wwPDB X-ray Structure Validation Summary Report (i)

Jan 28, 2025 – 12:08 pm GMT

PDB ID	:	9GU5
Title	:	Crystal Structure of Hfq V22A
Authors	:	McQuail, J.; Krepl, M.; Katsuya-Gaviria, K.; Tabib-Salazar, A.; Burchell, L.; Bischler, T.; Grafenhan, T.; Brear, P.; Luisi, B.
Deposited on	:	2024-09-18
Resolution	:	2.90 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:

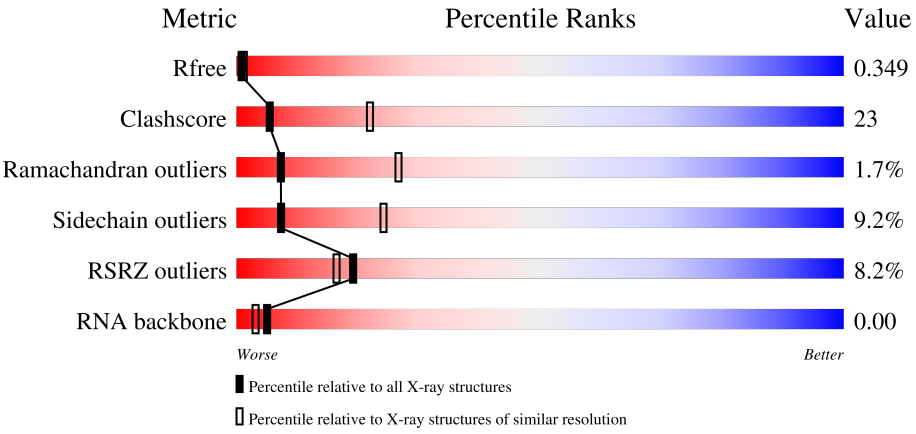
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.90 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	2335 (2.90-2.90)
Clashscore	180529	2564 (2.90-2.90)
Ramachandran outliers	177936	2514 (2.90-2.90)
Sidechain outliers	177891	2516 (2.90-2.90)
RSRZ outliers	164620	2337 (2.90-2.90)
RNA backbone	3690	1039 (3.10-2.70)

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	102	7% 35% 25% • • 35%
1	B	102	8% 35% 25% 6% 33%
1	C	102	6% 41% 21% • 36%
1	D	102	3% 33% 25% • 37%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	E	102	32% 29% 36%
1	F	102	6% 33% 27% 36%
1	G	102	3% 30% 29% 5% 35%
1	H	102	3% 34% 25% 37%
1	I	102	8% 27% 29% 5% 38%
1	J	102	6% 39% 24% 33%
1	K	102	4% 34% 24% 37%
1	L	102	11% 37% 25% 33%
2	M	4	50% 50%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 12531 atoms, of which 6275 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 RNA-binding protein Hfq.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	66	Total	C	H	N	O	S	0	0	0
			1025	332	508	91	93	1			
1	B	68	Total	C	H	N	O	S	0	0	0
			1048	338	523	93	93	1			
1	C	65	Total	C	H	N	O	S	0	0	0
			1017	325	515	87	89	1			
1	D	64	Total	C	H	N	O	S	0	0	0
			1005	321	508	86	89	1			
1	E	65	Total	C	H	N	O	S	0	0	0
			1033	330	521	90	91	1			
1	F	65	Total	C	H	N	O	S	0	0	0
			1025	327	519	87	91	1			
1	G	66	Total	C	H	N	O	S	0	0	0
			1045	332	530	91	91	1			
1	H	64	Total	C	H	N	O	S	0	0	0
			1014	324	511	89	89	1			
1	I	63	Total	C	H	N	O	S	0	0	0
			1025	324	524	88	88	1			
1	J	68	Total	C	H	N	O	S	0	0	0
			1067	339	539	93	95	1			
1	K	64	Total	C	H	N	O	S	0	0	0
			1004	324	504	86	89	1			
1	L	68	Total	C	H	N	O	S	0	0	0
			1052	338	528	92	93	1			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	22	ALA	VAL	engineered mutation	UNP P0A6X3
B	22	ALA	VAL	engineered mutation	UNP P0A6X3
C	22	ALA	VAL	engineered mutation	UNP P0A6X3
D	22	ALA	VAL	engineered mutation	UNP P0A6X3
E	22	ALA	VAL	engineered mutation	UNP P0A6X3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	22	ALA	VAL	engineered mutation	UNP P0A6X3
G	22	ALA	VAL	engineered mutation	UNP P0A6X3
H	22	ALA	VAL	engineered mutation	UNP P0A6X3
I	22	ALA	VAL	engineered mutation	UNP P0A6X3
J	22	ALA	VAL	engineered mutation	UNP P0A6X3
K	22	ALA	VAL	engineered mutation	UNP P0A6X3
L	22	ALA	VAL	engineered mutation	UNP P0A6X3

