wwPDB X-ray Structure Validation Summary Report (i)

Mar 8, 2026 – 10:13 AM UTC

PDB ID	:	7CGK / pdb_00007cgk
Title	:	Crystal Structure of PUF-8 in Complex with PBE-RNA
Authors	:	Zheng, X.; Yunyu, S.; Shouhong, G.
Deposited on	:	2020-07-01
Resolution	:	2.40 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 (1)) were used in the production of this report:

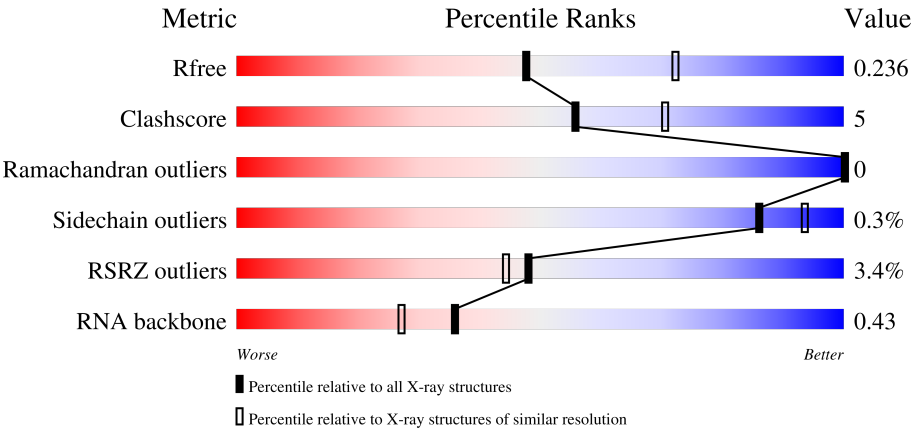
MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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	4912 (2.40-2.40)
Clashscore	190562	5391 (2.40-2.40)
Ramachandran outliers	187476	5320 (2.40-2.40)
Sidechain outliers	187428	5321 (2.40-2.40)
RSRZ outliers	180081	4916 (2.40-2.40)
RNA backbone	3983	1155 (2.70-2.10)

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% 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	360	3%, 87%, 9%, .
2	B	8	62%, 25%, 12%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2997 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 PUM-HD domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	346	Total	C	N	O	S	0	1	0
			2768	1756	488	506	18			

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	163	MET	-	initiating methionine	UNP Q09487
A	164	GLY	-	expression tag	UNP Q09487
A	165	HIS	-	expression tag	UNP Q09487
A	166	HIS	-	expression tag	UNP Q09487
A	167	HIS	-	expression tag	UNP Q09487
A	168	HIS	-	expression tag	UNP Q09487
A	169	HIS	-	expression tag	UNP Q09487
A	170	HIS	-	expression tag	UNP Q09487
A	171	MET	-	expression tag	UNP Q09487
A	318	CYS	ASN	engineered mutation	UNP Q09487
A	319	ARG	HIS	engineered mutation	UNP Q09487

- Molecule 2 is a RNA chain called PBE-5A (5'-R(P*UP*GP*UP*AP*AP*AP*UP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	8	Total	C	N	O	P	0	0	0
			171	77	31	55	8			

- Molecule 3 is CHLORIDE ION (CCD ID: CL) (formula: Cl) (labeled as "Ligand of Interest" by depositor).

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

- Molecule 4 is water.

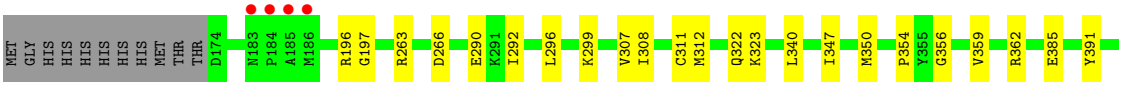
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	51	Total O 51 51	0	0
4	B	6	Total O 6 6	0	0

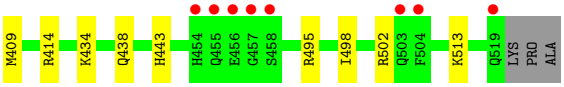
3 Residue-property plots (i)

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: PUM-HD domain-containing protein

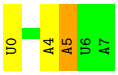
Chain A:





• Molecule 2: PBE-5A (5'-R(P*UP*GP*UP*AP*AP*AP*UP*A)-3')

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	89.80Å 89.80Å 117.55Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.00 – 2.40 38.00 – 2.40	Depositor EDS
% Data completeness (in resolution range)	100.0 (38.00-2.40) 99.9 (38.00-2.40)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.98 (at 2.39Å)	Xtriage
Refinement program	PHENIX 1.10-2155	Depositor
R, R_{free}	0.187 , 0.235 0.190 , 0.236	Depositor DCC
R_{free} test set	982 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	32.5	Xtriage
Anisotropy	0.288	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 27.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	2997	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.20% 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.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CL

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	0.36	0/2826	0.55	0/3822
2	B	0.43	0/191	0.70	0/295
All	All	0.37	0/3017	0.56	0/4117

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

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	2768	0	2711	25	0
2	B	171	0	86	3	0
3	A	1	0	0	0	0
4	A	51	0	0	1	0
4	B	6	0	0	0	0
All	All	2997	0	2797	27	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 5.

