



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 18, 2022 – 01:49 pm BST

PDB ID : 7QR8
Title : SpCas9 bound to PTPRC off-target1 DNA substrate
Authors : Pacesa, M.; Jinek, M.
Deposited on : 2022-01-10
Resolution : 2.75 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) : 1.13
EDS : 2.31.2
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0267
CCP4 : 7.1.010 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.31.2

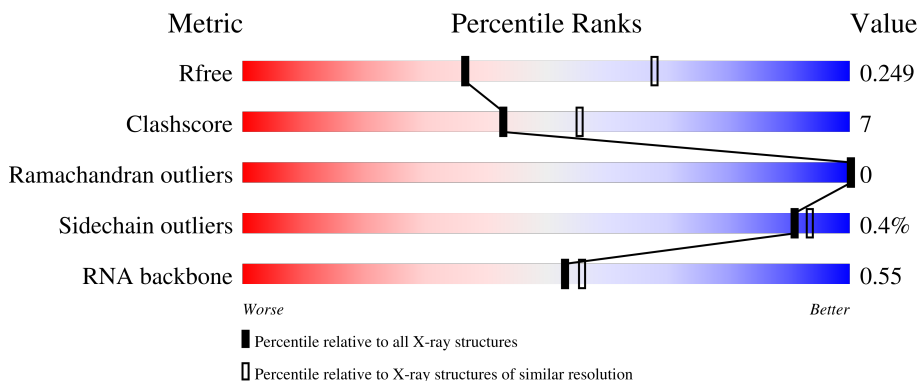
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.75 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}	130704	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RNA backbone	3102	1060 (3.02-2.50)

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	85	49% (green), 39% (yellow), 12% (orange)
2	B	1368	82% (green), 15% (yellow), 3% (orange), 0% (red), 0% (grey)
3	C	28	64% (green), 36% (yellow)
4	D	12	42% (green), 42% (yellow), 17% (grey)

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 13644 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 PTPRC sgRNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	A	85	1775	794	324	573	84	0	0	1

- Molecule 2 is a protein called CRISPR-associated endonuclease Cas9/Csn1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	1332	10883	6933	1889	2039	22	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	10	ALA	ASP	engineered mutation	UNP Q99ZW2
B	840	ALA	HIS	engineered mutation	UNP Q99ZW2

- Molecule 3 is a DNA chain called PTPRC off-target1 target strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	C	28	578	276	105	170	27	0	0	0

- Molecule 4 is a DNA chain called PTPRC off-target1 non-target strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
4	D	10	205	98	34	63	10	0	0	0

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	2	Total	Mg	0	0
			2	2		

- Molecule 6 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	3	Total K 3 3	0	0
6	B	6	Total K 6 6	0	0
6	C	1	Total K 1 1	0	0

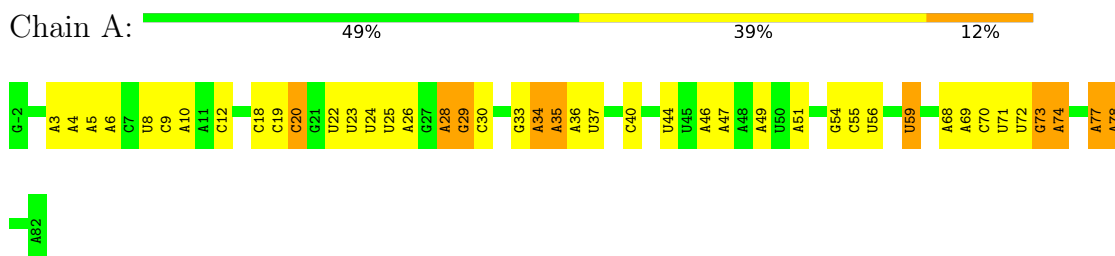
- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	50	Total O 50 50	0	0
7	B	126	Total O 126 126	0	0
7	C	11	Total O 11 11	0	0
7	D	4	Total O 4 4	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. 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. 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: PTPRC sgRNA



- Molecule 2: CRISPR-associated endonuclease Cas9/Csn1



- Molecule 3: PTPRC off-target1 target strand

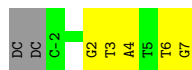




- Molecule 4: PTPRC off-target1 non-target strand

Chain D:
42% 42% 17%

A horizontal bar representing the quality score for Chain D. The bar is divided into three segments: a green segment (42%), a yellow segment (42%), and a grey segment (17%).



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	178.28Å 68.03Å 186.74Å 90.00° 110.74° 90.00°	Depositor
Resolution (Å)	48.12 – 2.75 48.12 – 2.75	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.12-2.75) 99.9 (48.12-2.75)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.33 (at 2.77Å)	Xtriage
Refinement program	PHENIX 1.20_4459	Depositor
R, R_{free}	0.222 , 0.250 0.221 , 0.249	Depositor DCC
R_{free} test set	2741 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å ²)	58.8	Xtriage
Anisotropy	0.649	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	13644	wwPDB-VP
Average B, all atoms (Å ²)	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.38% 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](#)

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

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	0.44	4/1987 (0.2%)	0.81	0/3095
2	B	0.26	0/11075	0.42	0/14881
3	C	0.74	1/648 (0.2%)	0.99	0/1001
4	D	0.57	0/228	1.02	0/350
All	All	0.33	5/13938 (0.0%)	0.56	0/19327

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	18	C	O3'-P	-6.31	1.53	1.61
1	A	20	C	O3'-P	-6.05	1.53	1.61
1	A	19	C	O3'-P	-5.99	1.53	1.61
1	A	6	A	O3'-P	-5.40	1.54	1.61
3	C	1	DG	O3'-P	-5.12	1.55	1.61

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	1775	0	894	31	0
2	B	10883	0	11054	139	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	578	0	319	8	0
4	D	205	0	115	7	0
5	A	2	0	0	0	0
6	A	3	0	0	0	0
6	B	6	0	0	0	0
6	C	1	0	0	0	0
7	A	50	0	0	0	0
7	B	126	0	0	6	0
7	C	11	0	0	0	0
7	D	4	0	0	0	0
All	All	13644	0	12382	174	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 7.

The worst 5 of 174 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)
2:B:1057:ILE:HG23	2:B:1058:ARG:H	1.16	1.11
2:B:1057:ILE:HG23	2:B:1058:ARG:N	1.87	0.90
2:B:178:ASN:HB2	2:B:184:LEU:HD11	1.65	0.76
1:A:3:A:H2'	1:A:4:A:C8	2.22	0.74
2:B:253:LYS:HA	2:B:258:LEU:HD12	1.70	0.73

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
2	B	1324/1368 (97%)	1278 (96%)	46 (4%)	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
2	B	1194/1225 (98%)	1189 (100%)	5 (0%)	91 93

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	182	ASP
2	B	309	ASN
2	B	586	ARG
2	B	763	MET
2	B	870	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	265	GLN
2	B	309	ASN
2	B	402	GLN
2	B	844	GLN
2	B	1252	ASN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	82/85 (96%)	21 (25%)	0

5 of 21 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	8	U
1	A	9	C

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	10	A
1	A	12	C
1	A	20	C

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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 12 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 [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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.