wwPDB X-ray Structure Validation Summary Report (i)

Aug 6, 2023 – 01:41 AM EDT

PDB ID	:	1K3X
Title	:	Crystal structure of a trapped reaction intermediate of the DNA repair enzyme Endonuclease VIII with Brominated-DNA
Authors	:	Golan, G.; Zharkov, D.O.; Gilboa, R.; Fernandes, A.S.; Kycia, J.H.; Gerchman, S.E.; Rieger, R.A.; Grollman, A.P.; Shoham, G.
Deposited on	:	2001-10-04
Resolution	:	1.25 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:

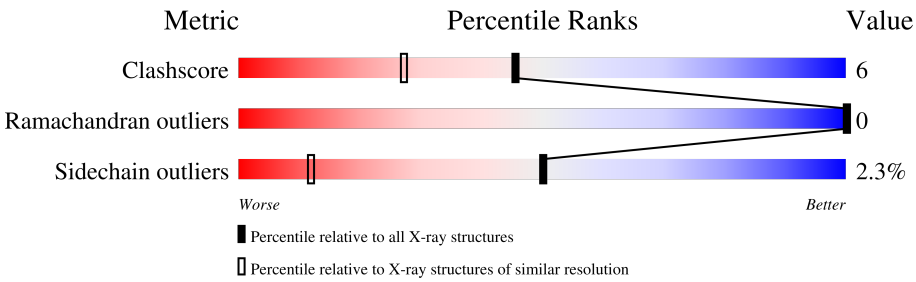
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 1.25 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))
Clashscore	141614	1060 (1.28-1.24)
Ramachandran outliers	138981	1029 (1.28-1.24)
Sidechain outliers	138945	1028 (1.28-1.24)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	B	13	31% 46% 23%
2	C	13	8% 54% 15% 23%
3	A	262	84% 11% • •

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
1	BRU	B	404	X	-	-	-
6	GOL	A	511	-	X	X	-

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 2998 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 DNA chain called 5'-D(*GP*GP*CP*(BRU)P*(BRU)P*CP*AP*(BRU)P*CP*CP*(BRU)P*GP*G)-3'.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	Br	C	N	O	P			
1	B	10	199	4	92	30	63	10	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	C	10	202	94	43	55	10	0	0	0

- Molecule 3 is a protein called Endonuclease VIII.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	253	2059	1316	370	368	5	0	8	0

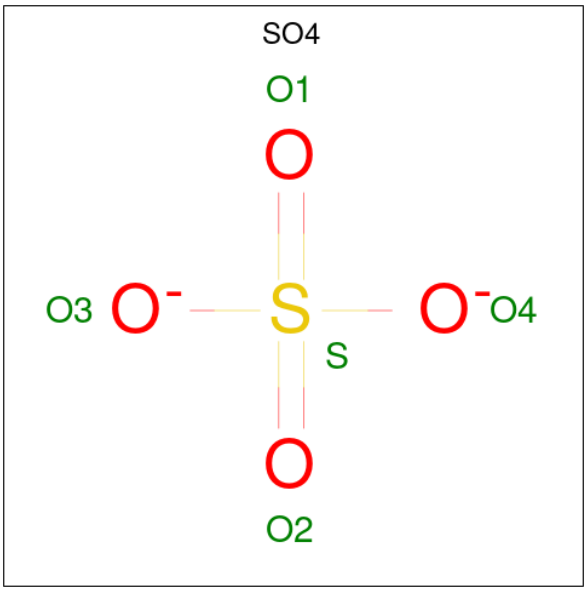
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	34	THR	PRO	SEE REMARK 999	UNP P50465
A	112	ARG	THR	SEE REMARK 999	UNP P50465

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

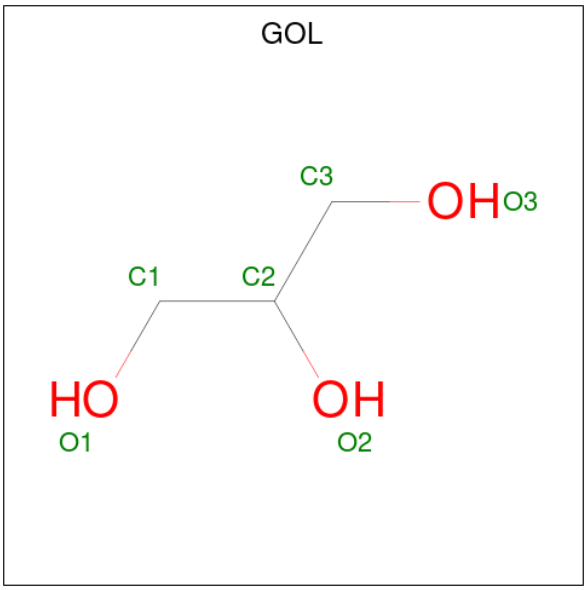
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
4	A	1	1	1	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total O S 5 4 1	0	0
5	A	1	Total O S 5 4 1	0	0
5	A	1	Total O S 5 4 1	0	0
5	A	1	Total O S 5 4 1	0	0
5	A	1	Total O S 5 4 1	0	0

