WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 10, 2023 – 06:58 AM EDT

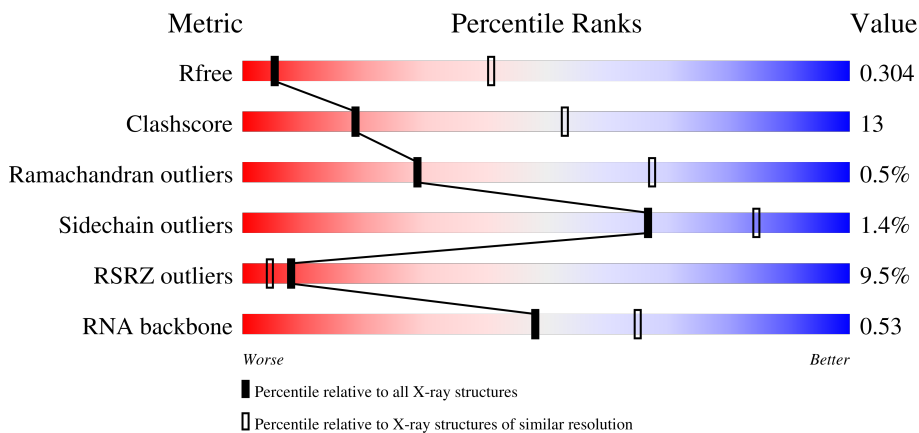|		
PDB ID	:	6WC0
Title	:	Crystal structure of AceCas9 bound with guide RNA and DNA with 5'-NNNTC-3' PAM
Authors	:	Li, H.; Das, A.
Deposited on	:	2020-03-28
Resolution	:	3.61 Å(reported)

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 3.61 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	Percentile Ranks	Value
Rfree		0.304
Clashscore		13
Ramachandran outliers		0.5%
Sidechain outliers		1.4%
RSRZ outliers		9.5%
RNA backbone		0.53

Worse → *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1290 (3.74-3.50)
Clashscore	141614	1387 (3.74-3.50)
Ramachandran outliers	138981	1339 (3.74-3.50)
Sidechain outliers	138945	1339 (3.74-3.50)
RSRZ outliers	127900	1191 (3.74-3.50)
RNA backbone	3102	1018 (4.22-3.00)

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 <=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	983	10% — 68% / 26% / · 5%
2	B	94	.% — 37% / 43% / 16% / · ·
3	C	30	33% / 67%
4	D	10	50% / 50%

WORLDWIDE PDB PROTEIN DATA BANK

2   Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 10131 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 CRISPR-associated endonuclease, Csn1 family.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	938	Total	C	N	O	S	0	0	0
			7371	4633	1373	1352	13			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	518	GLY	-	linker	UNP A0LWB3
A	680	GLY	-	linker	UNP A0LWB3
A	681	GLY	-	linker	UNP A0LWB3
A	682	SER	-	linker	UNP A0LWB3
A	683	ALA	-	linker	UNP A0LWB3
A	684	GLY	-	linker	UNP A0LWB3
A	1139	HIS	-	expression tag	UNP A0LWB3
A	1140	HIS	-	expression tag	UNP A0LWB3
A	1141	HIS	-	expression tag	UNP A0LWB3
A	1142	HIS	-	expression tag	UNP A0LWB3
A	1143	HIS	-	expression tag	UNP A0LWB3
A	1144	HIS	-	expression tag	UNP A0LWB3

- Molecule 2 is a RNA chain called sgRNA (95-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	91	Total	C	N	O	P	9	0	0
			1952	866	348	645	93			

- Molecule 3 is a DNA chain called DNA (30-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	30	Total	C	N	O	P	0	0	0
			605	290	109	177	29			

- Molecule 4 is a DNA chain called DNA (5'-D(*AP*TP*AP*CP*TP*TP*GP*GP*CP*G)-3

').

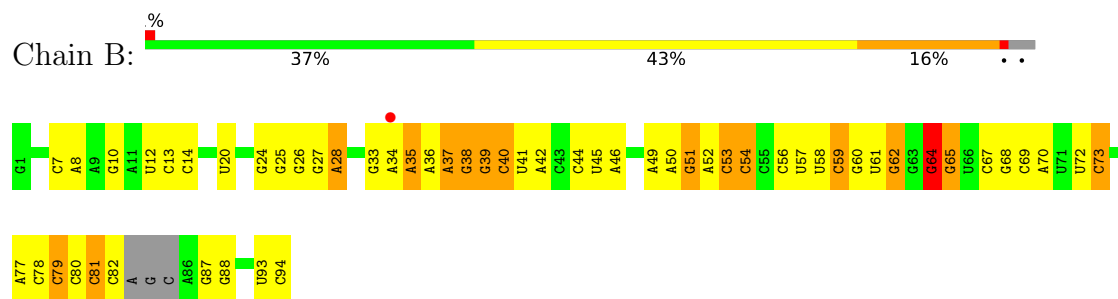| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
4	D	10	203	98	37	59	9	0	0	0

