C1'-N1-C2

There are no ring outliers.

2 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	W	6	5CM	5	0
2	F	6	5CM	4	0

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)

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	134/134 (100%)	0.29	4 (2%) 50 40	41, 72, 113, 120	0
2	F	11/12 (91%)	-0.55	0 100 100	66, 72, 88, 94	0
2	W	11/12 (91%)	-0.43	0 100 100	56, 66, 89, 97	0
All	All	156/158 (98%)	0.18	4 (2%) 56 46	41, 72, 112, 120	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	33	THR	4.0
1	A	32	LEU	2.8
1	A	38	ALA	2.5
1	A	36	GLY	2.4

6.2 Non-standard residues in protein, DNA, RNA chains (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(Å^2)	Q<0.9
2	5CM	W	6	20/21	0.82	0.16	69,84,101,101	0
2	5CM	F	6	20/21	0.87	0.17	53,73,83,89	0

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(Å^2)	Q<0.9
3	ZN	A	200	1/1	0.98	0.21	56,56,56,56	0

6.5 Other polymers (i)

There are no such residues in this entry.