wwPDB X-ray Structure Validation Summary Report (i)

Jun 24, 2024 – 07:10 AM EDT

PDB ID	:	5XWP
Title	:	Crystal structure of LbuCas13a-crRNA-target RNA ternary complex
Authors	:	Liu, L.; Li, X.; Li, Z.; Wang, Y.
Deposited on	:	2017-06-30
Resolution	:	3.09 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 (i) 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 (1)) were used in the production of this report:

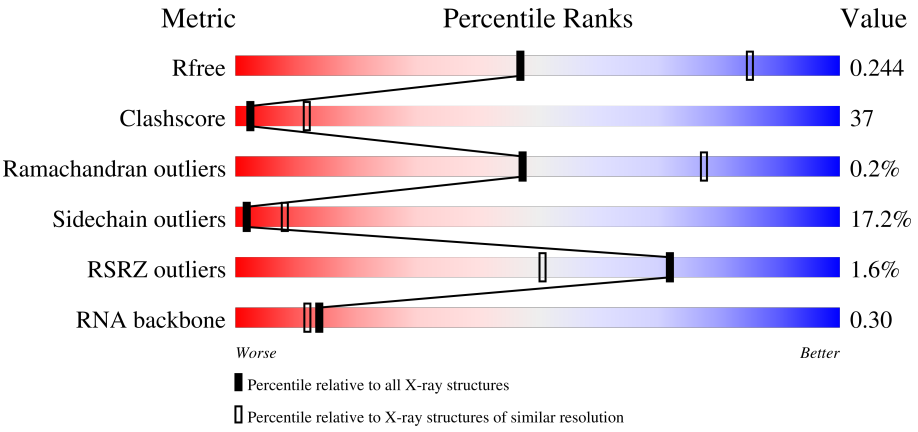
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.09 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.



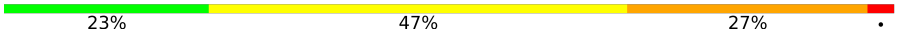
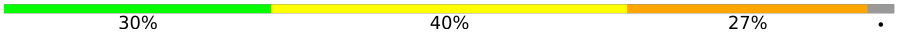
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1447 (3.10-3.06)
Clashscore	141614	1546 (3.10-3.06)
Ramachandran outliers	138981	1487 (3.10-3.06)
Sidechain outliers	138945	1486 (3.10-3.06)
RSRZ outliers	127900	1416 (3.10-3.06)
RNA backbone	3102	1063 (3.36-2.80)

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 <=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	1160	46% 42% 9%
1	B	1160	44% 42% 9%
2	C	59	20% 37% 37%
2	E	59	25% 42% 29%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	30	
3	F	30	

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 21838 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 Uncharacterized protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	1125	Total	C	N	O	S	Se	0	0	0
			9160	5877	1539	1720	3	21			
1	B	1114	Total	C	N	O	S	Se	0	0	0
			8940	5718	1508	1690	3	21			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	expression tag	UNP C7NBY4
A	1048	ALA	ARG	engineered mutation	UNP C7NBY4
A	1053	ALA	HIS	engineered mutation	UNP C7NBY4
B	0	SER	-	expression tag	UNP C7NBY4
B	1048	ALA	ARG	engineered mutation	UNP C7NBY4
B	1053	ALA	HIS	engineered mutation	UNP C7NBY4

- Molecule 2 is a RNA chain called RNA (59-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	58	Total	C	N	O	P	0	0	0
			1229	553	224	394	58			
2	E	58	Total	C	N	O	P	0	0	0
			1229	553	224	394	58			

- Molecule 3 is a RNA chain called RNA (30-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	30	Total	C	N	O	P	0	0	0
			644	288	116	210	30			
3	F	29	Total	C	N	O	P	0	0	0
			624	279	113	203	29			

- Molecule 4 is water.