- Molecule 2 is a RNA chain called RNA (5'-R(P*AP*AP*AP*A)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	M	4	Total	C	H	N	O	P	0	0	0
			133	40	45	20	24	4			

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	K	1	Total	Mg	0	0
			1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	4	Total	O	0	0
			4	4		
4	B	3	Total	O	0	0
			3	3		
4	C	4	Total	O	0	0
			4	4		
4	D	1	Total	O	0	0
			1	1		
4	E	4	Total	O	0	0
			4	4		
4	F	5	Total	O	0	0
			5	5		
4	G	3	Total	O	0	0
			3	3		
4	H	5	Total	O	0	0
			5	5		
4	J	2	Total	O	0	0
			2	2		

Continued on next page...

Continued from previous page...

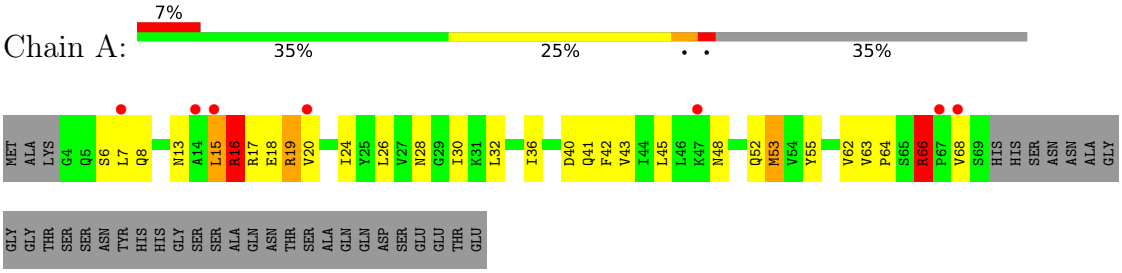
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	K	1	Total O 1 1	0	0
4	L	2	Total O 2 2	0	0
4	M	3	Total O 3 3	0	0

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.

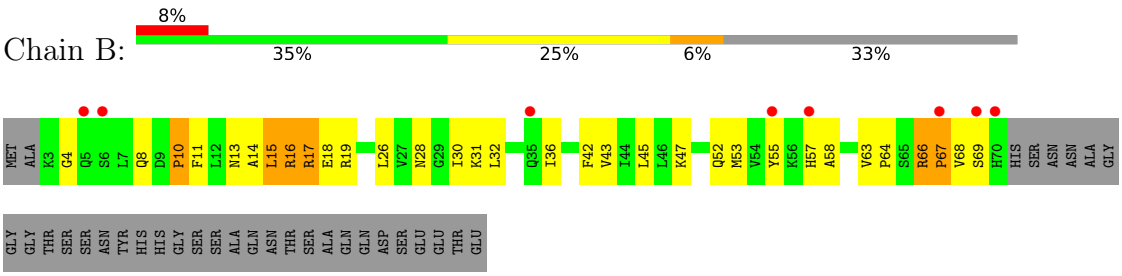
• Molecule 1: RNA-binding protein Hfq

Chain A:



