



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 23, 2024 – 04:17 PM JST

PDB ID : 8Y08  
Title : Crystal structure of LbCas12a in complex with crRNA and 14nt target DNA  
Authors : Lin, X.; Chen, J.; Liu, L.  
Deposited on : 2024-01-22  
Resolution : 3.64 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

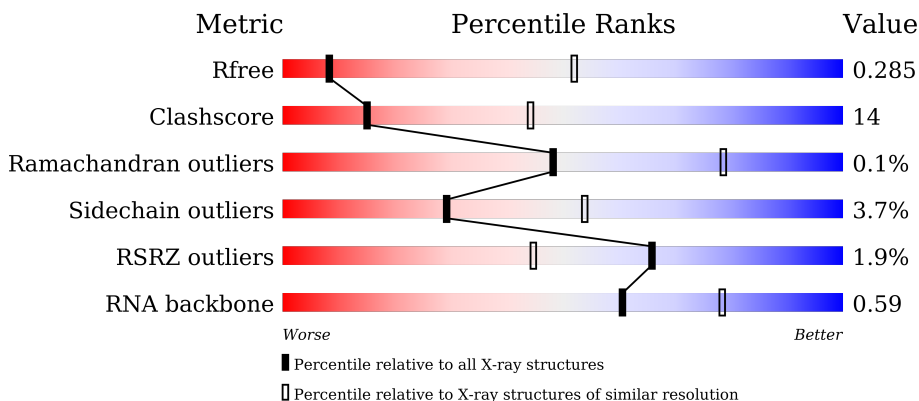
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

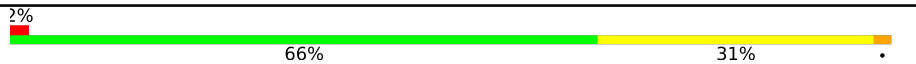



The reported resolution of this entry is 3.64 Å.

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))
$R_{free}$	164625	1000 (3.76-3.52)
Clashscore	180529	1046 (3.76-3.52)
Ramachandran outliers	177936	1031 (3.76-3.52)
Sidechain outliers	177891	1029 (3.76-3.52)
RSRZ outliers	164620	1682 (3.78-3.50)
RNA backbone	3690	1110 (4.26-3.00)

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	 2% 66% 31%
1	E	1228	 2% 71% 27%
2	B	40	 45% 40% 12%
2	F	40	 40% 45% 12%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	C	23	
3	G	23	
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

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 22796 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
			Total	C	N	O	S			
1	A	1223	9964	6412	1623	1900	29	0	0	0
1	E	1223	9964	6412	1623	1900	29	0	0	0

- Molecule 2 is a RNA chain called RNA (35-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	35	739	333	131	241	34	0	0	0
2	F	35	739	333	131	241	34	0	0	0

- 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
			Total	C	N	O	P			
3	C	23	476	226	89	138	23	0	0	0
3	G	23	476	226	89	138	23	0	0	0

- 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
			Total	C	N	O	P			
4	D	11	217	107	31	69	10	0	0	0
4	H	11	217	107	31	69	10	0	0	0

