



# wwPDB X-ray Structure Validation Summary Report

Mar 6, 2026 – 06:08 PM UTC

PDB ID : 6PPV / pdb\_00006ppv  
Title : Structure of *S. pombe* Lsm1-7 with RNA, polyuridine with 3' guanosine  
Authors : Montemayor, E.J.; Butcher, S.E.  
Deposited on : 2019-07-08  
Resolution : 2.05 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](mailto: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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4-5-2 with Phenix2.0  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

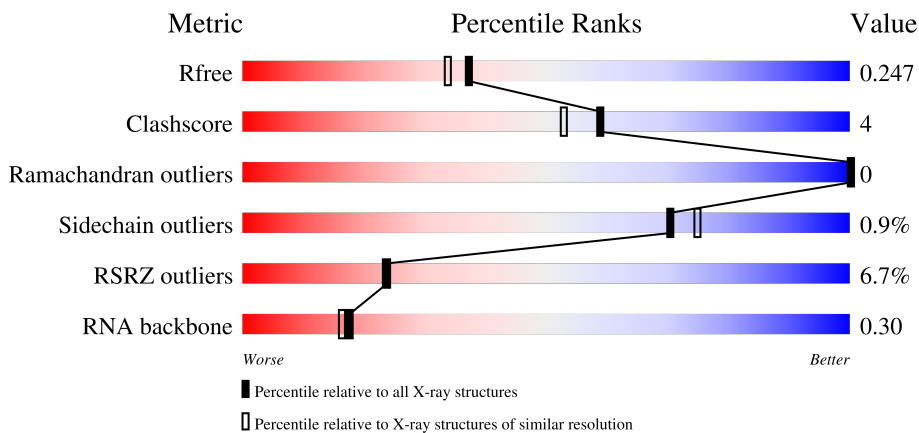
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.05 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.



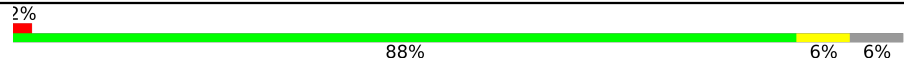

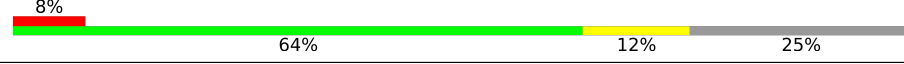

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	180053	2260 (2.04-2.04)
Clashscore	190562	2333 (2.04-2.04)
Ramachandran outliers	187476	2318 (2.04-2.04)
Sidechain outliers	187428	2318 (2.04-2.04)
RSRZ outliers	180081	2260 (2.04-2.04)
RNA backbone	3983	1085 (2.40-1.68)

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  $\geq 3$ , 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	A	86	 11% 70% 10% 20%
2	B	96	 4% 78% 19%
3	C	95	 7% 71% 8% 21%
4	D	129	 7% 51% 12% 36%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	E	80	 <p>2% 88% 6% 6%</p>
6	F	77	 <p>% 86% 6% 8%</p>
7	G	113	 <p>8% 64% 12% 25%</p>
8	H	6	 <p>67% 33%</p>

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 4577 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 U6 snRNA-associated Sm-like protein LSm1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	69	556	360	91	103	2	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP P87173
A	0	SER	-	expression tag	UNP P87173

- Molecule 2 is a protein called U6 snRNA-associated Sm-like protein LSm2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	93	727	469	121	132	5	0	0	0

- Molecule 3 is a protein called Probable U6 snRNA-associated Sm-like protein LSm3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	75	599	383	104	110	2	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-1	GLY	-	expression tag	UNP Q9Y7M4
C	0	SER	-	expression tag	UNP Q9Y7M4

- Molecule 4 is a protein called Probable U6 snRNA-associated Sm-like protein LSm4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	82	630	397	111	117	5	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	122	TRP	-	expression tag	UNP O14352
D	123	SER	-	expression tag	UNP O14352
D	124	HIS	-	expression tag	UNP O14352
D	125	PRO	-	expression tag	UNP O14352
D	126	GLN	-	expression tag	UNP O14352
D	127	PHE	-	expression tag	UNP O14352
D	128	GLU	-	expression tag	UNP O14352
D	129	LYS	-	expression tag	UNP O14352

- Molecule 5 is a protein called U6 snRNA-associated Sm-like protein LSm5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	75	572	366	89	111	6	0	0	0

- Molecule 6 is a protein called U6 snRNA-associated Sm-like protein LSm6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	F	71	544	346	91	105	2	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-1	GLY	-	expression tag	UNP Q9UUI1
F	0	SER	-	expression tag	UNP Q9UUI1

- Molecule 7 is a protein called U6 snRNA-associated Sm-like protein LSm7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
7	G	85	637	401	112	122	2	0	0	0

- Molecule 8 is a RNA chain called RNA (5'-R(\*AP\*UP\*UP\*UP\*UP\*G)-3').

