



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 24, 2021 – 01:01 pm BST

PDB ID : 6XYJ
Title : Hfq from E.coli with inserted long loop L4 sequence
Authors : Nikulin, A.D.; Alipov, A.A.; Lekontseva, N.V.
Deposited on : 2020-01-30
Resolution : 2.77 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 ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.1
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)	:	2.23.1

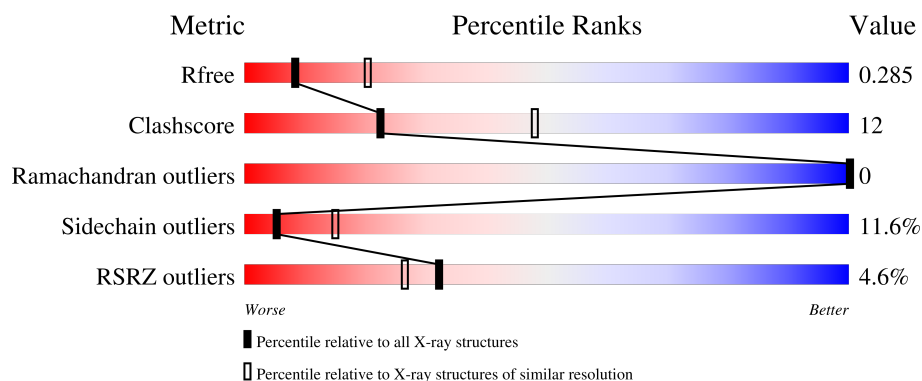
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.77 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	4107 (2.80-2.76)
Clashscore	141614	4575 (2.80-2.76)
Ramachandran outliers	138981	4487 (2.80-2.76)
Sidechain outliers	138945	4489 (2.80-2.76)
RSRZ outliers	127900	4027 (2.80-2.76)

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 $\leq 5\%$. 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	AAA	95	<div> <div>11%</div> <div>61%</div> <div>20%</div> <div>•</div> <div>16%</div> </div>
1	BBB	95	<div> <div>%</div> <div>60%</div> <div>23%</div> <div>•</div> <div>14%</div> </div>
1	CCC	95	<div> <div>3%</div> <div>53%</div> <div>26%</div> <div>•</div> <div>17%</div> </div>
1	DDD	95	<div> <div>2%</div> <div>54%</div> <div>27%</div> <div>•</div> <div>16%</div> </div>
1	EEE	95	<div> <div>%</div> <div>63%</div> <div>19%</div> <div>•</div> <div>16%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	FFF	95	<div><div>5%</div><div><div></div><div>58%</div><div>21%</div><div>5%</div><div>16%</div></div></div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3862 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 RNA-binding protein Hfq.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	80	Total	C	N	O	S	0	0	0
			643	412	114	116	1			
1	BBB	82	Total	C	N	O	S	0	0	0
			656	420	117	118	1			
1	CCC	79	Total	C	N	O	S	0	0	0
			633	406	111	115	1			
1	DDD	80	Total	C	N	O	S	0	0	0
			643	412	114	116	1			
1	EEE	80	Total	C	N	O	S	0	0	0
			643	412	114	116	1			
1	FFF	80	Total	C	N	O	S	0	0	0
			643	412	114	116	1			

There are 78 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	50	ALA	-	insertion	UNP A6VD57
AAA	51	GLU	-	insertion	UNP A6VD57
AAA	52	GLU	-	insertion	UNP A6VD57
AAA	53	ILE	-	insertion	UNP A6VD57
AAA	54	GLN	-	insertion	UNP A6VD57
AAA	55	ASN	-	insertion	UNP A6VD57
AAA	56	GLY	-	insertion	UNP A6VD57
AAA	57	GLU	-	insertion	UNP A6VD57
AAA	58	VAL	-	insertion	UNP A6VD57
AAA	59	VAL	-	insertion	UNP A6VD57
AAA	60	ARG	-	insertion	UNP A6VD57
AAA	61	LYS	-	insertion	UNP A6VD57
AAA	62	VAL	-	insertion	UNP A6VD57
BBB	50	ALA	-	insertion	UNP A6VD57
BBB	51	GLU	-	insertion	UNP A6VD57
BBB	52	GLU	-	insertion	UNP A6VD57
BBB	53	ILE	-	insertion	UNP A6VD57

