



wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 10:36 PM GMT

PDB ID : 3M7N
Title : archaeoglobus fulgidus exosome with RNA bound to the active site
Authors : Hartung, S.; Hopfner, K.-P.
Deposited on : 2010-03-16
Resolution : 2.40 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

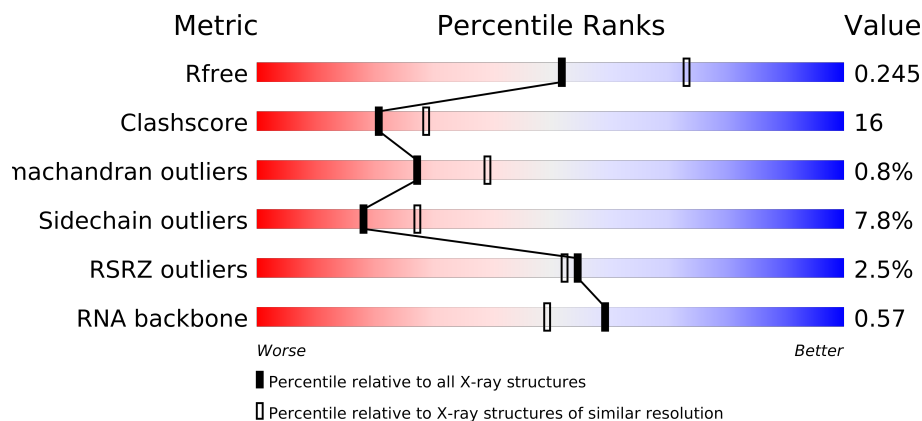
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.40 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}	66092	2207 (2.40-2.40)
Clashscore	79885	2789 (2.40-2.40)
Ramachandran outliers	78287	2736 (2.40-2.40)
Sidechain outliers	78261	2737 (2.40-2.40)
RSRZ outliers	66119	2210 (2.40-2.40)
RNA backbone	1838	1029 (3.00-1.80)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	179	
1	B	179	
1	C	179	
2	D	258	
2	E	258	
2	F	258	
3	G	259	
3	H	259	
3	I	259	
4	X	6	
4	Y	6	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	Z	6	 A horizontal bar chart representing the quality of chain 4. The bar is divided into three segments: a red segment on the left (approximately 15% of the total length), a green segment in the middle (approximately 75% of the total length), and a grey segment on the right (approximately 10% of the total length).

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 16859 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Putative uncharacterized protein AF_0206.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	179	Total	C	N	O	S	28	0	0
			1371	855	242	266	8			
1	B	179	Total	C	N	O	S	44	0	0
			1371	855	242	266	8			
1	C	179	Total	C	N	O	S	31	0	0
			1371	855	242	266	8			

- Molecule 2 is a protein called Probable exosome complex exonuclease 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	243	Total	C	N	O	S	12	0	0
			1902	1198	334	357	13			
2	E	248	Total	C	N	O	S	7	0	0
			1944	1224	340	367	13			
2	F	246	Total	C	N	O	S	16	0	0
			1926	1213	337	363	13			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	65	GLU	ARG	ENGINEERED	UNP O29757
E	65	GLU	ARG	ENGINEERED	UNP O29757
F	65	GLU	ARG	ENGINEERED	UNP O29757

- Molecule 3 is a protein called Probable exosome complex exonuclease 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	254	Total	C	N	O	S	9	0	0
			1971	1245	327	394	5			
3	H	258	Total	C	N	O	S	14	0	0
			2005	1266	331	403	5			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	I	258	Total	C	N	O	S	11	0	0
			2005	1266	331	403	5			

- Molecule 4 is a RNA chain called 5'-R(*C*UP*CP*CP*CP*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	X	5	Total	C	N	O	P	0	0	0
			97	45	14	34	4			
4	Y	5	Total	C	N	O	P	0	0	0
			100	45	14	36	5			
4	Z	4	Total	C	N	O	P	0	0	0
			77	36	12	26	3			

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	Zn	0	0
			1	1		
5	A	1	Total	Zn	0	0
			1	1		
5	C	1	Total	Zn	0	0
			1	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	30	Total	O	0	0
			30	30		
6	B	7	Total	O	0	0
			7	7		
6	C	38	Total	O	0	0
			38	38		
6	D	63	Total	O	0	0
			63	63		
6	E	99	Total	O	0	0
			99	99		
6	F	136	Total	O	0	0
			136	136		
6	G	66	Total	O	0	0
			66	66		
6	H	126	Total	O	0	0
			126	126		

Continued on next page...

