



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2026 – 04:51 AM UTC

PDB ID : 6A42 / pdb_00006a42
Title : R1EN(5-223)-ubiquitin fusion
Authors : Maita, N.
Deposited on : 2018-06-19
Resolution : 1.70 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
Mogul : 2022.3.0, CSD as543be (2022)
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

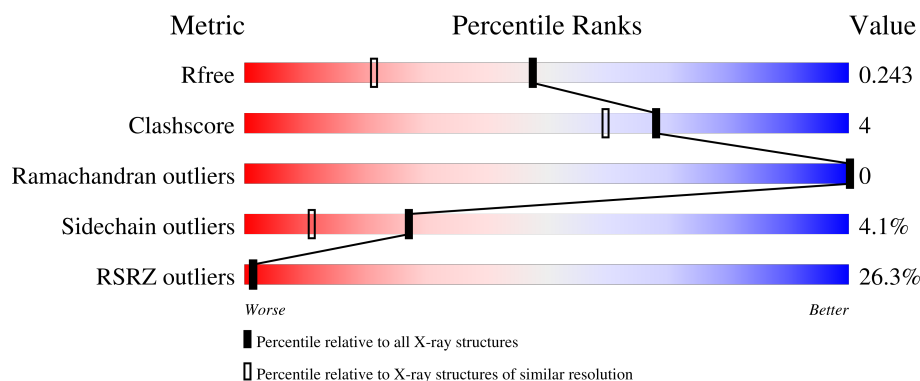
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.70 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	5551 (1.70-1.70)
Clashscore	190562	5924 (1.70-1.70)
Ramachandran outliers	187476	5846 (1.70-1.70)
Sidechain outliers	187428	5846 (1.70-1.70)
RSRZ outliers	180081	5554 (1.70-1.70)

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	298	26% 86% 12% ..

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 2515 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.

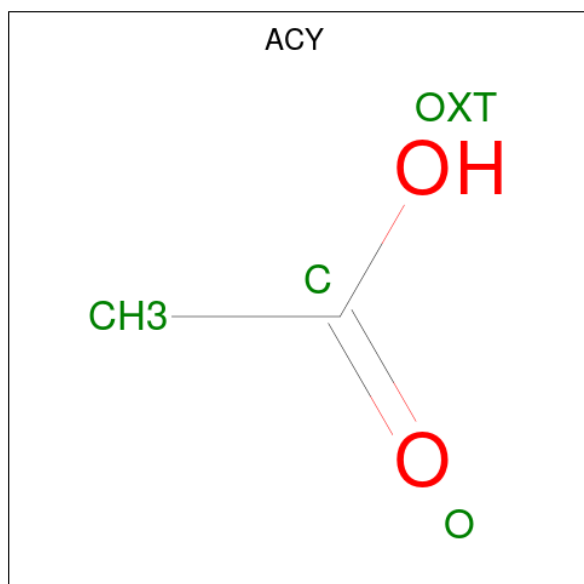
- Molecule 1 is a protein called RNA-directed DNA polymerase homolog (R1),Polyubiquitin-C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	293	Total	C	N	O	S	0	7	0
			2333	1467	429	429	8			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	4	HIS	-	expression tag	UNP Q7M4J4
A	44	ALA	GLU	engineered mutation	UNP Q7M4J4
A	224	GLY	-	linker	UNP Q7M4J4
A	225	SER	-	linker	UNP Q7M4J4

- Molecule 2 is ACETIC ACID (CCD ID: ACY) (formula: C₂H₄O₂).



