



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 7, 2026 – 01:44 AM UTC

PDB ID : 4QYZ / pdb_00004qyz
Title : Crystal structure of a CRISPR RNA-guided surveillance complex, Cascade, bound to a ssDNA target
Authors : Mulepati, S.; Bailey, S.
Deposited on : 2014-07-26
Resolution : 3.03 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

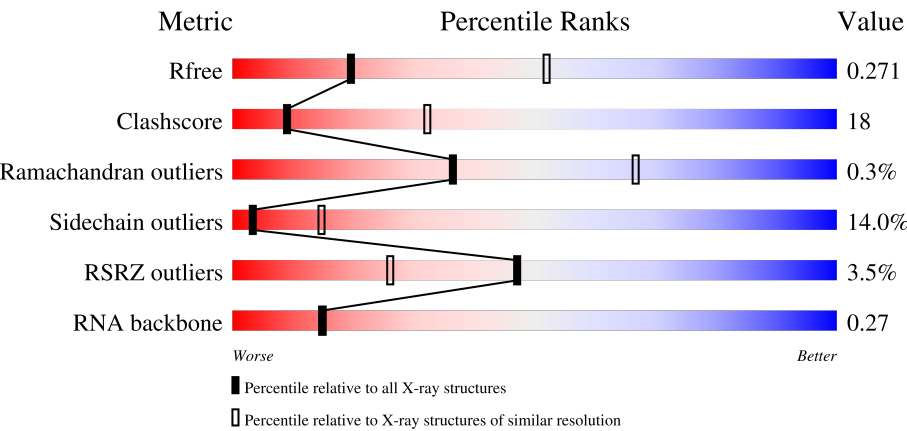
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.03 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	3685 (3.08-3.00)
Clashscore	190562	4007 (3.08-3.00)
Ramachandran outliers	187476	3834 (3.08-3.00)
Sidechain outliers	187428	3836 (3.08-3.00)
RSRZ outliers	180081	3684 (3.08-3.00)
RNA backbone	3983	1163 (3.28-2.80)

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	502	
2	B	160	
2	C	160	
3	D	363	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	E	363	2% 61% 32% 6% 2%
3	F	363	1% 60% 34% 5% 1%
3	G	363	1% 59% 35% 5% 1%
3	H	363	2% 61% 33% 5% 1%
3	I	363	2% 42% 24% 29% 1%
4	J	224	5% 34% 40% 12% 13%
5	K	199	17% 41% 32% 24% 1%
6	L	61	15% 49% 26% 10%
7	M	40	35% 48% 18%

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 26543 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 system Cascade subunit CasA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	468	Total	C	N	O	S	0	0	0
			3702	2368	653	662	19			

- Molecule 2 is a protein called CRISPR system Cascade subunit CasB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	150	Total	C	N	O	S	0	0	0
			1238	777	239	215	7			
2	C	150	Total	C	N	O	S	0	0	0
			1238	777	239	215	7			

- Molecule 3 is a protein called CRISPR system Cascade subunit CasC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	340	Total	C	N	O	S	0	0	0
			2630	1643	469	502	16			
3	E	361	Total	C	N	O	S	0	0	0
			2794	1747	496	535	16			
3	F	362	Total	C	N	O	S	0	0	0
			2803	1752	497	538	16			
3	G	360	Total	C	N	O	S	0	0	0
			2788	1742	495	535	16			
3	H	359	Total	C	N	O	S	0	0	0
			2781	1738	494	533	16			
3	I	256	Total	C	N	O	S	0	0	0
			1978	1231	354	380	13			

- Molecule 4 is a protein called CRISPR system Cascade subunit CasD.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	J	195	Total	C	N	O	S	0	0	0
			1545	987	273	276	9			

- Molecule 5 is a protein called CRISPR system Cascade subunit CasE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	K	152	Total	C	N	O	S	0	0	0
			1197	777	209	206	5			

- Molecule 6 is a RNA chain called RNA (55-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	L	55	Total	C	N	O	P	0	0	0
			1178	526	216	382	54			

- Molecule 7 is a DNA chain called DNA (33-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	M	33	Total	C	N	O	P	0	0	0
			670	319	122	196	33			