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
BBB	54	GLN	-	insertion	UNP A6VD57
BBB	55	ASN	-	insertion	UNP A6VD57
BBB	56	GLY	-	insertion	UNP A6VD57
BBB	57	GLU	-	insertion	UNP A6VD57
BBB	58	VAL	-	insertion	UNP A6VD57
BBB	59	VAL	-	insertion	UNP A6VD57
BBB	60	ARG	-	insertion	UNP A6VD57
BBB	61	LYS	-	insertion	UNP A6VD57
BBB	62	VAL	-	insertion	UNP A6VD57
CCC	50	ALA	-	insertion	UNP A6VD57
CCC	51	GLU	-	insertion	UNP A6VD57
CCC	52	GLU	-	insertion	UNP A6VD57
CCC	53	ILE	-	insertion	UNP A6VD57
CCC	54	GLN	-	insertion	UNP A6VD57
CCC	55	ASN	-	insertion	UNP A6VD57
CCC	56	GLY	-	insertion	UNP A6VD57
CCC	57	GLU	-	insertion	UNP A6VD57
CCC	58	VAL	-	insertion	UNP A6VD57
CCC	59	VAL	-	insertion	UNP A6VD57
CCC	60	ARG	-	insertion	UNP A6VD57
CCC	61	LYS	-	insertion	UNP A6VD57
CCC	62	VAL	-	insertion	UNP A6VD57
DDD	50	ALA	-	insertion	UNP A6VD57
DDD	51	GLU	-	insertion	UNP A6VD57
DDD	52	GLU	-	insertion	UNP A6VD57
DDD	53	ILE	-	insertion	UNP A6VD57
DDD	54	GLN	-	insertion	UNP A6VD57
DDD	55	ASN	-	insertion	UNP A6VD57
DDD	56	GLY	-	insertion	UNP A6VD57
DDD	57	GLU	-	insertion	UNP A6VD57
DDD	58	VAL	-	insertion	UNP A6VD57
DDD	59	VAL	-	insertion	UNP A6VD57
DDD	60	ARG	-	insertion	UNP A6VD57
DDD	61	LYS	-	insertion	UNP A6VD57
DDD	62	VAL	-	insertion	UNP A6VD57
EEE	50	ALA	-	insertion	UNP A6VD57
EEE	51	GLU	-	insertion	UNP A6VD57
EEE	52	GLU	-	insertion	UNP A6VD57
EEE	53	ILE	-	insertion	UNP A6VD57
EEE	54	GLN	-	insertion	UNP A6VD57
EEE	55	ASN	-	insertion	UNP A6VD57
EEE	56	GLY	-	insertion	UNP A6VD57

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
EEE	57	GLU	-	insertion	UNP A6VD57
EEE	58	VAL	-	insertion	UNP A6VD57
EEE	59	VAL	-	insertion	UNP A6VD57
EEE	60	ARG	-	insertion	UNP A6VD57
EEE	61	LYS	-	insertion	UNP A6VD57
EEE	62	VAL	-	insertion	UNP A6VD57
FFF	50	ALA	-	insertion	UNP A6VD57
FFF	51	GLU	-	insertion	UNP A6VD57
FFF	52	GLU	-	insertion	UNP A6VD57
FFF	53	ILE	-	insertion	UNP A6VD57
FFF	54	GLN	-	insertion	UNP A6VD57
FFF	55	ASN	-	insertion	UNP A6VD57
FFF	56	GLY	-	insertion	UNP A6VD57
FFF	57	GLU	-	insertion	UNP A6VD57
FFF	58	VAL	-	insertion	UNP A6VD57
FFF	59	VAL	-	insertion	UNP A6VD57
FFF	60	ARG	-	insertion	UNP A6VD57
FFF	61	LYS	-	insertion	UNP A6VD57
FFF	62	VAL	-	insertion	UNP A6VD57

