



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 26, 2026 – 03:41 AM EDT

PDB ID : 8Y0D / pdb_00008y0d
Title : Crystal structure of SauCas9 in complex with sgRNA and 20nt ssDNA target
Authors : Chen, Y.; Chen, J.; Liu, L.
Deposited on : 2024-01-22
Resolution : 3.92 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 : 4-5-2 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

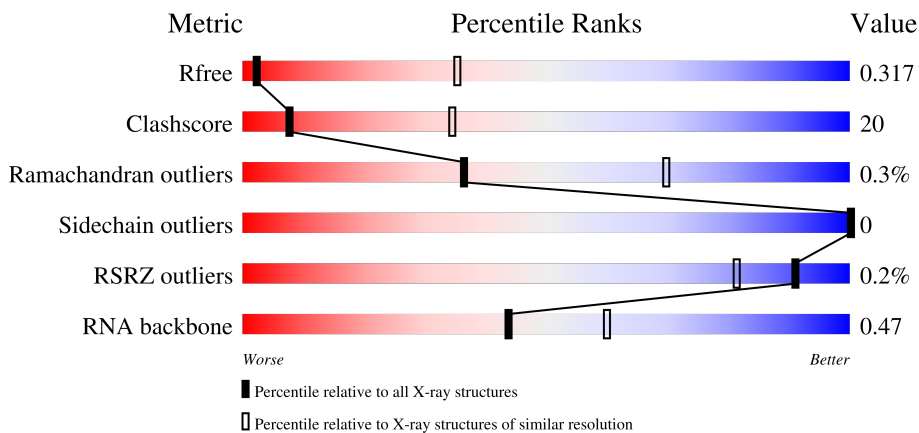
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.92 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	1033 (4.10-3.74)
Clashscore	190562	1070 (4.10-3.74)
Ramachandran outliers	187476	1017 (4.10-3.74)
Sidechain outliers	187428	1010 (4.10-3.74)
RSRZ outliers	180081	1033 (4.10-3.74)
RNA backbone	3983	1014 (4.70-3.10)

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 $\leq 5\%$. 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	74	
2	C	20	
3	A	1053	

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 10023 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 (73-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	73	Total	C	N	O	P	0	0	0
			1530	684	265	508	73			

- Molecule 2 is a DNA chain called DNA (5'-D(P*GP*TP*AP*AP*AP*GP*TP*TP*AP*AP*AP*TP*AP*GP*CP*AP*GP*AP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	19	Total	C	N	O	P	0	0	0
			395	188	79	109	19			

- Molecule 3 is a protein called CRISPR-associated endonuclease Cas9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	1009	Total	C	N	O	S	0	0	0
			8093	5154	1376	1550	13			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	580	ALA	ASN	conflict	UNP J7RUA5
A	946	ALA	CYS	conflict	UNP J7RUA5

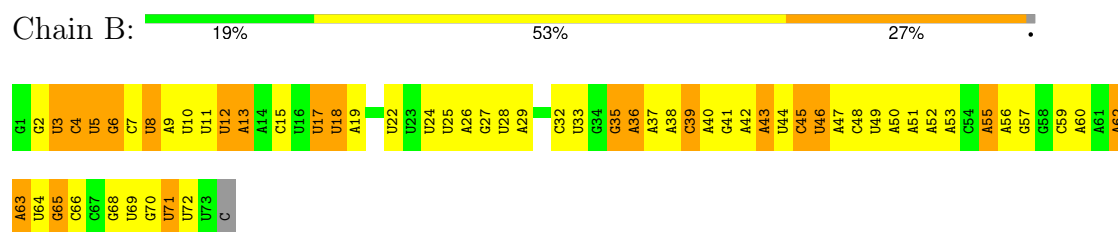
- Molecule 4 is water.

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

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.

