



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 18, 2024 – 01:01 PM EDT

PDB ID : 5SUI  
Title : Ribosome assembly factor NSA1: C-terminal truncation  
Authors : Lo, Y.H.; Stanley, R.E.  
Deposited on : 2016-08-03  
Resolution : 1.30 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
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)	:	2.37.1

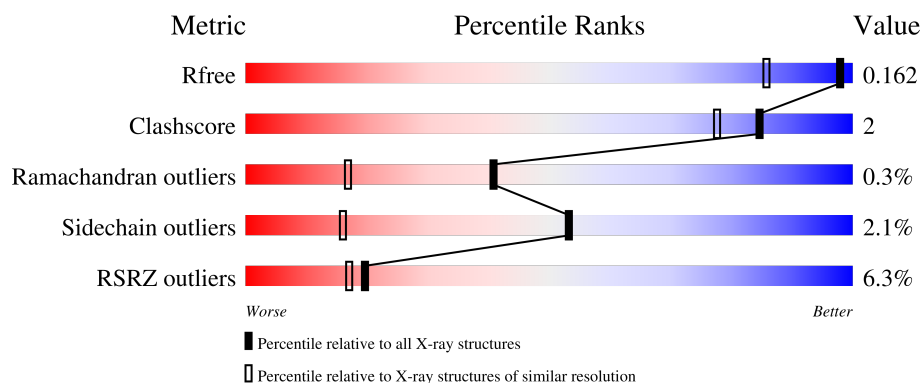
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.30 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	1058 (1.30-1.30)
Clashscore	141614	1101 (1.30-1.30)
Ramachandran outliers	138981	1058 (1.30-1.30)
Sidechain outliers	138945	1058 (1.30-1.30)
RSRZ outliers	127900	1029 (1.30-1.30)

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 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	439	<div> <div>5%</div> <div>72%</div> <div>12%</div> <div>13%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3447 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 Ribosome biogenesis protein NSA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	382	Total	C	N	O	S	0	0	0
			3007	1925	506	565	11			

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	expression tag	UNP P53136
A	2	ALA	-	expression tag	UNP P53136
A	3	MET	-	expression tag	UNP P53136
A	4	GLY	-	expression tag	UNP P53136
A	5	SER	-	expression tag	UNP P53136

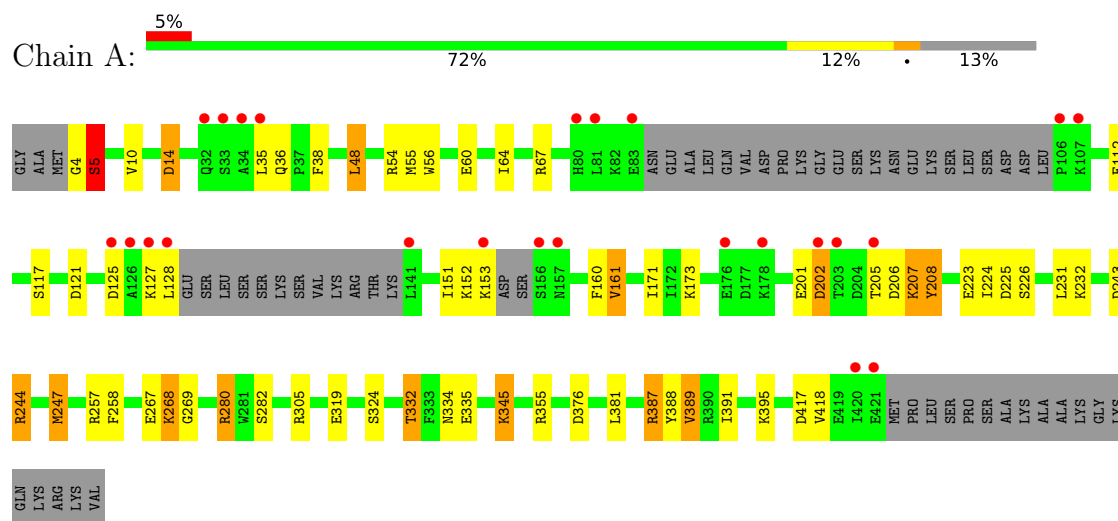
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	440	Total	O	0	0
			440	440		

### 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: Ribosome biogenesis protein NSA1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	54.33Å 81.05Å 87.37Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	59.42 – 1.30 32.48 – 1.30	Depositor EDS
% Data completeness (in resolution range)	99.9 (59.42-1.30) 99.9 (32.48-1.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.04 (at 1.30Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, $R_{free}$	0.129 , 0.162 0.130 , 0.162	Depositor DCC
$R_{free}$ test set	1994 reflections (2.09%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	15.7	Xtriage
Anisotropy	0.169	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 47.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	3447	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.82% 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

