Full wwPDB X-ray Structure Validation Report ⓘ

Oct 2, 2023 – 12:54 PM EDT

PDB ID	:	6NFL
Title	:	Crystal Structure of the Cancer Genomic DNA Mutator APOBEC3B with loop 7 from APOBEC3G complexed with 2-HP
Authors	:	Shi, K.; Orellana, K.; Aihara, H.
Deposited on	:	2018-12-20
Resolution	:	1.73 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 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	:	FAILED
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
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

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 1.73 Å.

There are no overall percentile quality scores available for this entry.

MolProbity and EDS failed to run properly - the sequence quality summary graphics cannot be shown.

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
3	71O	A	402	-	X	-	-

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 3165 atoms, of which 1458 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 dC->dU-editing enzyme APOBEC-3B.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	182	2978	974	1432	275	284	13	0	5	0

There are 28 discrepancies between the modelled and reference sequences:

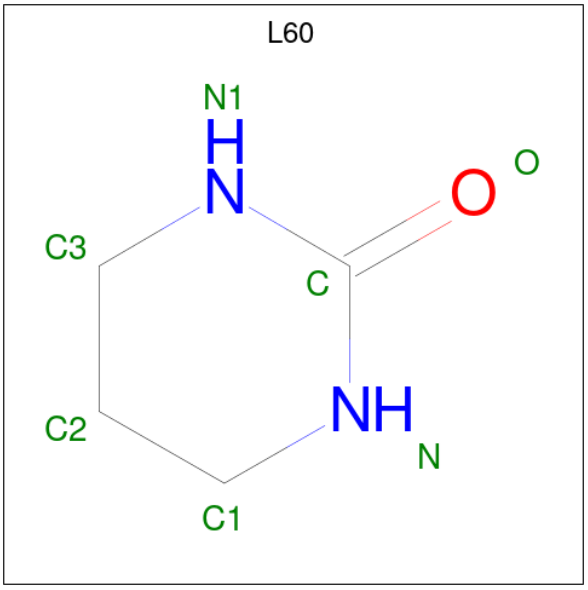
Chain	Residue	Modelled	Actual	Comment	Reference
A	186	MET	-	initiating methionine	UNP Q9UH17
A	200	SER	PHE	engineered mutation	UNP Q9UH17
A	228	SER	TRP	engineered mutation	UNP Q9UH17
A	230	LYS	LEU	engineered mutation	UNP Q9UH17
A	250	SER	ALA	engineered mutation	UNP Q9UH17
A	?	-	LYS	deletion	UNP Q9UH17
A	?	-	ASN	deletion	UNP Q9UH17
A	?	-	LEU	deletion	UNP Q9UH17
A	?	-	LEU	deletion	UNP Q9UH17
A	?	-	CYS	deletion	UNP Q9UH17
A	?	-	GLY	deletion	UNP Q9UH17
A	?	-	PHE	deletion	UNP Q9UH17
A	?	-	TYR	deletion	UNP Q9UH17
A	308	LYS	PHE	engineered mutation	UNP Q9UH17
A	315	ASP	TYR	engineered mutation	UNP Q9UH17
A	316	GLN	ASP	engineered mutation	UNP Q9UH17
A	317	GLY	PRO	engineered mutation	UNP Q9UH17
A	318	ARG	LEU	engineered mutation	UNP Q9UH17
A	319	CYS	TYR	engineered mutation	UNP Q9UH17
A	320	GLN	LYS	engineered mutation	UNP Q9UH17
A	379	LEU	-	expression tag	UNP Q9UH17
A	380	GLU	-	expression tag	UNP Q9UH17
A	381	HIS	-	expression tag	UNP Q9UH17
A	382	HIS	-	expression tag	UNP Q9UH17
A	383	HIS	-	expression tag	UNP Q9UH17
A	384	HIS	-	expression tag	UNP Q9UH17
A	385	HIS	-	expression tag	UNP Q9UH17

Continued on next page...

Continued from previous page...

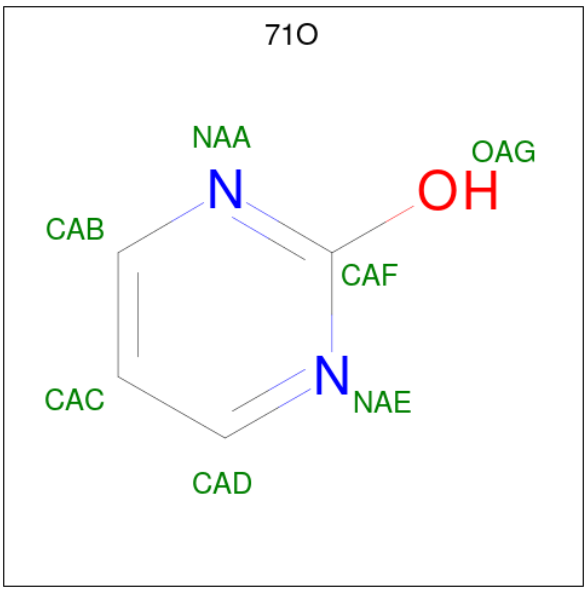
Chain	Residue	Modelled	Actual	Comment	Reference
A	386	HIS	-	expression tag	UNP Q9UH17

- Molecule 2 is 1,3-diazinan-2-one (three-letter code: L60) (formula: $C_4H_8N_2O$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	N	O		
2	A	1	11	4	4	2	1	0	0

- Molecule 3 is pyrimidin-2-ol (three-letter code: 71O) (formula: $C_4H_4N_2O$).

