



Full wwPDB X-ray Structure Validation Report ⓘ

May 17, 2020 – 12:24 pm BST

PDB ID : 6PNQ
Title : Crystal structure of the SS2A splice insert-containing neurexin-1 LNS2 domain in complex with neurexophilin-1
Authors : Wilson, S.C.; White, K.I.; Zhou, Q.; Brunger, A.T.
Deposited on : 2019-07-02
Resolution : 1.95 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

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
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

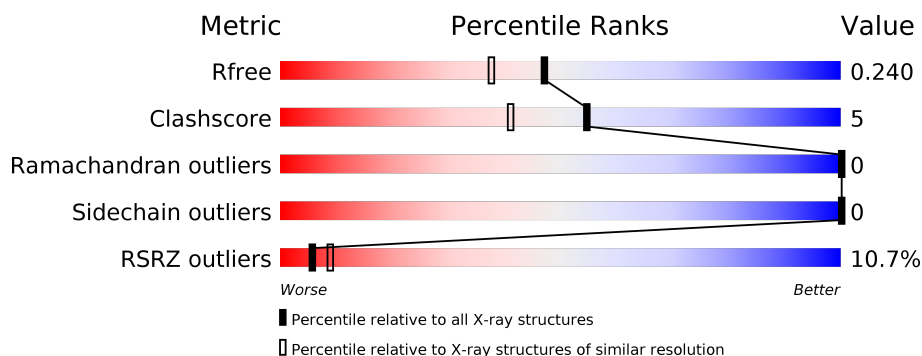
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.95 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	4310 (1.96-1.92)
Clashscore	141614	1023 (1.94-1.94)
Ramachandran outliers	138981	1007 (1.94-1.94)
Sidechain outliers	138945	1007 (1.94-1.94)
RSRZ outliers	127900	4250 (1.96-1.92)

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 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	202	
2	B	168	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5340 atoms, of which 2549 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 Neurexin-1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	194	Total	C	H	N	O	S	0	6	0
			3006	953	1490	255	299	9			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	280	ASP	-	expression tag	UNP Q9CS84
A	281	ALA	-	expression tag	UNP Q9CS84
A	282	SER	-	expression tag	UNP Q9CS84
A	?	-	VAL	deletion	UNP Q9CS84
A	?	-	ASN	deletion	UNP Q9CS84
A	?	-	LYS	deletion	UNP Q9CS84
A	?	-	LEU	deletion	UNP Q9CS84
A	?	-	HIS	deletion	UNP Q9CS84
A	?	-	CYS	deletion	UNP Q9CS84
A	?	-	SER	deletion	UNP Q9CS84
A	481	ALA	-	expression tag	UNP Q9CS84
A	482	ALA	-	expression tag	UNP Q9CS84
A	483	LEU	-	expression tag	UNP Q9CS84
A	484	GLU	-	expression tag	UNP Q9CS84
A	485	VAL	-	expression tag	UNP Q9CS84
A	486	LEU	-	expression tag	UNP Q9CS84
A	487	PHE	-	expression tag	UNP Q9CS84
A	488	GLN	-	expression tag	UNP Q9CS84

- Molecule 2 is a protein called Neurexophilin-1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	133	Total	C	H	N	O	S	0	4	0
			2154	706	1059	175	208	6			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	110	ASP	-	expression tag	UNP Q63366
B	111	ALA	-	expression tag	UNP Q63366
B	112	ALA	-	expression tag	UNP Q63366
B	113	GLN	-	expression tag	UNP Q63366
B	114	PRO	-	expression tag	UNP Q63366
B	115	ALA	-	expression tag	UNP Q63366
B	116	ALA	-	expression tag	UNP Q63366
B	117	ARG	-	expression tag	UNP Q63366
B	118	ASP	-	expression tag	UNP Q63366
B	146	ASP	ASN	engineered mutation	UNP Q63366
B	156	ASP	ASN	engineered mutation	UNP Q63366
B	162	ASP	ASN	engineered mutation	UNP Q63366
B	272	LEU	-	expression tag	UNP Q63366
B	273	GLU	-	expression tag	UNP Q63366
B	274	VAL	-	expression tag	UNP Q63366
B	275	LEU	-	expression tag	UNP Q63366
B	276	PHE	-	expression tag	UNP Q63366
B	277	GLN	-	expression tag	UNP Q63366