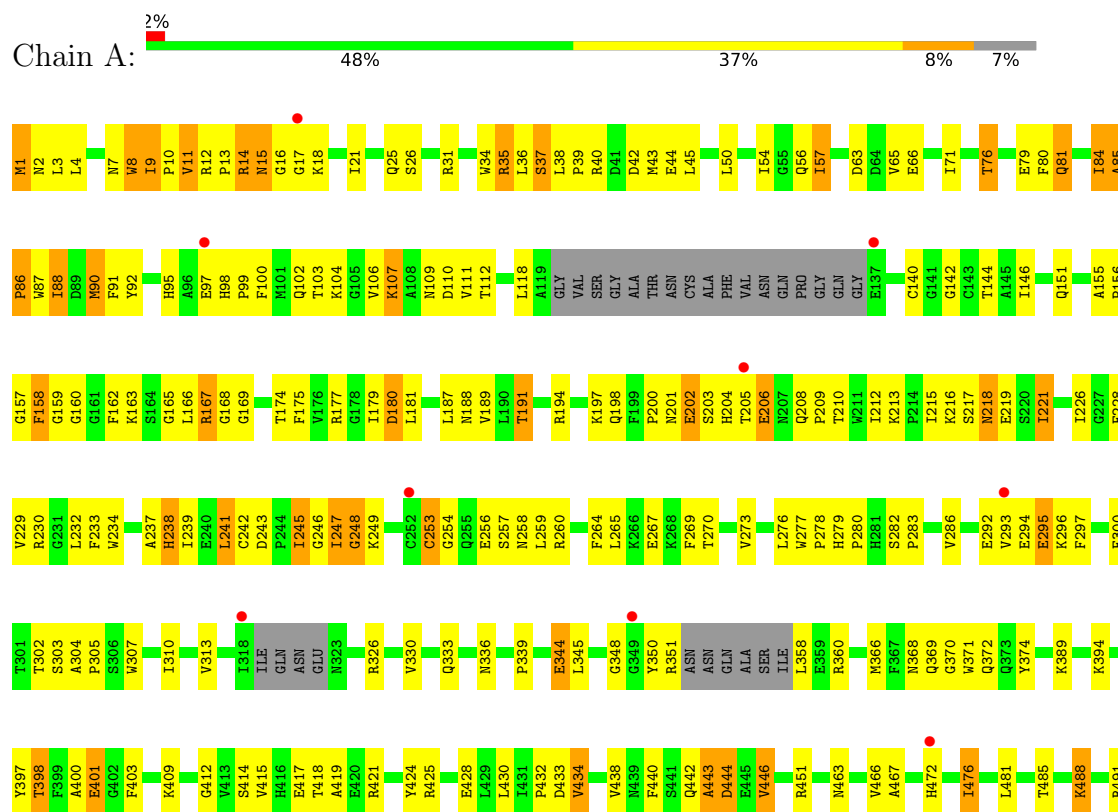
- Molecule 8 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	1	Total	Zn	0	0
			1	1		

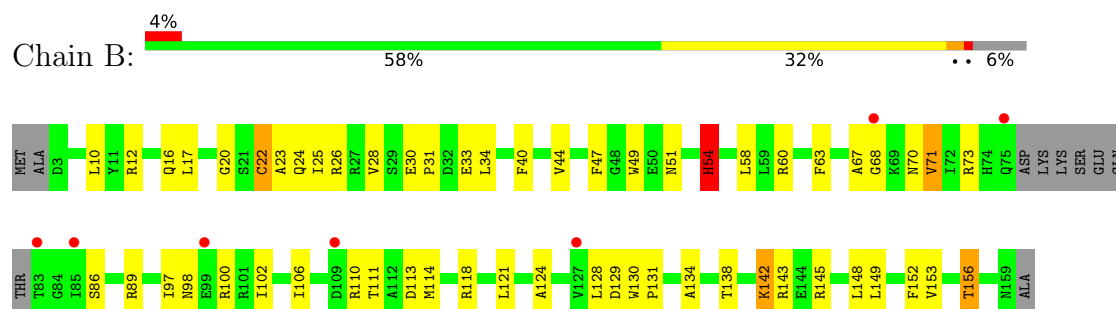
3 Residue-property plots

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.