3 Residue-property plots ⓘ

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: CRISPR-associated endonuclease, Csn1 family

• Molecule 2: sgRNA (95-MER)

Chain B:

%

37%   43%   16%

G1 C7 A8 A9 G10 A11 U12 C13 C14 U20 G24 G25 G26 G27 A28 G33 A34 A35 A36 A37 G38 G39 C40 U41 A42 C43 C44 U45 A46 A49 A50 G51 A52 C53 C54 C55 C56 U57 U58 C59 G60 U61 G62 G63 C64 G65 U66 C67 G68 C69 A70 U71 U72 C73

A77 C78 C79 C80 C81 C82 A G C A86 G87 G88 U93 C94

• Molecule 3: DNA (30-MER)

Chain C:

33%   67%

C-10 G-9 C-8 C-7 A-6 G-5 A-4 T-3 A-2 C5 A6 G7 G8 A9 T10 C11 T12 T13 G14 C15 C16 A17 T18 C19 C20
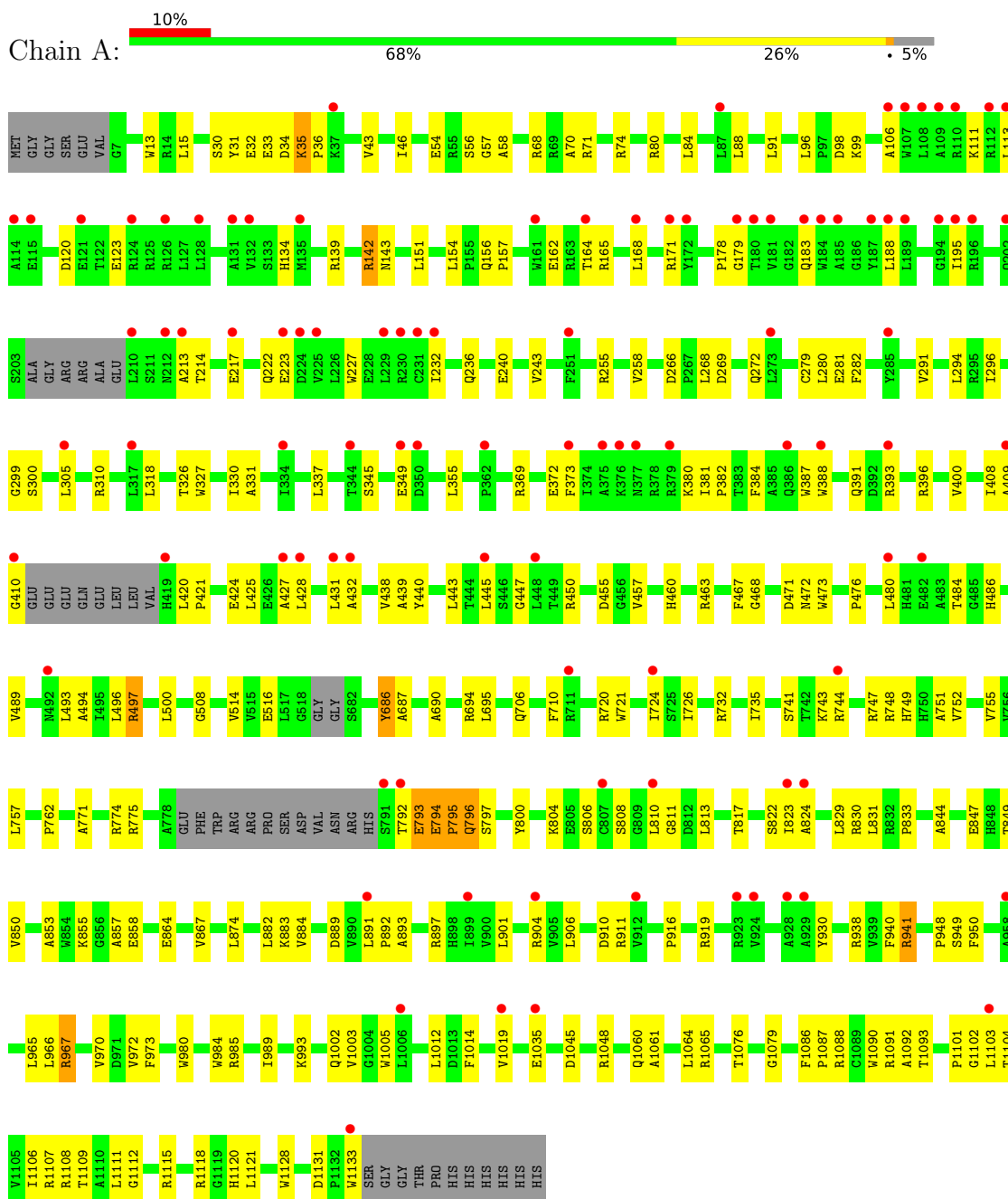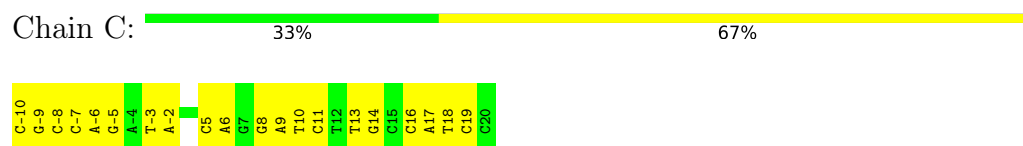
• Molecule 4: DNA (5'-D(*AP*TP*AP*CP*TP*TP*GP*GP*CP*G)-3')