Continued from previous page...

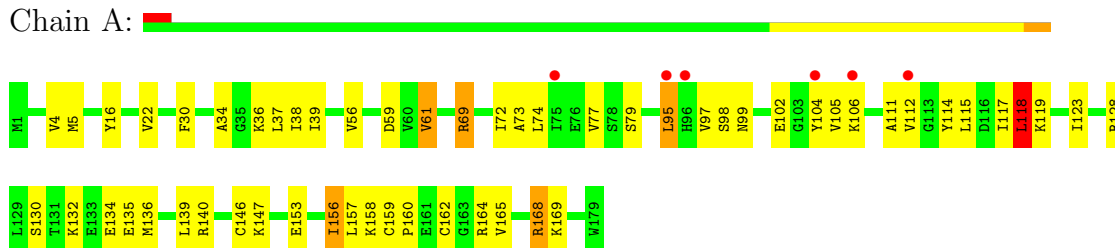
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	I	136	Total 136	O 136	0	0
6	X	5	Total 5	O 5	0	0
6	Y	7	Total 7	O 7	0	0
6	Z	3	Total 3	O 3	0	0

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

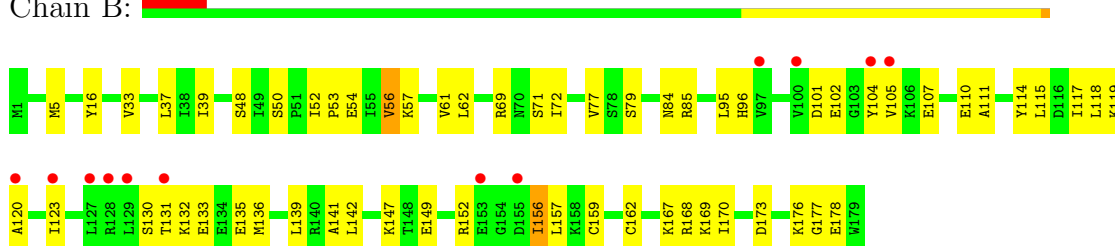
- Molecule 1: Putative uncharacterized protein AF_0206

Chain A:



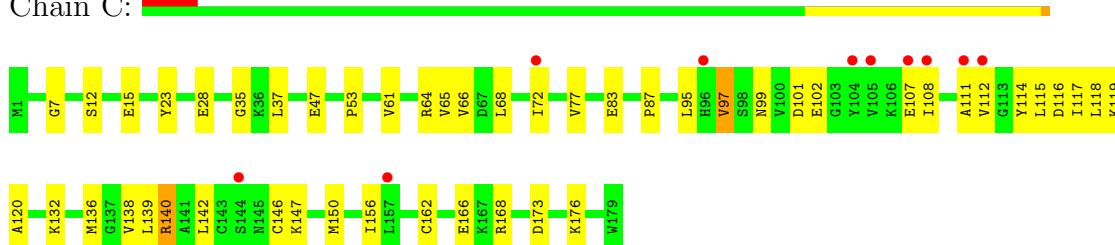
- Molecule 1: Putative uncharacterized protein AF_0206

Chain B:



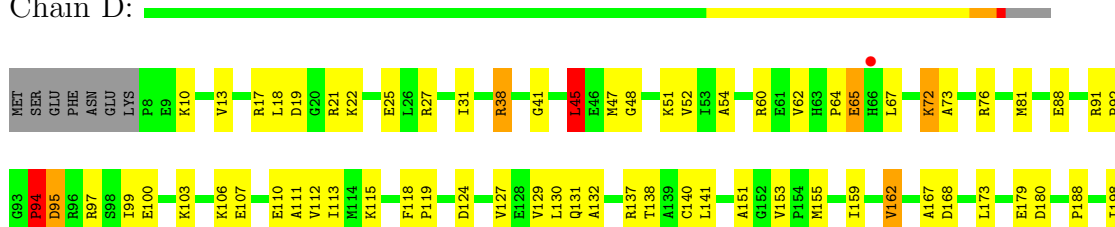
- Molecule 1: Putative uncharacterized protein AF_0206

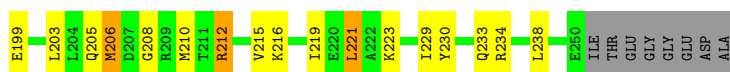
Chain C:



