



# Full wwPDB X-ray Structure Validation Report ⓘ

May 21, 2020 – 07:20 pm BST

PDB ID : 6DU2  
Title : Structure of Scp1 D96N bound to REST-pS861/4 peptide  
Authors : Burkholder, N.T.; Mayfield, J.E.; Yu, X.; Irani, S.; Arce, D.K.; Jiang, F.;  
Matthews, W.; Xue, Y.; Zhang, Y.J.  
Deposited on : 2018-06-19  
Resolution : 2.50 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
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

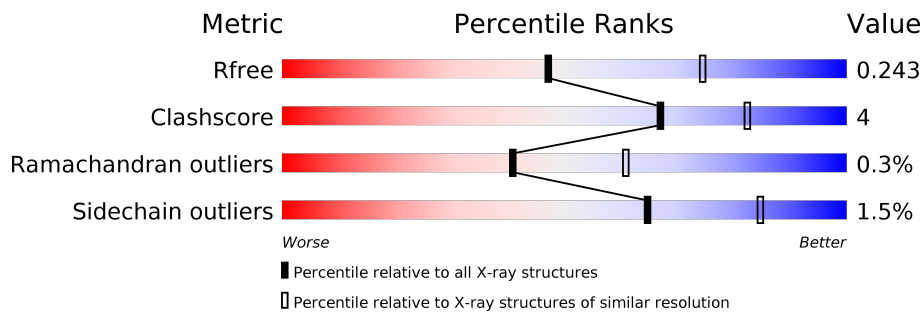
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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  $\geq 3$ , 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	180	93% (green), 7% (yellow), 0% (orange), 0% (red), 0% (grey)
1	B	180	86% (green), 14% (yellow), 0% (orange), 0% (red), 0% (grey)
2	C	12	33% (green), 17% (yellow), 50% (grey)
2	D	12	42% (green), 8% (yellow), 50% (grey)

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 3011 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 carboxy-terminal domain RNA polymerase II polypeptide A small phosphatase 1 isoform X2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	180	1459	938	249	267	5	0	0	0
1	B	180	1449	932	245	267	5	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	96	ASN	ASP	engineered mutation	UNP A0A1S2ZIN9
B	96	ASN	ASP	engineered mutation	UNP A0A1S2ZIN9

- Molecule 2 is a protein called REST-pS861/4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	C	6	49	27	6	14	2	0	0	0
2	D	6	42	25	6	10	1	0	0	0

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Mg	0	0
			1	1		
3	A	1	Total	Mg	0	0
			1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	4	Total O 4 4	0	0
4	B	2	Total O 2 2	0	0
4	C	2	Total O 2 2	0	0
4	D	2	Total O 2 2	0	0

### 3 Residue-property plots [i](#)


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. 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. 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: carboxy-terminal domain RNA polymerase II polypeptide A small phosphatase 1 isoform X2

Chain A:  93% 7%




- Molecule 1: carboxy-terminal domain RNA polymerase II polypeptide A small phosphatase 1 isoform X2

Chain B:  86% 14%



- Molecule 2: REST-pS861/4

Chain C:  33% 17% 50%



- Molecule 2: REST-pS861/4

Chain D:  42% 8% 50%



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	125.33Å 78.57Å 62.88Å 90.00° 112.13° 90.00°	Depositor
Resolution (Å)	39.46 – 2.50 39.46 – 2.50	Depositor EDS
% Data completeness (in resolution range)	98.5 (39.46-2.50) 98.2 (39.46-2.50)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.20 (at 2.51Å)	Xtrriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	(Not available) , (Not available) 0.204 , 0.243	Depositor DCC
$R_{free}$ test set	1929 reflections (10.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	43.9	Xtrriage
Anisotropy	0.573	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 43.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	3011	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.58% 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](#)

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

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	0.33	0/1495	0.48	0/2033
1	B	0.29	0/1485	0.47	0/2022
2	C	0.32	0/29	0.57	0/36
2	D	0.24	0/32	0.46	0/42
All	All	0.31	0/3041	0.48	0/4133

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	A	1459	0	1429	7	0
1	B	1449	0	1407	18	0
2	C	49	0	38	1	0
2	D	42	0	32	1	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	4	0	0	0	0
4	B	2	0	0	0	0
4	C	2	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	2	0	0	0	0
All	All	3011	0	2906	23	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.