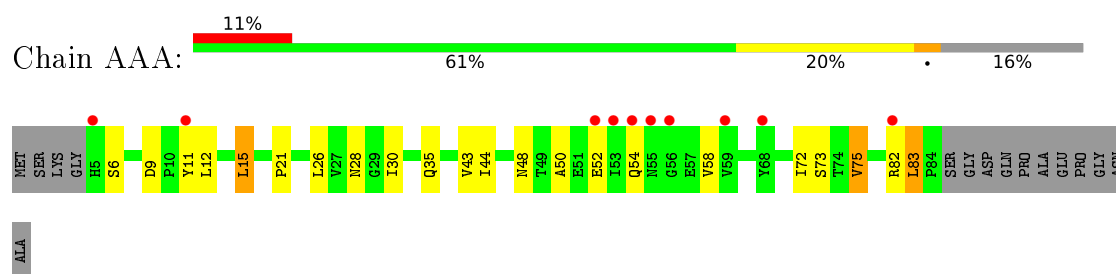
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	BBB	1	Total O 1 1	0	0

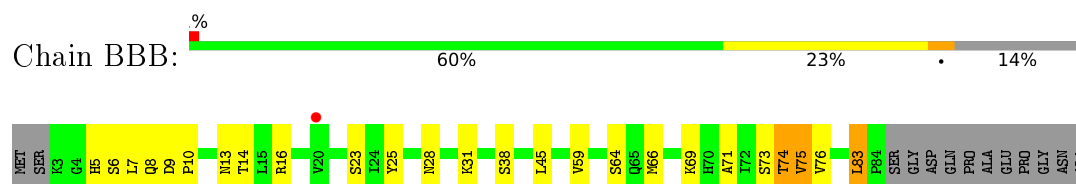
3 Residue-property plots [i](#)

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.

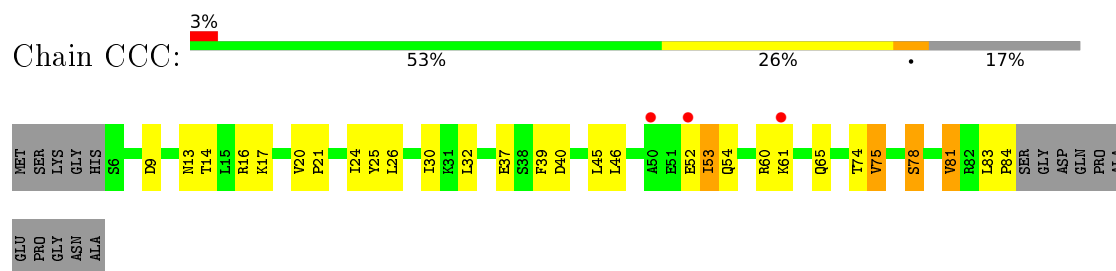
- Molecule 1: RNA-binding protein Hfq



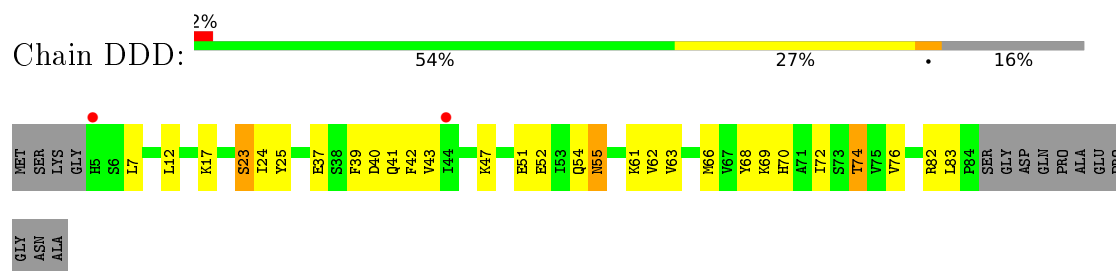
- Molecule 1: RNA-binding protein Hfq



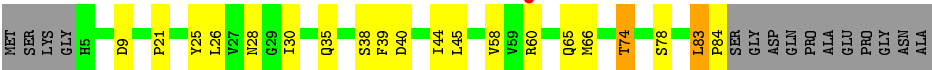
- Molecule 1: RNA-binding protein Hfq



- Molecule 1: RNA-binding protein Hfq



- Molecule 1: RNA-binding protein Hfq



● Molecule 1: RNA-binding protein Hfq



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	76.81Å 112.62Å 78.05Å 90.00° 119.43° 90.00°	Depositor
Resolution (Å)	43.40 – 2.77 43.36 – 2.77	Depositor EDS
% Data completeness (in resolution range)	98.1 (43.40-2.77) 98.2 (43.36-2.77)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.02 (at 2.77Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.211 , 0.285 0.213 , 0.285	Depositor DCC
R_{free} test set	715 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å ²)	72.9	Xtriage
Anisotropy	0.413	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 70.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	3862	wwPDB-VP
Average B, all atoms (Å ²)	105.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.30% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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](#)

