



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 5, 2026 – 03:20 PM UTC

PDB ID : 8EK8 / pdb_00008ek8
Title : Human PU.1 ETS-Domain (165-270) Bound to d(AATAAAAGGAGAAGGG)
Authors : Terrell, J.R.; Poon, G.M.K.
Deposited on : 2022-09-20
Resolution : 2.63 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

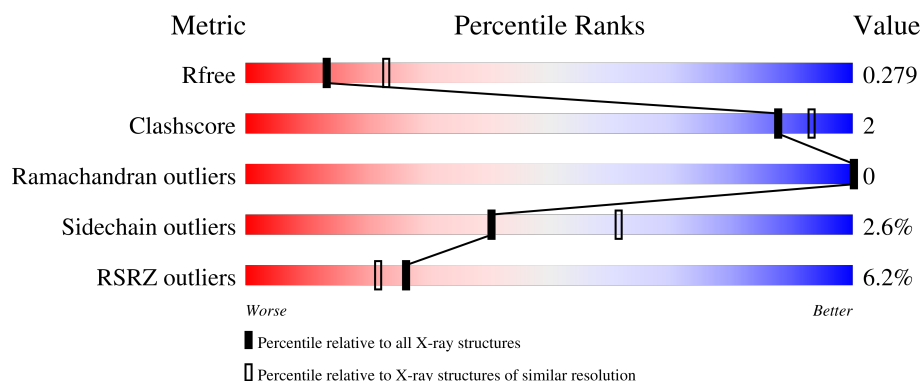
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.63 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	2053 (2.66-2.62)
Clashscore	190562	2097 (2.66-2.62)
Ramachandran outliers	187476	2066 (2.66-2.62)
Sidechain outliers	187428	2066 (2.66-2.62)
RSRZ outliers	180081	2052 (2.66-2.62)

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	16	94% 6%
1	C	16	100%
2	B	15	73% 27%
2	D	15	100%
3	E	106	7% 74% 11% 15%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	106	8% 82% • 15%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4952 atoms, of which 2227 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 DNA (5'-D(*AP*TP*AP*AP*AP*GP*GP*AP*GP*AP*AP*GP*GP*G)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	C	16	Total	C	H	N	O	P	0	0	1
			488	150	167	72	84	15			
1	A	15	Total	C	H	N	O	P	0	0	0
			484	150	167	72	81	14			

- Molecule 2 is a DNA chain called DNA (5'-D(P*CP*CP*CP*TP*TP*CP*TP*CP*CP*TP*TP*TP*AP*T)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	D	15	Total	C	H	N	O	P	0	0	0
			469	144	174	39	97	15			
2	B	15	Total	C	H	N	O	P	0	0	0
			467	144	175	39	95	14			

- Molecule 3 is a protein called Transcription factor PU.1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	F	90	Total	C	H	N	O	S	10	1	0
			1522	482	772	139	126	3			
3	E	90	Total	C	H	N	O	S	5	1	0
			1522	482	772	139	126	3			

3 Residue-property plots [i](#)

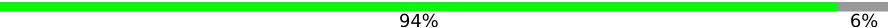
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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 (5'-D(*AP*TP*AP*AP*AP*AP*GP*GP*AP*GP*AP*AP*GP*GP*G)-3')

Chain C:  100%

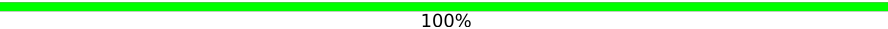
There are no outlier residues recorded for this chain.

- Molecule 1: DNA (5'-D(*AP*TP*AP*AP*AP*AP*GP*GP*AP*GP*AP*AP*GP*GP*G)-3')

Chain A:  94% 6%



- Molecule 2: DNA (5'-D(P*CP*CP*CP*TP*TP*CP*TP*CP*CP*TP*TP*TP*TP*AP*T)-3')

Chain D:  100%

There are no outlier residues recorded for this chain.

