W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 12:01 am BST

PDB ID	:	5AFF
Title	:	Symportin 1 chaperones 5S RNP assembly during ribosome biogenesis by occupying an essential rRNA binding site
Authors	:	Calvino, F.R.; Kharde, S.; Wild, K.; Bange, G.; Sinning, I.
Deposited on	:	2015-01-21
Resolution	:	3.40 Å(reported)

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

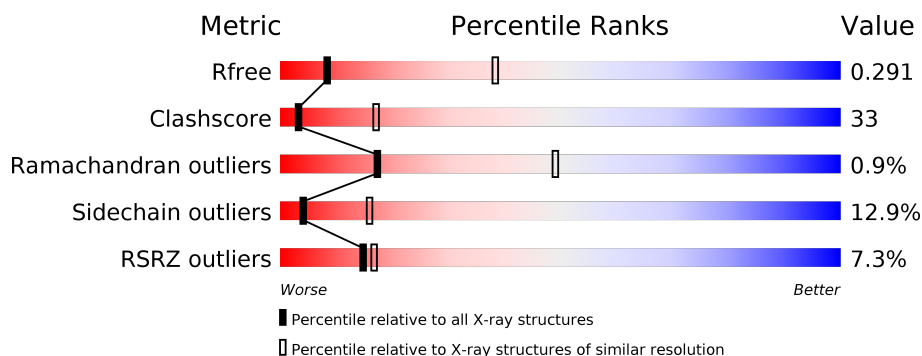|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 3.40 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	Percentile Ranks	Value
Rfree		0.291
Clashscore		33
Ramachandran outliers		0.9%
Sidechain outliers		12.9%
RSRZ outliers		7.3%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1026 (3.48-3.32)
Clashscore	141614	1055 (3.48-3.32)
Ramachandran outliers	138981	1038 (3.48-3.32)
Sidechain outliers	138945	1038 (3.48-3.32)
RSRZ outliers	127900	2173 (3.50-3.30)

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 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	654	4% 42% 39% 6% 13%
2	B	48	23% 23% 10% 44%
3	C	173	17% 25% 34% 9% 32%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5606 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 SYMPORTIN 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	568	Total	C	N	O	S	0	0	0
			4414	2800	756	845	13			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	23	MET	-	expression tag	UNP G0S5S6

- Molecule 2 is a protein called RIBOSOMAL PROTEIN L5.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	27	Total	C	N	O	0	0	0
			245	161	45	39			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-6	MET	-	expression tag	UNP G0SEG2
B	-5	HIS	-	expression tag	UNP G0SEG2
B	-4	HIS	-	expression tag	UNP G0SEG2
B	-3	HIS	-	expression tag	UNP G0SEG2
B	-2	HIS	-	expression tag	UNP G0SEG2
B	-1	HIS	-	expression tag	UNP G0SEG2
B	0	HIS	-	expression tag	UNP G0SEG2

- Molecule 3 is a protein called RIBOSOMAL PROTEIN L11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	117	Total	C	N	O	S	0	0	0
			947	601	169	173	4			