- Molecule 2: Probable exosome complex exonuclease 1

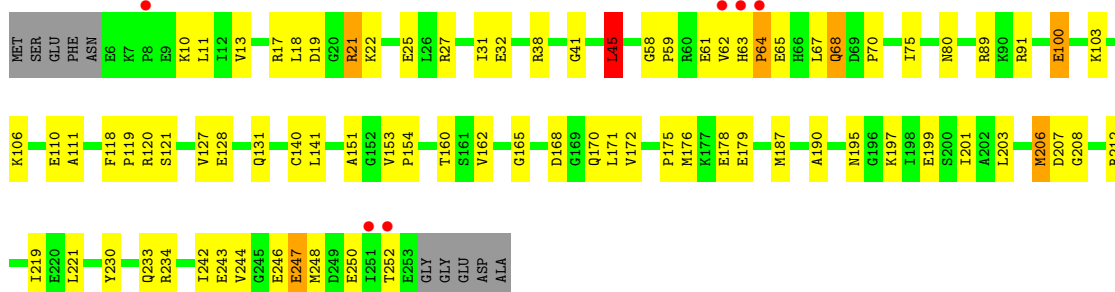
Chain D:





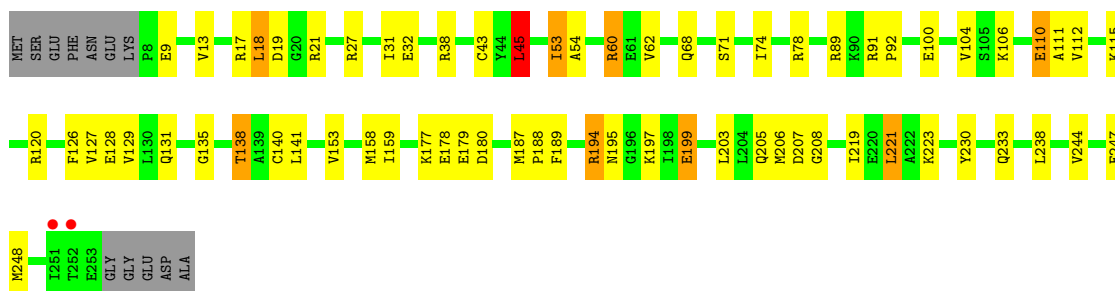
• Molecule 2: Probable exosome complex exonuclease 1

Chain E:



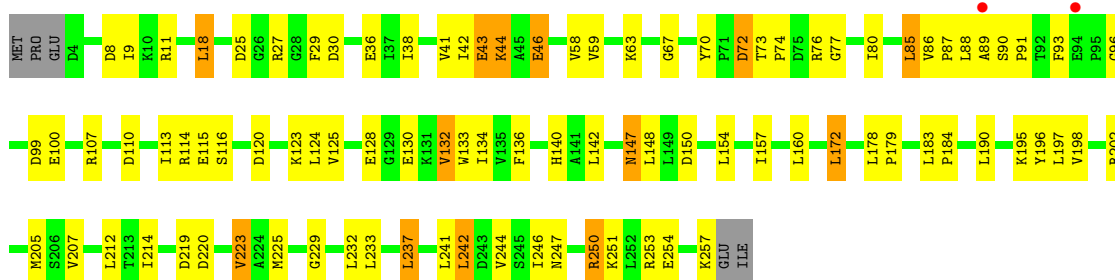
• Molecule 2: Probable exosome complex exonuclease 1

Chain F:



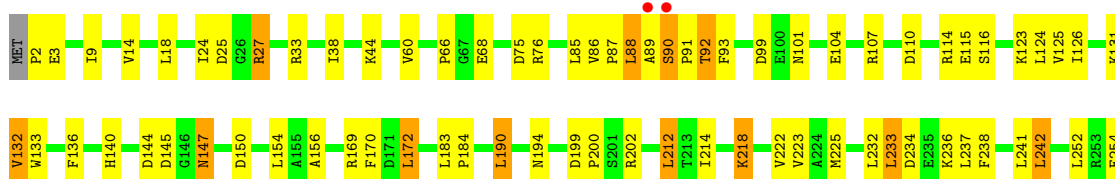
• Molecule 3: Probable exosome complex exonuclease 2

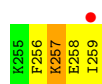
Chain G:



