



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 23, 2024 – 04:17 PM JST

PDB ID : 8Y0B
Title : Crystal structure of FnCas12a in complex with pre-crRNA and 12nt target DNA
Authors : Lin, X.; Chen, J.; Liu, L.
Deposited on : 2024-01-22
Resolution : 2.30 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 ⓘ 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 : 4.02b-467
Xtriage (Phenix) : 1.21
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.004 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.40

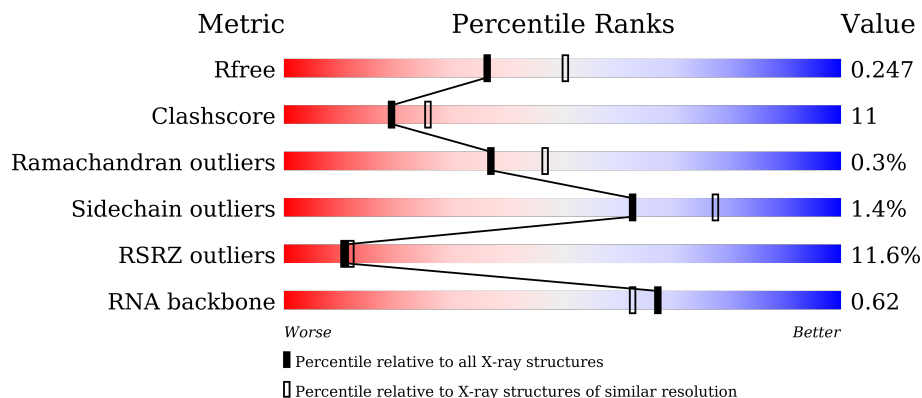
1 Overall quality at a glance

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}	164625	5963 (2.30-2.30)
Clashscore	180529	6698 (2.30-2.30)
Ramachandran outliers	177936	6640 (2.30-2.30)
Sidechain outliers	177891	6640 (2.30-2.30)
RSRZ outliers	164620	5963 (2.30-2.30)
RNA backbone	3690	1032 (2.60-2.00)

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	B	58	 9% 60% 22% 7% • 9%
2	C	21	 81% 19%
3	D	18	 28% 44% 28%
4	A	1300	 12% 78% 19% ••

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 12785 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 RNA chain called RNA (53-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	B	53	1120	504	194	370	52	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	C	21	430	206	79	124	21	0	0	0

- Molecule 3 is a DNA chain called DNA (5'-D(P*AP*GP*TP*CP*CP*TP*TP*TP*AP*GP*AP*TP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	D	13	266	128	46	79	13	0	0	0

- Molecule 4 is a protein called CRISPR-associated endonuclease Cas12a.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	A	1264	10399	6679	1714	1984	22	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1006	ALA	GLU	conflict	UNP A0Q7Q2

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total Mg 1 1	0	0
5	A	1	Total Mg 1 1	0	0

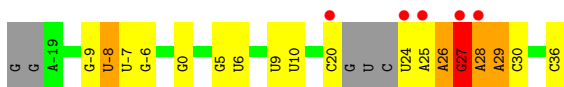
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	91	Total O 91 91	0	0
6	C	44	Total O 44 44	0	0
6	D	16	Total O 16 16	0	0
6	A	417	Total O 417 417	0	0

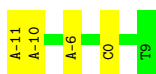
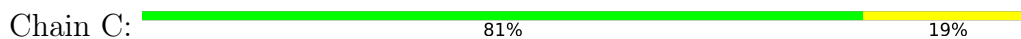
3 Residue-property plots [i](#)

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 (53-MER)



