



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 12:48 PM GMT

PDB ID : 3S6N
Title : Crystal Structure of the Gemin2-binding domain of SMN, Gemin2 in Complex with SmD1/D2/F/E/G from Human
Authors : Zhang, R.; So, B.R.; Li, P.; Yong, J.; Glisovic, T.; Wan, L.; Dreyfuss, G.
Deposited on : 2011-05-25
Resolution : 2.50 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
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

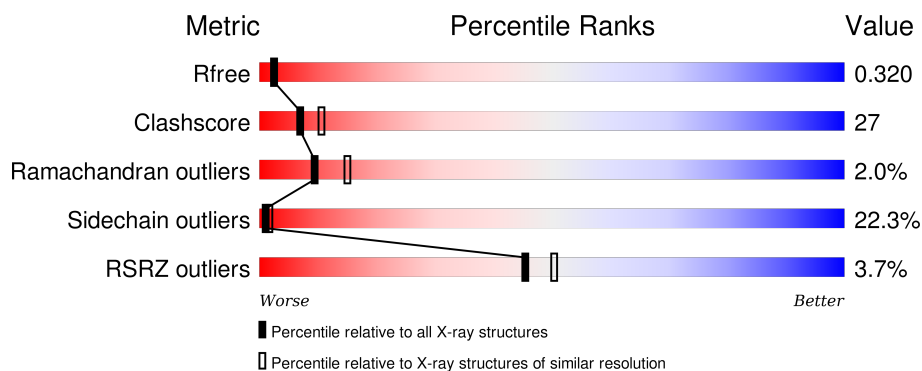
1 Overall quality at a glance

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}	91344	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	2	280	<div> <div>5%</div> <div>37% 21% 8% . 33%</div> </div>
2	A	119	<div> <div>45% 16% 7% . 32%</div> </div>
3	B	118	<div> <div>28% 31% 8% . 31%</div> </div>
4	E	92	<div> <div>5%</div> <div>39% 28% 15% . 16%</div> </div>
5	F	86	<div> <div>49% 30% 7% 14%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	G	76	<div><div><div></div><div></div><div></div><div></div></div><div><div>3%</div><div>30%</div><div>34%</div><div>8%</div><div>28%</div></div></div>
7	M	37	<div><div><div></div><div></div><div></div><div></div></div><div><div>3%</div><div>19%</div><div>19%</div><div>•</div><div>59%</div></div></div>

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 4586 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 Survival of motor neuron protein-interacting protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	2	188	Total	C	N	O	S	0	0	0
			1515	958	271	277	9			

- Molecule 2 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	81	Total	C	N	O	S	0	0	0
			641	409	112	116	4			

- Molecule 3 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	81	Total	C	N	O	S	0	0	0
			654	412	121	116	5			

- Molecule 4 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	77	Total	C	N	O	S	0	0	0
			638	405	113	115	5			

- Molecule 5 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	74	Total	C	N	O	S	0	0	0
			576	372	95	104	5			

- Molecule 6 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	G	55	Total	C	N	O	S	0	0	0
			415	262	74	74	5			

- Molecule 7 is a protein called Survival motor neuron protein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
7	M	15	Total	C	N	O	0	0	0
			114	76	18	20			

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	2	14	Total	O	0	0
			14	14		
8	A	3	Total	O	0	0
			3	3		
8	B	5	Total	O	0	0
			5	5		
8	E	5	Total	O	0	0
			5	5		
8	F	4	Total	O	0	0
			4	4		
8	G	1	Total	O	0	0
			1	1		
8	M	1	Total	O	0	0
			1	1		



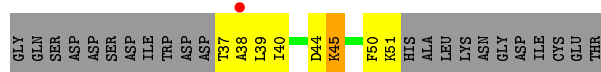
- Molecule 5: Small nuclear ribonucleoprotein F



- Molecule 6: Small nuclear ribonucleoprotein G



- Molecule 7: Survival motor neuron protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	82.83Å 84.60Å 104.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.22 – 2.50 39.22 – 2.50	Depositor EDS
% Data completeness (in resolution range)	76.2 (39.22-2.50) 76.2 (39.22-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.95 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.5.0110	Depositor
R, R_{free}	0.253 , 0.332 0.249 , 0.320	Depositor DCC
R_{free} test set	1015 reflections (5.39%)	DCC
Wilson B-factor (Å ²)	38.1	Xtriage
Anisotropy	0.726	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 51.0	EDS
Estimated twinning fraction	0.019 for k,h,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 19831 reflections	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	4586	wwPDB-VP
Average B, all atoms (Å ²)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.23% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	2	0.68	0/1543	0.82	3/2090 (0.1%)
2	A	0.76	0/649	0.83	1/877 (0.1%)
3	B	0.72	0/661	0.81	0/888
4	E	0.57	0/646	0.72	0/867
5	F	0.61	0/588	0.77	0/795
6	G	0.50	0/417	0.69	0/556
7	M	1.03	0/115	0.93	0/153
All	All	0.67	0/4619	0.79	4/6226 (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	2	0	1
4	E	0	1
All	All	0	2