All (23) 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:B:203:LEU:HD21	1:B:250:VAL:HG21	1.57	0.83
1:A:251:TYR:HB3	1:A:256:GLN:HA	1.73	0.68
1:B:141:ARG:NH1	1:B:144:GLU:OE1	2.27	0.66
1:B:155:LEU:HD12	2:D:859:ASP:HA	1.81	0.62
1:A:253:VAL:O	1:A:255:ARG:NH2	2.33	0.60
1:B:200:ARG:O	1:B:250:VAL:HG12	2.06	0.55
1:A:116:ILE:HD13	1:A:161:PRO:HB2	1.88	0.54
1:B:244:LEU:HD22	1:B:250:VAL:HG23	1.90	0.54
1:B:190:LYS:NZ	1:B:206:ASP:OD2	2.40	0.54
1:B:132:ARG:HG2	1:B:133:PRO:HD2	1.90	0.52
1:A:209:PRO:HA	1:A:212:TYR:CZ	2.44	0.52
1:A:246:ARG:HA	1:B:107:LYS:HE3	1.95	0.48
1:A:89:ASP:HB3	1:B:107:LYS:HB2	1.97	0.47
2:C:860:LEU:O	2:C:862:PRO:HD3	2.16	0.46
1:B:231:ASP:OD1	1:B:232:THR:N	2.49	0.45
1:B:210:ALA:O	1:B:213:VAL:HG23	2.18	0.43
1:B:132:ARG:NH2	1:B:231:ASP:OD2	2.51	0.43
1:B:209:PRO:HA	1:B:212:TYR:CZ	2.54	0.43
1:B:97:LEU:HB2	1:B:150:LEU:HD11	2.01	0.42
1:B:141:ARG:NE	1:B:242:GLU:OE2	2.46	0.42
1:B:183:PHE:CZ	1:B:186:GLY:HA2	2.55	0.42
1:A:201:ARG:HH22	1:B:107:LYS:NZ	2.18	0.41
1:B:234:LEU:HD23	1:B:234:LEU:HA	1.93	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	178/180 (99%)	173 (97%)	4 (2%)	1 (1%)	25	43
1	B	178/180 (99%)	171 (96%)	7 (4%)	0	100	100
2	C	2/12 (17%)	2 (100%)	0	0	100	100
2	D	3/12 (25%)	2 (67%)	1 (33%)	0	100	100
All	All	361/384 (94%)	348 (96%)	12 (3%)	1 (0%)	41	61

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	255	ARG

### 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	162/163 (99%)	160 (99%)	2 (1%)	71	88
1	B	160/163 (98%)	157 (98%)	3 (2%)	57	80
2	C	4/10 (40%)	4 (100%)	0	100	100
2	D	3/10 (30%)	3 (100%)	0	100	100
All	All	329/346 (95%)	324 (98%)	5 (2%)	65	85

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	82	GLU
1	A	151	PHE
1	B	107	LYS
1	B	151	PHE
1	B	166	LEU

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

Mol	Chain	Res	Type
1	B	103	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

3 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	SEP	C	861	3,2	8,9,10	1.55	1 (12%)	8,12,14	1.78	2 (25%)
2	SEP	D	861	3,2	8,9,10	1.63	1 (12%)	8,12,14	1.55	2 (25%)
2	SEP	C	864	2	8,9,10	2.47	4 (50%)	8,12,14	3.49	5 (62%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SEP	C	861	3,2	-	4/5/8/10	-
2	SEP	D	861	3,2	-	1/5/8/10	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SEP	C	864	2	-	0/5/8/10	-

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	864	SEP	P-O1P	4.70	1.65	1.50
2	C	864	SEP	P-O2P	3.83	1.69	1.54
2	D	861	SEP	P-O1P	3.62	1.62	1.50
2	C	861	SEP	P-O1P	3.34	1.61	1.50
2	C	864	SEP	P-O3P	2.35	1.63	1.54
2	C	864	SEP	P-OG	2.05	1.66	1.60

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	C	864	SEP	OG-CB-CA	5.92	113.91	108.14
2	C	864	SEP	P-OG-CB	-5.06	104.37	118.30
2	C	864	SEP	O2P-P-OG	4.49	118.69	106.73
2	C	861	SEP	OG-CB-CA	3.41	111.46	108.14
2	C	861	SEP	P-OG-CB	-3.24	109.38	118.30
2	D	861	SEP	OG-CB-CA	2.94	111.01	108.14
2	C	864	SEP	O3P-P-OG	2.92	114.52	106.73
2	C	864	SEP	O3P-P-O2P	-2.66	97.49	107.64
2	D	861	SEP	OG-P-O1P	2.46	113.38	106.47

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	861	SEP	N-CA-CB-OG
2	C	861	SEP	CB-OG-P-O1P
2	C	861	SEP	CB-OG-P-O2P
2	D	861	SEP	N-CA-CB-OG
2	C	861	SEP	CB-OG-P-O3P

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are 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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.