- 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

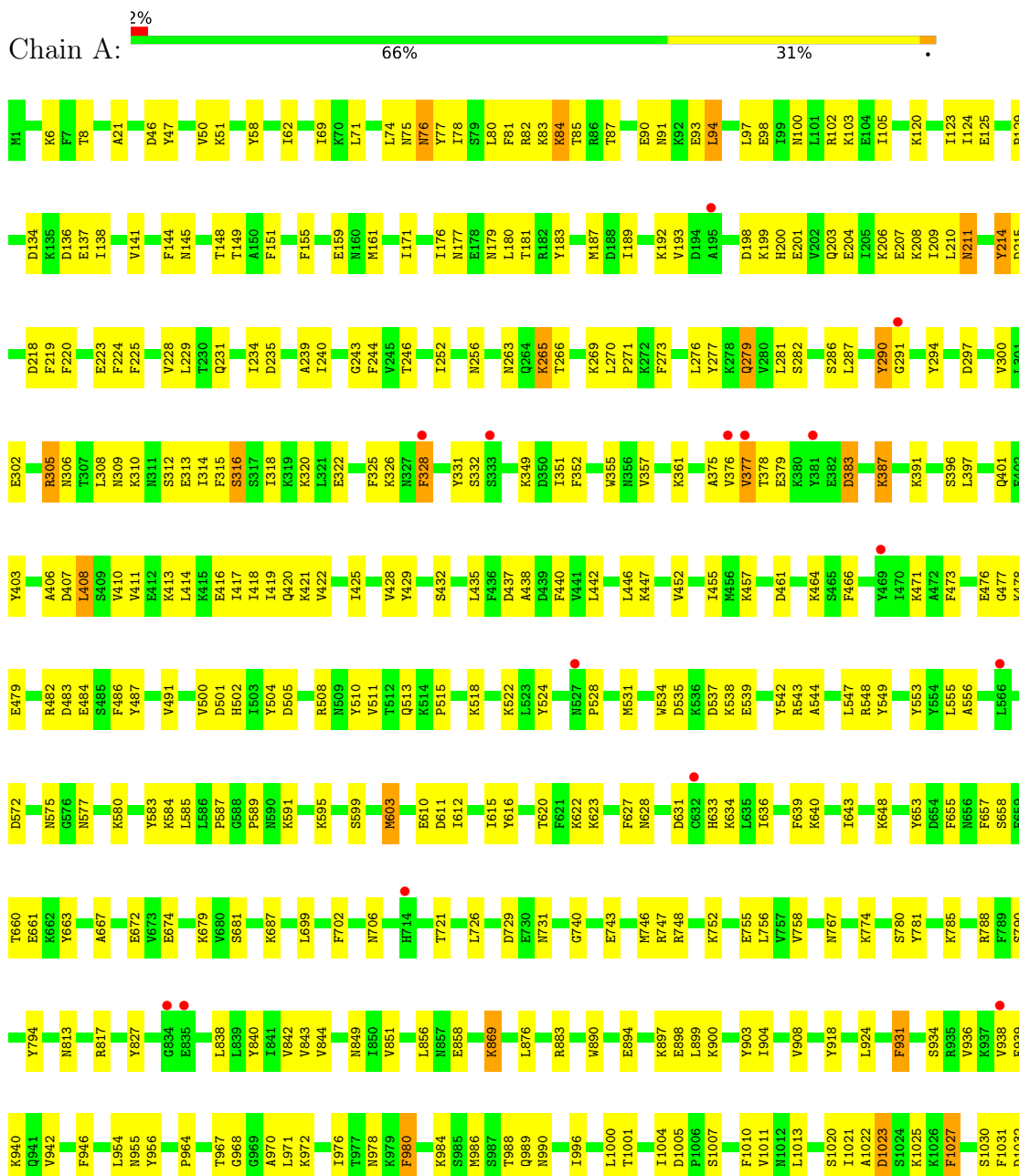
- Molecule 6 is water.

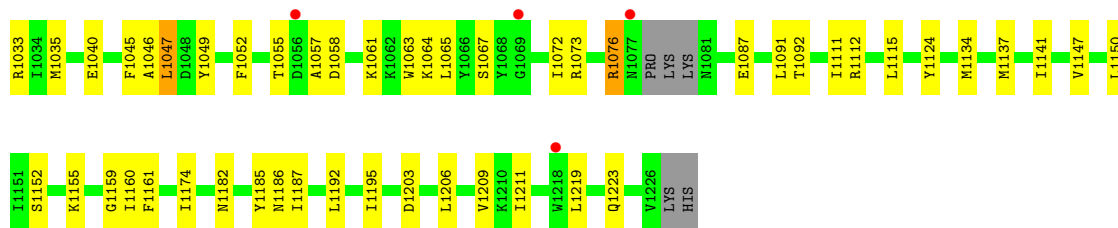
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	E	1	Total 1	O 1	0	0

