



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 25, 2017 – 09:06 PM EST

PDB ID : 5NNI
Title : Dimer structure of Sortilin ectodomain crystal form 2, 3.2 Angstrom
Authors : Leloup, N.O.L.; Janssen, B.J.C.
Deposited on : unknown
Resolution : 3.21 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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20030345
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
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) : rb-20030345

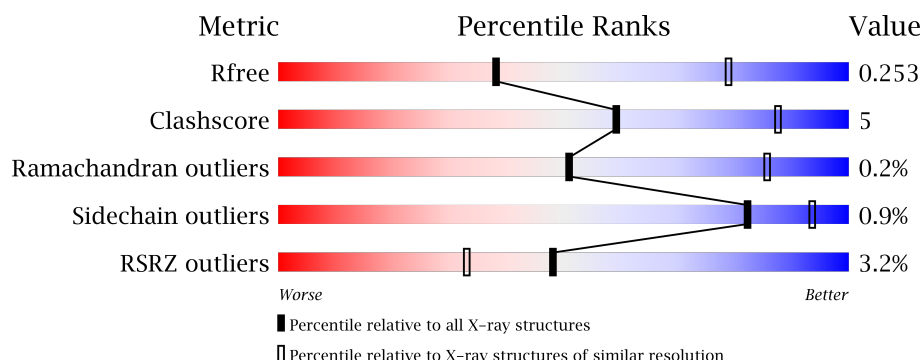
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.21 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}	100719	1036 (3.24-3.20)
Clashscore	112137	1161 (3.24-3.20)
Ramachandran outliers	110173	1140 (3.24-3.20)
Sidechain outliers	110143	1139 (3.24-3.20)
RSRZ outliers	101464	1040 (3.24-3.20)

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	A	732	<div> <div>2%</div> <div>76%</div> <div>13%</div> <div>11%</div> </div>
1	B	732	<div> <div>4%</div> <div>76%</div> <div>12%</div> <div>11%</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	NAG	B	802	-	-	-	X

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 10409 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 Sortilin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	651	Total	C	N	O	S	0	0	0
			5120	3237	861	993	29			
1	B	651	Total	C	N	O	S	0	0	0
			5116	3236	861	990	29			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	723	ALA	-	expression tag	UNP Q6PHU5
A	724	ALA	-	expression tag	UNP Q6PHU5
A	725	ALA	-	expression tag	UNP Q6PHU5
A	726	HIS	-	expression tag	UNP Q6PHU5
A	727	HIS	-	expression tag	UNP Q6PHU5
A	728	HIS	-	expression tag	UNP Q6PHU5
A	729	HIS	-	expression tag	UNP Q6PHU5
A	730	HIS	-	expression tag	UNP Q6PHU5
A	731	HIS	-	expression tag	UNP Q6PHU5
B	723	ALA	-	expression tag	UNP Q6PHU5
B	724	ALA	-	expression tag	UNP Q6PHU5
B	725	ALA	-	expression tag	UNP Q6PHU5
B	726	HIS	-	expression tag	UNP Q6PHU5
B	727	HIS	-	expression tag	UNP Q6PHU5
B	728	HIS	-	expression tag	UNP Q6PHU5
B	729	HIS	-	expression tag	UNP Q6PHU5
B	730	HIS	-	expression tag	UNP Q6PHU5
B	731	HIS	-	expression tag	UNP Q6PHU5

- Molecule 2 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: C₈H₁₅NO₆).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 3 is BETA-D-MANNOSE (three-letter code: BMA) (formula: C₆H₁₂O₆).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			11	6	5		
3	A	1	Total	C	O	0	0
			11	6	5		
3	B	1	Total	C	O	0	0
			11	6	5		

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 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.

Chain A:

2% 76% 13% 11%

GLY GLN ASP ARG LEU ASP PRO PRO PRO PRO ALA PRO LEU LEU ARG TRP ALA GLY VAL GLY VAL SER GLY LEU ARG ALA ALA PRO GLY PRO VAL VAL ARG GLY ARG TRP ARG GLY PRO ASP GLN ASP

DY3 L94 V102 I103 Q108 DI16 F121 I128 N129 W130 T131 F132 I133 A140 K148 V149 I150 R166 D169 M187 D195 L198 L206 L206 N211 F212 G213 E214 H220 V223 W228 G229 I233 F236 S243 L253 T266