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

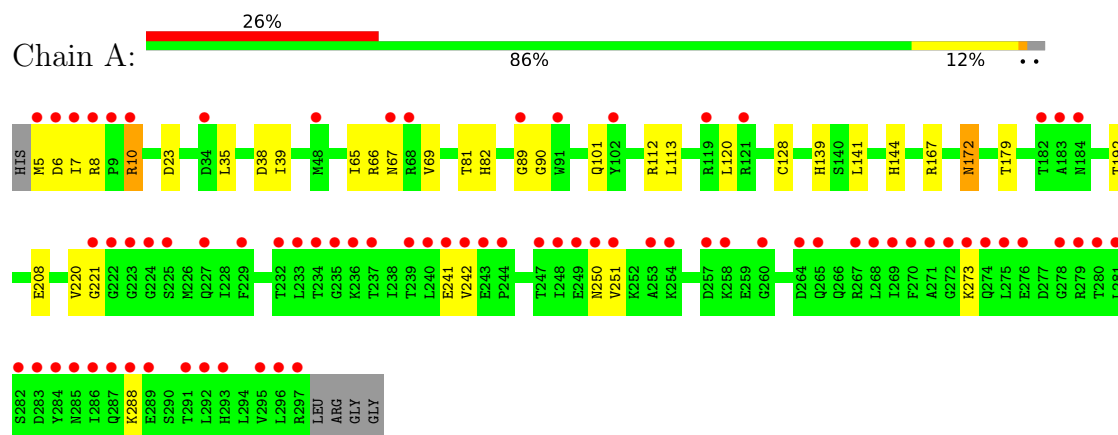
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	178	Total 178	O 178	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: RNA-directed DNA polymerase homolog (R1),Polyubiquitin-C



4 Data and refinement statistics

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, α , β , γ	141.21Å 141.21Å 37.91Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	36.23 – 1.70 36.23 – 1.70	Depositor EDS
% Data completeness (in resolution range)	95.8 (36.23-1.70) 95.8 (36.23-1.70)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.47 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.8.0230	Depositor
R, R_{free}	0.191 , 0.220 (Not available) , 0.243	Depositor DCC
R_{free} test set	2156 reflections (4.51%)	wwPDB-VP
Wilson B-factor (Å ²)	24.1	Xtriage
Anisotropy	0.005	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 31.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.035 for -h,-k,l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	2515	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 5.33% 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: ACY

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	1.08	6/2388 (0.3%)	1.07	2/3240 (0.1%)

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	A	0	4

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	139	HIS	CE1-NE2	7.89	1.40	1.32
1	A	113	LEU	N-CA	-6.01	1.39	1.46
1	A	128	CYS	N-CA	-5.67	1.39	1.46
1	A	120	LEU	N-CA	-5.36	1.40	1.46
1	A	65	ILE	N-CA	-5.09	1.40	1.46
1	A	39	ILE	C-O	5.03	1.29	1.23

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	23	ASP	CA-CB-CG	5.49	118.09	112.60
1	A	251	VAL	N-CA-C	-5.35	107.57	111.90

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	10[A]	ARG	Sidechain
1	A	10[B]	ARG	Sidechain
1	A	167	ARG	Sidechain
1	A	66	ARG	Sidechain

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	2333	0	2322	17	0
2	A	4	0	3	0	0
3	A	178	0	0	4	0
All	All	2515	0	2325	17	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 4.

All (17) 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:67:ASN:OD1	1:A:69:VAL:HG22	1.83	0.78
1:A:172:ASN:HD21	1:A:192:THR:H	1.34	0.76
1:A:144:HIS:HE1	3:A:634:HOH:O	1.76	0.69
1:A:81:THR:OG1	1:A:82:HIS:HD2	1.82	0.62
1:A:141:LEU:H	1:A:144:HIS:CD2	2.23	0.57
1:A:10[B]:ARG:HB2	1:A:220:VAL:O	2.05	0.57
1:A:172:ASN:HD22	1:A:172:ASN:H	1.52	0.56
1:A:82:HIS:HE1	1:A:101:GLN:OE1	1.95	0.50
1:A:35:LEU:HD21	1:A:208:GLU:HG3	1.95	0.49
1:A:179:THR:HG21	3:A:560:HOH:O	2.13	0.48
1:A:221:GLY:HA3	3:A:630:HOH:O	2.15	0.47
1:A:172:ASN:HD22	1:A:172:ASN:N	2.15	0.45
1:A:172:ASN:ND2	1:A:192:THR:H	2.07	0.44
1:A:179:THR:CG2	3:A:560:HOH:O	2.65	0.44
1:A:8:ARG:NH1	1:A:38:ASP:OD2	2.51	0.43
1:A:141:LEU:H	1:A:144:HIS:HD2	1.65	0.41
1:A:89[B]:GLY:HA2	1:A:90[B]:GLY:HA2	1.84	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	298/298 (100%)	289 (97%)	9 (3%)	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	
1	A	251/251 (100%)	241 (96%)	10 (4%)	28	12