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
6	A	1	6	3	3	0	0
			Total	C	O		
6	A	1	6	3	3	0	0
			Total	C	O		
6	A	1	6	3	3	0	0
			Total	C	O		
6	A	1	6	3	3	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	B	24	24	24	0	0
			Total	O		
7	C	23	23	23	0	0
			Total	O		
7	A	441	441	441	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. 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. 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.

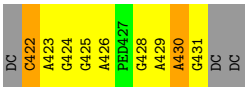
Note EDS was not executed.

• Molecule 1: 5'-D(*GP*GP*CP*(BRU)P*(BRU)P*CP*AP*(BRU)P*CP*CP*(BRU)P*GP*G)-3'

Chain B: 31% 46% 23%

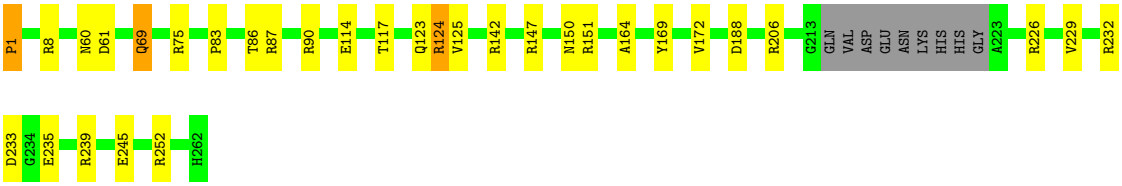
• Molecule 2: 5'-D(*CP*CP*AP*GP*GP*AP*(PED)P*GP*AP*AP*GP*CP*C)-3'

Chain C: 8% 54% 15% 23%



• Molecule 3: Endonuclease VIII

Chain A: 84% 11%



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 43 2 2	Depositor
Cell constants a, b, c, α, β, γ	76.50Å 76.50Å 164.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 1.25	Depositor
% Data completeness (in resolution range)	97.4 (10.00-1.25)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
Refinement program	SHELXL-97	Depositor
R, R_{free}	0.149 , 0.181	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2998	wwPDB-VP
Average B, all atoms (Å^2)	22.0	wwPDB-VP

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, PED, BRU, GOL, SO4

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	B	1.81	3/129 (2.3%)	5.89	44/188 (23.4%)
2	C	2.08	9/215 (4.2%)	6.05	92/328 (28.0%)
3	A	0.74	0/2136	1.29	21/2898 (0.7%)
All	All	1.01	12/2480 (0.5%)	2.62	157/3414 (4.6%)

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	B	1	0

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	412	DG	C5-C4	6.66	1.43	1.38
2	C	431	DG	C5-C4	6.30	1.42	1.38
2	C	431	DG	C5-C6	5.92	1.48	1.42
2	C	425	DG	C5-C6	5.47	1.47	1.42
2	C	431	DG	N3-C4	5.36	1.39	1.35

The worst 5 of 157 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	430	DA	C6-N1-C2	24.49	133.30	118.60
1	B	412	DG	N3-C4-C5	-22.55	117.33	128.60
2	C	431	DG	N7-C8-N9	21.99	124.09	113.10
2	C	424	DG	N7-C8-N9	20.22	123.21	113.10
2	C	431	DG	C8-N9-C4	-20.20	98.32	106.40

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	B	404	BRU	C4'

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	B	199	0	103	9	0
2	C	202	0	108	5	0
3	A	2059	0	2057	18	0
4	A	1	0	0	0	0
5	A	25	0	0	0	0
6	A	24	0	32	7	0
7	A	441	0	0	8	0
7	B	24	0	0	1	0
7	C	23	0	0	1	0
All	All	2998	0	2300	31	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 6.