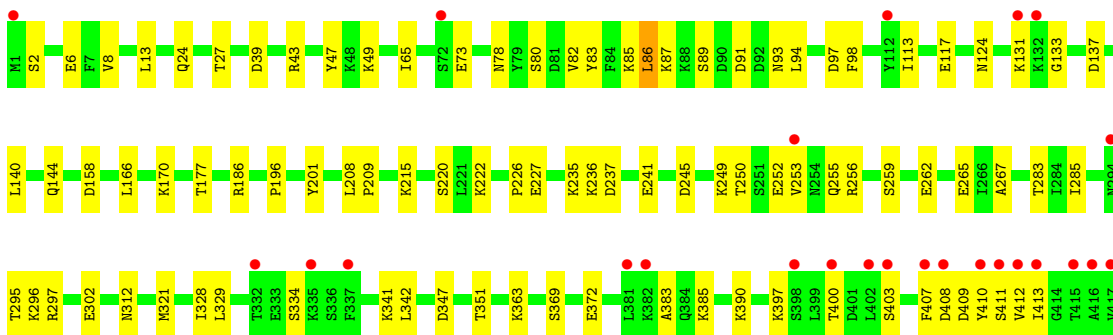
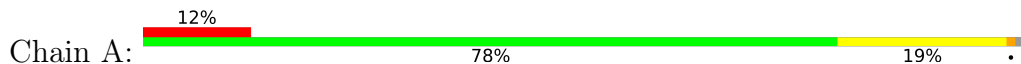
- Molecule 2: DNA (5'-D(P*AP*AP*AP*TP*GP*AP*CP*TP*TP*CP*TP*CP*TP*AP*AP*A P*GP*GP*AP*CP*T)-3')

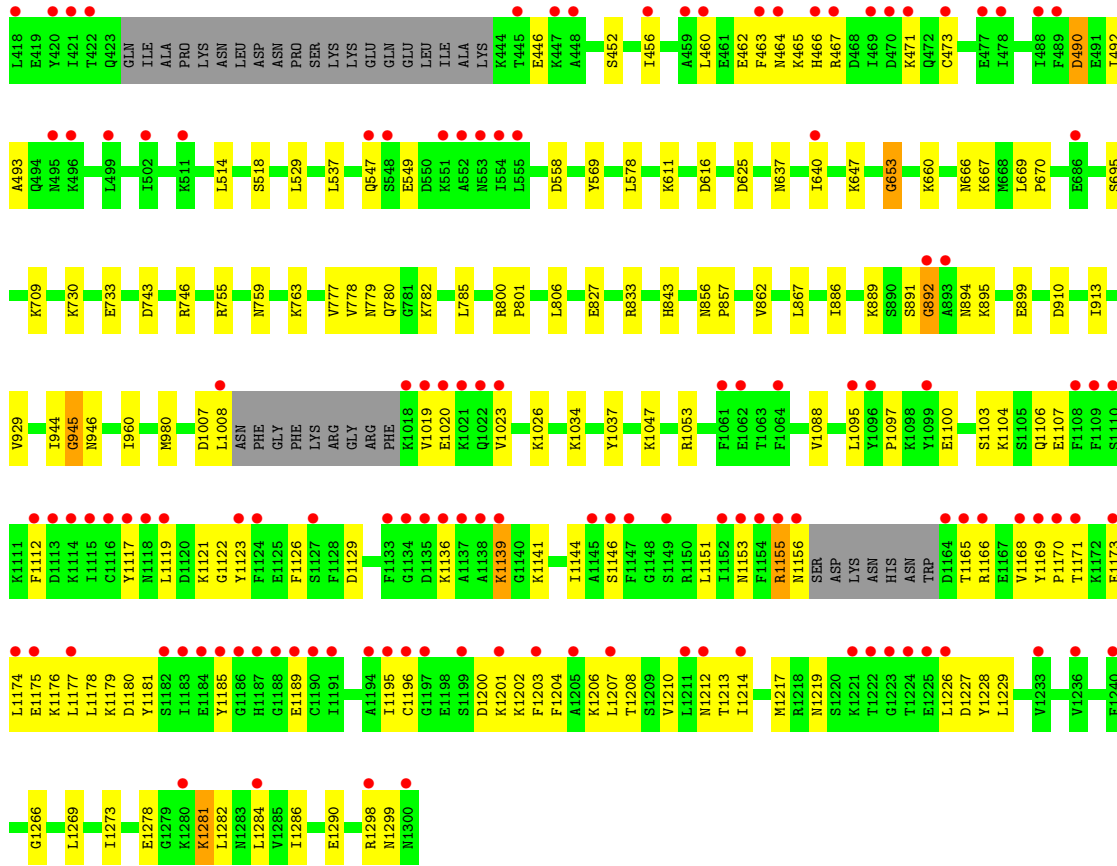


- Molecule 3: DNA (5'-D(P*AP*GP*TP*CP*CP*TP*TP*TP*AP*GP*AP*TP*A)-3')