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	AAA	0.64	0/654	0.78	0/887
1	BBB	0.63	0/667	0.77	0/903
1	CCC	0.62	0/643	0.78	0/872
1	DDD	0.63	0/654	0.81	0/887
1	EEE	0.63	0/654	0.80	0/887
1	FFF	0.66	0/654	0.83	0/887
All	All	0.63	0/3926	0.80	0/5323

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	AAA	643	0	674	17	0
1	BBB	656	0	690	22	0
1	CCC	633	0	667	18	0
1	DDD	643	0	674	23	0
1	EEE	643	0	674	10	0
1	FFF	643	0	674	13	0
2	BBB	1	0	0	0	0
All	All	3862	0	4053	92	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 12.

The worst 5 of 92 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:DDD:41:GLN:C	1:DDD:69:LYS:NZ	1.79	1.33
1:DDD:41:GLN:O	1:DDD:69:LYS:NZ	1.77	1.10
1:DDD:41:GLN:C	1:DDD:69:LYS:HZ1	1.51	1.03
1:DDD:41:GLN:C	1:DDD:69:LYS:HZ2	1.54	1.00
1:BBB:6:SER:O	1:BBB:10:PRO:HG2	1.77	0.84

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	AAA	78/95 (82%)	74 (95%)	4 (5%)	0	100	100
1	BBB	80/95 (84%)	74 (92%)	6 (8%)	0	100	100
1	CCC	77/95 (81%)	72 (94%)	5 (6%)	0	100	100
1	DDD	78/95 (82%)	69 (88%)	9 (12%)	0	100	100
1	EEE	78/95 (82%)	73 (94%)	5 (6%)	0	100	100
1	FFF	78/95 (82%)	74 (95%)	4 (5%)	0	100	100
All	All	469/570 (82%)	436 (93%)	33 (7%)	0	100	100

There are no Ramachandran outliers to report.

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	AAA	75/85 (88%)	67 (89%)	8 (11%)	6	18
1	BBB	76/85 (89%)	69 (91%)	7 (9%)	9	24
1	CCC	74/85 (87%)	64 (86%)	10 (14%)	4	10
1	DDD	75/85 (88%)	66 (88%)	9 (12%)	5	14
1	EEE	75/85 (88%)	68 (91%)	7 (9%)	9	24
1	FFF	75/85 (88%)	64 (85%)	11 (15%)	3	8
All	All	450/510 (88%)	398 (88%)	52 (12%)	5	15

5 of 52 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	DDD	52	GLU
1	EEE	35	GLN
1	FFF	74	THR
1	DDD	54	GLN
1	DDD	74	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

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 [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, 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	AAA	80/95 (84%)	0.78	10 (12%) 3 2	77, 100, 156, 170	0
1	BBB	82/95 (86%)	0.29	1 (1%) 79 76	79, 97, 134, 160	0
1	CCC	79/95 (83%)	0.52	3 (3%) 40 35	80, 105, 146, 159	0
1	DDD	80/95 (84%)	0.38	2 (2%) 57 52	79, 100, 132, 149	0
1	EEE	80/95 (84%)	0.34	1 (1%) 77 75	76, 99, 132, 140	0
1	FFF	80/95 (84%)	0.47	5 (6%) 20 15	80, 101, 151, 162	0
All	All	481/570 (84%)	0.46	22 (4%) 32 26	76, 101, 144, 170	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	53	ILE	7.7
1	AAA	56	GLY	7.0
1	DDD	5	HIS	3.8
1	AAA	52	GLU	3.6
1	AAA	54	GLN	3.4

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 [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.