### 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	1.51	40/3061 (1.3%)	1.39	41/4124 (1.0%)

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	5

All (40) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	4	GLY	C-O	14.88	1.47	1.23
1	A	201	GLU	CD-OE2	13.09	1.40	1.25
1	A	335	GLU	CD-OE1	11.90	1.38	1.25
1	A	112	GLU	CD-OE1	11.40	1.38	1.25
1	A	201	GLU	CD-OE1	10.98	1.37	1.25
1	A	280	ARG	CZ-NH2	10.91	1.47	1.33
1	A	223	GLU	CD-OE1	10.64	1.37	1.25
1	A	335	GLU	CD-OE2	-9.65	1.15	1.25
1	A	152	LYS	N-CA	8.85	1.64	1.46
1	A	201	GLU	CG-CD	8.85	1.65	1.51
1	A	388	TYR	CE2-CZ	-8.51	1.27	1.38
1	A	324	SER	CB-OG	-8.38	1.31	1.42
1	A	201	GLU	C-O	7.82	1.38	1.23
1	A	152	LYS	CA-CB	7.77	1.71	1.53
1	A	4	GLY	N-CA	7.51	1.57	1.46
1	A	161	VAL	CB-CG2	-7.38	1.37	1.52
1	A	280	ARG	NE-CZ	-7.17	1.23	1.33
1	A	14	ASP	CG-OD2	-6.72	1.09	1.25
1	A	257	ARG	NE-CZ	-6.37	1.24	1.33
1	A	152	LYS	CA-C	6.32	1.69	1.52

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	324	SER	CA-CB	6.32	1.62	1.52
1	A	269	GLY	CA-C	-6.32	1.41	1.51
1	A	112	GLU	CG-CD	6.24	1.61	1.51
1	A	319	GLU	CD-OE1	6.22	1.32	1.25
1	A	201	GLU	CB-CG	6.16	1.63	1.52
1	A	389	VAL	CB-CG2	-6.16	1.40	1.52
1	A	226	SER	CA-CB	-5.76	1.44	1.52
1	A	345	LYS	CE-NZ	5.73	1.63	1.49
1	A	36	GLN	CD-NE2	5.64	1.47	1.32
1	A	201	GLU	CA-CB	-5.64	1.41	1.53
1	A	5	SER	CB-OG	5.56	1.49	1.42
1	A	153	LYS	C-O	5.44	1.33	1.23
1	A	117	SER	CB-OG	-5.42	1.35	1.42
1	A	60	GLU	CD-OE2	5.38	1.31	1.25
1	A	38	PHE	CE1-CZ	5.29	1.47	1.37
1	A	38	PHE	CB-CG	-5.05	1.42	1.51
1	A	232	LYS	CB-CG	-5.05	1.39	1.52
1	A	305	ARG	CB-CG	-5.05	1.39	1.52
1	A	56	TRP	CD1-NE1	-5.04	1.29	1.38
1	A	247	MET	CG-SD	-5.01	1.68	1.81

