



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 23, 2024 – 04:19 PM JST

PDB ID : 8Y09  
Title : Crystal structure of LbCas12a in complex with crRNA and 15nt target DNA  
Authors : Lin, X.; Chen, J.; Liu, L.  
Deposited on : 2024-01-22  
Resolution : 2.87 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:

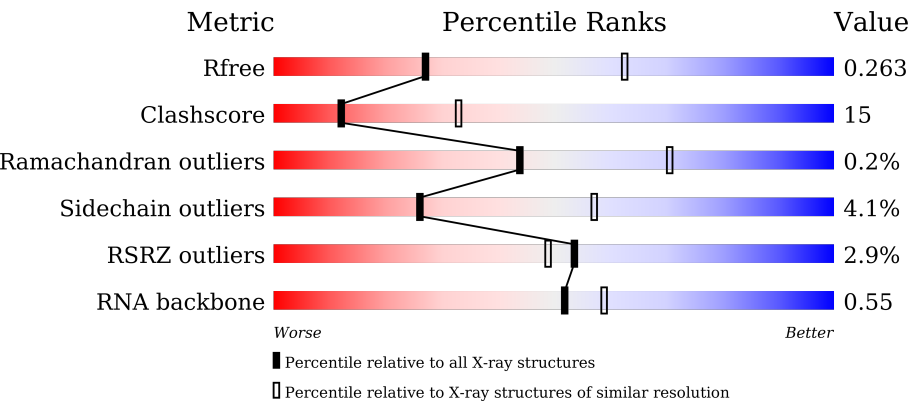
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

# 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 2.87 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 <sub>free</sub>	164625	3316 (2.90-2.86)
Clashscore	180529	3609 (2.90-2.86)
Ramachandran outliers	177936	3529 (2.90-2.86)
Sidechain outliers	177891	3532 (2.90-2.86)
RSRZ outliers	164620	3319 (2.90-2.86)
RNA backbone	3690	1088 (3.10-2.66)

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	1228	<div><div>2%</div><div>68%</div><div>29%</div><div>..</div></div>
1	E	1228	<div><div>4%</div><div>66%</div><div>30%</div><div>..</div></div>
2	B	40	<div><div>2%</div><div>28%</div><div>40%</div><div>15%</div><div>15%</div></div>
2	F	40	<div><div>2%</div><div>38%</div><div>30%</div><div>18%</div><div>15%</div></div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	C	24	
3	G	24	
4	D	11	
4	H	11	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	LI	A	1301	-	-	-	X
5	LI	A	1302	-	-	-	X
5	LI	B	101	-	-	-	X
5	LI	E	1301	-	-	-	X

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 22752 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 LbCas12a.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1214	Total	C	N	O	S	0	0	0
			9949	6402	1626	1892	29			
1	E	1217	Total	C	N	O	S	0	0	0
			9964	6409	1627	1899	29			

- Molecule 2 is a RNA chain called RNA (34-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	34	Total	C	N	O	P	0	0	0
			719	324	128	234	33			
2	F	34	Total	C	N	O	P	0	0	0
			719	324	128	234	33			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	23	Total	C	N	O	P	0	0	0
			476	226	89	138	23			
3	G	23	Total	C	N	O	P	0	0	0
			476	226	89	138	23			

- Molecule 4 is a DNA chain called DNA (5'-D(\*CP\*GP\*TP\*CP\*CP\*TP\*TP\*TP\*AP\*TP\*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	11	Total	C	N	O	P	0	0	0
			217	107	31	69	10			
4	H	11	Total	C	N	O	P	0	0	0
			217	107	31	69	10			

- Molecule 5 is LITHIUM ION (three-letter code: LI) (formula: Li).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	2	Total 2	Li 2	0	0
5	B	1	Total 1	Li 1	0	0
5	E	1	Total 1	Li 1	0	0
5	F	1	Total 1	Li 1	0	0