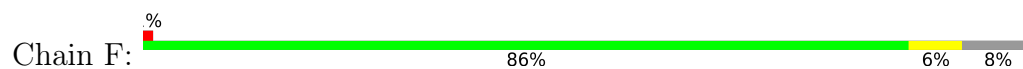
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
8	H	6	122	56	18	43	5	0	0	0

- Molecule 9 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	31	Total O 31 31	0	0
9	B	19	Total O 19 19	0	0
9	C	36	Total O 36 36	0	0
9	D	20	Total O 20 20	0	0
9	E	34	Total O 34 34	0	0
9	F	19	Total O 19 19	0	0
9	G	20	Total O 20 20	0	0
9	H	11	Total O 11 11	0	0



- Molecule 6: U6 snRNA-associated Sm-like protein LSm6



- Molecule 7: U6 snRNA-associated Sm-like protein LSm7



- Molecule 8: RNA (5'-R(\*AP\*UP\*UP\*UP\*UP\*G)-3')



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.02Å 69.02Å 296.07Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	42.07 – 2.05 42.07 – 2.05	Depositor EDS
% Data completeness (in resolution range)	99.9 (42.07-2.05) 100.0 (42.07-2.05)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.99 (at 2.05Å)	Xtrriage
Refinement program	PHENIX 1.15.2_3472	Depositor
R, $R_{free}$	0.209 , 0.245 0.211 , 0.247	Depositor DCC
$R_{free}$ test set	2003 reflections (3.79%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	45.7	Xtrriage
Anisotropy	0.150	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 42.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	0.039 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	4577	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	50.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.31% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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	A	0.27	0/562	0.60	1/758 (0.1%)
2	B	0.24	0/738	0.49	0/998
3	C	0.26	0/607	0.51	0/820
4	D	0.28	0/639	0.53	0/867
5	E	0.29	0/581	0.53	0/787
6	F	0.26	0/551	0.55	0/744
7	G	0.24	0/643	0.48	0/871
8	H	0.32	0/135	0.58	0/208
All	All	0.26	0/4456	0.53	1/6053 (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
5	E	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	59	ASP	CB-CA-C	-5.51	110.20	116.54

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	E	40	VAL	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	556	0	570	4	0
2	B	727	0	736	11	0
3	C	599	0	610	5	0
4	D	630	0	615	11	0
5	E	572	0	573	3	0
6	F	544	0	542	2	0
7	G	637	0	648	9	0
8	H	122	0	64	0	0
9	A	31	0	0	0	0
9	B	19	0	0	2	0
9	C	36	0	0	0	0
9	D	20	0	0	0	0
9	E	34	0	0	1	0
9	F	19	0	0	0	0
9	G	20	0	0	0	0
9	H	11	0	0	0	0
All	All	4577	0	4358	38	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 4.

The worst 5 of 38 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)
3:C:68:LEU:HG	3:C:69:LYS:HG3	1.72	0.72
5:E:78:LYS:NZ	9:E:101:HOH:O	2.27	0.68
4:D:62:ARG:HD2	7:G:93:LEU:O	1.97	0.64
7:G:93:LEU:HD13	7:G:96:ILE:HD11	1.79	0.64
4:D:58:GLU:OE1	7:G:30:ARG:NH2	2.34	0.59

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	67/86 (78%)	64 (96%)	3 (4%)	0	100	100
2	B	91/96 (95%)	89 (98%)	2 (2%)	0	100	100
3	C	71/95 (75%)	69 (97%)	2 (3%)	0	100	100
4	D	80/129 (62%)	76 (95%)	4 (5%)	0	100	100
5	E	73/80 (91%)	70 (96%)	3 (4%)	0	100	100
6	F	69/77 (90%)	67 (97%)	2 (3%)	0	100	100
7	G	83/113 (74%)	79 (95%)	4 (5%)	0	100	100
All	All	534/676 (79%)	514 (96%)	20 (4%)	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	A	61/76 (80%)	59 (97%)	2 (3%)	33	28
2	B	79/87 (91%)	79 (100%)	0	100	100
3	C	65/83 (78%)	65 (100%)	0	100	100
4	D	66/110 (60%)	66 (100%)	0	100	100
5	E	65/71 (92%)	65 (100%)	0	100	100
6	F	59/66 (89%)	57 (97%)	2 (3%)	32	27
7	G	68/98 (69%)	68 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	463/591 (78%)	459 (99%)	4 (1%)	70 75

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	29	ARG
1	A	53	ILE
6	F	34	LEU
6	F	38	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
7	G	37	GLN
7	G	45	GLN
7	G	67	GLN
4	D	34	ASN
4	D	65	ASN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
8	H	5/6 (83%)	2 (40%)	0

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
8	H	96	U
8	H	98	U

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](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	69/86 (80%)	0.17	1 (1%) 73 75	34, 44, 58, 72	0
2	B	93/96 (96%)	0.79	11 (11%) 9 8	36, 54, 79, 95	0
3	C	75/95 (78%)	0.50	4 (5%) 32 32	35, 50, 70, 90	0
4	D	82/129 (63%)	0.82	9 (10%) 10 10	34, 50, 77, 104	0
5	E	75/80 (93%)	0.23	2 (2%) 56 58	36, 44, 60, 79	0
6	F	71/77 (92%)	0.26	1 (1%) 73 75	35, 46, 61, 76	0
7	G	85/113 (75%)	0.66	9 (10%) 11 11	36, 48, 82, 97	0
8	H	6/6 (100%)	-0.33	0 100 100	41, 45, 58, 85	0
All	All	556/682 (81%)	0.51	37 (6%) 24 23	34, 48, 76, 104	0

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	53	PHE	4.9
2	B	93	ASN	4.8
4	D	51	ASP	4.7
3	C	57	PHE	4.5
2	B	76	TYR	4.1

### 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 oligosaccharides in this entry.

## 6.4 Ligands

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.