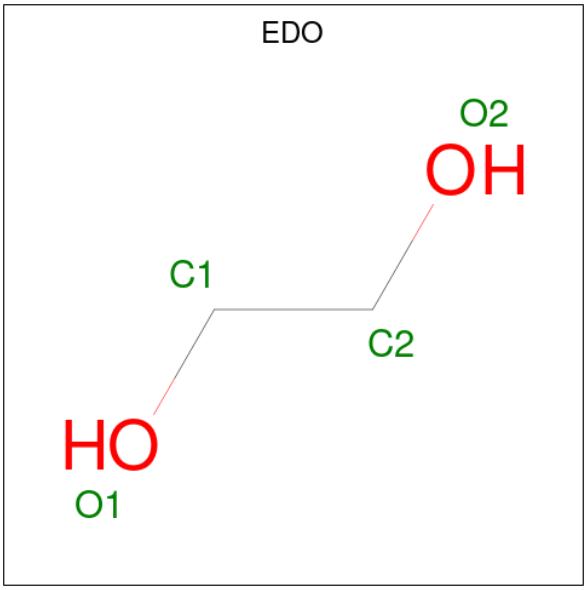


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	H	N	O	0	0
			11	4	4	2	1		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	H	O		0	0
			10	2	6	2			
5	A	1	Total	C	H	O		0	0
			10	2	6	2			
5	A	1	Total	C	H	O		0	0
			10	2	6	2			

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	134	Total	O	0	0
			134	134		

MolProbity and EDS failed to run properly - this section is therefore empty.

3 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	50.65Å 50.65Å 149.25Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.91 – 1.73	Depositor
% Data completeness (in resolution range)	97.7 (41.91-1.73)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.52 (at 1.73Å)	Xtriage
Refinement program	PHENIX (dev_3366: ???)	Depositor
R, R_{free}	0.162 , 0.194	Depositor
Wilson B-factor (Å^2)	21.0	Xtriage
Anisotropy	0.291	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3165	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.83% 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.

4 Model quality

4.1 Standard geometry

MolProbity failed to run properly - this section is therefore empty.

4.2 Too-close contacts

MolProbity failed to run properly - this section is therefore empty.

4.3 Torsion angles

4.3.1 Protein backbone

MolProbity failed to run properly - this section is therefore empty.

4.3.2 Protein sidechains

MolProbity failed to run properly - this section is therefore empty.

4.3.3 RNA

MolProbity failed to run properly - this section is therefore empty.

4.4 Non-standard residues in protein, DNA, RNA chains

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

4.5 Carbohydrates

There are no monosaccharides in this entry.

4.6 Ligand geometry

Of 6 ligands modelled in this entry, 1 is monoatomic - leaving 5 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	$\#\|Z\| > 2$
5	EDO	A	405	-	3,3,3	0.45	0	2,2,2	0.32	0
5	EDO	A	406	-	3,3,3	0.47	0	2,2,2	0.31	0
5	EDO	A	404	-	3,3,3	0.46	0	2,2,2	0.22	0
3	71O	A	402	-	6,7,7	1.76	2 (33%)	8,8,8	3.89	5 (62%)
2	L60	A	401	-	7,7,7	3.58	6 (85%)	7,8,8	2.91	3 (42%)

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
5	EDO	A	405	-	-	0/1/1/1	-
5	EDO	A	406	-	-	0/1/1/1	-
5	EDO	A	404	-	-	0/1/1/1	-
3	71O	A	402	-	-	-	0/1/1/1
2	L60	A	401	-	-	-	0/1/1/1

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	401	L60	C1-N	-5.89	1.33	1.46
2	A	401	L60	C3-N1	-5.11	1.35	1.46
2	A	401	L60	C2-C3	-3.00	1.35	1.49
2	A	401	L60	C-N1	2.87	1.37	1.34
3	A	402	71O	CAD-NAE	2.83	1.40	1.34
3	A	402	71O	CAB-NAA	2.77	1.40	1.34
2	A	401	L60	C2-C1	-2.45	1.37	1.49
2	A	401	L60	C-N	2.25	1.37	1.34

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	402	71O	CAD-NAE-CAF	6.30	120.79	114.55
3	A	402	71O	CAB-NAA-CAF	6.23	120.72	114.55
2	A	401	L60	C2-C1-N	5.58	122.48	110.24
3	A	402	71O	NAA-CAF-NAE	-5.35	120.75	127.99
2	A	401	L60	C2-C3-N1	4.34	119.78	110.24

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	402	71O	CAC-CAD-NAE	-2.45	119.43	123.43
3	A	402	71O	CAC-CAB-NAA	-2.38	119.53	123.43
2	A	401	L60	C1-N-C	-2.02	119.02	123.97

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

4.7 Other polymers

There are no such residues in this entry.

4.8 Polymer linkage issues

There are no chain breaks in this entry.

5 Fit of model and data

5.1 Protein, DNA and RNA chains

EDS failed to run properly - this section is therefore empty.

5.2 Non-standard residues in protein, DNA, RNA chains

EDS failed to run properly - this section is therefore empty.

5.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

5.4 Ligands

EDS failed to run properly - this section is therefore empty.

5.5 Other polymers

EDS failed to run properly - this section is therefore empty.