



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 10, 2026 – 11:47 AM UTC

PDB ID : 5XJR / pdb_00005xjr
Title : Crystal Structure of the Gemin2-binding domain of SMN, Gemin2dN39 in Complex with SmD1(1-82)/D2/F/E/G from Human
Authors : Yi, H.; Zhang, R.
Deposited on : 2017-05-04
Resolution : 3.12 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

<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

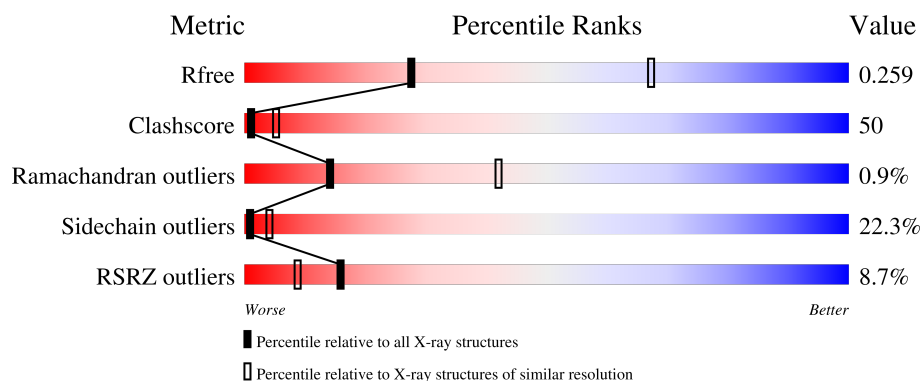
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.12 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	1816 (3.14-3.10)
Clashscore	190562	1906 (3.14-3.10)
Ramachandran outliers	187476	1802 (3.14-3.10)
Sidechain outliers	187428	1802 (3.14-3.10)
RSRZ outliers	180081	1816 (3.14-3.10)

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 ≥ 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	2	241	5% 36% 35% 16% 14%
2	A	82	65% 27% 7%
3	B	118	5% 49% 22% 8% 21%
4	E	92	10% 45% 34% 16%
5	F	86	3% 37% 35% 14% 14%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	G	76	
7	M	37	

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 4890 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 Gem-associated protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	2	208	Total	C	N	O	S	0	0	0
			1672	1055	296	313	8			

- 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	93	Total	C	N	O	S	0	0	0
			754	473	139	137	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.

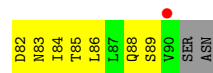
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	74	Total	C	N	O	S	0	0	0
			576	372	95	104	5			

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

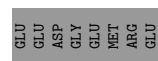
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	G	61	Total	C	N	O	S	0	0	0
			469	297	85	81	6			

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

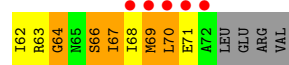
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
7	M	18	Total	C	N	O	0	0	0
			140	90	23	27			



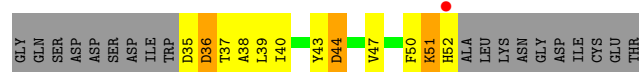
Chain F:



Chain G:



Chain M:



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	82.96Å 114.12Å 130.14Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	60.00 – 3.12 60.00 – 3.12	Depositor EDS
% Data completeness (in resolution range)	83.1 (60.00-3.12) 83.1 (60.00-3.12)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.86 (at 3.12Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.168 , 0.242 0.220 , 0.259	Depositor DCC
R_{free} test set	953 reflections (4.19%)	wwPDB-VP
Wilson B-factor (Å ²)	55.8	Xtriage
Anisotropy	0.025	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 66.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	4890	wwPDB-VP
Average B, all atoms (Å ²)	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.48% 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, 0.333 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	2	1.16	1/1709 (0.1%)	1.04	6/2319 (0.3%)
2	A	0.91	0/649	0.97	0/877
3	B	1.06	0/762	1.12	3/1022 (0.3%)
4	E	0.74	0/646	0.94	0/867
5	F	1.03	0/588	1.05	1/795 (0.1%)
6	G	0.90	1/472 (0.2%)	0.90	0/626
7	M	1.20	2/142 (1.4%)	1.35	1/190 (0.5%)
All	All	1.03	4/4968 (0.1%)	1.03	11/6696 (0.2%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	G	64	GLY	C-N	12.13	1.50	1.33
7	M	44	ASP	CA-C	-5.64	1.45	1.52
1	2	208	THR	CA-C	-5.19	1.47	1.52
7	M	38	ALA	CA-C	-5.03	1.46	1.52