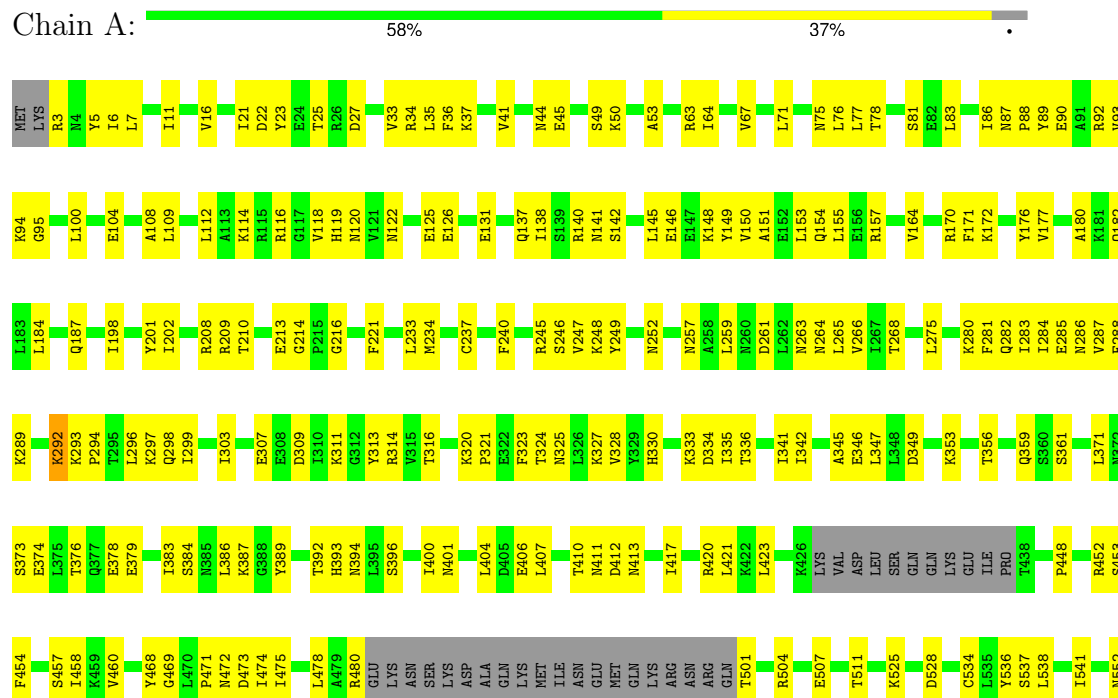
• Molecule 1: RNA (73-MER)



• Molecule 2: DNA (5'-D(P*GP*TP*AP*AP*AP*GP*TP*TP*AP*AP*AP*TP*AP*GP*CP*A P*GP*AP*C)-3')



