



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 18, 2026 – 05:13 AM UTC

PDB ID : 6AEB / pdb_00006aeb
Title : Crystal structure of xCas9 in complex with sgRNA and target DNA (AAG PAM)
Authors : Guo, M.; Ren, K.; Zhu, Y.; Huang, Z.
Deposited on : 2018-08-04
Resolution : 3.00 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

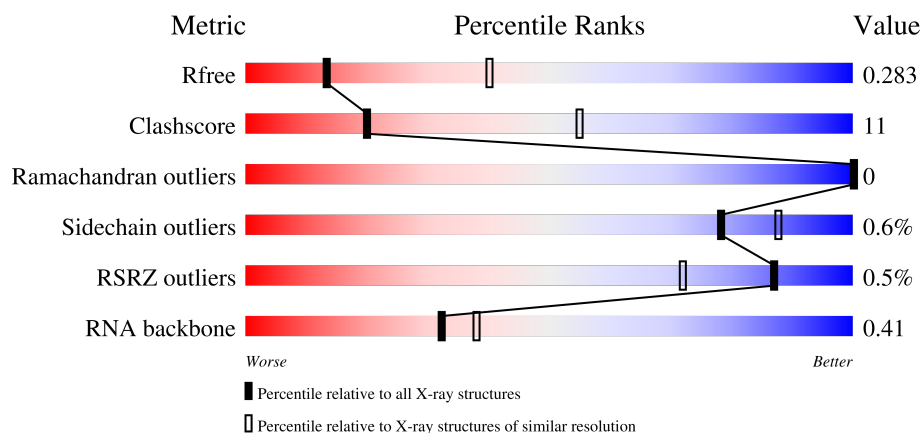
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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	2672 (3.00-3.00)
Clashscore	190562	2977 (3.00-3.00)
Ramachandran outliers	187476	2877 (3.00-3.00)
Sidechain outliers	187428	2880 (3.00-3.00)
RSRZ outliers	180081	2671 (3.00-3.00)
RNA backbone	3983	1109 (3.20-2.80)

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	A	100	 45% 35% 15% 5%
1	E	100	 44% 39% 12% 5%
2	C	28	 43% 46% 11%
2	G	28	 64% 25% 11%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	11	55% 45%
3	H	11	45% 55%
4	B	1368	% 71% 25% • •
4	F	1368	72% 25% •

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 26351 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 (95-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	95	Total	C	N	O	P	0	0	0
			2032	911	372	654	95			
1	E	95	Total	C	N	O	P	0	0	0
			2032	911	372	654	95			

- Molecule 2 is a DNA chain called DNA (25-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	25	Total	C	N	O	P	0	0	0
			506	247	83	152	24			
2	G	25	Total	C	N	O	P	0	0	0
			506	247	83	152	24			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	11	Total	C	N	O	P	0	0	0
			227	110	46	61	10			
3	H	11	Total	C	N	O	P	0	0	0
			227	110	46	61	10			

- Molecule 4 is a protein called DNA Nuclease.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	1322	Total	C	N	O	S	0	0	0
			10411	6656	1811	1923	21			
4	F	1322	Total	C	N	O	S	0	0	0
			10409	6655	1811	1922	21			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	10	ALA	ASP	engineered mutation	UNP Q99ZW2
B	262	THR	ALA	engineered mutation	UNP Q99ZW2
B	324	LEU	ARG	engineered mutation	UNP Q99ZW2
B	409	ILE	SER	engineered mutation	UNP Q99ZW2
B	480	LYS	GLU	engineered mutation	UNP Q99ZW2
B	543	ASP	GLU	engineered mutation	UNP Q99ZW2
B	694	ILE	MET	engineered mutation	UNP Q99ZW2
B	840	ALA	HIS	engineered mutation	UNP Q99ZW2
B	1219	VAL	GLU	engineered mutation	UNP Q99ZW2
F	10	ALA	ASP	engineered mutation	UNP Q99ZW2
F	262	THR	ALA	engineered mutation	UNP Q99ZW2
F	324	LEU	ARG	engineered mutation	UNP Q99ZW2
F	409	ILE	SER	engineered mutation	UNP Q99ZW2
F	480	LYS	GLU	engineered mutation	UNP Q99ZW2
F	543	ASP	GLU	engineered mutation	UNP Q99ZW2
F	694	ILE	MET	engineered mutation	UNP Q99ZW2
F	840	ALA	HIS	engineered mutation	UNP Q99ZW2
F	1219	VAL	GLU	engineered mutation	UNP Q99ZW2