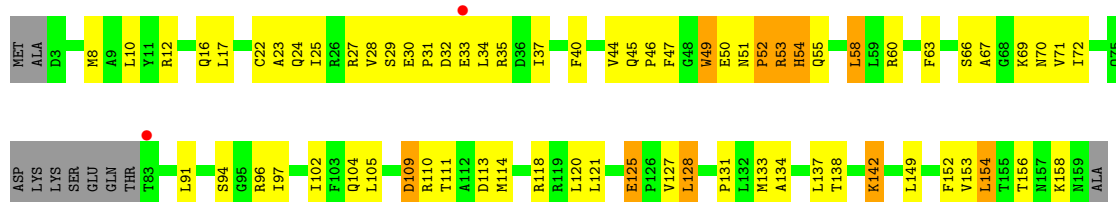
• Molecule 1: CRISPR system Cascade subunit CasA



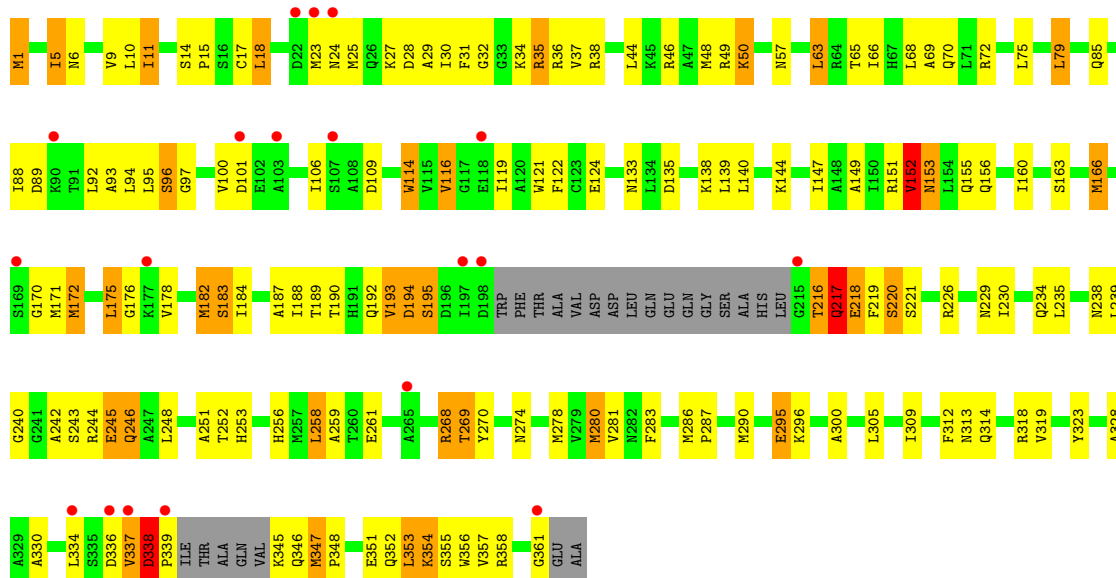
• Molecule 2: CRISPR system Cascade subunit CasB



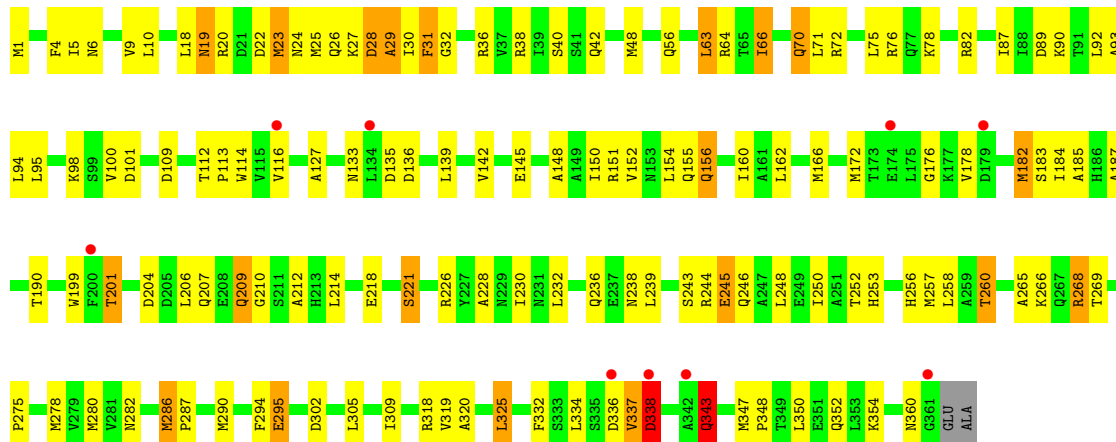
- Molecule 2: CRISPR system Cascade subunit CasB