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	F	41	ASN	N-CA-C	-8.51	99.95	110.41
1	2	256	VAL	CB-CA-C	-7.36	106.55	114.35
7	M	38	ALA	N-CA-C	-7.07	104.69	113.38
1	2	121	ARG	N-CA-C	-6.66	104.80	113.12
1	2	45	PRO	CA-C-N	-6.43	113.65	120.66

There are no chirality outliers.

There are no planarity outliers.

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	1672	0	1657	153	0
2	A	641	0	689	41	0
3	B	754	0	784	53	0
4	E	638	0	657	62	0
5	F	576	0	581	70	0
6	G	469	0	491	146	0
7	M	140	0	138	19	0
All	All	4890	0	4997	492	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 50.

The worst 5 of 492 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:66:SER:O	4:E:67:LYS:HG2	1.39	1.20
6:G:32:ARG:CD	6:G:43:ASP:HB2	1.72	1.20
1:2:113:VAL:O	1:2:117:VAL:HG23	1.39	1.20
1:2:135:MET:SD	1:2:182:ILE:CD1	2.29	1.20
5:F:8:LYS:HE3	5:F:12:ASN:OD1	1.35	1.19

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	200/241 (83%)	190 (95%)	8 (4%)	2 (1%)	12	39
2	A	79/82 (96%)	74 (94%)	4 (5%)	1 (1%)	9	33
3	B	89/118 (75%)	81 (91%)	8 (9%)	0	100	100
4	E	75/92 (82%)	73 (97%)	1 (1%)	1 (1%)	9	33
5	F	72/86 (84%)	71 (99%)	1 (1%)	0	100	100
6	G	57/76 (75%)	49 (86%)	8 (14%)	0	100	100
7	M	16/37 (43%)	14 (88%)	1 (6%)	1 (6%)	1	6
All	All	588/732 (80%)	552 (94%)	31 (5%)	5 (1%)	14	42

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	51	GLU
4	E	67	LYS
1	2	136	PRO
1	2	177	PRO
7	M	36	ASP

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	189/216 (88%)	138 (73%)	51 (27%)	0	2
2	A	76/77 (99%)	67 (88%)	9 (12%)	5	20
3	B	87/110 (79%)	73 (84%)	14 (16%)	2	10
4	E	72/84 (86%)	60 (83%)	12 (17%)	2	10
5	F	62/74 (84%)	51 (82%)	11 (18%)	2	8
6	G	51/66 (77%)	26 (51%)	25 (49%)	0	0
7	M	14/30 (47%)	13 (93%)	1 (7%)	13	39
All	All	551/657 (84%)	428 (78%)	123 (22%)	1	4

5 of 123 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	66	ARG
6	G	40	LEU
3	B	71	LYS
6	G	32	ARG
6	G	66	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
3	B	48	ASN
6	G	65	ASN
4	E	19	ASN
5	F	68	ASN
3	B	62	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](#)

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, 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	208/241 (86%)	0.27	11 (5%) 32 17	12, 46, 89, 124	0
2	A	81/82 (98%)	-0.04	0 100 100	16, 33, 63, 79	0
3	B	93/118 (78%)	0.10	6 (6%) 25 13	14, 34, 107, 127	0
4	E	77/92 (83%)	0.74	9 (11%) 9 5	25, 59, 108, 137	0
5	F	74/86 (86%)	-0.01	3 (4%) 41 23	21, 43, 74, 100	0
6	G	61/76 (80%)	1.81	23 (37%) 1 0	65, 101, 141, 178	0
7	M	18/37 (48%)	0.56	1 (5%) 30 16	35, 56, 102, 120	0
All	All	612/732 (83%)	0.39	53 (8%) 16 9	12, 46, 111, 178	0

The worst 5 of 53 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	2	42	PRO	4.9
6	G	35	ASP	4.8
6	G	12	PHE	4.6
6	G	69	MET	3.9
6	G	10	LYS	3.8

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.