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	5	MET
1	A	6	ASP
1	A	7	ILE
1	A	112	ARG
1	A	172	ASN
1	A	241	GLU
1	A	242	VAL
1	A	250	ASN
1	A	273	LYS
1	A	288	LYS

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	82	HIS
1	A	144	HIS
1	A	172	ASN
1	A	256	GLN
1	A	274	GLN

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 ⓘ

1 ligand is 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	ACY	A	401	-	3,3,3	0.87	0	3,3,3	0.56	0

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	293/298 (98%)	1.15	77 (26%) 1 1	12, 24, 44, 72	85 (29%)

All (77) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	271	ALA	8.7
1	A	7	ILE	6.8
1	A	288	LYS	6.4
1	A	282	SER	6.1
1	A	269	ILE	6.0
1	A	244	PRO	5.8
1	A	281	LEU	5.7
1	A	224	GLY	5.6
1	A	222	GLY	5.5
1	A	225	SER	5.5
1	A	273	LYS	5.4
1	A	182	THR	5.2
1	A	5	MET	5.1
1	A	183	ALA	4.8
1	A	221	GLY	4.5
1	A	9	PRO	4.4
1	A	253	ALA	4.4
1	A	270	PHE	4.4
1	A	296	LEU	4.4
1	A	272	GLY	4.4
1	A	250	ASN	4.2
1	A	297	ARG	4.2
1	A	287	GLN	4.1
1	A	233	LEU	4.1
1	A	275	LEU	4.1
1	A	239	THR	4.0
1	A	229	PHE	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	223	GLY	3.9
1	A	257	ASP	3.8
1	A	242	VAL	3.8
1	A	240	LEU	3.7
1	A	227	GLN	3.6
1	A	235	GLY	3.5
1	A	274	GLN	3.5
1	A	276	GLU	3.5
1	A	8	ARG	3.5
1	A	234	THR	3.4
1	A	260	GLY	3.4
1	A	184	ASN	3.4
1	A	247	THR	3.3
1	A	232	THR	3.2
1	A	280	THR	3.2
1	A	279	ARG	3.2
1	A	264	ASP	3.2
1	A	102	TYR	3.1
1	A	121	ARG	3.1
1	A	236	LYS	3.1
1	A	249	GLU	3.1
1	A	295	VAL	3.0
1	A	293	HIS	3.0
1	A	286	ILE	2.9
1	A	278	GLY	2.9
1	A	265	GLN	2.9
1	A	283	ASP	2.9
1	A	67	ASN	2.9
1	A	6	ASP	2.9
1	A	91[A]	TRP	2.9
1	A	119	ARG	2.8
1	A	284	TYR	2.8
1	A	291	THR	2.8
1	A	248	ILE	2.8
1	A	285	ASN	2.8
1	A	243	GLU	2.8
1	A	289	GLU	2.7
1	A	237	THR	2.6
1	A	292	LEU	2.6
1	A	241	GLU	2.6
1	A	254	LYS	2.6
1	A	89[A]	GLY	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	34	ASP	2.5
1	A	258	LYS	2.5
1	A	68	ARG	2.4
1	A	267	ARG	2.3
1	A	268	LEU	2.3
1	A	251	VAL	2.3
1	A	10[A]	ARG	2.1
1	A	48	MET	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(Å ²)	Q<0.9
2	ACY	A	401	4/4	0.97	0.06	22,23,25,27	0

6.5 Other polymers [i](#)

There are no such residues in this entry.