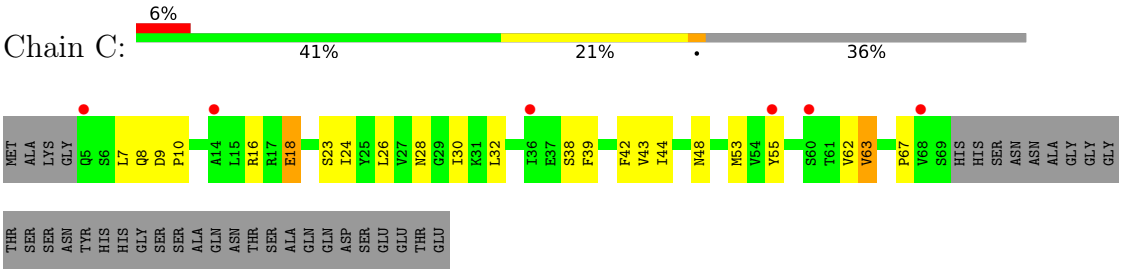
• Molecule 1: RNA-binding protein Hfq

Chain B:



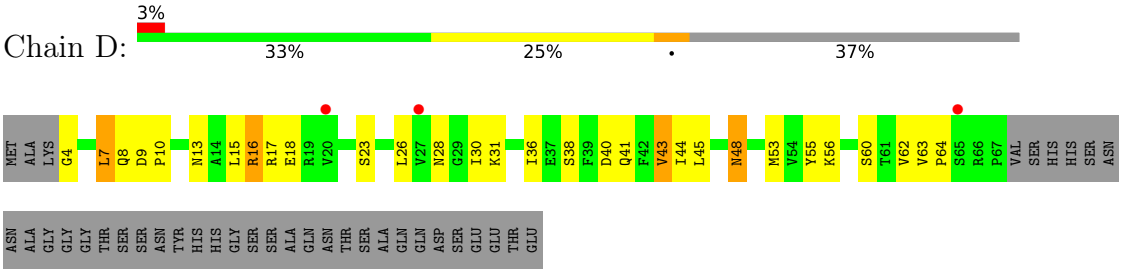
• Molecule 1: RNA-binding protein Hfq

Chain C:



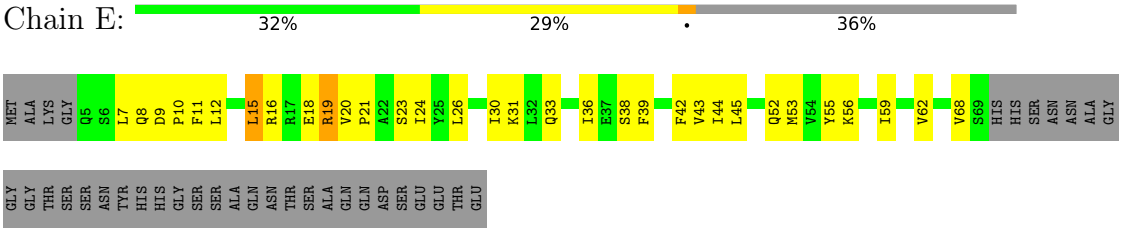
• Molecule 1: RNA-binding protein Hfq

Chain D:



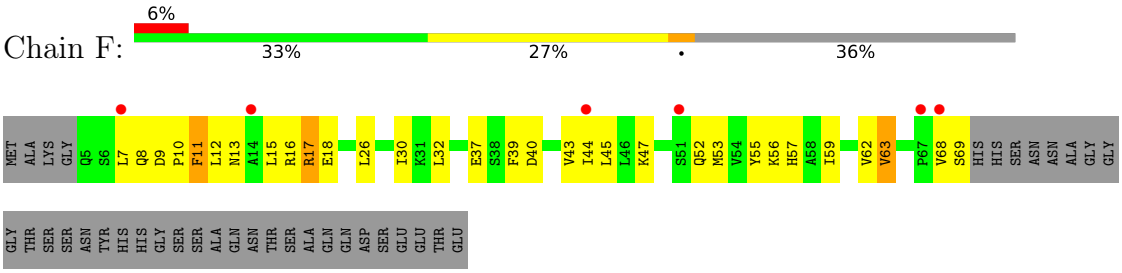
• Molecule 1: RNA-binding protein Hfq

Chain E:



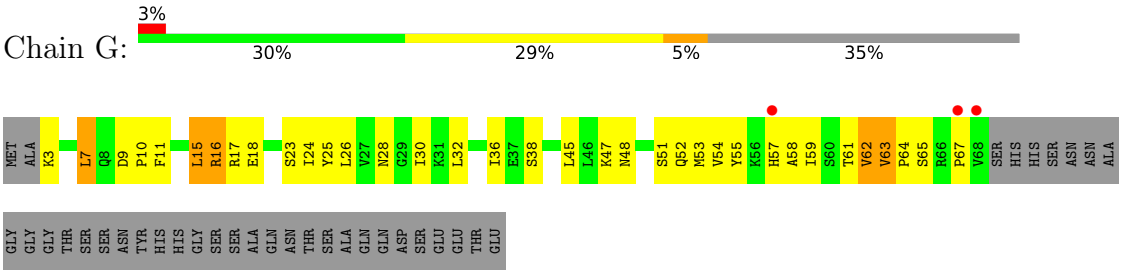
• Molecule 1: RNA-binding protein Hfq

Chain F:



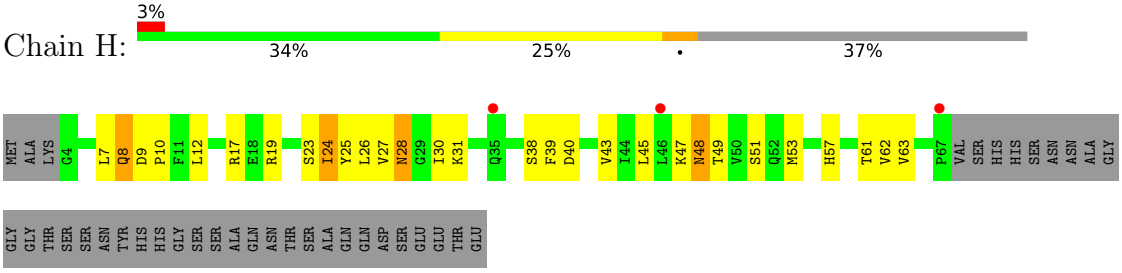
• Molecule 1: RNA-binding protein Hfq

Chain G:



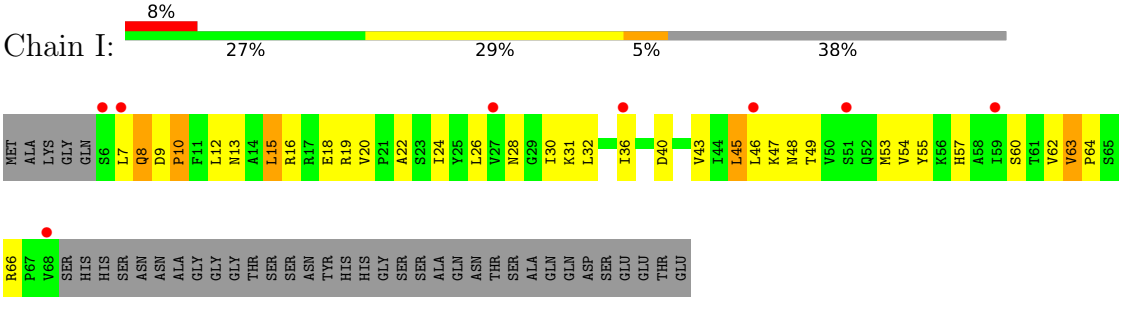
• Molecule 1: RNA-binding protein Hfq

Chain H:



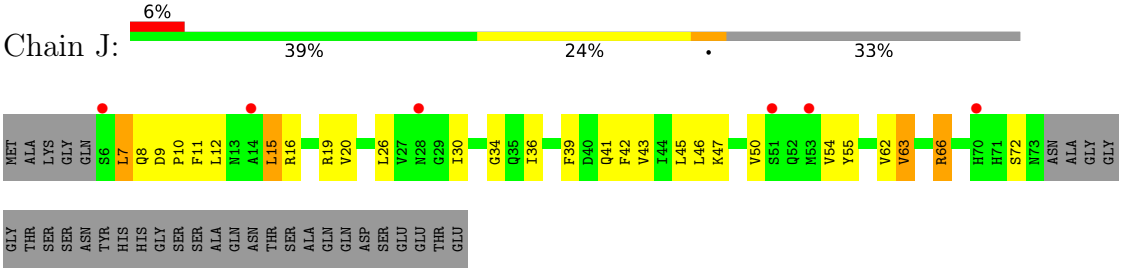
• Molecule 1: RNA-binding protein Hfq

Chain I:



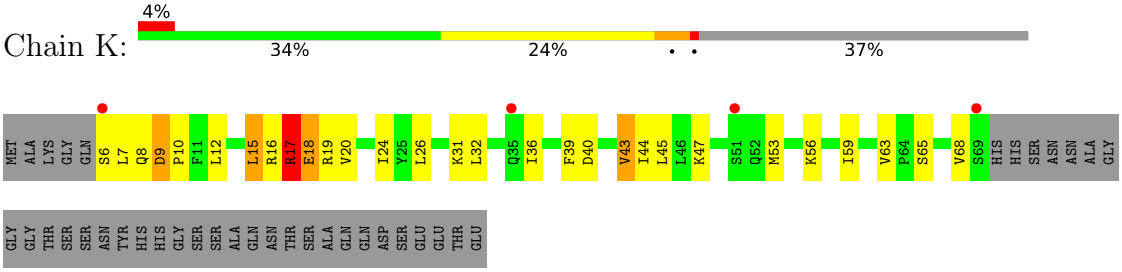
• Molecule 1: RNA-binding protein Hfq

Chain J:



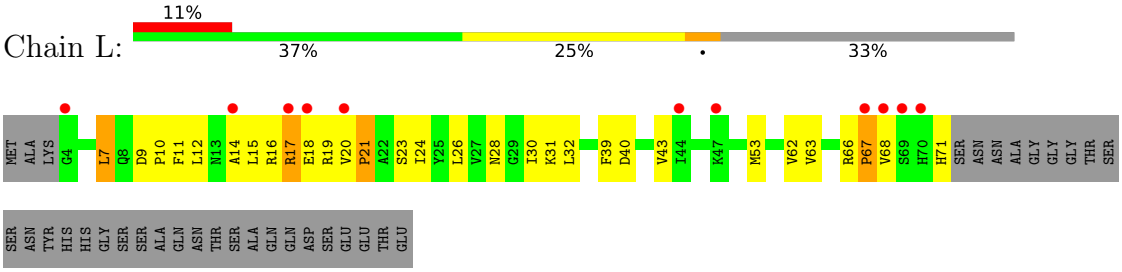
• Molecule 1: RNA-binding protein Hfq

Chain K:



• Molecule 1: RNA-binding protein Hfq

Chain L:



• Molecule 2: RNA (5'-R(P*AP*AP*AP*A)-3')

Chain M:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	69.18Å 69.21Å 73.19Å 63.54° 89.30° 60.04°	Depositor
Resolution (Å)	57.69 – 2.90 57.69 – 2.90	Depositor EDS
% Data completeness (in resolution range)	98.3 (57.69-2.90) 99.0 (57.69-2.90)	Depositor EDS
R_{merge}	0.23	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.18 (at 2.65Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, R_{free}	0.274 , 0.387 0.290 , 0.349	Depositor DCC
R_{free} test set	1028 reflections (4.62%)	wwPDB-VP
Wilson B-factor (Å^2)	55.4	Xtriage
Anisotropy	0.767	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.43 , 86.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.065 for -h+k,-h,-k+l 0.065 for -k,h-k,h-k+l 0.086 for h,h-k,-l 0.076 for -h+k,k,k-l 0.066 for -k,-h,-h+k-l	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	12531	wwPDB-VP
Average B, all atoms (Å^2)	75.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.41% 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