### 3 Residue-property plots

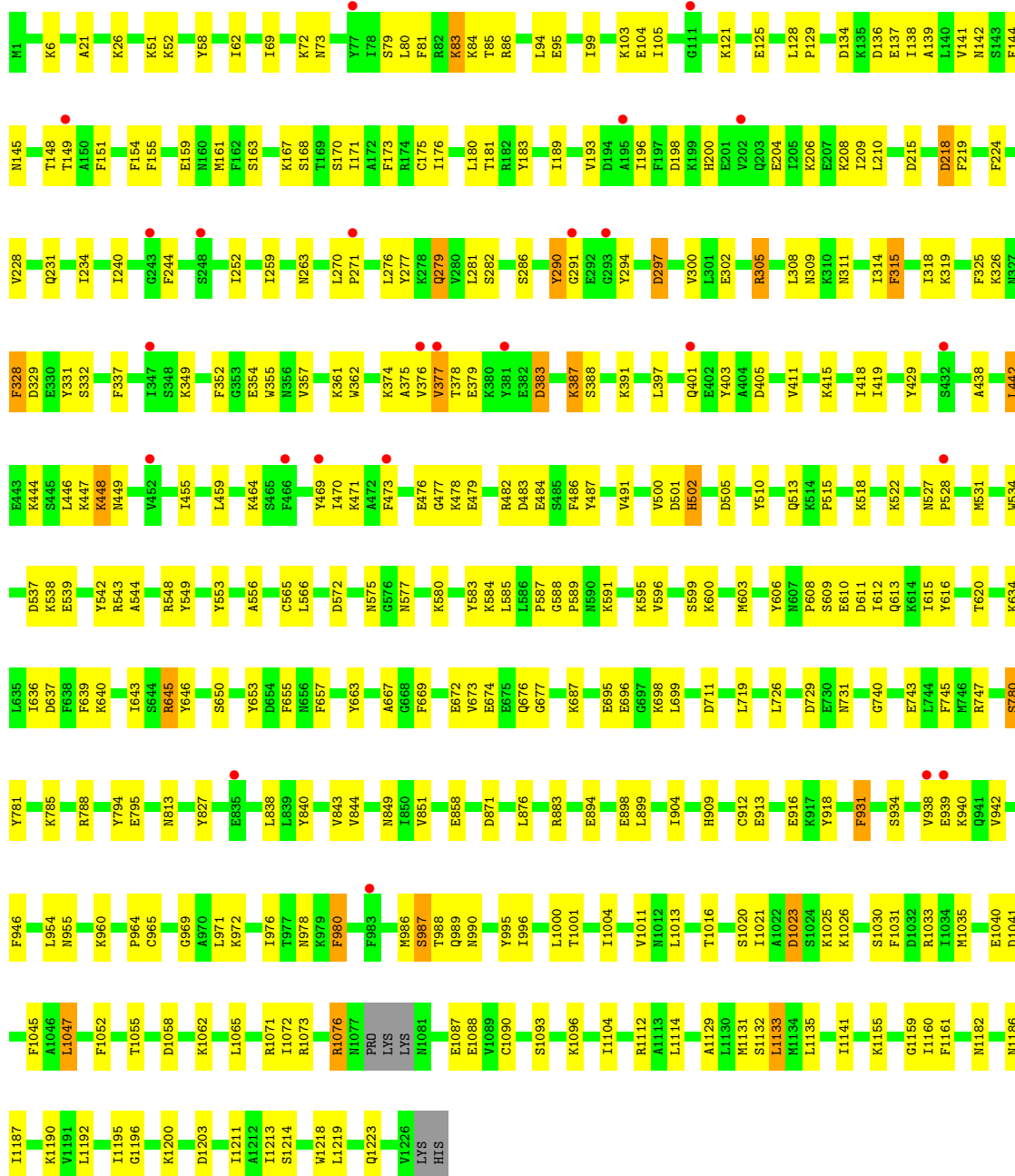
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density.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



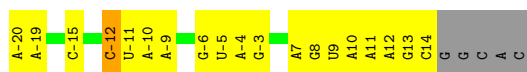
• Molecule 2: RNA (35-MER)

Chain B:  45% 40% 12%



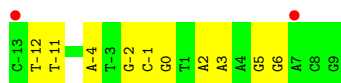
- Molecule 2: RNA (35-MER)

Chain F:  40% 45% 12%



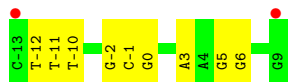
- Molecule 3: DNA (5'-D(P\*CP\*TP\*TP\*TP\*AP\*CP\*TP\*GP\*GP\*AP\*TP\*GP\*CP\*GP\*TP\*A P\*AP\*AP\*GP\*GP\*AP\*CP\*G)-3')

Chain C:  9% 57% 43%



- 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')