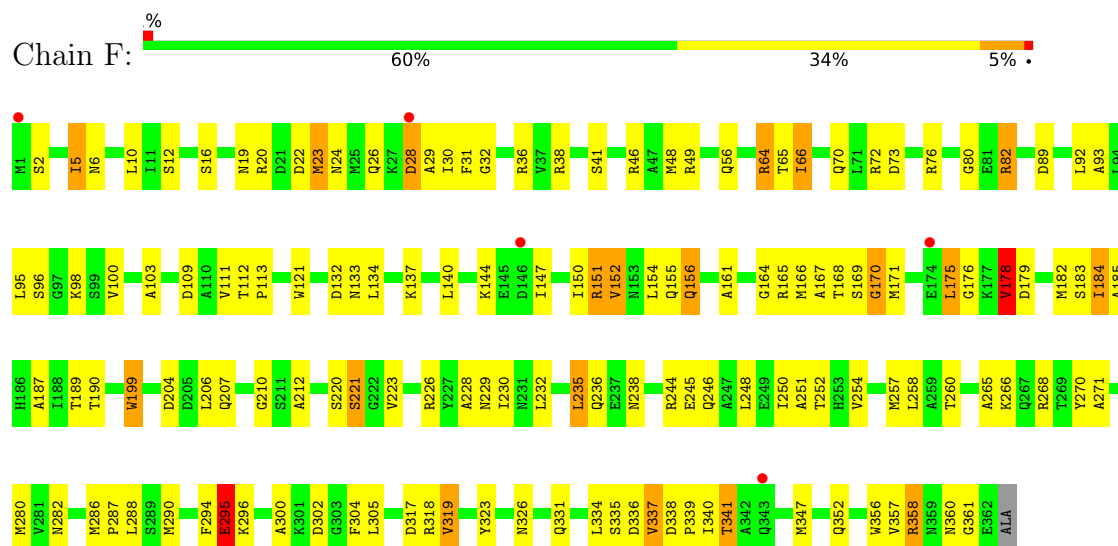
- Molecule 3: CRISPR system Cascade subunit CasC



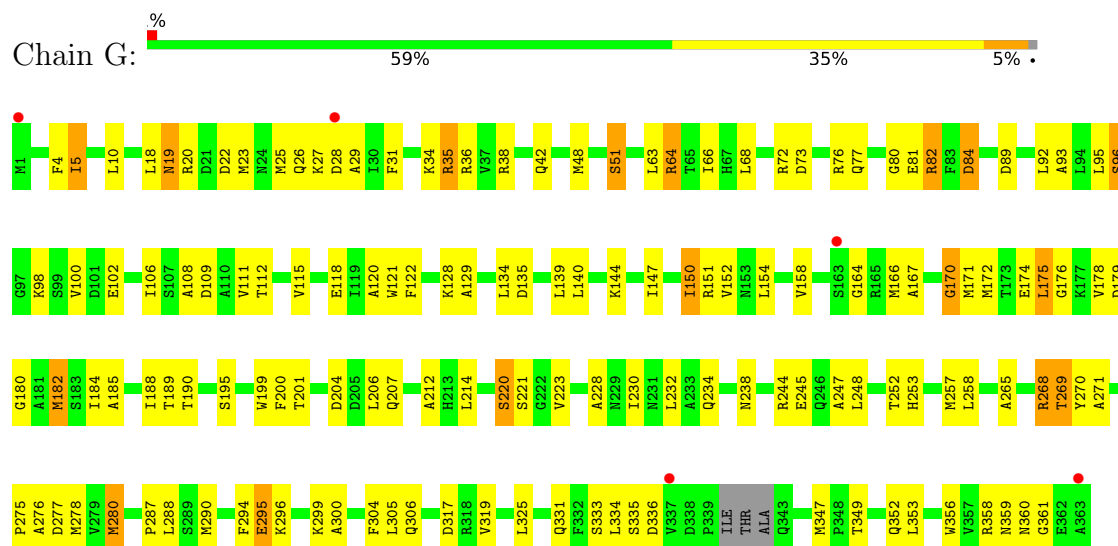
- Molecule 3: CRISPR system Cascade subunit CasC