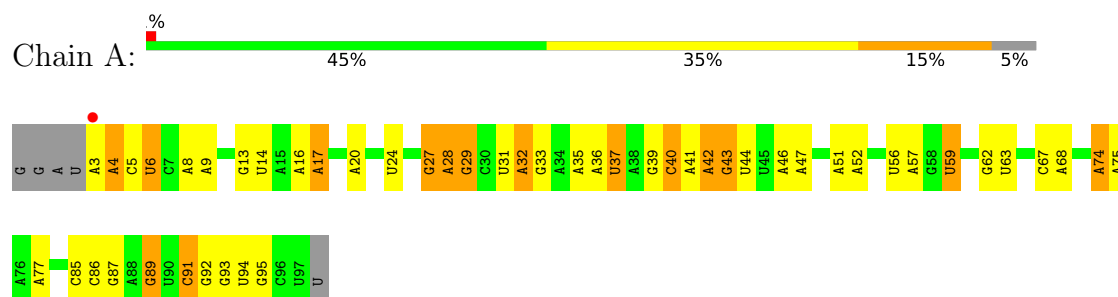
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total O 1 1	0	0

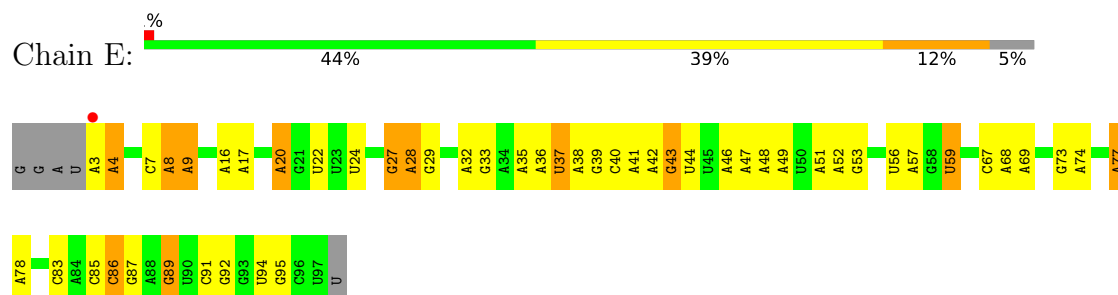
3 Residue-property plots [i](#)

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.

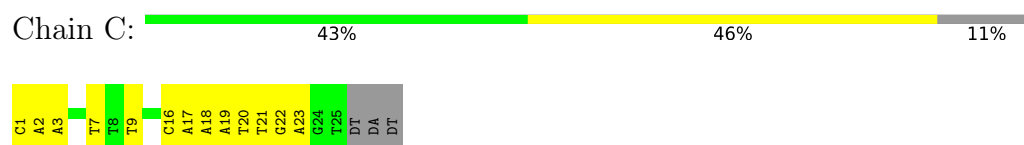
- Molecule 1: RNA (95-MER)



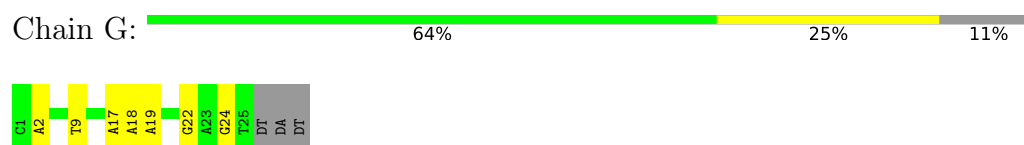
- Molecule 1: RNA (95-MER)



- Molecule 2: DNA (25-MER)



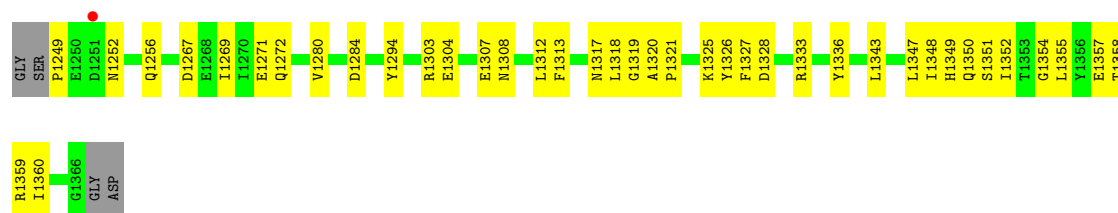
- Molecule 2: DNA (25-MER)



- Molecule 3: DNA (5'-D(*AP*AP*AP*AP*AP*GP*TP*AP*TP*TP*G)-3')