The worst 5 of 27 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:350:MET:HE3	1:A:356:GLY:HA3	1.57	0.86
1:A:312:MET:HE2	1:A:350:MET:HE2	1.64	0.77
1:A:290:GLU:OE1	1:A:323:LYS:NZ	2.21	0.61
1:A:322:GLN:HG2	1:A:362:ARG:HG3	1.89	0.55
2:B:0:U:P	2:B:0:U:H6	2.32	0.53

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	345/360 (96%)	336 (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	299/326 (92%)	298 (100%)	1 (0%)	86 93

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	347	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	206	GLN
1	A	278	GLN
1	A	430	ASN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	7/8 (87%)	2 (28%)	0

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	4	A
2	B	5	A

There are no RNA pucker outliers to report.

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)

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 (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, 95^{th} 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	346/360 (96%)	-0.14	12 (3%) 47 43	18, 34, 67, 90	1 (0%)
2	B	8/8 (100%)	-0.17	0 100 100	32, 37, 45, 57	0
All	All	354/368 (96%)	-0.14	12 (3%) 48 44	18, 34, 67, 90	1 (0%)

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	456	GLU	4.3
1	A	455	GLN	3.6
1	A	454	HIS	3.6
1	A	183	ASN	3.1
1	A	504	PHE	3.1

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

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

6.3 Carbohydrates

There are no oligosaccharides in this entry.

6.4 Ligands

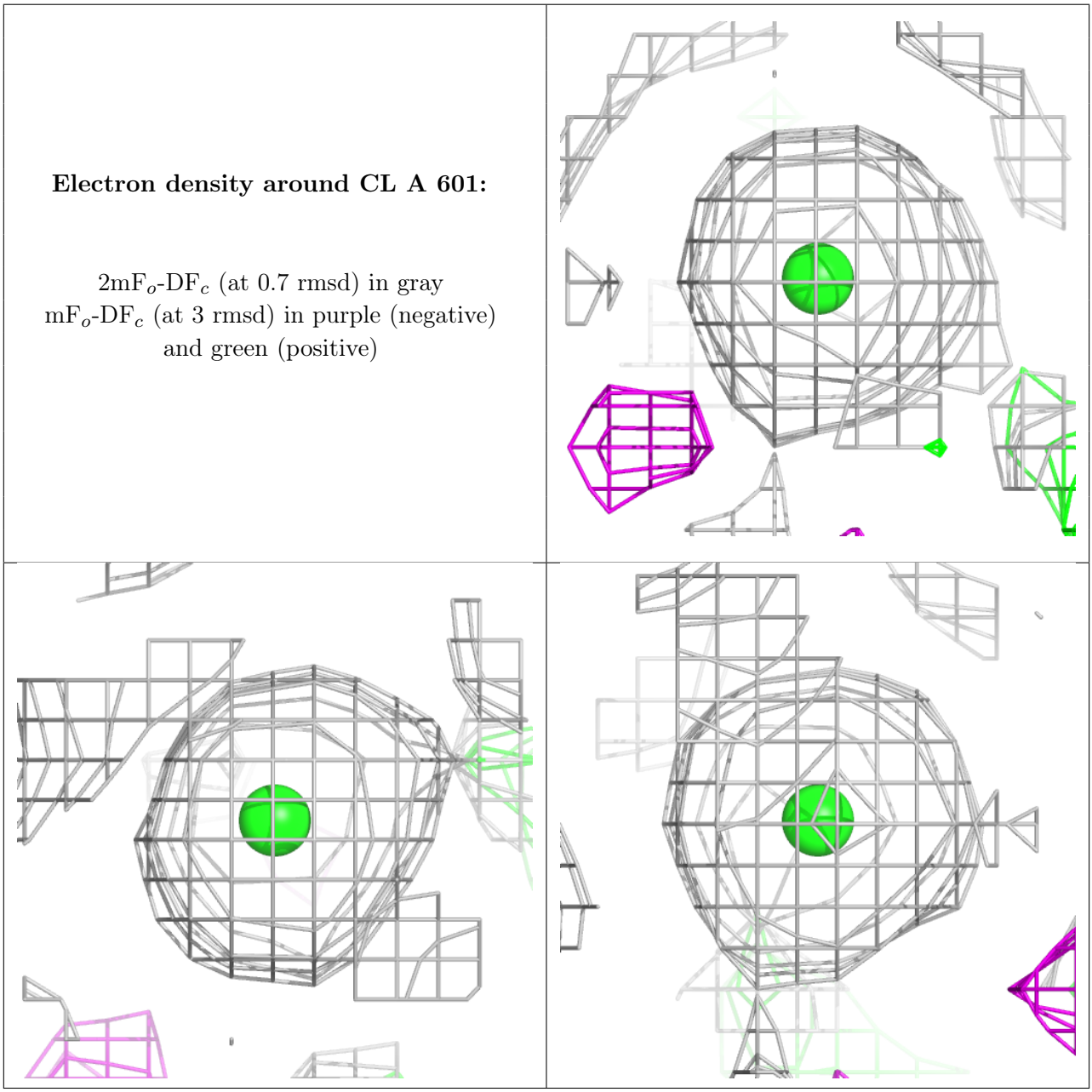
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
3	CL	A	601	1/1	0.98	0.04	33,33,33,33	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around CL A 601:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



6.5 Other polymers (i)

There are no such residues in this entry.