• Molecule 3: CRISPR system Cascade subunit CasC



• Molecule 3: CRISPR system Cascade subunit CasC

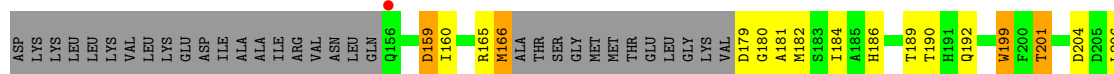
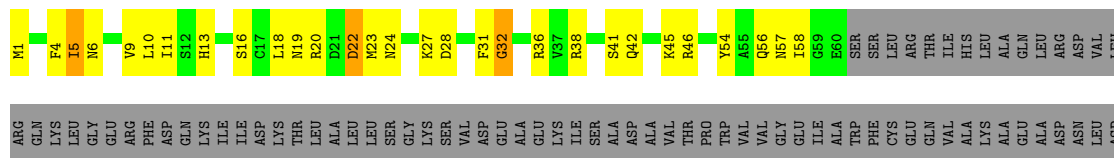
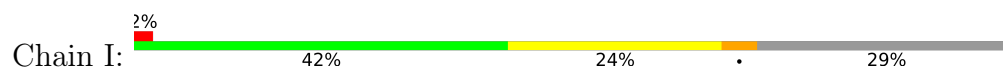


• Molecule 3: CRISPR system Cascade subunit CasC

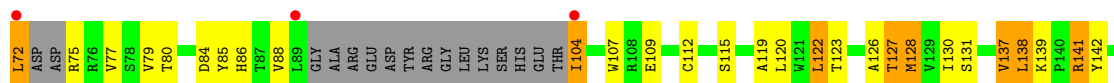
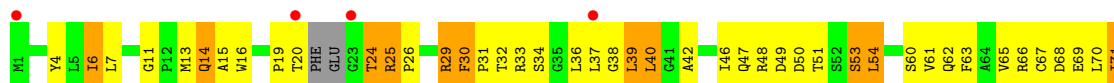




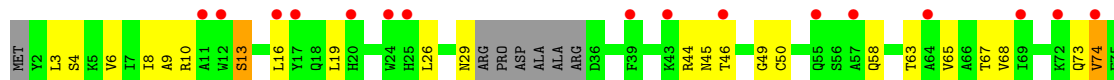
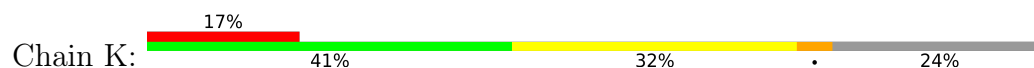
• Molecule 3: CRISPR system Cascade subunit CasC

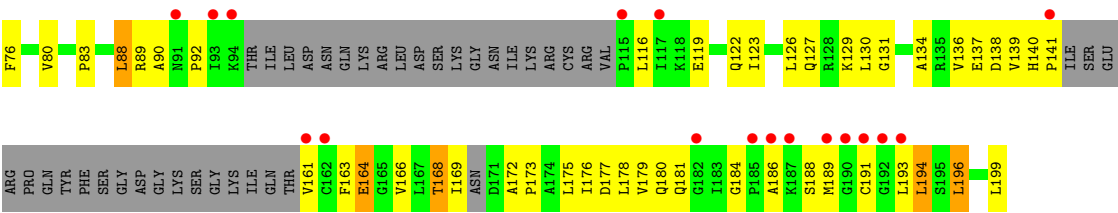


• Molecule 4: CRISPR system Cascade subunit CasD

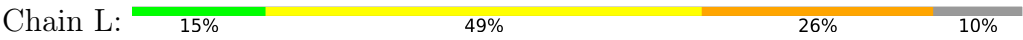


• Molecule 5: CRISPR system Cascade subunit CasE

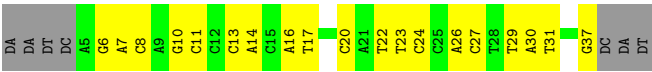




● Molecule 6: RNA (55-MER)