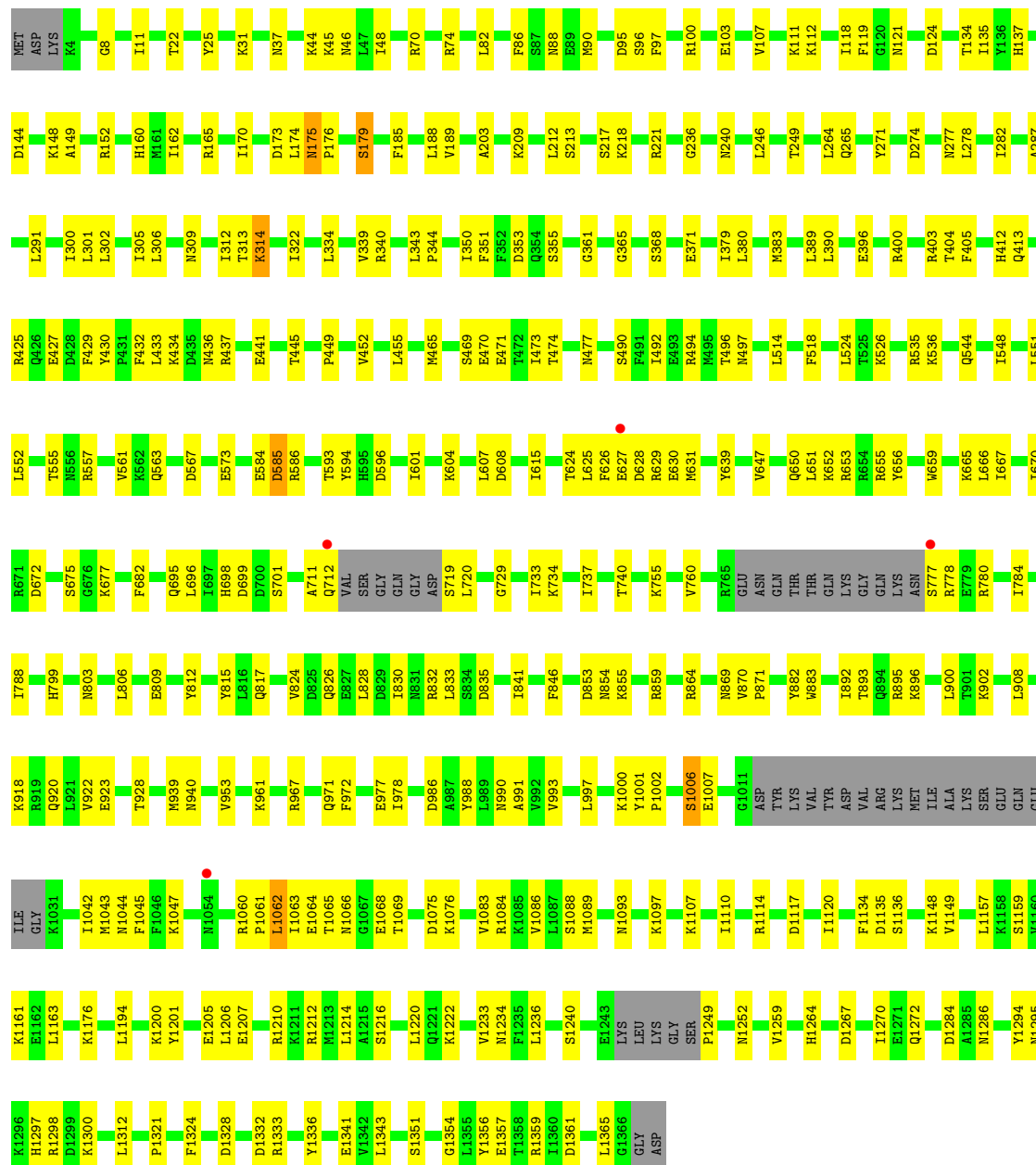






• Molecule 4: DNA Nuclease

Chain F: 72% 25% .