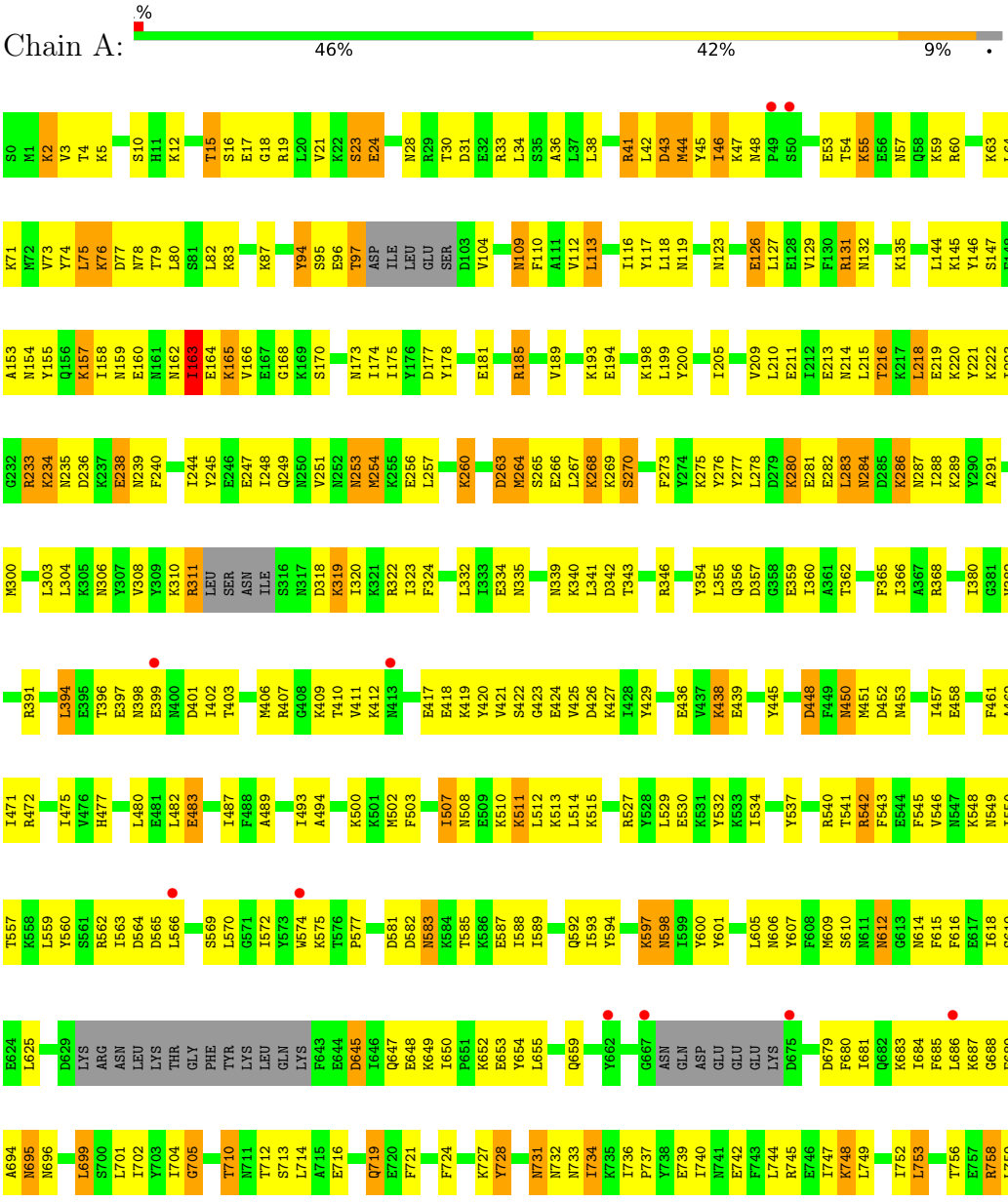
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	10	Total O 10 10	0	0
4	B	2	Total O 2 2	0	0