● Molecule 7: DNA (33-MER)



4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	223.87Å 223.87Å 290.65Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	39.47 – 3.03 39.47 – 3.03	Depositor EDS
% Data completeness (in resolution range)	96.2 (39.47-3.03) 96.3 (39.47-3.03)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.36 (at 3.01Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.224 , 0.267 0.230 , 0.271	Depositor DCC
R_{free} test set	7932 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å ²)	86.4	Xtriage
Anisotropy	0.268	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 76.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.016 for -h,-k,l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	26543	wwPDB-VP
Average B, all atoms (Å ²)	106.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.29% 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

5.1 Standard geometry

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

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	0.66	2/3789 (0.1%)	1.14	31/5138 (0.6%)
2	B	0.61	0/1262	1.07	2/1704 (0.1%)
2	C	0.72	1/1262 (0.1%)	1.11	5/1704 (0.3%)
3	D	0.56	0/2669	0.98	6/3595 (0.2%)
3	E	0.68	1/2839 (0.0%)	1.10	10/3832 (0.3%)
3	F	0.71	0/2848	1.10	13/3844 (0.3%)
3	G	0.76	0/2832	1.07	5/3820 (0.1%)
3	H	0.69	1/2824 (0.0%)	1.10	12/3808 (0.3%)
3	I	0.71	1/2013 (0.0%)	1.11	8/2716 (0.3%)
4	J	0.66	0/1578	1.16	10/2136 (0.5%)
5	K	0.47	0/1222	0.93	2/1654 (0.1%)
6	L	0.87	16/1317 (1.2%)	0.82	0/2051
7	M	0.44	0/750	0.66	0/1153
All	All	0.68	22/27205 (0.1%)	1.06	104/37155 (0.3%)

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
1	A	0	2
3	D	0	1
3	E	0	1
All	All	0	4

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	L	57	C	N1-C2	10.47	1.61	1.40
6	L	50	G	N9-C4	7.15	1.52	1.38

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	L	50	G	C2-N3	7.14	1.47	1.32
6	L	57	C	C2-O2	6.36	1.37	1.24
6	L	57	C	C4-C5	6.13	1.55	1.43

The worst 5 of 104 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	338	ASP	CA-C-N	-13.63	105.27	119.99
3	E	338	ASP	C-N-CA	-13.63	105.27	119.99
3	D	218	GLU	N-CA-C	11.74	127.25	110.23
2	C	110	ARG	N-CA-C	10.99	124.42	111.02
2	B	54	HIS	N-CA-C	-10.54	102.61	114.62

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	157	GLY	Peptide
1	A	442	GLN	Peptide
3	D	217	GLN	Peptide
3	E	31	PHE	Peptide

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	3702	0	3706	137	0
2	B	1238	0	1242	37	0
2	C	1238	0	1242	52	0
3	D	2630	0	2626	117	1
3	E	2794	0	2780	110	0
3	F	2803	0	2786	113	0
3	G	2788	0	2767	117	1
3	H	2781	0	2765	98	0
3	I	1978	0	1918	82	0
4	J	1545	0	1562	87	0
5	K	1197	0	1232	45	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	L	1178	0	597	66	0
7	M	670	0	371	19	0
8	A	1	0	0	0	0
All	All	26543	0	25594	955	1

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 18.

The worst 5 of 955 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:E:337:VAL:N	3:E:338:ASP:HB2	1.40	1.35
3:E:337:VAL:H	3:E:338:ASP:CB	1.61	1.13
1:A:4:LEU:HB3	1:A:84:ILE:HD11	1.37	1.06
3:E:337:VAL:CA	3:E:338:ASP:HB2	1.85	1.06
3:H:185:ALA:HB2	3:I:271:ALA:HB3	1.41	1.02

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:318:ARG:NH1	3:G:80:GLY:O[2_655]	2.10	0.10

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	460/502 (92%)	414 (90%)	42 (9%)	4 (1%)	14	44
2	B	146/160 (91%)	139 (95%)	7 (5%)	0	100	100
2	C	146/160 (91%)	136 (93%)	8 (6%)	2 (1%)	9	33