All (41) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	280	ARG	NE-CZ-NH1	-19.16	110.72	120.30
1	A	202	ASP	CB-CG-OD2	-14.15	105.56	118.30
1	A	247	MET	CG-SD-CE	-12.96	79.47	100.20
1	A	305	ARG	NE-CZ-NH1	-11.05	114.77	120.30
1	A	55	MET	CG-SD-CE	10.75	117.41	100.20
1	A	151	ILE	C-N-CA	10.42	147.74	121.70
1	A	417	ASP	CB-CG-OD1	10.02	127.32	118.30
1	A	38	PHE	CB-CG-CD2	-9.57	114.10	120.80
1	A	332	THR	CA-CB-CG2	-9.13	99.62	112.40
1	A	280	ARG	NH1-CZ-NH2	8.49	128.74	119.40
1	A	4	GLY	N-CA-C	-8.42	92.04	113.10
1	A	395	LYS	CD-CE-NZ	8.37	130.96	111.70
1	A	14	ASP	CB-CG-OD1	-7.99	111.11	118.30
1	A	54	ARG	NE-CZ-NH1	-7.85	116.37	120.30
1	A	223	GLU	OE1-CD-OE2	7.61	132.43	123.30
1	A	48	LEU	CB-CG-CD2	7.60	123.92	111.00
1	A	201	GLU	OE1-CD-OE2	-7.58	114.20	123.30
1	A	257	ARG	NE-CZ-NH1	-7.56	116.52	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	268	LYS	CD-CE-NZ	-7.47	94.52	111.70
1	A	244	ARG	NE-CZ-NH1	7.42	124.01	120.30
1	A	14	ASP	CB-CG-OD2	7.13	124.72	118.30
1	A	125	ASP	CB-CG-OD1	-6.95	112.05	118.30
1	A	201	GLU	CB-CA-C	6.70	123.80	110.40
1	A	152	LYS	N-CA-CB	-6.66	98.61	110.60
1	A	38	PHE	CB-CG-CD1	6.53	125.37	120.80
1	A	206	ASP	CB-CG-OD1	-6.44	112.50	118.30
1	A	258	PHE	CB-CG-CD2	6.42	125.29	120.80
1	A	201	GLU	N-CA-CB	-6.33	99.20	110.60
1	A	243	ASP	CB-CG-OD2	-6.32	112.61	118.30
1	A	67	ARG	NE-CZ-NH2	-6.08	117.26	120.30
1	A	225	ASP	CB-CG-OD1	6.01	123.70	118.30
1	A	173	LYS	CA-CB-CG	5.93	126.46	113.40
1	A	417	ASP	CB-CG-OD2	-5.79	113.09	118.30
1	A	208	TYR	CB-CG-CD2	-5.76	117.54	121.00
1	A	5	SER	N-CA-CB	-5.74	101.89	110.50
1	A	201	GLU	CG-CD-OE2	5.42	129.13	118.30
1	A	14	ASP	OD1-CG-OD2	-5.26	113.31	123.30
1	A	376	ASP	CB-CG-OD2	-5.06	113.75	118.30
1	A	244	ARG	NE-CZ-NH2	-5.04	117.78	120.30
1	A	387	ARG	NE-CZ-NH1	5.03	122.82	120.30
1	A	121	ASP	CB-CG-OD1	-5.00	113.80	118.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	14	ASP	Sidechain
1	A	207	LYS	Mainchain
1	A	267	GLU	Mainchain
1	A	268	LYS	Mainchain
1	A	387	ARG	Sidechain

## 5.2 Too-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 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	3007	0	3028	14	0
2	A	440	0	0	2	0
All	All	3447	0	3028	14	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 2.

All (14) 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:161:VAL:HG12	1:A:171:ILE:CD1	2.18	0.72
1:A:280:ARG:NE	2:A:501:HOH:O	2.15	0.65
1:A:355:ARG:HG2	2:A:810:HOH:O	2.01	0.60
1:A:5:SER:OG	1:A:418:VAL:CG2	2.57	0.53
1:A:332:THR:HG22	1:A:334:ASN:H	1.75	0.52
1:A:161:VAL:HG12	1:A:171:ILE:HD13	1.94	0.50
1:A:10:VAL:HG21	1:A:389:VAL:HG21	1.96	0.47
1:A:231:LEU:HD12	1:A:231:LEU:C	2.34	0.47
1:A:418:VAL:O	1:A:418:VAL:HG23	2.13	0.47
1:A:64:ILE:HD11	1:A:160:PHE:CZ	2.50	0.47
1:A:127:LYS:N	1:A:128:LEU:HA	2.29	0.47
1:A:381:LEU:HD22	1:A:391:ILE:HD12	2.02	0.42
1:A:202:ASP:OD2	1:A:205:THR:HG23	2.20	0.41
1:A:208:TYR:HB2	1:A:224:ILE:HG23	2.03	0.41

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	374/439 (85%)	360 (96%)	13 (4%)	1 (0%)	41 17

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	345	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	332/388 (86%)	325 (98%)	7 (2%)	53 16

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	5	SER
1	A	35	LEU
1	A	48	LEU
1	A	207	LYS
1	A	244	ARG
1	A	247	MET
1	A	282	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	157	ASN

### 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 [i](#)

There are no monosaccharides 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	382/439 (87%)	-0.09	24 (6%) 20 17	11, 18, 41, 56	0

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	141	LEU	5.1
1	A	420	ILE	4.9
1	A	203	THR	4.6
1	A	125	ASP	4.1
1	A	176	GLU	3.8
1	A	127	LYS	3.6
1	A	128	LEU	3.5
1	A	421	GLU	3.4
1	A	34	ALA	3.4
1	A	156	SER	3.4
1	A	106	PRO	3.3
1	A	126	ALA	3.2
1	A	107	LYS	3.1
1	A	202	ASP	3.0
1	A	35	LEU	2.9
1	A	33	SER	2.8
1	A	81	LEU	2.4
1	A	80	HIS	2.4
1	A	157	ASN	2.3
1	A	153	LYS	2.2
1	A	178	LYS	2.2
1	A	32	GLN	2.2
1	A	83	GLU	2.1
1	A	205	THR	2.0

## 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.