5.1 Standard geometry

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

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	0.59	0/526	0.72	0/713
1	B	0.66	0/534	0.78	0/725
1	C	0.73	0/511	0.82	1/695 (0.1%)
1	D	0.64	0/506	0.74	0/687
1	E	0.60	0/521	0.72	0/707
1	F	0.66	0/515	0.79	0/700
1	G	0.69	0/524	0.84	0/711
1	H	0.63	0/512	0.78	0/694
1	I	0.60	0/510	0.71	0/692
1	J	0.65	0/537	0.71	0/729
1	K	0.62	0/509	0.72	0/692
1	L	0.59	0/534	0.69	0/726
2	M	0.68	0/99	1.32	2/152 (1.3%)
All	All	0.64	0/6338	0.77	3/8623 (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
1	A	0	3
1	B	0	2
1	D	0	2
1	F	0	1
1	G	0	1
1	H	0	2
1	I	0	1
1	J	0	3
1	K	0	1
1	L	0	2

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
All	All	0	18

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	M	1	A	P-O3'-C3'	-7.66	110.51	119.70
2	M	2	A	P-O3'-C3'	-6.24	112.21	119.70
1	C	18	GLU	OE1-CD-OE2	-5.36	116.87	123.30

There are no chirality outliers.

5 of 18 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	16	ARG	Sidechain
1	A	19	ARG	Sidechain
1	A	66	ARG	Sidechain
1	B	16	ARG	Sidechain
1	B	17	ARG	Sidechain

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	517	508	536	34	7
1	B	525	523	534	27	1
1	C	502	515	515	20	1
1	D	497	508	508	23	8
1	E	512	521	530	27	2
1	F	506	519	519	30	7
1	G	515	530	530	41	9
1	H	503	511	519	32	0
1	I	501	524	523	36	4
1	J	528	539	539	32	0
1	K	500	504	514	29	4
1	L	524	528	528	23	2

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	M	88	45	45	3	0
3	K	1	0	0	0	0
4	A	4	0	0	0	0
4	B	3	0	0	0	0
4	C	4	0	0	0	1
4	D	1	0	0	0	0
4	E	4	0	0	0	0
4	F	5	0	0	0	0
4	G	3	0	0	1	0
4	H	5	0	0	2	0
4	J	2	0	0	0	0
4	K	1	0	0	0	0
4	L	2	0	0	0	0
4	M	3	0	0	0	0
All	All	6256	6275	6340	294	23

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 23.

The worst 5 of 294 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:20:VAL:HG21	1:A:68:VAL:HG22	1.53	0.89
1:H:57:HIS:ND1	4:H:201:HOH:O	2.05	0.85
1:D:41:GLN:OE1	1:D:41:GLN:N	2.14	0.80
1:H:7:LEU:HD11	1:I:45:LEU:HD22	1.61	0.79
1:I:28:ASN:O	1:I:28:ASN:ND2	2.20	0.75

The worst 5 of 23 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:D:4:GLY:O	1:G:3:LYS:CA[1_554]	0.78	1.42
1:A:18:GLU:OE1	1:F:18:GLU:OE1[1_455]	0.89	1.31
1:A:18:GLU:OE2	1:F:18:GLU:OE2[1_455]	1.06	1.14
1:D:4:GLY:O	1:G:3:LYS:CB[1_554]	1.38	0.82
1:D:4:GLY:O	1:G:3:LYS:N[1_554]	1.46	0.74