V284 D287 LYS ASP T290 R293 G301 S311 VAL GLY Q314 Y318 D326 E331 V332 D333 F344 T345 S346 D347 D348 Y353 L361 D370 F371 T372 R378 L386 D399 L407 W412 D416 ALA THR ALA Y420 M423 A431 S432 Y432

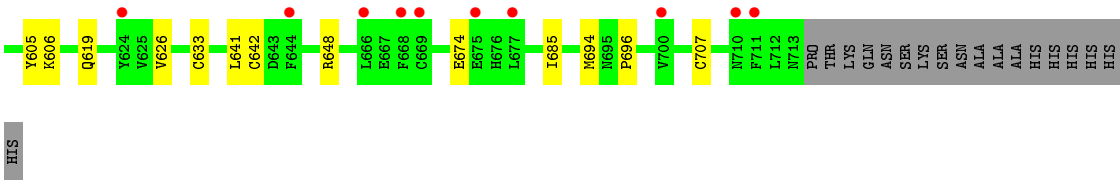
S434 I435 S436 Q437 P445 L446 A451 I454 V455 I456 V461 Q462 V472 W482 L486 W490 L495 D519 C523 S541 A550 E557 S561 P562 I562 R562 Q563 D570 H597 K598 L603 R606 K614 S615 S616 A627 C633 S634

F640 L641 G642 R648 E665 L666 E667 E674 E675 H676 G681 Y682 L703 K704 K705 K706 K707 T708 S709 I712 N713 PRO THR LYS GLN ASN SER LYS SER ASN PHE ALA ALA ALA HIS HIS HIS HIS HIS HIS

Chain B:

76% 12% 11%

GLY GLN ASP ARG LEU ASP ALA PRO PRO PRO LEU ARG TRP ALA GLY PRO VAL GLY VAL SER TRP GLY LEU ARG ALA ALA ALA PRO GLY PRO GLY VAL VAL ARG ALA GLY ARG TRP ARG ARG GLY ALA PRO ALA ASP GLN ASP C93 G54 R55 L56 P57 D58



HIS

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	79.84Å 137.17Å 147.56Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	62.20 – 3.21 62.20 – 3.21	Depositor EDS
% Data completeness (in resolution range)	99.5 (62.20-3.21) 99.5 (62.20-3.21)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.62 (at 3.19Å)	Xtriage
Refinement program	PHENIX (1.11_2567: ???)	Depositor
R, R_{free}	0.227 , 0.255 0.225 , 0.253	Depositor DCC
R_{free} test set	1363 reflections (5.04%)	DCC
Wilson B-factor (Å ²)	94.3	Xtriage
Anisotropy	0.384	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 56.1	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	0.93	EDS
Total number of atoms	10409	wwPDB-VP
Average B, all atoms (Å ²)	114.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.36% 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

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

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	0.24	0/5240	0.43	0/7095
1	B	0.25	0/5237	0.43	0/7092
All	All	0.24	0/10477	0.43	0/14187

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

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	5120	0	4903	55	0
1	B	5116	0	4905	56	0
2	A	70	0	61	0	0
2	B	70	0	62	6	0
3	A	22	0	20	0	0
3	B	11	0	10	0	0
All	All	10409	0	9961	109	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 5.

The worst 5 of 109 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:324:ASN:HB3	2:B:802:NAG:H62	1.70	0.74
1:B:324:ASN:HB2	2:B:803:NAG:O7	1.92	0.70
1:B:94:LEU:HD11	1:B:133:ILE:HG13	1.73	0.69
1:B:140:ALA:HB3	1:B:150:ILE:HB	1.76	0.66
1:B:148:LYS:NZ	1:B:195:ASP:OD1	2.30	0.65

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	641/732 (88%)	609 (95%)	32 (5%)	0	100	100
1	B	643/732 (88%)	599 (93%)	42 (6%)	2 (0%)	44	81
All	All	1284/1464 (88%)	1208 (94%)	74 (6%)	2 (0%)	51	86

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	340	PHE
1	B	626	VAL

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	568/626 (91%)	565 (100%)	3 (0%)	91	96
1	B	567/626 (91%)	560 (99%)	7 (1%)	75	91
All	All	1135/1252 (91%)	1125 (99%)	10 (1%)	82	93

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	412	ASN
1	B	495	LEU
1	B	605	TYR
1	B	220	HIS
1	B	582	ASP

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	423	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