The worst 5 of 31 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)
3:A:147:ARG:HH21	6:A:511:GOL:H11	1.26	1.00
1:B:410:DC:H2"	1:B:411:BRU:BR	2.35	0.80
2:C:422:DC:H2'	2:C:422:DC:OP2	1.82	0.79
3:A:69:GLN:HG3	7:A:789:HOH:O	1.85	0.76
3:A:150:ASN:HB2	7:A:894:HOH:O	1.87	0.72

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
3	A	257/262 (98%)	247 (96%)	10 (4%)	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
3	A	224/226 (99%)	218 (97%)	6 (3%)	44 8

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

Mol	Chain	Res	Type
3	A	124[A]	ARG
3	A	124[B]	ARG
3	A	169	TYR
3	A	69	GLN
3	A	1	PRO

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	68	ASN
3	A	184	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)

4 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	$\#\|Z\| > 2$
1	BRU	B	411	1,2	18,21,22	1.48	1 (5%)	26,30,33	2.66	14 (53%)
1	BRU	B	405	1,2	18,21,22	0.94	1 (5%)	26,30,33	2.26	9 (34%)
1	BRU	B	404	1,2	18,21,22	1.29	3 (16%)	26,30,33	4.49	13 (50%)
1	BRU	B	408	1,2	18,21,22	1.19	2 (11%)	26,30,33	1.63	6 (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
1	BRU	B	411	1,2	-	0/7/21/22	0/2/2/2
1	BRU	B	405	1,2	-	0/7/21/22	0/2/2/2
1	BRU	B	404	1,2	1/1/4/4	2/7/21/22	0/2/2/2
1	BRU	B	408	1,2	-	0/7/21/22	0/2/2/2

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	411	BRU	C2-N3	4.88	1.46	1.38
1	B	404	BRU	C2-N3	2.82	1.43	1.38
1	B	408	BRU	C2-N3	2.59	1.42	1.38
1	B	404	BRU	C4-N3	-2.42	1.34	1.38
1	B	408	BRU	O4'-C1'	2.10	1.47	1.42

The worst 5 of 42 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	404	BRU	BR-C5-C4	-14.20	101.55	118.03
1	B	404	BRU	BR-C5-C6	11.30	136.45	120.64
1	B	404	BRU	O2-C2-N1	6.82	131.85	122.79
1	B	411	BRU	C1'-N1-C6	-6.50	109.57	120.77
1	B	405	BRU	C6-C5-C4	6.43	127.20	120.67

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	B	404	BRU	C4'

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	B	404	BRU	C4'-C5'-O5'-P
1	B	404	BRU	C3'-C4'-C5'-O5'

There are no ring outliers.

3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	411	BRU	5	0
1	B	405	BRU	2	0
1	B	404	BRU	2	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 10 ligands modelled in this entry, 1 is monoatomic - leaving 9 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	$\#\|Z\| > 2$
6	GOL	A	512	-	5,5,5	0.53	0	5,5,5	0.88	0
6	GOL	A	511	-	5,5,5	0.96	0	5,5,5	2.29	2 (40%)
5	SO4	A	554	-	4,4,4	0.17	0	6,6,6	0.13	0
5	SO4	A	552	-	4,4,4	0.48	0	6,6,6	0.52	0
6	GOL	A	514	-	5,5,5	0.62	0	5,5,5	0.39	0
6	GOL	A	513	-	5,5,5	1.03	1 (20%)	5,5,5	1.35	1 (20%)
5	SO4	A	551	-	4,4,4	0.29	0	6,6,6	0.69	0
5	SO4	A	555	-	4,4,4	0.29	0	6,6,6	0.76	0
5	SO4	A	553	-	4,4,4	0.25	0	6,6,6	0.32	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
6	GOL	A	514	-	-	4/4/4/4	-
6	GOL	A	513	-	-	2/4/4/4	-
6	GOL	A	512	-	-	0/4/4/4	-
6	GOL	A	511	-	-	4/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	513	GOL	C1-C2	2.08	1.60	1.51

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	511	GOL	O3-C3-C2	4.20	130.34	110.20
6	A	511	GOL	C3-C2-C1	-2.81	100.79	111.70
6	A	513	GOL	O3-C3-C2	2.29	121.16	110.20

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	511	GOL	O1-C1-C2-C3
6	A	513	GOL	O1-C1-C2-C3
6	A	514	GOL	O1-C1-C2-C3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
6	A	513	GOL	O1-C1-C2-O2
6	A	514	GOL	C1-C2-C3-O3

There are no ring outliers.

3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	511	GOL	4	0
6	A	514	GOL	2	0
6	A	513	GOL	1	0

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)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.