• Molecule 3: CRISPR-associated endonuclease Cas9



I1034	I1035	G1036	N1037	Q1048	I1049	I1050	K1051	K1052	GLY	I934	K935	K936	E937	N938	Y939	Y940	E941	V942	K954	K955	I956	F962	I963	A964	S965	L971	I972	K973	I974	R980	N986	D987	L988	L989	V994	N995	I999	T1000	E1003	Y1004	L1005	E1006	N1007	M1008	N1009	D1010	K1011	R1012	I1016	I1017	K1018	S1022	I1027	K1028	K1029	Y1030	D1033	Y553	H557	I558	I559	P560	R561	F565	D566	N567	S568	V573	L574	V575	E579	K583	R586	Y591	I599	S600	Y601	E602	K605	S619	K622	Y625	R630	D631	I632	F635	S636	V637	Q638	K639	I642	N643	R644	N645	L646	V647	D648	THR	ARG	TYR	A652	T653	M657	R661	S662	Y663	F664	R665	V671	K674	S675	I676	M677	G678	G679	F680	F683	L684	R685	R686	K687	W688	K691	R694	N695	K699	H700	H701	A702	E703	D704	I707	I708	A709	D712	F715	W718	K719	K720	M728	E729	F733	GLU	GLU	LYS	GLN	ALA	A652	GLU	SER	R741	P742	E749	I753	I760	I763	F766	K767	D768	Y771	S772	H773	R774	R781	E782	L783	T791	R792	K793	D794	N798	T799	L800	I801	V802	N803	D814	R815	L816	K817	L827	L828	R829	Y830	D833	P834	Q835	T836	Y837	Q838	K839	L840	I843	M844	E845	Q846	Y847	R851	N852	P853	L854	Y855	K856	Y857	Y858	E859	E860	T861	S862	M863	Y864	L865	T866	K867	Y868	K879	I880	K881	Y882	Y883	G884	L887	H890	L891	D892	I893	T894	Y897	S900	K903	L907	S908	P911	Y912	D915	V916	Y917	L918	V922	Y923	K924	R929	V933
-------	-------	-------	-------	-------	-------	-------	-------	-------	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	------	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	68.96Å 145.57Å 97.53Å 90.00° 93.99° 90.00°	Depositor
Resolution (Å)	48.65 – 3.92 48.65 – 3.92	Depositor EDS
% Data completeness (in resolution range)	77.0 (48.65-3.92) 74.5 (48.65-3.92)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.86 (at 3.88Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.286 , 0.318 0.286 , 0.317	Depositor DCC
R_{free} test set	1325 reflections (7.67%)	wwPDB-VP
Wilson B-factor (Å ²)	-1.3	Xtriage
Anisotropy	6.044	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.10 , 0.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.38$, $\langle L^2 \rangle = 0.21$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.86	EDS
Total number of atoms	10023	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.57% 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	B	0.50	9/1709 (0.5%)	0.43	0/2657
2	C	0.49	0/445	0.50	0/685
3	A	0.21	0/8231	0.45	0/11115
All	All	0.30	9/10385 (0.1%)	0.45	0/14457

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	12	U	C1'-N1	6.88	1.58	1.48
1	B	5	U	C1'-N1	6.59	1.58	1.48
1	B	18	U	C1'-N1	6.57	1.58	1.48
1	B	6	G	C1'-N9	-6.44	1.38	1.48
1	B	8	U	C1'-N1	6.35	1.57	1.48

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	B	1530	0	769	73	0
2	C	395	0	214	26	0
3	A	8093	0	7933	317	0
4	A	4	0	0	0	0
4	B	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	10023	0	8916	376	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 20.

The worst 5 of 376 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:A:907:LEU:HD13	3:A:908:SER:H	1.34	0.92
3:A:284:ILE:HA	3:A:288:PHE:HB2	1.51	0.90
1:B:45:C:OP2	3:A:792:ARG:NH1	2.07	0.88
3:A:911:PRO:HA	3:A:929:LYS:HA	1.57	0.87
3:A:44:ASN:CB	3:A:44:ASN:C	2.49	0.85

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	A	998/1053 (95%)	938 (94%)	57 (6%)	3 (0%)	36 70

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	292	LYS
3	A	908	SER
3	A	964	ALA

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	A	864/970 (89%)	864 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
3	A	607	HIS
3	A	645	ASN
3	A	1037	ASN
3	A	838	GLN
3	A	264	ASN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	B	71/74 (95%)	13 (18%)	2 (2%)

5 of 13 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	B	17	U
1	B	36	A
1	B	39	C
1	B	43	A
1	B	45	C

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	B	35	G
1	B	62	A

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 [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	73/74 (98%)	-1.36	0	100 100	19, 39, 90, 115	0
2	C	19/20 (95%)	-0.84	0	100 100	21, 34, 50, 52	0
3	A	1009/1053 (95%)	-1.00	2 (0%)	91 81	16, 39, 60, 149	0
All	All	1101/1147 (95%)	-1.02	2 (0%)	91 81	16, 39, 62, 149	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	937	GLU	3.3
3	A	691	LYS	2.2

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.