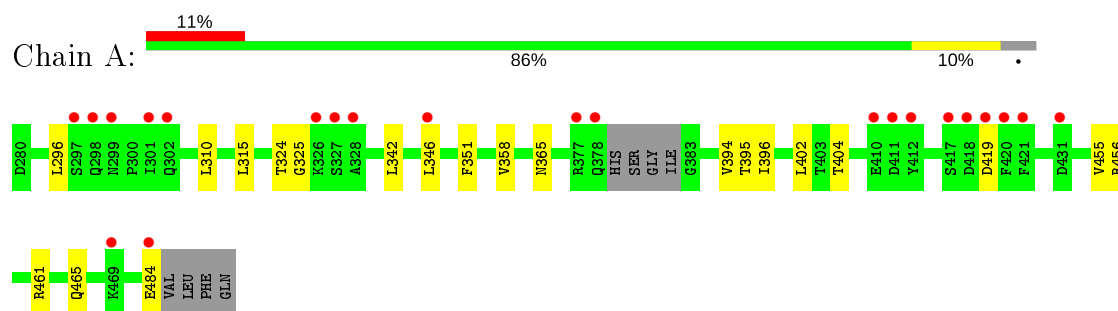
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	87	Total O 87 87	0	0
3	B	93	Total O 93 93	0	0

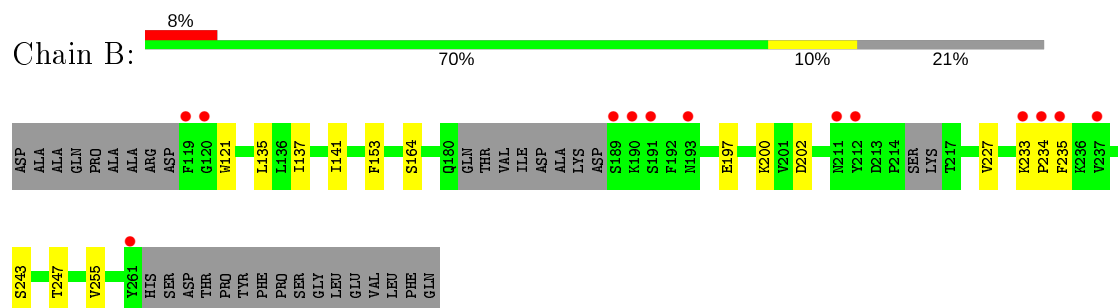
3 Residue-property plots [i](#)

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 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: Neurexin-1



• Molecule 2: Neurexophilin-1



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	71.96Å 61.44Å 79.39Å 90.00° 106.22° 90.00°	Depositor
Resolution (Å)	45.92 – 1.95 45.92 – 1.95	Depositor EDS
% Data completeness (in resolution range)	83.2 (45.92-1.95) 83.2 (45.92-1.95)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.70 (at 1.95Å)	Xtriage
Refinement program	PHENIX 1.13 _2998	Depositor
R, R_{free}	0.196 , 0.240 0.196 , 0.240	Depositor DCC
R_{free} test set	1997 reflections (9.81%)	wwPDB-VP
Wilson B-factor (Å ²)	20.7	Xtriage
Anisotropy	0.138	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 49.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	5340	wwPDB-VP
Average B, all atoms (Å ²)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.35% 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 [i](#)

5.1 Standard geometry [i](#)

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	0.29	0/1550	0.52	0/2097
2	B	0.32	0/1128	0.52	0/1528
All	All	0.30	0/2678	0.52	0/3625

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	1516	1490	1473	17	0
2	B	1095	1059	1058	12	0
3	A	87	0	0	2	1
3	B	93	0	0	2	2
All	All	2791	2549	2531	26	2

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.