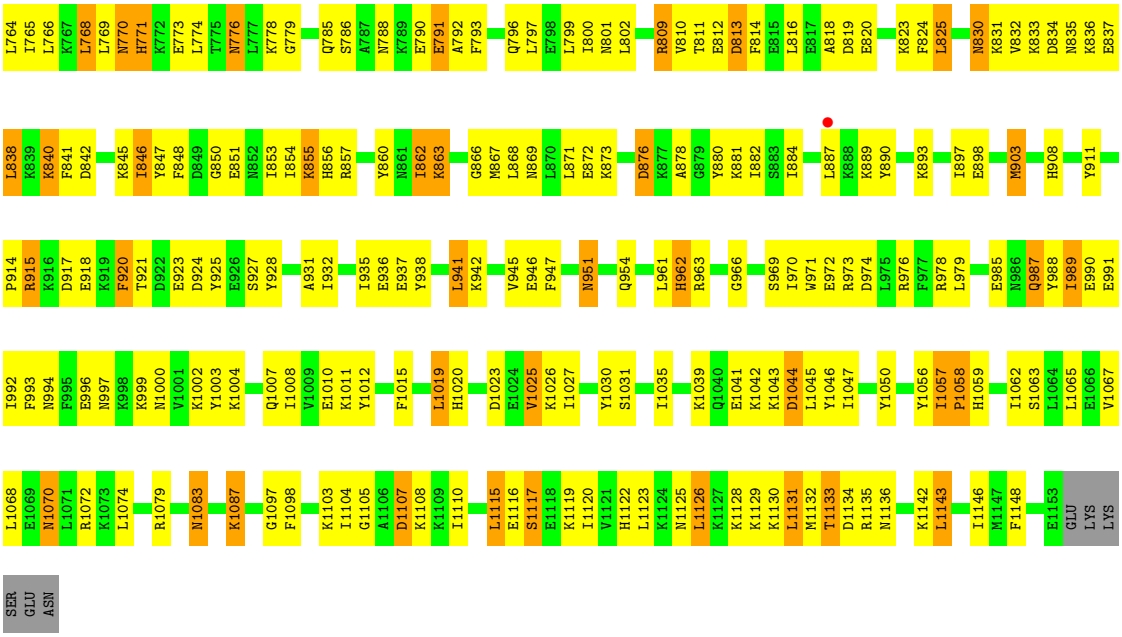
3 Residue-property plots (i)

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: Uncharacterized protein

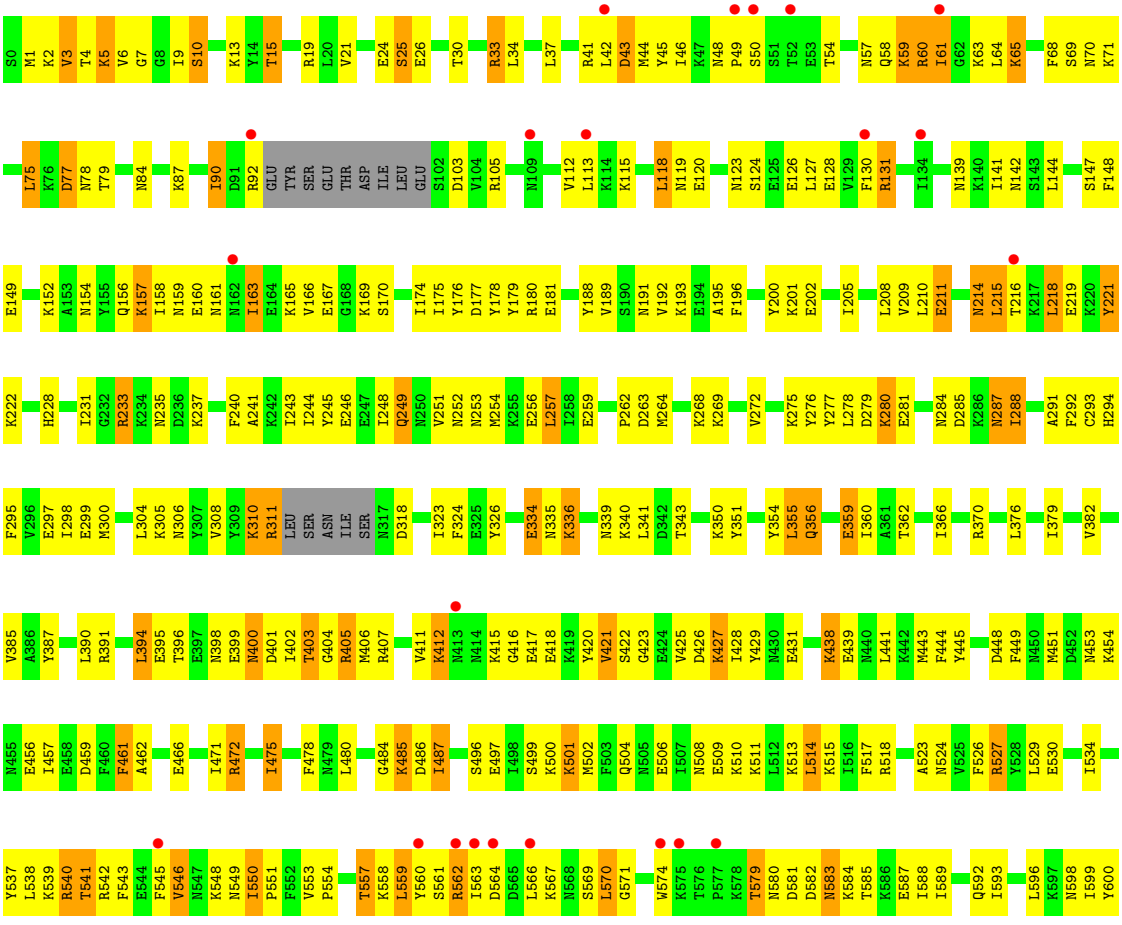
Chain A:

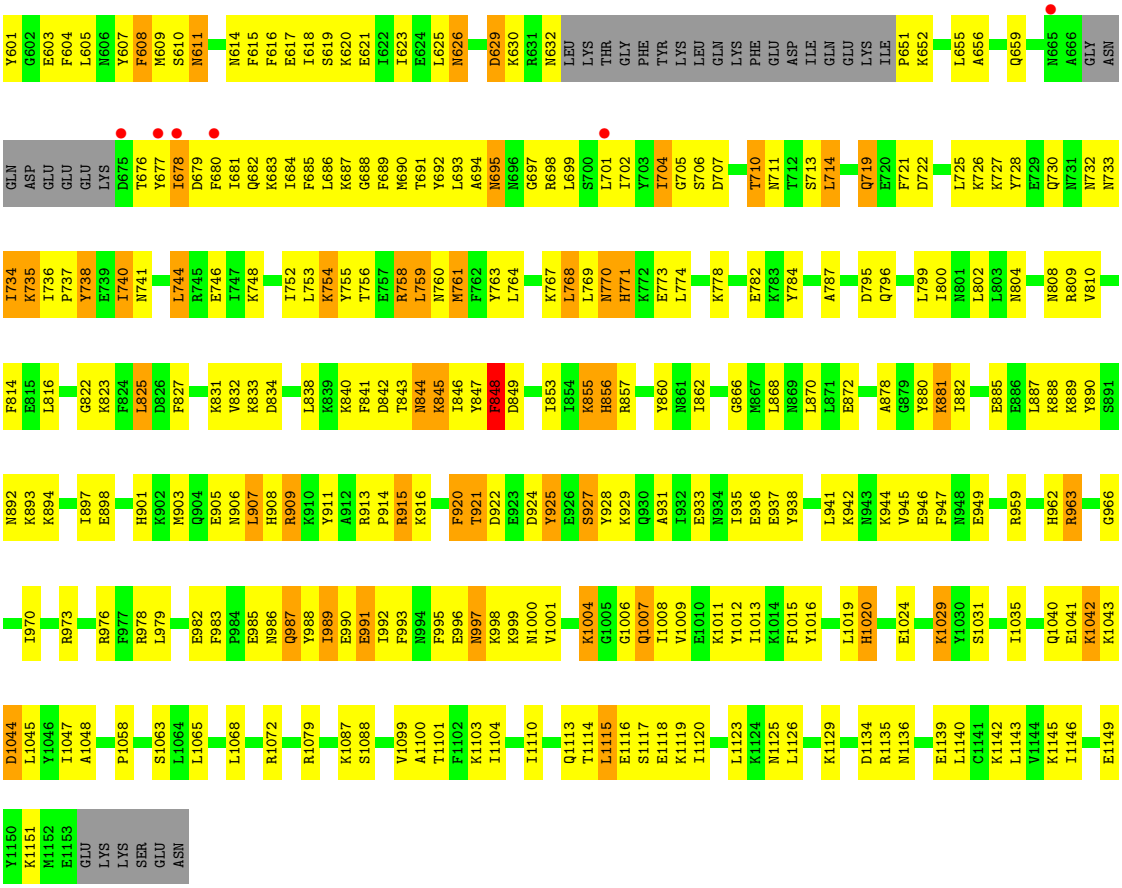




• Molecule 1: Uncharacterized protein

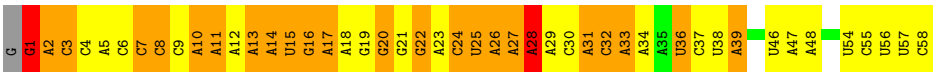
Chain B:





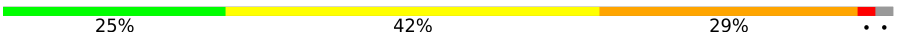
• Molecule 2: RNA (59-MER)

Chain C:



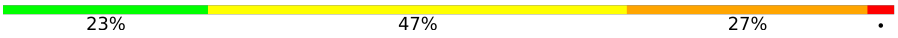
• Molecule 2: RNA (59-MER)

Chain E:



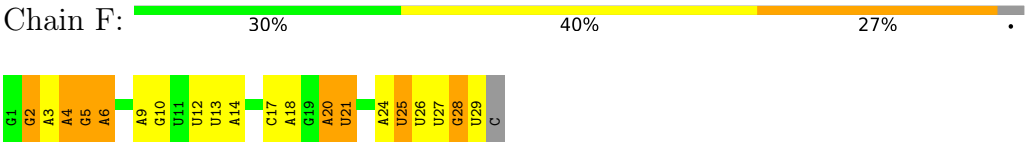
• Molecule 3: RNA (30-MER)

Chain D:



• Molecule 3: RNA (30-MER)

Chain F:


30%
40%
27%
G1
G2
A3
A4
G5
A6
A9
G10
U11
U12
U13
A14
C17
A18
G19
A20
U21
A24
U25
U26
U27
G28
U29
C

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	95.58Å 132.87Å 139.96Å 90.00° 90.80° 90.00°	Depositor
Resolution (Å)	48.18 – 3.09 48.18 – 3.09	Depositor EDS
% Data completeness (in resolution range)	95.1 (48.18-3.09) 91.3 (48.18-3.09)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	6.18 (at 3.07Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155)	Depositor
R, R_{free}	0.218 , 0.244 0.220 , 0.244	Depositor DCC
R_{free} test set	2881 reflections (4.70%)	wwPDB-VP
Wilson B-factor (Å^2)	58.5	Xtriage
Anisotropy	0.074	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , -23.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.36$, $< L^2 > = 0.18$	Xtriage
Estimated twinning fraction	0.186 for h,-k,-l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	21838	wwPDB-VP
Average B, all atoms (Å^2)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.45% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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	0.70	0/9290	0.79	10/12431 (0.1%)
1	B	0.67	0/9063	0.83	15/12146 (0.1%)
2	C	0.86	4/1375 (0.3%)	0.81	1/2137 (0.0%)
2	E	0.79	1/1375 (0.1%)	0.78	0/2137
3	D	0.72	3/721 (0.4%)	0.76	1/1122 (0.1%)
3	F	0.75	2/699 (0.3%)	0.75	1/1088 (0.1%)
All	All	0.71	10/22523 (0.0%)	0.81	28/31061 (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
1	A	0	2
1	B	0	2
All	All	0	4

