



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2026 – 07:53 AM UTC

PDB ID : 3E1U / pdb_00003e1u
Title : The Crystal Structure of the Anti-Viral APOBEC3G Catalytic Domain
Authors : Holden, L.; Prochnow, C.; Chang, Y.P.; Bransteitter, R.; Chelico, L.; Sen, U.;
Stevens, R.C.; Goodman, R.F.; Chen, X.S.
Deposited on : 2008-08-04
Resolution : 2.30 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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

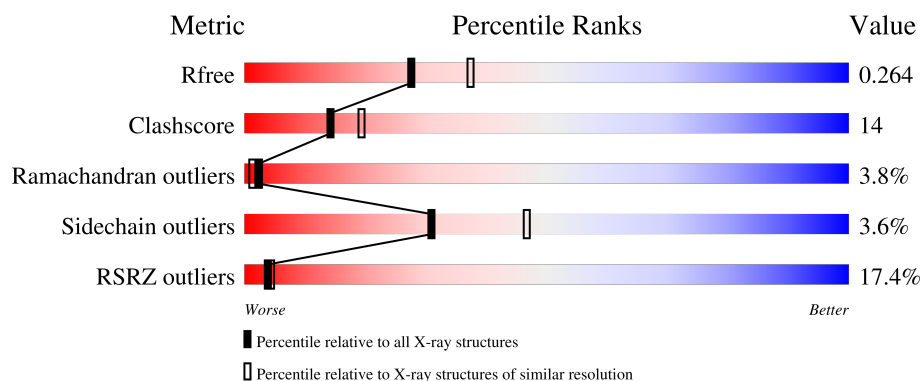
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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}	180053	6319 (2.30-2.30)
Clashscore	190562	6919 (2.30-2.30)
Ramachandran outliers	187476	6854 (2.30-2.30)
Sidechain outliers	187428	6854 (2.30-2.30)
RSRZ outliers	180081	6325 (2.30-2.30)

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	189	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 1554 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 dC->dU-editing enzyme APOBEC-3G.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	184	Total	C	N	O	S	0	0	0
			1523	967	268	274	14			

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	192	GLY	-	expression tag	UNP Q9HC16
A	193	PRO	-	expression tag	UNP Q9HC16
A	194	LEU	-	expression tag	UNP Q9HC16
A	195	GLY	-	expression tag	UNP Q9HC16
A	196	SER	-	expression tag	UNP Q9HC16

- Molecule 2 is ZINC ION (CCD ID: ZN) (formula: Zn).

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

- Molecule 3 is water.

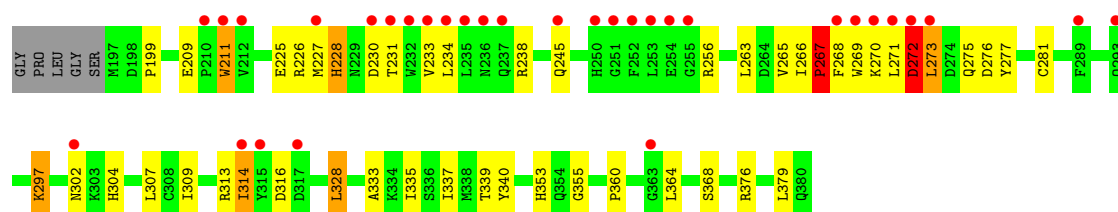
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	30	Total	O	0	0
			30	30		

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 dC->dU-editing enzyme APOBEC-3G

Chain A: 



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	83.46Å 57.33Å 40.58Å 90.00° 96.46° 90.00°	Depositor
Resolution (Å)	30.00 – 2.30 30.00 – 2.30	Depositor EDS
% Data completeness (in resolution range)	(Not available) (30.00-2.30) 94.5 (30.00-2.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.16 (at 2.20Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.251 , 0.267 0.256 , 0.264	Depositor DCC
R_{free} test set	1511 reflections (8.07%)	wwPDB-VP
Wilson B-factor (Å ²)	22.2	Xtriage
Anisotropy	0.785	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 47.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	1554	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.64% 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

5.1 Standard geometry

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	0.63	0/1570	0.87	5/2125 (0.2%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	209	GLU	CA-C-N	6.48	127.94	119.84
1	A	209	GLU	C-N-CA	6.48	127.94	119.84
1	A	199	PRO	N-CA-C	5.64	117.58	110.70
1	A	271	LEU	N-CA-C	5.64	115.99	108.38
1	A	272	ASP	N-CA-C	5.29	122.06	110.80

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	1523	0	1426	42	0
2	A	1	0	0	0	0
3	A	30	0	0	1	0
All	All	1554	0	1426	42	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 14.