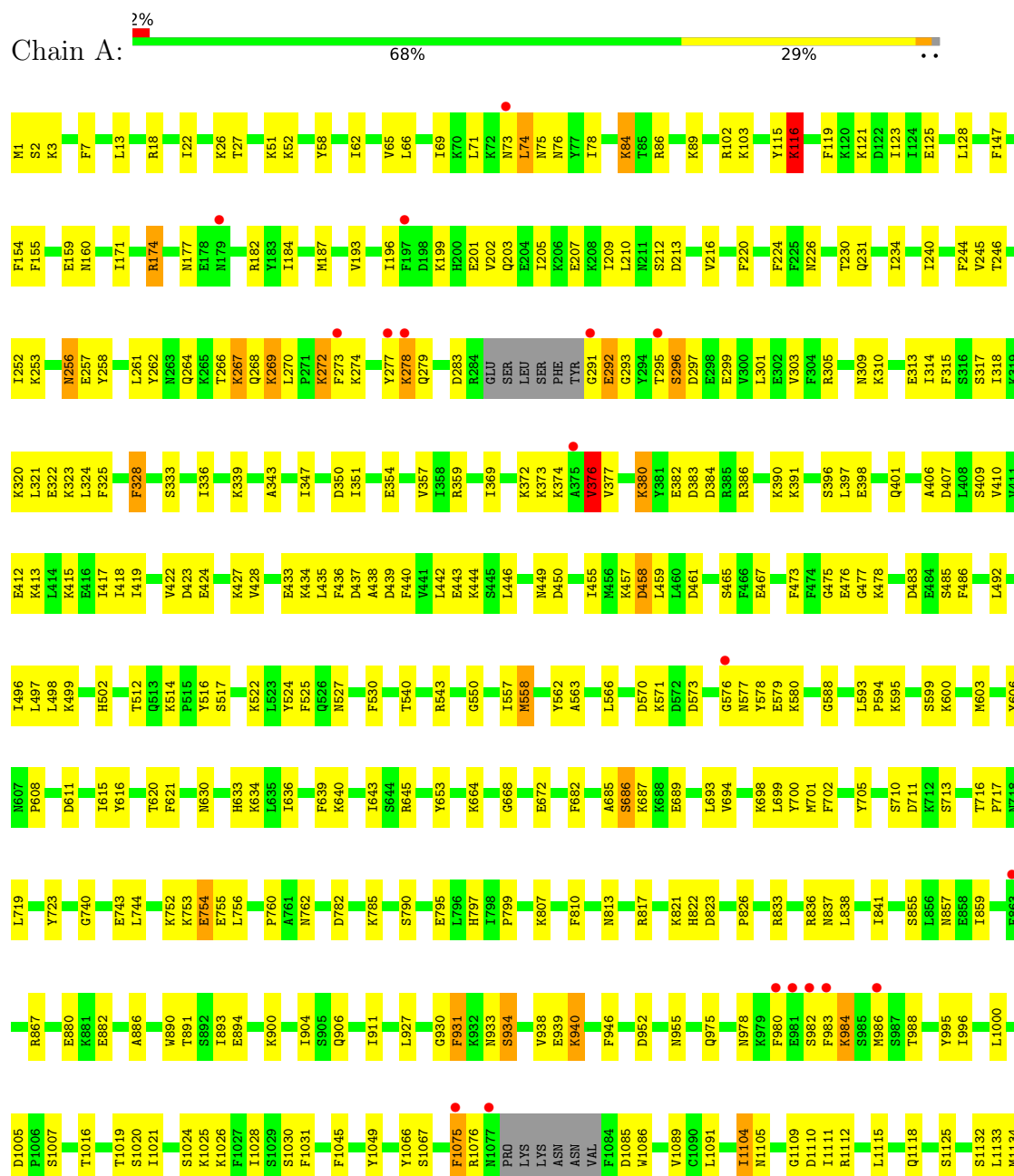
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	8	Total 8	O 8	0	0
6	E	2	Total 2	O 2	0	0