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	2	A	O3'-P	-9.91	1.49	1.61
2	C	2	A	O3'-P	-9.17	1.50	1.61
2	C	1	G	O3'-P	-6.05	1.53	1.61
3	F	12	U	O3'-P	-5.86	1.54	1.61
3	D	2	G	C4'-O4'	5.58	1.52	1.45

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	403	THR	N-CA-C	-14.91	70.75	111.00
1	B	570	LEU	N-CA-C	12.38	144.44	111.00
1	B	678	ILE	CB-CA-C	-8.31	94.97	111.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	570	LEU	CB-CA-C	-8.22	94.59	110.20
1	B	65	LYS	CB-CA-C	-7.21	95.99	110.40

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1025	VAL	Peptide
1	A	705	GLY	Peptide
1	B	421	VAL	Peptide
1	B	695	ASN	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	9160	0	9025	680	0
1	B	8940	0	8665	681	0
2	C	1229	0	627	87	0
2	E	1229	0	627	95	0
3	D	644	0	321	40	0
3	F	624	0	310	36	0
4	A	10	0	0	0	0
4	B	2	0	0	0	0
All	All	21838	0	19575	1546	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 37.

The worst 5 of 1546 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:566:LEU:HD12	1:A:686:LEU:CD1	1.58	1.30
1:A:406:MSE:CE	1:A:461:PHE:HB3	1.62	1.30
1:B:920:PHE:O	1:B:921:THR:CG2	1.82	1.24

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:265:SER:HB3	1:A:268:LYS:HG2	1.24	1.19
1:B:846:ILE:HD11	1:B:887:LEU:HD21	1.21	1.15

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	1115/1160 (96%)	1079 (97%)	34 (3%)	2 (0%)	47 77
1	B	1104/1160 (95%)	1064 (96%)	38 (3%)	2 (0%)	47 77
All	All	2219/2320 (96%)	2143 (97%)	72 (3%)	4 (0%)	47 77

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	77	ASP
1	B	921	THR
1	B	1024	GLU
1	A	914	PRO

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.

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	967/1061 (91%)	800 (83%)	167 (17%)	2 8
1	B	925/1061 (87%)	767 (83%)	158 (17%)	2 8
All	All	1892/2122 (89%)	1567 (83%)	325 (17%)	2 8

5 of 325 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	443	MSE
1	B	799	LEU
1	B	499	SER
1	B	608	PHE
1	B	885	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 54 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	70	ASN
1	B	271	GLN
1	B	908	HIS
1	B	84	ASN
1	B	214	ASN

5.3.3 RNA

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	C	58/59 (98%)	26 (44%)	5 (8%)
2	E	57/59 (96%)	22 (38%)	4 (7%)
3	D	29/30 (96%)	10 (34%)	0
3	F	28/30 (93%)	10 (35%)	2 (7%)
All	All	172/178 (96%)	68 (39%)	11 (6%)

5 of 68 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	C	3	C
2	C	4	C
2	C	5	A

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	C	7	C
2	C	8	C

5 of 11 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	E	31	A
2	E	32	C
3	F	4	A
3	F	2	G
2	C	30	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 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 (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^{th} 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	1104/1160 (95%)	-0.30	11 (0%) 82 66	17, 43, 82, 138	0
1	B	1093/1160 (94%)	-0.17	28 (2%) 56 31	20, 51, 113, 149	0
2	C	58/59 (98%)	-0.49	0 100 100	22, 36, 68, 81	0
2	E	58/59 (98%)	-0.55	0 100 100	29, 43, 82, 92	0
3	D	30/30 (100%)	-0.53	0 100 100	26, 36, 75, 105	0
3	F	29/30 (96%)	-0.63	0 100 100	30, 40, 82, 122	0
All	All	2372/2498 (94%)	-0.26	39 (1%) 72 51	17, 46, 98, 149	0

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	574	TRP	10.1
1	B	563	ILE	6.8
1	A	49	PRO	6.7
1	B	675	ASP	6.4
1	B	665	ASN	6.0

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

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.