- Molecule 2: DNA (5'-D(P*CP*CP*CP*TP*TP*CP*TP*CP*CP*TP*TP*TP*TP*AP*T)-3')

Chain B:  73% 27%



- Molecule 3: Transcription factor PU.1

Chain F:  8% 82% 15%



- Molecule 3: Transcription factor PU.1



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	44.06Å 61.06Å 154.25Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.81 – 2.63 34.81 – 2.63	Depositor EDS
% Data completeness (in resolution range)	98.8 (34.81-2.63) 98.7 (34.81-2.63)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.62 (at 2.65Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.237 , 0.280 0.236 , 0.279	Depositor DCC
R_{free} test set	1297 reflections (9.98%)	wwPDB-VP
Wilson B-factor (Å ²)	63.8	Xtriage
Anisotropy	0.778	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 41.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	4952	wwPDB-VP
Average B, all atoms (Å ²)	87.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.75% 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 [i](#)

5.1 Standard geometry [i](#)

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	0.17	0/360	0.38	0/556
1	C	0.18	0/364	0.38	0/563
2	B	0.23	0/322	0.51	0/493
2	D	0.23	0/325	0.53	0/497
3	E	0.11	0/768	0.24	0/1021
3	F	0.08	0/768	0.24	0/1021
All	All	0.15	0/2907	0.36	0/4151

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

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	A	317	167	168	0	0
1	C	321	167	167	0	0
2	B	292	175	175	2	0
2	D	295	174	174	0	0
3	E	750	772	774	6	0
3	F	750	772	774	1	0
All	All	2725	2227	2232	9	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 2.

All (9) 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:E:170:ILE:HG21	3:E:216:GLN:NE2	2.20	0.57
3:E:243:LYS:NZ	3:E:244:VAL:O	2.39	0.55
3:E:170:ILE:HG21	3:E:216:GLN:HE21	1.81	0.43
3:E:224:THR:HG22	3:E:225:TYR:N	2.33	0.43
2:B:29:DT:H2'	2:B:30:DT:H72	2.00	0.43
3:E:239:GLY:O	3:E:253:SER:HB3	2.18	0.43
3:E:253:SER:HB2	3:E:255:GLU:OE1	2.18	0.42
3:F:185:MET:HE1	3:F:209:LEU:HA	2.01	0.42
2:B:20:DC:H4'	2:B:21:DC:OP1	2.19	0.42

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	E	89/106 (84%)	87 (98%)	2 (2%)	0	100	100
3	F	89/106 (84%)	87 (98%)	2 (2%)	0	100	100
All	All	178/212 (84%)	174 (98%)	4 (2%)	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	E	77/90 (86%)	74 (96%)	3 (4%)	28	47
3	F	77/90 (86%)	76 (99%)	1 (1%)	61	76
All	All	154/180 (86%)	150 (97%)	4 (3%)	40	61

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	F	222	LYS
3	E	186	LYS
3	E	222	LYS
3	E	226	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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

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	15/16 (93%)	0.46	0	100	100	73, 94, 105, 109	0
1	C	16/16 (100%)	0.36	0	100	100	68, 95, 110, 131	0
2	B	15/15 (100%)	0.30	0	100	100	73, 84, 101, 112	0
2	D	15/15 (100%)	0.46	0	100	100	69, 80, 109, 113	0
3	E	90/106 (84%)	0.86	7 (7%)	19	15	55, 93, 117, 128	3 (3%)
3	F	90/106 (84%)	0.75	8 (8%)	15	12	45, 75, 96, 109	4 (4%)
All	All	241/274 (87%)	0.70	15 (6%)	26	22	45, 85, 112, 131	7 (2%)

All (15) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	F	258	GLY	5.1
3	F	225	TYR	3.6
3	E	169	LYS	3.5
3	E	220	ARG	3.2
3	E	258	GLY	3.1
3	E	186	LYS	2.8
3	F	169	LYS	2.5
3	F	246	LYS	2.3
3	F	221	LYS	2.2
3	F	194	LYS	2.2
3	F	219	ASN	2.1
3	F	241	VAL	2.1
3	E	241	VAL	2.1
3	E	232	LEU	2.0
3	E	209	LEU	2.0

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.