



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 22, 2023 – 11:54 AM JST

PDB ID : 7VK9  
Title : Crystal structure of xCas9 P411T  
Authors : Bao, R.; Liu, H.Y.; Luo, Y.Z.; Song, Y.J.  
Deposited on : 2021-09-29  
Resolution : 2.90 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](mailto: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 : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

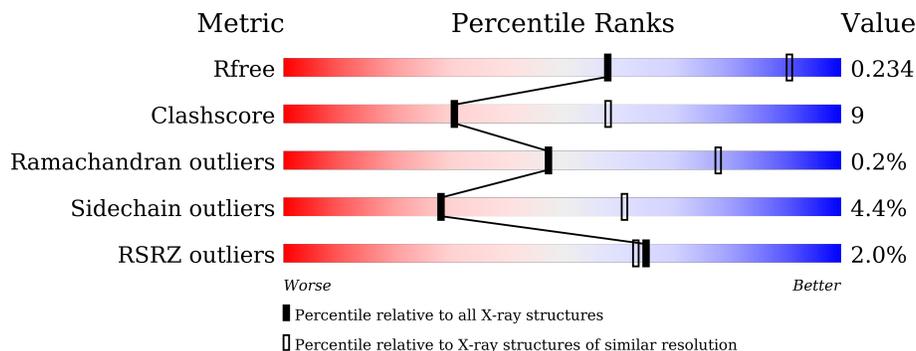
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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  $\geq 3$ , 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	A	1368	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 10529 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 Cas9/Csn1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1279	10437	6653	1804	1960	20	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	282	THR	ALA	engineered mutation	UNP Q99ZW2
A	344	LEU	ARG	engineered mutation	UNP Q99ZW2
A	429	ILE	SER	engineered mutation	UNP Q99ZW2
A	431	THR	PRO	engineered mutation	UNP Q99ZW2
A	500	LYS	GLU	engineered mutation	UNP Q99ZW2
A	563	ASP	GLU	engineered mutation	UNP Q99ZW2
A	714	ILE	MET	engineered mutation	UNP Q99ZW2
A	1239	VAL	GLU	engineered mutation	UNP Q99ZW2

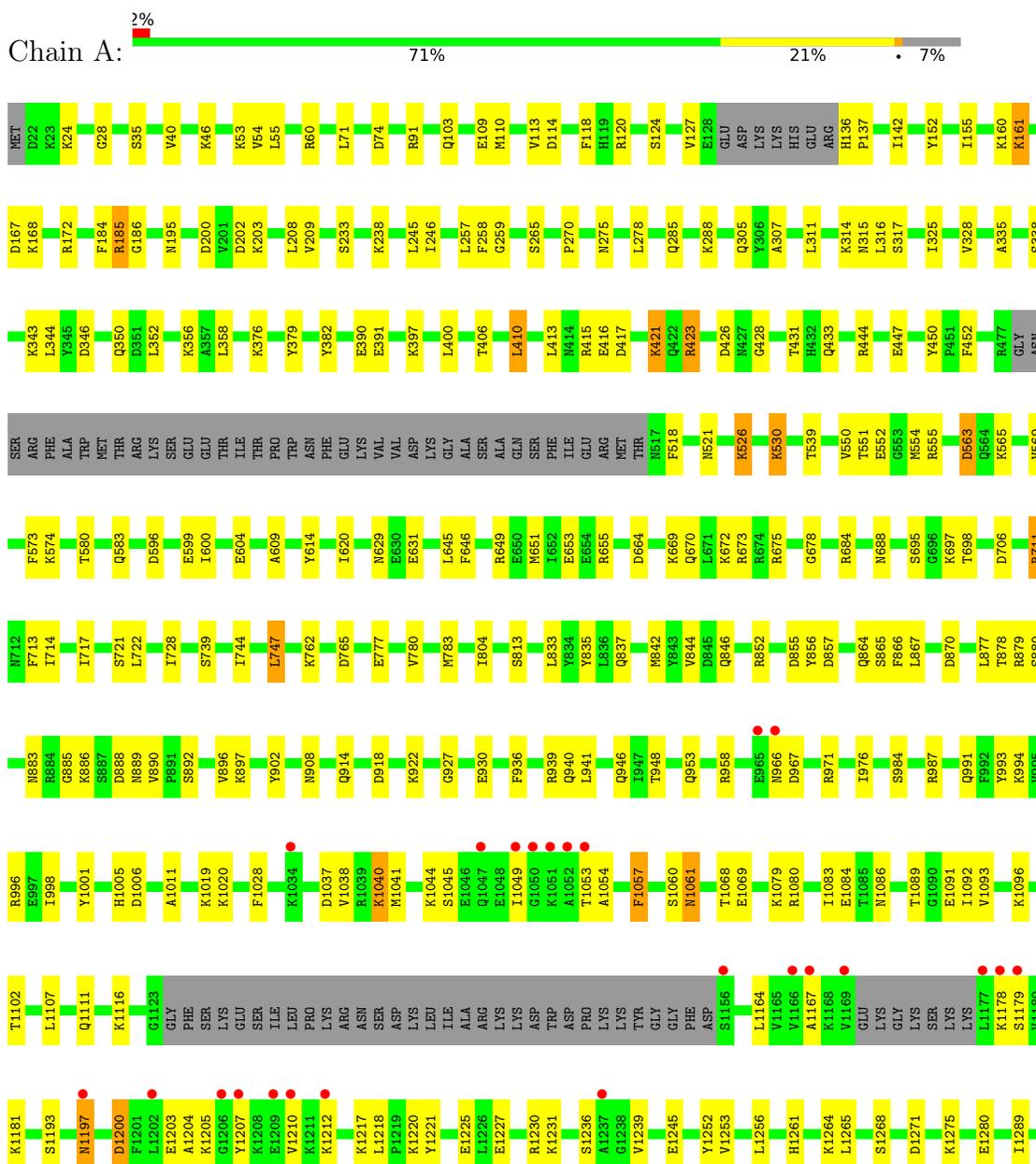
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	92	Total	O	0	0
			92	92		