### 3 Residue-property plots

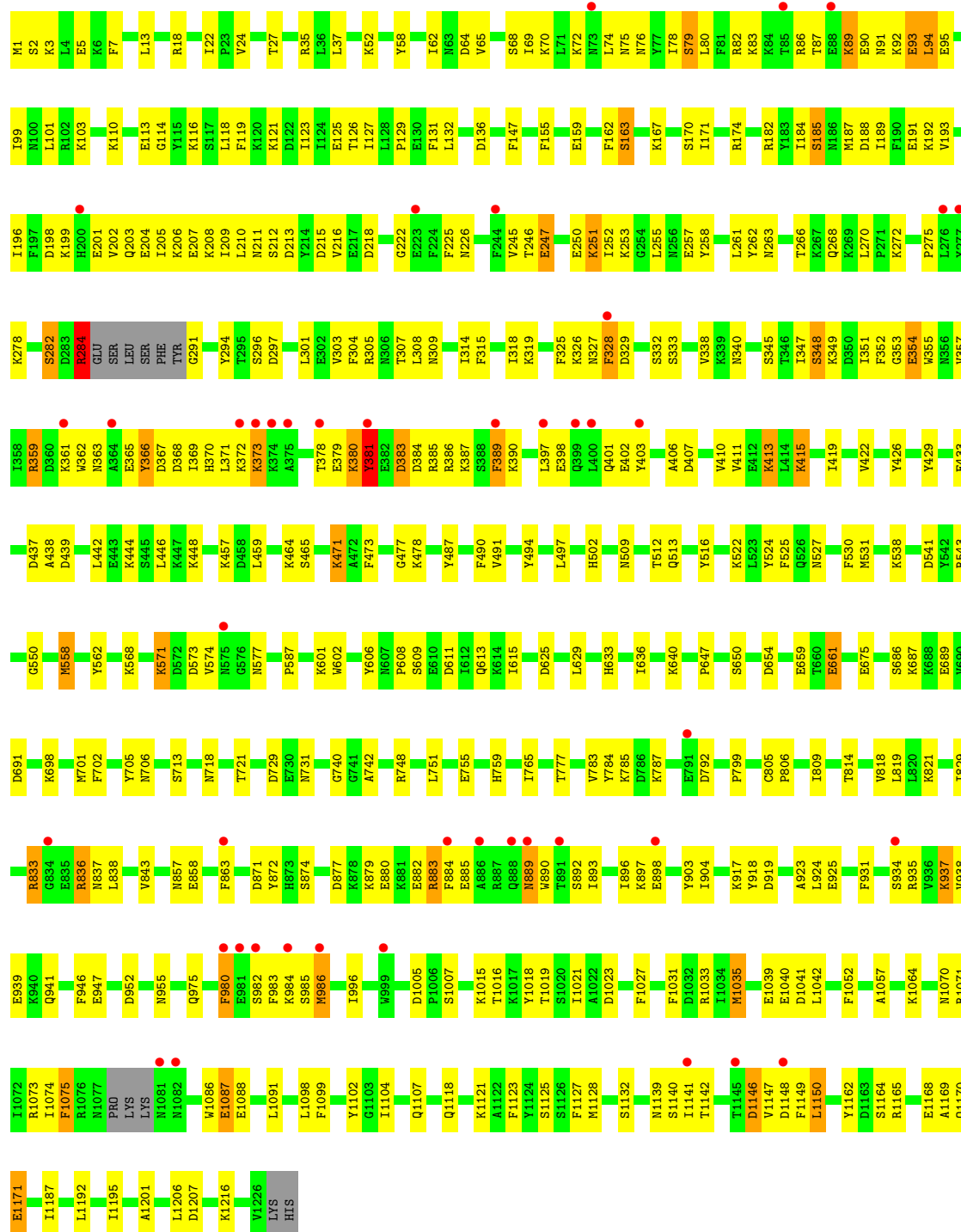
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: LbCas12a