Chain D:

50%   50%

A1 G7 G8 C9 G10

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	81.64Å 111.14Å 177.62Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	94.22 – 3.61 94.22 – 3.61	Depositor EDS
% Data completeness (in resolution range)	99.0 (94.22-3.61) 82.2 (94.22-3.61)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.61 (at 3.58Å)	Xtriage
Refinement program	PHENIX 1.18rc4_3812	Depositor
R, R_{free}	0.237 , 0.304 0.237 , 0.304	Depositor DCC
R_{free} test set	1903 reflections (9.99%)	wwPDB-VP
Wilson B-factor (Å²)	140.0	Xtriage
Anisotropy	0.303	Xtriage
Bulk solvent k_{sol}(e/Å³), B_{sol}(Å²)	0.30 , 125.9	EDS
L-test for twinning[2]	$< |L| > = 0.46$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	10131	wwPDB-VP
Average B, all atoms (Å²)	192.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.12% 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: GTP

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	0.27	0/7538	0.53	0/10245
2	B	0.27	0/2144	0.86	1/3338 (0.0%)
3	C	0.60	0/677	0.96	0/1041
4	D	0.53	0/227	0.98	0/349
All	All	0.31	0/10586	0.67	1/14973 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	64	G	P-O3'-C3'	5.14	125.87	119.70

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	7371	0	7343	195	0
2	B	1952	0	984	40	0
3	C	605	0	339	18	0
4	D	203	0	115	6	0
All	All	10131	0	8781	238	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 13.

The worst 5 of 238 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:43:VAL:HG11	1:A:1112:GLY:HA3	1.43	1.00
1:A:817:THR:HG23	1:A:822:SER:HB2	1.61	0.80
1:A:724:ILE:HD11	1:A:811:GLY:HA2	1.63	0.80
4:D:9:DC:H2'	4:D:10:DG:C8	2.20	0.77
1:A:486:HIS:HD2	1:A:489:VAL:HG23	1.51	0.76

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	928/983 (94%)	870 (94%)	53 (6%)	5 (0%)	29 67

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	57	GLY
1	A	794	GLU
1	A	795	PRO
1	A	796	GLN
1	A	468	GLY

5.3.2 Protein sidechains ⓘ

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	770/806 (96%)	759 (99%)	11 (1%)	67 84

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	904	ARG
1	A	941	ARG
1	A	1131	ASP
1	A	967	ARG
1	A	686	TYR

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	472	ASN
1	A	486	HIS
1	A	706	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	88/94 (93%)	26 (29%)	2 (2%)

5 of 26 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	20	U
2	B	28	A
2	B	35	A
2	B	36	A
2	B	37	A

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	B	36	A
2	B	64	G

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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(Å²)	Q<0.9
1	A	938/983 (95%)	0.54	100 (10%)	6	3	118, 180, 259, 323	0
2	B	90/94 (95%)	-0.23	1 (1%)	80	68	122, 186, 333, 354	0
3	C	30/30 (100%)	-0.77	0	100	100	147, 180, 270, 284	0
4	D	10/10 (100%)	-0.69	0	100	100	179, 216, 306, 315	0
All	All	1068/1117 (95%)	0.43	101 (9%)	8	4	118, 181, 269, 354	0

The worst 5 of 101 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	110	ARG	10.8
1	A	195	ILE	9.6
1	A	791	SER	7.9
1	A	350	ASP	7.2
1	A	107	TRP	6.7

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 monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.