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 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.
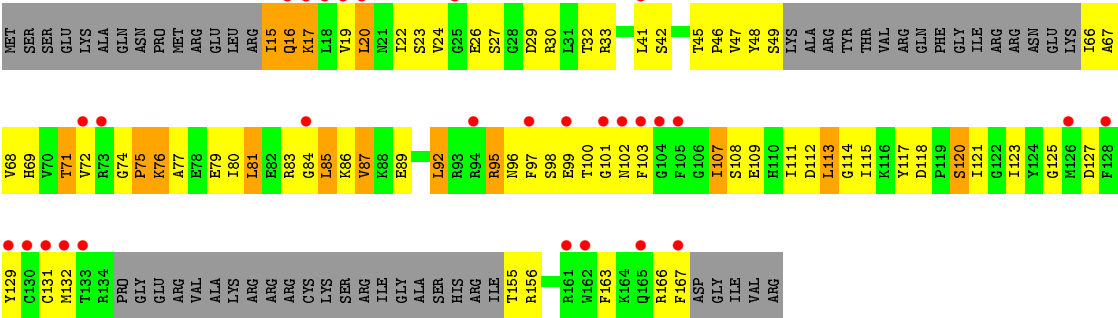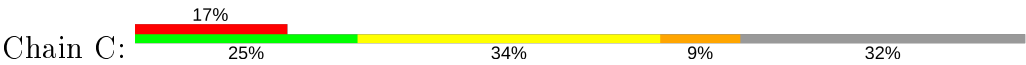
• Molecule 1: SYMPORTIN 1
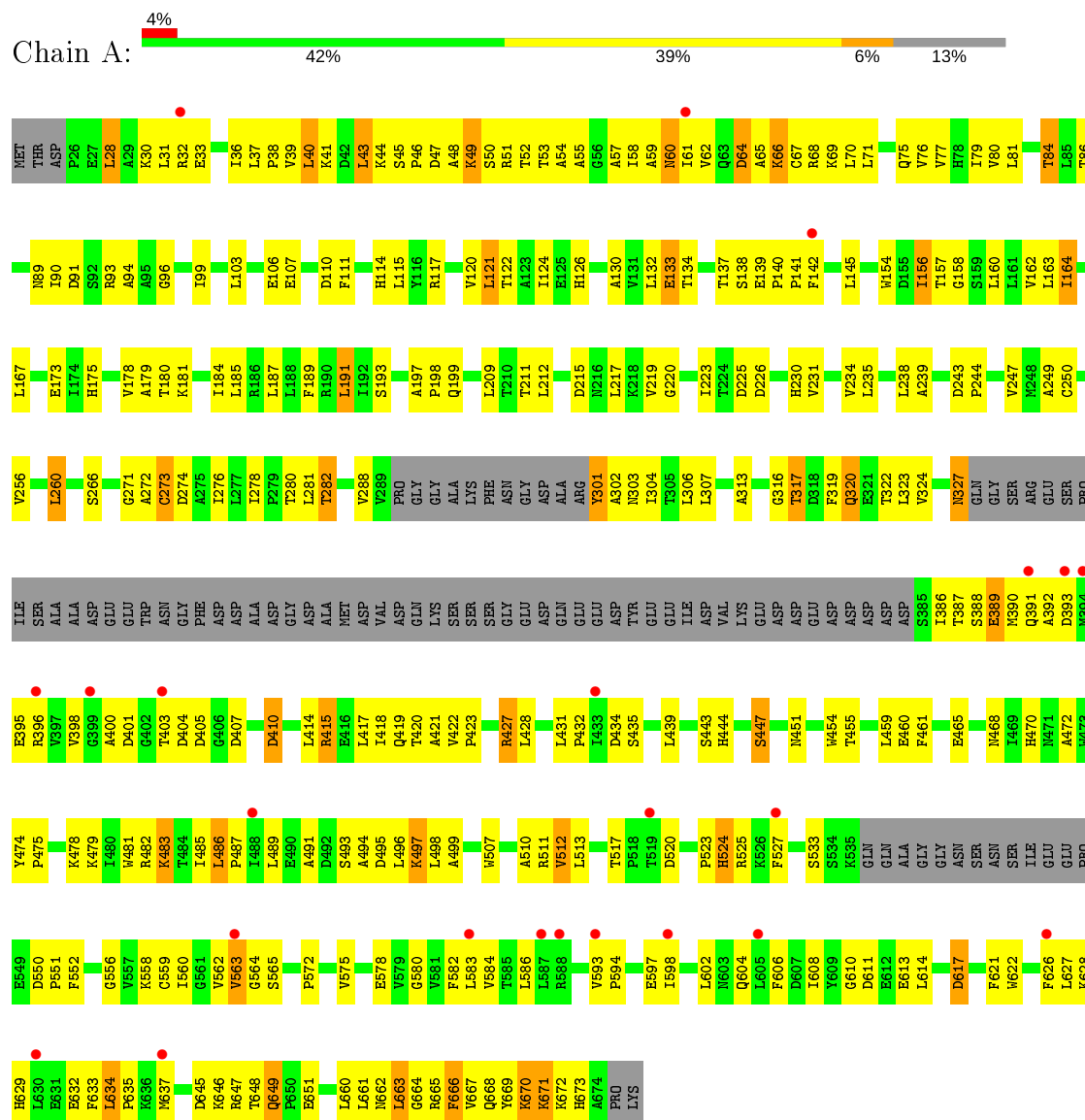
• Molecule 2: RIBOSOMAL PROTEIN L5

MET HIS HIS HIS HIS HIS MET A2 F3 H4 K5 L6 V7 K8 N9 S10 Y13 S14 R15 F16 Q17 K21 R22 R23 ARG GLN GLY K26 T27 D28 Y29 Y30 ALA ARG LYS ARG LEU ILE THR GLN ALA LYS

• Molecule 3: RIBOSOMAL PROTEIN L11

Chain C:

17%

25%　　　　　34%　　　　　9%　　　　　32%

MET SER SER GLU LYS ALA GLN ASN PRO MET ARG GLU LEU ARG I15 Q16 K17 L18 V19 L20 N21 I22 S23 V24 G25 E26 S27 G28 D29 R30 L31 T32 R33 L41 S42 T45 P46 V47 Y48 S49 LYS ALA ARG TYR THR VAL ARG GLN PHE GLY ILE ARG ARG ASN GLU LYS I66 A67

V68 H69 V70 T71 V72 R73 G74 P75 K76 A77 E78 E79 I80 L81 E82 R83 G84 L85 K86 V87 K88 E89 L92 R93 R94 R95 N96 F97 S98 E99 T100 G101 N102 F103 G104 F105 G106 I107 S108 E109 H110 I111 D112 L113 G114 I115 K116 Y117 D118 P119 S120 I121 G122 I123 Y124 G125 M126 D127 F128

Y129 C130 C131 M132 T133 R134 PRO GLY GLU ARG VAL ALA LYS ARG ARG CYS LYS SER ARG ILE GLY ALA SER HIS ARG ILE T155 R156 R161 W162 F163 K164 Q165 R166 F167 ASP GLY ILE VAL ARG

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	60.09Å 105.96Å 147.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.26 – 3.40 49.26 – 3.40	Depositor EDS
% Data completeness (in resolution range)	99.8 (49.26-3.40) 99.8 (49.26-3.40)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.04 (at 3.40Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.236 , 0.300 0.234 , 0.291	Depositor DCC
R_{free} test set	665 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å^2)	113.3	Xtriage
Anisotropy	0.284	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 115.0	EDS
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	5606	wwPDB-VP
Average B, all atoms (Å^2)	141.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.08% 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 ⓘ

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	0.27	0/4487	0.54	1/6101 (0.0%)
2	B	0.53	0/251	0.69	0/331
3	C	0.57	0/962	1.10	4/1288 (0.3%)
All	All	0.35	0/5700	0.67	5/7720 (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	1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	92	LEU	CA-CB-CG	-7.02	99.16	115.30
3	C	108	SER	N-CA-C	5.92	126.98	111.00
1	A	666	PHE	CB-CG-CD1	-5.29	117.09	120.80
3	C	95	ARG	CB-CA-C	-5.09	100.22	110.40
3	C	114	GLY	N-CA-C	5.07	125.77	113.10

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	494	ALA	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	4414	0	4488	299	1
2	B	245	0	244	24	0
3	C	947	0	953	88	0
All	All	5606	0	5685	372	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 33.

The worst 5 of 372 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)
2:B:7:VAL:HG12	2:B:8:LYS:HG3	1.18	1.10
3:C:24:VAL:HG11	3:C:30:ARG:HG2	1.35	1.07
3:C:29:ASP:OD2	3:C:33:ARG:NH2	1.88	1.06
1:A:670:LYS:HA	1:A:670:LYS:HE2	1.40	1.00
1:A:47:ASP:HB2	1:A:50:SER:HB3	1.48	0.95

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:A:114:HIS:ND1	1:A:225:ASP:OD2[4_445]	2.11	0.09

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	560/654 (86%)	509 (91%)	49 (9%)	2 (0%)	34	67
2	B	23/48 (48%)	21 (91%)	1 (4%)	1 (4%)	2	17
3	C	111/173 (64%)	97 (87%)	11 (10%)	3 (3%)	5	26
All	All	694/875 (79%)	627 (90%)	61 (9%)	6 (1%)	17	49

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	272	ALA
2	B	29	TYR
3	C	87	VAL
3	C	75	PRO
1	A	107	GLU

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	485/556 (87%)	428 (88%)	57 (12%)	5	19
2	B	25/43 (58%)	20 (80%)	5 (20%)	1	3
3	C	102/150 (68%)	85 (83%)	17 (17%)	2	8
All	All	612/749 (82%)	533 (87%)	79 (13%)	4	16

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

Mol	Chain	Res	Type
1	A	405	ASP
1	A	512	VAL
3	C	85	LEU
1	A	407	ASP
1	A	447	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	60	ASN
3	C	102	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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 carbohydrates 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 (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(Å²)	Q<0.9
1	A	568/654 (86%)	0.14	23 (4%)	38	37	34, 119, 221, 299	0
2	B	27/48 (56%)	0.08	0	100	100	101, 146, 233, 257	0
3	C	117/173 (67%)	1.37	29 (24%)	0	0	109, 196, 282, 392	0
All	All	712/875 (81%)	0.34	52 (7%)	15	17	34, 137, 241, 392	0

The worst 5 of 52 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	72	VAL	10.2
3	C	103	PHE	8.4
3	C	18	LEU	6.4
3	C	132	MET	5.9
3	C	167	PHE	5.7

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 carbohydrates 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.