• Molecule 1: LbCas12a



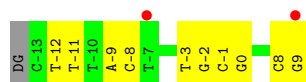
• Molecule 2: RNA (34-MER)



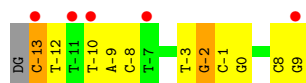
• Molecule 2: RNA (34-MER)



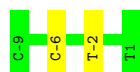
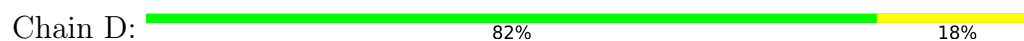
• Molecule 3: DNA (5'-D(P\*CP\*TP\*TP\*TP\*AP\*CP\*TP\*GP\*GP\*AP\*TP\*GP\*CP\*GP\*TP\*A  
P\*AP\*AP\*GP\*GP\*AP\*CP\*G)-3')



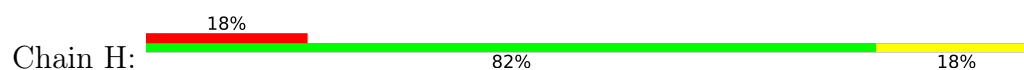
• Molecule 3: DNA (5'-D(P\*CP\*TP\*TP\*TP\*AP\*CP\*TP\*GP\*GP\*AP\*TP\*GP\*CP\*GP\*TP\*A  
P\*AP\*AP\*GP\*GP\*AP\*CP\*G)-3')



• Molecule 4: DNA (5'-D(\*CP\*GP\*TP\*CP\*CP\*TP\*TP\*TP\*AP\*TP\*T)-3')



• Molecule 4: DNA (5'-D(\*CP\*GP\*TP\*CP\*CP\*TP\*TP\*TP\*AP\*TP\*T)-3')





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	122.15Å 143.40Å 203.53Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.28 – 2.87 49.28 – 2.87	Depositor EDS
% Data completeness (in resolution range)	84.5 (49.28-2.87) 84.7 (49.28-2.87)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.64 (at 2.86Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, $R_{free}$	0.204 , 0.267 0.203 , 0.263	Depositor DCC
$R_{free}$ test set	4087 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	50.2	Xtriage
Anisotropy	0.051	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 49.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	22752	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.14% 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](#)

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

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	0.56	4/10158 (0.0%)	0.74	21/13664 (0.2%)
1	E	0.60	5/10172 (0.0%)	0.80	28/13683 (0.2%)
2	B	0.78	0/804	1.39	11/1250 (0.9%)
2	F	0.77	0/804	1.32	7/1250 (0.6%)
3	C	1.04	1/534 (0.2%)	1.13	1/823 (0.1%)
3	G	1.06	2/534 (0.4%)	1.14	0/823
4	D	1.12	0/240	1.27	1/368 (0.3%)
4	H	1.21	0/240	1.22	0/368
All	All	0.64	12/23486 (0.1%)	0.86	69/32229 (0.2%)

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	1
1	E	0	5
All	All	0	6

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	754	GLU	CB-CG	-8.68	1.35	1.52
1	E	985	SER	CA-C	-7.45	1.33	1.52
1	A	84	LYS	CD-CE	7.25	1.69	1.51
1	A	391	LYS	CD-CE	6.80	1.68	1.51
3	G	-13	DC	C1'-N1	6.72	1.57	1.49

The worst 5 of 69 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	754	GLU	CA-CB-CG	11.32	138.31	113.40
1	A	754	GLU	N-CA-CB	-11.29	90.28	110.60
1	E	381	TYR	CB-CG-CD2	-10.98	114.41	121.00
1	E	261	LEU	CA-CB-CG	9.86	137.97	115.30
1	E	359	ARG	CD-NE-CZ	9.27	136.57	123.60

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	376	VAL	Peptide
1	E	284	ARG	Sidechain
1	E	381	TYR	Sidechain
1	E	705	TYR	Peptide
1	E	93	GLU	Sidechain

## 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	9949	0	9787	287	0
1	E	9964	0	9793	356	0
2	B	719	0	366	19	0
2	F	719	0	366	21	0
3	C	476	0	260	7	0
3	G	476	0	260	9	0
4	D	217	0	129	1	0
4	H	217	0	129	1	0
5	A	2	0	0	0	0
5	B	1	0	0	0	0
5	E	1	0	0	0	0
5	F	1	0	0	0	0
6	A	8	0	0	0	0
6	E	2	0	0	1	0
All	All	22752	0	21090	672	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 15.