All (42) 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:A:353:HIS:HD2	1:A:355:GLY:H	1.20	0.88
1:A:328:LEU:HB3	1:A:335:ILE:HD11	1.61	0.83
1:A:227:MET:HE2	1:A:276:ASP:HB3	1.68	0.76
1:A:267:PRO:O	1:A:269:TRP:N	2.20	0.74
1:A:245:GLN:HB2	1:A:256:ARG:HB3	1.72	0.72
1:A:353:HIS:CD2	1:A:355:GLY:H	2.10	0.68
1:A:267:PRO:C	1:A:269:TRP:H	2.02	0.68
1:A:228:HIS:HE1	1:A:230:ASP:HB2	1.66	0.60
1:A:265:VAL:O	1:A:269:TRP:HB2	2.01	0.59
1:A:272:ASP:HB3	1:A:277:TYR:OH	2.07	0.54
1:A:340:TYR:HE1	1:A:360:PRO:HG3	1.73	0.54
1:A:226:ARG:HH21	1:A:275:GLN:HE21	1.54	0.53
1:A:275:GLN:O	1:A:304:HIS:HB3	2.08	0.53
1:A:297:LYS:HB2	1:A:297:LYS:NZ	2.24	0.52
1:A:364:LEU:C	1:A:364:LEU:HD23	2.33	0.52
1:A:263:LEU:O	1:A:267:PRO:HD3	2.10	0.51
1:A:267:PRO:C	1:A:269:TRP:N	2.65	0.51
1:A:307:LEU:HD11	1:A:309:ILE:HD11	1.93	0.50
1:A:211:TRP:CE3	1:A:211:TRP:HA	2.48	0.48
1:A:228:HIS:CE1	1:A:230:ASP:HB2	2.46	0.48
1:A:275:GLN:HG3	3:A:27:HOH:O	2.14	0.48
1:A:314:ILE:HD11	1:A:337:ILE:HD11	1.96	0.48
1:A:238:ARG:HD2	1:A:270:LYS:HD2	1.96	0.47
1:A:353:HIS:HD2	1:A:355:GLY:N	2.00	0.47
1:A:275:GLN:HA	1:A:275:GLN:OE1	2.15	0.47
1:A:266:ILE:N	1:A:267:PRO:CD	2.78	0.46
1:A:211:TRP:HA	1:A:211:TRP:HE3	1.80	0.46
1:A:272:ASP:O	1:A:273:LEU:HB2	2.15	0.46
1:A:313:ARG:O	1:A:314:ILE:HB	2.15	0.45
1:A:225:GLU:HB3	1:A:234:LEU:HB3	1.98	0.45
1:A:238:ARG:CD	1:A:270:LYS:HD2	2.47	0.45
1:A:281:CYS:HB2	1:A:309:ILE:CD1	2.47	0.44
1:A:231:THR:OG1	1:A:233:VAL:HG22	2.18	0.44
1:A:281:CYS:HB2	1:A:309:ILE:HD13	2.00	0.43
1:A:266:ILE:N	1:A:267:PRO:HD2	2.33	0.43
1:A:272:ASP:OD2	1:A:273:LEU:N	2.52	0.43
1:A:340:TYR:CE1	1:A:360:PRO:HG3	2.54	0.43
1:A:302:ASN:C	1:A:304:HIS:H	2.25	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:328:LEU:HD22	1:A:333:ALA:CB	2.49	0.42
1:A:263:LEU:O	1:A:267:PRO:CD	2.68	0.42
1:A:339:THR:HG22	1:A:368:SER:OG	2.20	0.41
1:A:227:MET:HB2	1:A:276:ASP:O	2.20	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	182/189 (96%)	164 (90%)	11 (6%)	7 (4%)	2 1

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	268	PHE
1	A	272	ASP
1	A	316	ASP
1	A	267	PRO
1	A	228	HIS
1	A	273	LEU
1	A	314	ILE

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	165/168 (98%)	159 (96%)	6 (4%)	31 47

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	211	TRP
1	A	267	PRO
1	A	297	LYS
1	A	328	LEU
1	A	376	ARG
1	A	379	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	229	ASN
1	A	236	ASN
1	A	248	HIS
1	A	275	GLN
1	A	302	ASN
1	A	318	GLN
1	A	353	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

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

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

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	184/189 (97%)	1.00	32 (17%) 4 4	13, 27, 56, 72	0

All (32) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	212	VAL	7.1
1	A	269	TRP	6.3
1	A	268	PHE	5.9
1	A	211	TRP	5.5
1	A	271	LEU	5.3
1	A	273	LEU	4.8
1	A	210	PRO	4.5
1	A	314	ILE	3.9
1	A	272	ASP	3.8
1	A	235	LEU	3.5
1	A	302	ASN	3.4
1	A	236	ASN	3.3
1	A	237	GLN	3.2
1	A	231	THR	3.1
1	A	232	TRP	3.1
1	A	253	LEU	2.9
1	A	254	GLU	2.9
1	A	270	LYS	2.8
1	A	289	PHE	2.8
1	A	234	LEU	2.8
1	A	363	GLY	2.7
1	A	245	GLN	2.6
1	A	252	PHE	2.6
1	A	251	GLY	2.4
1	A	227	MET	2.4
1	A	233	VAL	2.3
1	A	230	ASP	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	255	GLY	2.2
1	A	317	ASP	2.2
1	A	315	TYR	2.1
1	A	293	GLN	2.1
1	A	250	HIS	2.1

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 oligosaccharides 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($\text{\AA}^2$)	Q<0.9
2	ZN	A	1001	1/1	0.99	0.04	31,31,31,31	0

6.5 Other polymers [i](#)

There are no such residues in this entry.