• Molecule 3: Probable exosome complex exonuclease 2

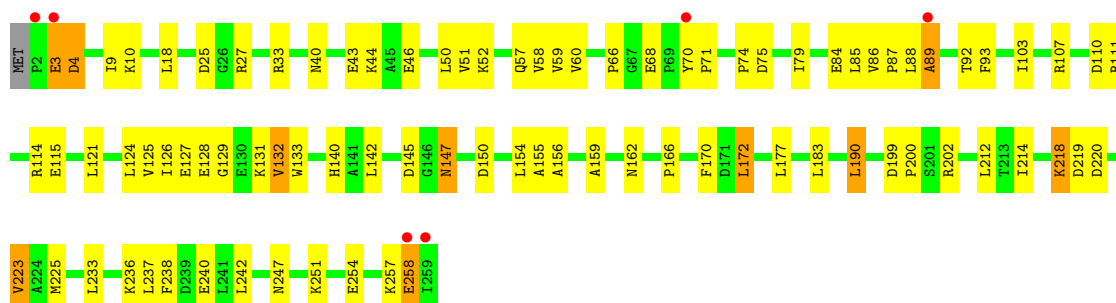
Chain H:





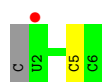
- Molecule 3: Probable exosome complex exonuclease 2

Chain I:



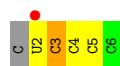
- Molecule 4: 5'-R(*C*UP*CP*CP*CP*C)-3'

Chain X:



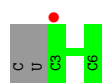
- Molecule 4: 5'-R(*C*UP*CP*CP*CP*C)-3'

Chain Y:



- Molecule 4: 5'-R(*C*UP*CP*CP*CP*C)-3'

Chain Z:



4 Data and refinement statistics

Property	Value	Source
Space group	P 43 2 2	Depositor
Cell constants a, b, c, α , β , γ	138.27Å 138.27Å 262.27Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.06 – 2.40 48.89 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.7 (48.06-2.40) 99.7 (48.89-2.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.88 (at 2.39Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.4.129)	Depositor
R, R_{free}	0.197 , 0.256 0.187 , 0.245	Depositor DCC
R_{free} test set	5022 reflections (5.05%)	DCC
Wilson B-factor (Å ²)	42.3	Xtriage
Anisotropy	0.206	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 41.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	2 of 99465 reflections (0.002%)	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	16859	wwPDB-VP
Average B, all atoms (Å ²)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.89% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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	0.31	0/1385	0.53	1/1855 (0.1%)
1	B	0.29	0/1385	0.48	0/1855
1	C	0.33	0/1385	0.52	0/1855
2	D	0.42	0/1929	0.61	1/2588 (0.0%)
2	E	0.45	0/1971	0.62	1/2645 (0.0%)
2	F	0.47	0/1953	0.63	1/2621 (0.0%)
3	G	0.39	0/1999	0.57	0/2713
3	H	0.46	0/2034	0.64	1/2759 (0.0%)
3	I	0.45	0/2034	0.62	0/2759
4	X	0.67	0/106	1.14	0/162
4	Y	0.67	0/109	1.20	0/166
4	Z	0.55	0/84	1.24	0/128
All	All	0.42	0/16374	0.61	5/22106 (0.0%)

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
2	D	0	1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	45	LEU	CA-CB-CG	8.24	134.26	115.30
2	F	45	LEU	CA-CB-CG	6.66	130.62	115.30
2	E	45	LEU	CA-CB-CG	6.38	129.97	115.30
1	A	118	LEU	CA-CB-CG	5.71	128.43	115.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	H	27	ARG	NE-CZ-NH1	-5.38	117.61	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	94	PRO	Peptide

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1371	0	1409	35	0
1	B	1371	0	1410	43	0
1	C	1371	0	1409	32	0
2	D	1902	0	1948	69	0
2	E	1944	0	1990	83	0
2	F	1926	0	1972	65	0
3	G	1971	0	1994	90	0
3	H	2005	0	2025	73	0
3	I	2005	0	2025	79	0
4	X	97	0	56	2	0
4	Y	100	0	55	4	0
4	Z	77	0	46	0	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
5	C	1	0	0	0	0
6	A	30	0	0	0	0
6	B	7	0	0	0	0
6	C	38	0	0	1	0
6	D	63	0	0	0	0
6	E	99	0	0	6	0
6	F	136	0	0	8	0
6	G	66	0	0	2	0
6	H	126	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	I	136	0	0	1	0
6	X	5	0	0	0	0
6	Y	7	0	0	1	0
6	Z	3	0	0	0	0
All	All	16859	0	16339	500	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 16.