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	D	334/363 (92%)	303 (91%)	29 (9%)	2 (1%)	21	53
3	E	359/363 (99%)	346 (96%)	12 (3%)	1 (0%)	36	67
3	F	360/363 (99%)	349 (97%)	11 (3%)	0	100	100
3	G	356/363 (98%)	349 (98%)	7 (2%)	0	100	100
3	H	355/363 (98%)	342 (96%)	13 (4%)	0	100	100
3	I	250/363 (69%)	245 (98%)	5 (2%)	0	100	100
4	J	183/224 (82%)	174 (95%)	8 (4%)	1 (0%)	24	57
5	K	142/199 (71%)	136 (96%)	6 (4%)	0	100	100
All	All	3091/3423 (90%)	2933 (95%)	148 (5%)	10 (0%)	36	67

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	52	PRO
3	E	338	ASP
1	A	369	GLN
2	C	53	ARG
1	A	253	CYS

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	401/426 (94%)	327 (82%)	74 (18%)	1	8
2	B	130/138 (94%)	118 (91%)	12 (9%)	8	30
2	C	130/138 (94%)	117 (90%)	13 (10%)	7	27
3	D	280/298 (94%)	227 (81%)	53 (19%)	1	7
3	E	297/298 (100%)	265 (89%)	32 (11%)	6	24
3	F	298/298 (100%)	266 (89%)	32 (11%)	6	24
3	G	296/298 (99%)	260 (88%)	36 (12%)	5	19
3	H	295/298 (99%)	262 (89%)	33 (11%)	6	22

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	I	208/298 (70%)	185 (89%)	23 (11%)	6	22
4	J	168/192 (88%)	123 (73%)	45 (27%)	0	2
5	K	129/170 (76%)	114 (88%)	15 (12%)	5	21
All	All	2632/2852 (92%)	2264 (86%)	368 (14%)	3	15

5 of 368 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	G	147	ILE
3	I	5	ILE
3	G	245	GLU
3	H	105	LYS
3	I	277	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 50 such sidechains are listed below:

Mol	Chain	Res	Type
3	G	26	GLN
3	H	42	GLN
5	K	127	GLN
3	G	57	ASN
3	G	234	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
6	L	53/61 (86%)	23 (43%)	6 (11%)

5 of 23 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
6	L	2	U
6	L	8	G
6	L	9	C
6	L	14	G
6	L	15	A

5 of 6 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
6	L	42	A
6	L	44	U
6	L	45	U
6	L	21	G
6	L	9	C

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](#)

Of 1 ligands modelled in this entry, 1 is 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 [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	468/502 (93%)	0.06	9 (1%) 66 42	63, 99, 129, 161	0
2	B	150/160 (93%)	0.29	7 (4%) 36 19	90, 117, 141, 158	0
2	C	150/160 (93%)	0.19	2 (1%) 75 52	78, 110, 139, 166	0
3	D	340/363 (93%)	0.32	19 (5%) 30 15	82, 124, 176, 195	0
3	E	361/363 (99%)	0.14	9 (2%) 58 34	68, 101, 155, 175	0
3	F	362/363 (99%)	-0.09	5 (1%) 73 50	60, 90, 125, 146	0
3	G	360/363 (99%)	-0.08	5 (1%) 73 50	56, 86, 121, 140	0
3	H	359/363 (98%)	-0.08	6 (1%) 69 45	56, 91, 131, 147	0
3	I	256/363 (70%)	0.10	6 (2%) 61 37	59, 92, 129, 149	0
4	J	195/224 (87%)	0.37	11 (5%) 30 15	72, 102, 137, 158	0
5	K	152/199 (76%)	1.23	33 (21%) 2 1	115, 174, 210, 234	0
6	L	55/61 (90%)	-0.13	0 100 100	72, 96, 195, 199	0
7	M	33/40 (82%)	-0.22	0 100 100	79, 102, 148, 168	0
All	All	3241/3524 (91%)	0.14	112 (3%) 47 26	56, 100, 162, 234	0

The worst 5 of 112 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	K	189	MET	7.0
4	J	23	GLY	6.1
5	K	191	CYS	5.3
3	G	363	ALA	4.8
4	J	167	PRO	4.1

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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
8	ZN	A	601	1/1	0.99	0.03	170,170,170,170	0

6.5 Other polymers [i](#)

There are no such residues in this entry.