Chain G:  9% 61% 39%



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

Chain D:  55% 45%



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

Chain H:  9% 64% 36%





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	121.39Å 142.29Å 208.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	26.07 – 3.64 26.07 – 3.64	Depositor EDS
% Data completeness (in resolution range)	90.5 (26.07-3.64) 90.5 (26.07-3.64)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.67 (at 3.67Å)	Xtrriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, $R_{free}$	0.252 , 0.285 0.252 , 0.285	Depositor DCC
$R_{free}$ test set	2081 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	45.5	Xtrriage
Anisotropy	0.169	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.25 , 9.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.41$ , $\langle L^2 \rangle = 0.24$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.86	EDS
Total number of atoms	22796	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	49.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 46.88 % 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 1.0752e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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	0.28	0/10174	0.48	0/13698
1	E	0.28	0/10174	0.46	0/13698
2	B	0.22	0/826	0.75	0/1284
2	F	0.22	0/826	0.76	0/1284
3	C	0.53	0/534	0.91	0/823
3	G	0.52	0/534	0.92	0/823
4	D	0.63	0/240	1.11	0/368
4	H	0.58	0/240	1.10	0/368
All	All	0.30	0/23548	0.55	0/32346

There are no bond length outliers.

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	9964	0	9743	317	0
1	E	9964	0	9743	248	1
2	B	739	0	377	14	0
2	F	739	0	377	15	0
3	C	476	0	260	10	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	G	476	0	260	12	0
4	D	217	0	129	6	0
4	H	217	0	129	4	0
5	A	2	0	0	0	0
5	B	1	0	0	0	0
6	E	1	0	0	0	0
All	All	22796	0	21018	587	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 14.

The worst 5 of 587 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:1141:ILE:HB	1:A:1147:VAL:HG21	1.56	0.85
1:E:538:LYS:NZ	3:G:3:DA:N7	2.26	0.83
1:A:538:LYS:NZ	3:C:3:DA:N7	2.26	0.82
1:A:522:LYS:HE3	1:A:743:GLU:HG2	1.60	0.81
1:E:636:ILE:HD13	1:E:663:TYR:HD2	1.46	0.80

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)
1:E:332:SER:OG	1:E:637:ASP:OD2[2_545]	2.14	0.06

## 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	1219/1228 (99%)	1187 (97%)	31 (2%)	1 (0%)	48 78

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	E	1219/1228 (99%)	1186 (97%)	32 (3%)	1 (0%)	48 78
All	All	2438/2456 (99%)	2373 (97%)	63 (3%)	2 (0%)	48 78

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	377	VAL
1	A	377	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	1076/1115 (96%)	1037 (96%)	39 (4%)	30 55
1	E	1076/1115 (96%)	1036 (96%)	40 (4%)	29 54
All	All	2152/2230 (96%)	2073 (96%)	79 (4%)	29 54

5 of 79 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	448	LYS
1	E	931	PHE
1	E	476	GLU
1	E	645	ARG
1	E	1023	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	990	ASN
1	E	527	ASN
1	E	75	ASN
1	A	1197	GLN
1	E	142	ASN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	34/40 (85%)	4 (11%)	0
2	F	34/40 (85%)	4 (11%)	0
All	All	68/80 (85%)	8 (11%)	0

5 of 8 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	-12	C
2	B	-10	A
2	B	-9	A
2	B	-6	G
2	F	-12	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 oligosaccharides in this entry.

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

Of 3 ligands modelled in this entry, 3 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

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	1223/1228 (99%)	0.24	19 (1%) 70 49	7, 47, 90, 133	0
1	E	1223/1228 (99%)	0.22	25 (2%) 64 44	8, 47, 104, 136	0
2	B	35/40 (87%)	0.31	0 100 100	11, 19, 44, 60	0
2	F	35/40 (87%)	0.33	0 100 100	10, 21, 50, 64	0
3	C	23/23 (100%)	0.74	2 (8%) 17 13	24, 44, 98, 133	0
3	G	23/23 (100%)	0.69	2 (8%) 17 13	26, 43, 101, 144	0
4	D	11/11 (100%)	0.72	0 100 100	42, 47, 102, 108	0
4	H	11/11 (100%)	1.02	1 (9%) 16 12	49, 56, 117, 124	0
All	All	2584/2604 (99%)	0.25	49 (1%) 66 45	7, 47, 100, 144	0

The worst 5 of 49 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	376	VAL	4.2
1	A	376	VAL	4.2
1	E	347	ILE	3.9
1	E	939	GLU	3.6
1	E	381	TYR	3.3

### 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

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(Å <sup>2</sup> )	Q<0.9
5	LI	A	1302	1/1	0.11	0.91	10,10,10,10	0
5	LI	A	1301	1/1	0.26	1.41	9,9,9,9	0
5	LI	B	101	1/1	0.87	1.01	5,5,5,5	0

## 6.5 Other polymers

There are no such residues in this entry.