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	64/102 (63%)	51 (80%)	12 (19%)	1 (2%)	8 28
1	B	66/102 (65%)	57 (86%)	6 (9%)	3 (4%)	2 8
1	C	63/102 (62%)	53 (84%)	8 (13%)	2 (3%)	3 13
1	D	62/102 (61%)	54 (87%)	8 (13%)	0	100 100
1	E	63/102 (62%)	56 (89%)	7 (11%)	0	100 100
1	F	63/102 (62%)	57 (90%)	6 (10%)	0	100 100
1	G	64/102 (63%)	56 (88%)	8 (12%)	0	100 100
1	H	62/102 (61%)	55 (89%)	6 (10%)	1 (2%)	8 28
1	I	61/102 (60%)	55 (90%)	4 (7%)	2 (3%)	3 13
1	J	66/102 (65%)	58 (88%)	6 (9%)	2 (3%)	3 15
1	K	62/102 (61%)	56 (90%)	6 (10%)	0	100 100
1	L	66/102 (65%)	56 (85%)	8 (12%)	2 (3%)	3 15
All	All	762/1224 (62%)	664 (87%)	85 (11%)	13 (2%)	7 27

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	67	PRO
1	B	10	PRO
1	B	4	GLY
1	J	7	LEU
1	J	72	SER

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	59/88 (67%)	54 (92%)	5 (8%)	8 27
1	B	57/88 (65%)	52 (91%)	5 (9%)	8 26
1	C	56/88 (64%)	53 (95%)	3 (5%)	18 49
1	D	55/88 (62%)	46 (84%)	9 (16%)	2 6
1	E	58/88 (66%)	54 (93%)	4 (7%)	13 37
1	F	57/88 (65%)	52 (91%)	5 (9%)	8 26
1	G	57/88 (65%)	52 (91%)	5 (9%)	8 26
1	H	56/88 (64%)	51 (91%)	5 (9%)	8 26
1	I	57/88 (65%)	49 (86%)	8 (14%)	3 9
1	J	59/88 (67%)	56 (95%)	3 (5%)	20 51
1	K	56/88 (64%)	48 (86%)	8 (14%)	2 8
1	L	57/88 (65%)	54 (95%)	3 (5%)	19 49
All	All	684/1056 (65%)	621 (91%)	63 (9%)	7 24

5 of 63 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	17	ARG
1	K	17	ARG
1	G	63	VAL
1	K	16	ARG
1	K	65	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	8	GLN
1	I	28	ASN
1	K	52	GLN
1	J	52	GLN
1	K	13	ASN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	M	3/4 (75%)	3 (100%)	0

All (3) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	M	2	A
2	M	3	A
2	M	4	A

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 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

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	66/102 (64%)	0.54	7 (10%) 13 11	30, 62, 134, 161	1 (1%)
1	B	68/102 (66%)	0.62	8 (11%) 10 9	40, 68, 146, 176	0
1	C	65/102 (63%)	0.52	6 (9%) 16 14	43, 62, 121, 179	0
1	D	64/102 (62%)	0.41	3 (4%) 37 31	33, 58, 120, 194	1 (1%)
1	E	65/102 (63%)	0.20	0 100 100	40, 65, 107, 193	0
1	F	65/102 (63%)	0.61	6 (9%) 16 14	35, 63, 137, 193	1 (1%)
1	G	66/102 (64%)	0.20	3 (4%) 39 32	42, 68, 112, 153	0
1	H	64/102 (62%)	0.23	3 (4%) 37 31	38, 70, 123, 207	0
1	I	63/102 (61%)	0.46	8 (12%) 9 8	36, 63, 121, 136	0
1	J	68/102 (66%)	0.59	6 (8%) 17 15	36, 64, 127, 192	0
1	K	64/102 (62%)	0.37	4 (6%) 27 23	40, 64, 103, 210	0
1	L	68/102 (66%)	0.87	11 (16%) 5 5	51, 80, 151, 269	0
2	M	4/4 (100%)	0.09	0 100 100	60, 76, 81, 91	0
All	All	790/1228 (64%)	0.47	65 (8%) 19 16	30, 66, 131, 269	3 (0%)

The worst 5 of 65 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	J	53	MET	5.8
1	L	20	VAL	4.9
1	F	14	ALA	4.1
1	F	7	LEU	3.9
1	C	60	SER	3.9

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 (i)

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^{th} 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^2)	Q<0.9
3	MG	K	201	1/1	0.68	0.17	24,24,24,24	0

6.5 Other polymers (i)

There are no such residues in this entry.