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	2	275	LEU	N-CA-C	5.93	127.01	111.00
2	A	47	LEU	CA-CB-CG	5.86	128.77	115.30
1	2	274	ASP	CB-CG-OD2	5.83	123.55	118.30
1	2	211	LEU	CA-CB-CG	5.41	127.73	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	2	274	ASP	Peptide
4	E	40	ASN	Peptide

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	2	1515	0	1516	83	0
2	A	641	0	689	27	0
3	B	654	0	683	43	0
4	E	638	0	657	51	0
5	F	576	0	581	49	0
6	G	415	0	434	29	0
7	M	114	0	123	8	0
8	2	14	0	0	0	0
8	A	3	0	0	0	0
8	B	5	0	0	1	0
8	E	5	0	0	0	0
8	F	4	0	0	0	0
8	G	1	0	0	0	0
8	M	1	0	0	0	0
All	All	4586	0	4683	246	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 27.

The worst 5 of 246 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)
4:E:37:GLU:CG	5:F:24:LYS:HB3	1.46	1.46
1:2:84:ILE:CG2	1:2:85:SER:H	1.33	1.27
1:2:84:ILE:HG22	1:2:85:SER:N	1.44	1.17
4:E:37:GLU:HG3	5:F:24:LYS:HB3	1.22	1.11
4:E:37:GLU:HG2	5:F:24:LYS:HB3	1.16	1.08

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	2	178/280 (64%)	162 (91%)	10 (6%)	6 (3%)	5	6
2	A	79/119 (66%)	67 (85%)	10 (13%)	2 (2%)	7	10
3	B	77/118 (65%)	67 (87%)	9 (12%)	1 (1%)	15	26
4	E	75/92 (82%)	70 (93%)	4 (5%)	1 (1%)	15	26
5	F	72/86 (84%)	68 (94%)	4 (6%)	0	100	100
6	G	51/76 (67%)	47 (92%)	3 (6%)	1 (2%)	9	15
7	M	13/37 (35%)	11 (85%)	2 (15%)	0	100	100
All	All	545/808 (68%)	492 (90%)	42 (8%)	11 (2%)	9	15

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	2	84	ILE
1	2	252	ASP
4	E	37	GLU
1	2	65	ASP
6	G	58	GLY

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	2	169/248 (68%)	131 (78%)	38 (22%)	1	1
2	A	76/101 (75%)	63 (83%)	13 (17%)	2	4

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	B	76/110 (69%)	58 (76%)	18 (24%)	1	1
4	E	72/84 (86%)	51 (71%)	21 (29%)	0	0
5	F	62/74 (84%)	51 (82%)	11 (18%)	2	4
6	G	46/66 (70%)	35 (76%)	11 (24%)	1	1
7	M	11/30 (37%)	9 (82%)	2 (18%)	2	3
All	All	512/713 (72%)	398 (78%)	114 (22%)	1	2

5 of 114 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	B	30	SER
3	B	94	ARG
6	G	38	MET
3	B	34	GLN
3	B	48	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
2	A	63	ASN
3	B	48	ASN
5	F	67	ASN
2	A	26	HIS
2	A	39	HIS

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 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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	2	188/280 (67%)	0.39	13 (6%) 20 22	23, 53, 104, 114	0
2	A	81/119 (68%)	-0.07	0 100 100	20, 35, 53, 62	0
3	B	81/118 (68%)	-0.09	0 100 100	19, 37, 64, 72	0
4	E	77/92 (83%)	0.49	5 (6%) 22 25	26, 55, 92, 109	0
5	F	74/86 (86%)	-0.08	0 100 100	21, 44, 71, 75	0
6	G	55/76 (72%)	0.57	2 (3%) 46 51	37, 69, 98, 121	0
7	M	15/37 (40%)	0.51	1 (6%) 21 23	24, 31, 44, 44	0
All	All	571/808 (70%)	0.23	21 (3%) 45 50	19, 47, 87, 121	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	2	144	TRP	7.0
4	E	37	GLU	4.6
1	2	147	PHE	4.6
4	E	38	GLN	4.3
1	2	100	LEU	4.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no carbohydrates 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.