All (26) 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:342:LEU:HD22	1:A:396:ILE:HD12	1.77	0.67

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:197:GLU:OE1	3:B:301:HOH:O	2.14	0.66
1:A:325:GLY:N	3:A:503:HOH:O	2.27	0.62
1:A:484:GLU:OE2	3:A:501:HOH:O	2.16	0.59
2:B:233:LYS:N	2:B:234:PRO:HD2	2.19	0.57
1:A:346:LEU:HD22	1:A:394[B]:VAL:CG1	2.34	0.56
2:B:153:PHE:HB2	2:B:227[B]:VAL:HG22	1.87	0.55
1:A:315:LEU:HD21	1:A:365:ASN:HB2	1.91	0.53
2:B:121:TRP:CZ2	2:B:141:ILE:HD11	2.45	0.52
1:A:351:PHE:CD2	2:B:255:VAL:HG22	2.45	0.52
2:B:235:PHE:N	3:B:307:HOH:O	2.45	0.50
1:A:395[B]:THR:HG22	1:A:404:THR:HG23	1.93	0.49
1:A:395[A]:THR:CG2	1:A:402:LEU:HD11	2.45	0.46
1:A:324:THR:O	1:A:324:THR:HG23	2.17	0.44
1:A:296:LEU:HD12	1:A:419:ASP:O	2.18	0.44
1:A:461:ARG:NH2	1:A:465:GLN:OE1	2.51	0.44
2:B:200:LYS:NZ	2:B:202:ASP:OD1	2.46	0.44
2:B:164:SER:HG	2:B:243[A]:SER:HB2	1.82	0.43
1:A:346:LEU:HD13	1:A:394[B]:VAL:HG12	2.02	0.42
1:A:455:VAL:HG22	1:A:456:ARG:N	2.33	0.42
1:A:346:LEU:HD22	1:A:394[A]:VAL:HG23	2.00	0.42
1:A:351:PHE:HD2	2:B:255:VAL:HG22	1.83	0.42
1:A:358:VAL:HG21	2:B:247:THR:O	2.19	0.42
2:B:135:LEU:HD21	2:B:137:ILE:HD11	2.03	0.41
2:B:121:TRP:CH2	2:B:141:ILE:HD11	2.56	0.41
1:A:310:LEU:C	1:A:310:LEU:HD12	2.41	0.41

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:534:HOH:O	3:B:362:HOH:O[3_445]	2.05	0.15
3:B:372:HOH:O	3:B:388:HOH:O[4_455]	2.07	0.13

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	196/202 (97%)	188 (96%)	8 (4%)	0	100	100
2	B	131/168 (78%)	125 (95%)	6 (5%)	0	100	100
All	All	327/370 (88%)	313 (96%)	14 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	169/170 (99%)	169 (100%)	0	100	100
2	B	128/153 (84%)	128 (100%)	0	100	100
All	All	297/323 (92%)	297 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	A	194/202 (96%)	0.42	22 (11%) 5 8	12, 34, 86, 120	0
2	B	133/168 (79%)	0.28	13 (9%) 7 11	8, 22, 81, 109	0
All	All	327/370 (88%)	0.36	35 (10%) 6 9	8, 28, 83, 120	0

All (35) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	377	ARG	5.7
1	A	299	ASN	5.1
1	A	326	LYS	4.5
2	B	237	VAL	4.4
1	A	420	PHE	4.2
1	A	410	GLU	4.1
2	B	234	PRO	4.1
1	A	327	SER	4.0
2	B	211	ASN	3.9
1	A	298	GLN	3.6
1	A	378	GLN	3.5
2	B	212	TYR	3.4
1	A	418	ASP	3.4
1	A	412	TYR	3.4
2	B	119	PHE	3.4
2	B	120	GLY	3.3
2	B	189	SER	3.3
1	A	301	ILE	3.2
2	B	190	LYS	3.2
1	A	417	SER	3.1
2	B	235	PHE	3.1
2	B	261	TYR	2.9
1	A	484	GLU	2.8
1	A	346	LEU	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	419	ASP	2.7
1	A	302	GLN	2.6
1	A	421	PHE	2.4
1	A	431	ASP	2.4
2	B	193	ASN	2.4
1	A	297	SER	2.3
2	B	191	SER	2.3
1	A	328	ALA	2.2
1	A	411	ASP	2.2
2	B	233	LYS	2.1
1	A	469	LYS	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 carbohydrates 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.