### 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 Cas9/Csn1



N1306	S1312	D1319	N1328	I1329	H1331	L1332	F1333	T1334	L1338	K1345	Y1346	F1347	D1348	T1349	T1350	I1351	K1354	S1368	T1369	K1360	E1361	V1362	T1368	I1372	E1377	T1378	R1379	T1380	D1381	L1382	S1383	Q1384	L1385	GLY	GLY	ASP
-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-----	-----	-----

## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	190.02Å 190.02Å 142.99Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	41.24 – 2.90 41.24 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.7 (41.24-2.90) 99.7 (41.24-2.90)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.51 (at 2.90Å)	Xtrriage
Refinement program	PHENIX 1.13_2998	Depositor
R, $R_{free}$	0.194 , 0.234 0.194 , 0.234	Depositor DCC
$R_{free}$ test set	2001 reflections (3.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	53.4	Xtrriage
Anisotropy	0.479	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 42.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.032 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	10529	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.31% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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	A	0.61	2/10614 (0.0%)	0.68	4/14273 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	379	TYR	CE2-CZ	-5.29	1.31	1.38
1	A	382	TYR	CE1-CZ	-5.11	1.31	1.38

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	604	GLU	N-CA-C	10.68	139.84	111.00
1	A	604	GLU	CB-CA-C	-6.77	96.86	110.40
1	A	1057	PHE	CB-CA-C	-6.03	98.33	110.40
1	A	1049	ILE	CG1-CB-CG2	-5.11	100.16	111.40

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	10437	0	10601	194	0
2	A	92	0	0	6	0
All	All	10529	0	10601	194	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 9.

The worst 5 of 194 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:1354:LYS:HE2	2:A:1428:HOH:O	1.50	1.12
1:A:136:HIS:N	1:A:137:PRO:CD	2.14	1.08
1:A:1230:ARG:NH2	1:A:1361:GLU:OE2	1.85	1.07
1:A:1345:LYS:NZ	1:A:1348:ASP:O	1.87	1.06
1:A:1275:LYS:NZ	2:A:1401:HOH:O	2.06	0.88

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	1269/1368 (93%)	1210 (95%)	56 (4%)	3 (0%)	47 78

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	996	ARG
1	A	1204	ALA
1	A	1264	LYS

### 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	1147/1228 (93%)	1097 (96%)	50 (4%)	28 61

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	897	LYS
1	A	1045	SER
1	A	1383	SER
1	A	908	ASN
1	A	984	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	433	GLN
1	A	629	ASN
1	A	883	ASN
1	A	1061	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 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](#)

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

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	1279/1368 (93%)	-0.09	25 (1%) 65 63	26, 47, 84, 131	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1050	GLY	6.7
1	A	1156	SER	5.9
1	A	1177	LEU	4.6
1	A	1052	ALA	4.6
1	A	1178	LYS	3.8

### 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](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.