The worst 5 of 500 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:E:61:GLU:HB3	2:E:63:HIS:HB3	1.34	1.06
2:E:62:VAL:H	2:E:63:HIS:HB2	1.25	0.99
2:E:19:ASP:OD2	2:E:21:ARG:HD3	1.63	0.98
3:G:43:GLU:O	3:G:44:LYS:HB2	1.65	0.96
3:G:214:ILE:HG12	3:G:225:MET:HE3	1.48	0.94

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	177/179 (99%)	167 (94%)	9 (5%)	1 (1%)	33	47
1	B	177/179 (99%)	151 (85%)	24 (14%)	2 (1%)	21	29
1	C	177/179 (99%)	163 (92%)	14 (8%)	0	100	100
2	D	241/258 (93%)	224 (93%)	13 (5%)	4 (2%)	14	17
2	E	246/258 (95%)	230 (94%)	15 (6%)	1 (0%)	43	61
2	F	244/258 (95%)	239 (98%)	4 (2%)	1 (0%)	43	61
3	G	252/259 (97%)	236 (94%)	13 (5%)	3 (1%)	19	26
3	H	256/259 (99%)	239 (93%)	15 (6%)	2 (1%)	27	39
3	I	256/259 (99%)	244 (95%)	9 (4%)	3 (1%)	19	26

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	2026/2088 (97%)	1893 (93%)	116 (6%)	17 (1%)	27	39

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	64	PRO
3	I	89	ALA
3	I	128	GLU
1	A	106	LYS
1	B	169	LYS

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	149/149 (100%)	133 (89%)	16 (11%)	10	13
1	B	149/149 (100%)	141 (95%)	8 (5%)	31	47
1	C	149/149 (100%)	141 (95%)	8 (5%)	31	47
2	D	202/214 (94%)	183 (91%)	19 (9%)	13	18
2	E	207/214 (97%)	192 (93%)	15 (7%)	21	31
2	F	205/214 (96%)	189 (92%)	16 (8%)	18	27
3	G	222/227 (98%)	202 (91%)	20 (9%)	14	20
3	H	226/227 (100%)	210 (93%)	16 (7%)	21	32
3	I	226/227 (100%)	209 (92%)	17 (8%)	19	29
All	All	1735/1770 (98%)	1600 (92%)	135 (8%)	18	27

5 of 135 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	199	GLU
2	F	138	THR
3	I	154	LEU
2	E	206	MET
2	F	53	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
2	E	228	GLN
2	F	228	GLN
3	I	147	ASN
2	E	233	GLN
2	F	131	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
4	X	4/6 (66%)	0	0
4	Y	4/6 (66%)	1 (25%)	0
4	Z	3/6 (50%)	0	0
All	All	11/18 (61%)	1 (9%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
4	Y	3	C

There are no RNA pucker outliers to report.

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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

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.

5.7 Other polymers ⓘ

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 ⓘ

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, 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	179/179 (100%)	0.16	6 (3%) 43 41	40, 62, 86, 120	9 (5%)
1	B	179/179 (100%)	0.40	12 (6%) 17 16	36, 83, 118, 135	12 (6%)
1	C	179/179 (100%)	0.25	10 (5%) 24 22	27, 57, 92, 117	8 (4%)
2	D	243/258 (94%)	-0.17	1 (0%) 90 90	28, 41, 69, 92	6 (2%)
2	E	248/258 (96%)	-0.22	6 (2%) 56 54	25, 37, 77, 109	2 (0%)
2	F	246/258 (95%)	-0.39	2 (0%) 83 82	22, 32, 71, 99	4 (1%)
3	G	254/259 (98%)	-0.20	2 (0%) 83 82	32, 48, 76, 88	2 (0%)
3	H	258/259 (99%)	-0.21	3 (1%) 75 75	22, 33, 74, 86	5 (1%)
3	I	258/259 (99%)	-0.12	6 (2%) 57 55	22, 34, 73, 103	3 (1%)
4	X	5/6 (83%)	0.64	1 (20%) 2 1	53, 55, 73, 104	0
4	Y	5/6 (83%)	1.43	1 (20%) 2 1	59, 65, 83, 112	0
4	Z	4/6 (66%)	1.38	1 (25%) 1 1	60, 60, 60, 65	0
All	All	2058/2106 (97%)	-0.08	51 (2%) 54 52	22, 43, 92, 135	51 (2%)

The worst 5 of 51 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	Y	2	U	6.8
3	I	259	ILE	6.3
1	B	104	TYR	5.4
1	A	106	LYS	5.0
3	H	259	ILE	4.9

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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
5	ZN	A	180	1/1	0.09	-0.90	59,59,59,59	0
5	ZN	B	180	1/1	0.10	-1.06	96,96,96,96	0
5	ZN	C	180	1/1	0.04	-2.41	71,71,71,71	0

6.5 Other polymers

There are no such residues in this entry.