- Molecule 4: CRISPR-associated endonuclease Cas12a





4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	124.73Å 124.73Å 267.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.21 – 2.30 49.21 – 2.30	Depositor EDS
% Data completeness (in resolution range)	97.1 (49.21-2.30) 97.1 (49.21-2.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.26 (at 2.29Å)	Xtrriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.203 , 0.247 0.204 , 0.247	Depositor DCC
R_{free} test set	4701 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	39.9	Xtrriage
Anisotropy	0.060	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 32.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	12785	wwPDB-VP
Average B, all atoms (Å ²)	53.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.70% 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: MG

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	B	0.73	1/1251 (0.1%)	1.35	15/1943 (0.8%)
2	C	1.02	0/482	1.07	0/741
3	D	1.06	0/297	1.16	1/456 (0.2%)
4	A	0.45	0/10604	0.60	5/14242 (0.0%)
All	All	0.54	1/12634 (0.0%)	0.76	21/17382 (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
4	A	0	3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	27	G	N9-C4	5.52	1.42	1.38

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	-8	U	C5-C4-O4	10.19	132.02	125.90
1	B	-8	U	C5-C6-N1	-10.04	117.68	122.70
1	B	27	G	N3-C4-C5	-9.92	123.64	128.60
1	B	-8	U	N3-C2-O2	-9.45	115.58	122.20
1	B	27	G	C4-N9-C1'	9.43	138.76	126.50

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	A	549	GLU	Peptide
4	A	86	LEU	Peptide
4	A	944	ILE	Peptide

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	B	1120	0	567	22	0
2	C	430	0	238	4	0
3	D	266	0	149	16	0
4	A	10399	0	10328	240	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
6	A	417	0	0	80	0
6	B	91	0	0	1	0
6	C	44	0	0	1	0
6	D	16	0	0	8	0
All	All	12785	0	11282	266	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 11.

The worst 5 of 266 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)
4:A:1117:TYR:OH	4:A:1119:LEU:HD13	1.34	1.22
4:A:1282:LEU:CA	6:A:1502:HOH:O	1.83	1.21
4:A:1281:LYS:HA	6:A:1520:HOH:O	1.39	1.17
3:D:-5:DC:C5'	6:D:101:HOH:O	1.93	1.15
4:A:1282:LEU:HA	6:A:1502:HOH:O	1.40	1.12

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
4	A	1256/1300 (97%)	1216 (97%)	36 (3%)	4 (0%)	37 47

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	411	SER
4	A	653	GLY
4	A	1019	VAL
4	A	892	GLY

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
4	A	1141/1183 (96%)	1125 (99%)	16 (1%)	62 77

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	A	1229	LEU
4	A	1155	ARG
4	A	666	ASN
4	A	1139	LYS
4	A	647	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
4	A	1219	ASN
4	A	1299	ASN
4	A	532	GLN
4	A	894	ASN
4	A	1118	ASN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	B	51/58 (87%)	8 (15%)	1 (1%)

5 of 8 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	B	-9	G
1	B	-7	U
1	B	-6	G
1	B	25	A
1	B	26	A

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	B	27	G

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 2 ligands modelled in this entry, 2 are 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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	B	53/58 (91%)	-0.19	5 (9%) 15 17	26, 35, 109, 138	0
2	C	21/21 (100%)	-0.58	0 100 100	28, 36, 65, 85	0
3	D	13/18 (72%)	-0.21	0 100 100	37, 40, 78, 94	0
4	A	1264/1300 (97%)	0.53	152 (12%) 10 11	25, 49, 104, 124	0
All	All	1351/1397 (96%)	0.48	157 (11%) 11 12	25, 48, 104, 138	0

The worst 5 of 157 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	A	499	LEU	5.4
4	A	469	ILE	5.3
4	A	552	ALA	5.3
4	A	1154	PHE	5.0
4	A	554	ILE	4.7

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 monosaccharides 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(\AA^2)	Q<0.9
5	MG	B	101	1/1	0.74	0.13	32,32,32,32	0
5	MG	A	1401	1/1	0.82	0.10	30,30,30,30	0

6.5 Other polymers [i](#)

There are no such residues in this entry.