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	361.85Å 71.13Å 198.32Å 90.00° 101.75° 90.00°	Depositor
Resolution (Å)	49.45 – 3.00 49.45 – 3.00	Depositor EDS
% Data completeness (in resolution range)	97.7 (49.45-3.00) 86.0 (49.45-3.00)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.74 (at 3.01Å)	Xtriage
Refinement program	PHENIX (1.11.1 _2575: ???)	Depositor
R, R_{free}	0.231 , 0.283 0.234 , 0.283	Depositor DCC
R_{free} test set	1981 reflections (1.99%)	wwPDB-VP
Wilson B-factor (Å ²)	62.2	Xtriage
Anisotropy	0.704	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 41.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	26351	wwPDB-VP
Average B, all atoms (Å ²)	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 41.30 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.4091e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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

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	0.20	0/2277	0.40	0/3546
1	E	0.21	0/2277	0.41	0/3546
2	C	0.35	0/565	0.57	0/870
2	G	0.35	0/565	0.57	0/870
3	D	0.33	0/256	0.45	0/394
3	H	0.32	0/256	0.50	0/394
4	B	0.41	5/10600 (0.0%)	0.56	24/14316 (0.2%)
4	F	0.36	4/10598 (0.0%)	0.52	17/14313 (0.1%)
All	All	0.36	9/27394 (0.0%)	0.52	41/38249 (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	B	0	1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	1319	GLY	C-O	12.27	1.30	1.24
4	B	1164	LEU	C-N	-8.89	1.27	1.33
4	B	119	PHE	C-N	-8.47	1.21	1.33
4	F	175	ASN	C-O	-7.31	1.18	1.25
4	F	175	ASN	CA-C	-5.55	1.45	1.52

The worst 5 of 41 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	119	PHE	N-CA-C	-11.72	86.98	108.65
4	B	118	ILE	CB-CA-C	-9.79	98.71	112.22
4	B	118	ILE	N-CA-C	9.45	120.26	110.72

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	F	1075	ASP	N-CA-C	-8.63	94.34	108.41
4	B	1083	VAL	CB-CA-C	-8.25	100.84	112.22

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	B	119	PHE	Mainchain

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	2032	0	1021	42	0
1	E	2032	0	1021	27	0
2	C	506	0	289	18	0
2	G	506	0	289	9	0
3	D	227	0	126	7	0
3	H	227	0	126	5	0
4	B	10411	0	10286	256	0
4	F	10409	0	10281	234	0
5	B	1	0	0	0	0
All	All	26351	0	23439	545	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 545 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:63:U:N3	4:B:65:LYS:NZ	1.95	1.13
4:F:1047:LYS:O	4:F:1076:LYS:NZ	1.89	1.05
4:B:150:ASP:OD1	4:B:152:ARG:HG2	1.69	0.93
4:F:1062:LEU:O	4:F:1076:LYS:HG3	1.69	0.92
4:B:561:VAL:HG23	4:B:585:ASP:O	1.70	0.91

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	B	1312/1368 (96%)	1188 (90%)	124 (10%)	0	100	100
4	F	1312/1368 (96%)	1181 (90%)	131 (10%)	0	100	100
All	All	2624/2736 (96%)	2369 (90%)	255 (10%)	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	
4	B	1086/1226 (89%)	1077 (99%)	9 (1%)	73	86
4	F	1085/1226 (88%)	1081 (100%)	4 (0%)	84	90
All	All	2171/2452 (88%)	2158 (99%)	13 (1%)	78	88

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	B	1325	LYS
4	B	1328	ASP
4	F	1063	ILE
4	F	179	SER
4	F	314	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
4	B	1308	ASN
4	F	1264	HIS
4	F	265	GLN
4	F	1308	ASN
4	F	930	HIS

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	94/100 (94%)	28 (29%)	4 (4%)
1	E	94/100 (94%)	28 (29%)	5 (5%)
All	All	188/200 (94%)	56 (29%)	9 (4%)

5 of 56 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	4	A
1	A	6	U
1	A	9	A
1	A	17	A
1	A	20	A

5 of 9 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	E	42	A
1	E	68	A
1	A	42	A
1	E	8	A
1	E	27	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 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	A	95/100 (95%)	-0.12	1 (1%) 78 57	51, 79, 116, 149	0
1	E	95/100 (95%)	-0.09	1 (1%) 78 57	49, 79, 118, 148	0
2	C	25/28 (89%)	0.22	0 100 100	58, 67, 105, 120	0
2	G	25/28 (89%)	0.25	0 100 100	56, 68, 112, 126	0
3	D	11/11 (100%)	0.32	0 100 100	53, 72, 117, 118	0
3	H	11/11 (100%)	0.42	0 100 100	55, 69, 115, 121	0
4	B	1322/1368 (96%)	0.09	9 (0%) 84 66	51, 79, 103, 138	0
4	F	1322/1368 (96%)	0.09	4 (0%) 90 79	46, 79, 104, 134	0
All	All	2906/3014 (96%)	0.09	15 (0%) 87 72	46, 79, 105, 149	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	F	777	SER	2.9
4	F	627	GLU	2.8
4	F	712	GLN	2.8
4	B	777	SER	2.7
4	B	1052	LEU	2.6

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.