The worst 5 of 672 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:E:361:LYS:HD2	1:E:403:TYR:CD1	1.47	1.46
1:A:278:LYS:HE2	2:B:6:C:OP1	1.22	1.36
1:E:983:PHE:CE1	1:E:986:MET:HE2	1.75	1.22
1:E:370:HIS:HB3	1:E:381:TYR:OH	1.44	1.15
1:A:174:ARG:NH2	1:A:277:TYR:HB3	1.62	1.12

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	1208/1228 (98%)	1174 (97%)	30 (2%)	4 (0%)	37	63
1	E	1211/1228 (99%)	1177 (97%)	32 (3%)	2 (0%)	44	71
All	All	2419/2456 (98%)	2351 (97%)	62 (3%)	6 (0%)	44	71

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	89	LYS
1	A	588	GLY
1	A	89	LYS
1	A	1105	ASN
1	A	376	VAL

### 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	1081/1115 (97%)	1037 (96%)	44 (4%)	26	57
1	E	1082/1115 (97%)	1037 (96%)	45 (4%)	25	56
All	All	2163/2230 (97%)	2074 (96%)	89 (4%)	26	57

5 of 89 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	329	ASP
1	E	629	LEU
1	E	348	SER
1	E	464	LYS
1	E	874	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	268	GLN
1	E	327	ASN
1	E	1108	GLN
1	E	401	GLN
1	A	577	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	33/40 (82%)	7 (21%)	0
2	F	33/40 (82%)	7 (21%)	0
All	All	66/80 (82%)	14 (21%)	0

5 of 14 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	-17	U
2	B	-10	A
2	B	-9	A
2	B	-8	G
2	B	-6	G

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 oligosaccharides in this entry.

#### 5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 5 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 ⓘ

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	1214/1228 (98%)	-0.23	19 (1%) 70 65	14, 44, 80, 138	0
1	E	1217/1228 (99%)	-0.09	44 (3%) 46 41	15, 47, 95, 130	1 (0%)
2	B	34/40 (85%)	-0.49	1 (2%) 54 49	18, 26, 157, 168	0
2	F	34/40 (85%)	-0.42	1 (2%) 54 49	20, 25, 196, 203	0
3	C	23/24 (95%)	0.15	2 (8%) 17 15	21, 68, 153, 161	0
3	G	23/24 (95%)	0.44	5 (21%) 3 3	27, 73, 180, 190	0
4	D	11/11 (100%)	-0.47	0 100 100	27, 30, 89, 105	0
4	H	11/11 (100%)	-0.04	2 (18%) 4 4	32, 35, 103, 118	0
All	All	2567/2606 (98%)	-0.16	74 (2%) 54 49	14, 45, 93, 203	1 (0%)

The worst 5 of 74 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	999	TRP	5.5
1	A	1077	ASN	4.2
1	A	982	SER	4.0
1	E	381	TYR	3.7
1	A	291	GLY	3.6

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates ⓘ

There are no monosaccharides 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, 95<sup>th</sup> 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
5	LI	A	1302	1/1	0.09	0.97	8,8,8,8	0
5	LI	E	1301	1/1	0.36	0.75	10,10,10,10	0
5	LI	B	101	1/1	0.61	1.76	2,2,2,2	0
5	LI	A	1301	1/1	0.67	0.81	3,3,3,3	0
5	LI	F	101	1/1	0.86	0.79	5,5,5,5	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.