13 ligands 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	NAG	A	801	1	14,14,15	0.37	0	15,19,21	1.13	1 (6%)
2	NAG	A	802	1,2	14,14,15	0.53	0	15,19,21	2.04	4 (26%)
2	NAG	A	803	3,2	14,14,15	0.23	0	15,19,21	2.18	6 (40%)
3	BMA	A	804	2	11,11,12	0.28	0	13,15,17	0.64	0
2	NAG	A	805	1,2	14,14,15	0.30	0	15,19,21	0.70	0
2	NAG	A	806	3,2	14,14,15	0.29	0	15,19,21	1.15	2 (13%)
3	BMA	A	807	2	11,11,12	0.30	0	13,15,17	0.89	0
2	NAG	B	801	1	14,14,15	0.27	0	15,19,21	0.61	0
2	NAG	B	802	1,2	14,14,15	0.44	0	15,19,21	1.45	3 (20%)
2	NAG	B	803	2	14,14,15	0.37	0	15,19,21	2.16	4 (26%)
2	NAG	B	804	1,2	14,14,15	0.36	0	15,19,21	1.66	2 (13%)
2	NAG	B	805	1,3,2	14,14,15	0.67	0	15,19,21	2.59	5 (33%)
3	BMA	B	806	2	11,11,12	0.34	0	13,15,17	1.52	2 (15%)

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	NAG	A	801	1	-	0/6/23/26	0/1/1/1
2	NAG	A	802	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	803	3,2	-	0/6/23/26	0/1/1/1
3	BMA	A	804	2	-	0/2/19/22	0/1/1/1
2	NAG	A	805	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	806	3,2	-	0/6/23/26	0/1/1/1
3	BMA	A	807	2	-	0/2/19/22	0/1/1/1
2	NAG	B	801	1	-	0/6/23/26	0/1/1/1
2	NAG	B	802	1,2	-	0/6/23/26	0/1/1/1
2	NAG	B	803	2	-	0/6/23/26	0/1/1/1
2	NAG	B	804	1,2	-	0/6/23/26	0/1/1/1
2	NAG	B	805	1,3,2	-	0/6/23/26	0/1/1/1
3	BMA	B	806	2	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	803	NAG	C1-O5-C5	-5.38	104.75	112.17
2	B	805	NAG	O4-C4-C3	-5.15	99.15	110.36
2	B	805	NAG	O4-C4-C5	-5.13	96.36	109.28
2	B	804	NAG	O4-C4-C3	-5.10	99.25	110.36
2	B	805	NAG	C1-O5-C5	-4.62	105.79	112.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	802	NAG	4	0
2	B	803	NAG	2	0

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	A	651/732 (88%)	0.16	13 (1%) 65 51	45, 96, 167, 240	0
1	B	651/732 (88%)	0.36	29 (4%) 34 22	71, 119, 193, 278	0
All	All	1302/1464 (88%)	0.26	42 (3%) 48 33	45, 108, 184, 278	0

The worst 5 of 42 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	668	PHE	6.2
1	A	423	ASN	4.8
1	B	56	LEU	4.6
1	B	700	VAL	4.5
1	B	677	LEU	4.3

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 carbohydrates in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	NAG	B	802	14/15	0.79	0.39	2.32	140,154,163,164	0
2	NAG	A	802	14/15	0.89	0.26	0.27	87,102,111,117	0
2	NAG	A	805	14/15	0.90	0.19	0.18	92,101,109,113	0
2	NAG	B	804	14/15	0.90	0.20	-0.94	107,124,135,143	0
2	NAG	A	801	14/15	0.84	0.24	-	149,157,170,172	0
2	NAG	A	806	14/15	0.92	0.18	-	84,109,116,121	0
3	BMA	A	807	11/12	0.56	0.23	-	127,132,137,142	0
2	NAG	B	803	14/15	0.66	0.36	-	145,155,158,159	0
3	BMA	B	806	11/12	0.77	0.24	-	162,165,167,169	0
2	NAG	A	803	14/15	0.91	0.18	-	92,116,126,134	0
2	NAG	B	805	14/15	0.74	0.29	-	149,159,164,165	0
3	BMA	A	804	11/12	0.80	0.29	-	142,149,153,153	0
2	NAG	B	801	14/15	0.87	0.19	-	110,118,125,125	0

6.5 Other polymers [i](#)

There are no such residues in this entry.