



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 17, 2026 – 11:01 PM UTC

PDB ID : 4AKV / pdb_00004akv
Title : Crystal structure of human sorting nexin 33 (SNX33)
Authors : Oberholzer, A.; Froese, D.S.; Krojer, T.; Shrestha, L.; Filippakopoulos, P.; Muniz, J.R.C.; von Delft, F.; Arrowsmith, C.; Weigelt, J.; Edwards, A.; Bountra, C.; Yue, W.W.
Deposited on : 2012-02-29
Resolution : 2.65 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
Mogul : 2022.3.0, CSD as543be (2022)
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

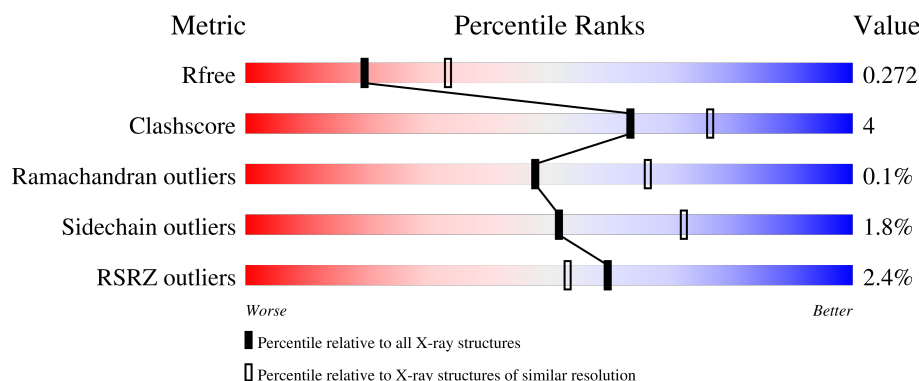
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

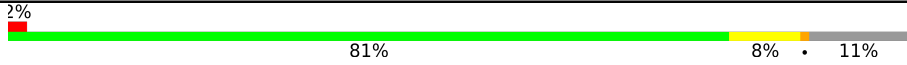
The reported resolution of this entry is 2.65 Å.

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	1110 (2.66-2.66)
Clashscore	190562	1141 (2.66-2.66)
Ramachandran outliers	187476	1126 (2.66-2.66)
Sidechain outliers	187428	1126 (2.66-2.66)
RSRZ outliers	180081	1110 (2.66-2.66)

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	A	386	
1	B	386	

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	SO4	B	602	-	-	X	-
2	SO4	B	603	-	-	X	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5805 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 SORTING NEXIN-33.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	345	Total	C	N	O	S	0	0	0
			2834	1801	504	509	20			
1	B	349	Total	C	N	O	S	0	0	0
			2879	1829	514	516	20			

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	189	MET	-	expression tag	UNP Q8WV41
A	190	HIS	-	expression tag	UNP Q8WV41
A	191	HIS	-	expression tag	UNP Q8WV41
A	192	HIS	-	expression tag	UNP Q8WV41
A	193	HIS	-	expression tag	UNP Q8WV41
A	194	HIS	-	expression tag	UNP Q8WV41
A	195	HIS	-	expression tag	UNP Q8WV41
A	196	SER	-	expression tag	UNP Q8WV41
A	197	SER	-	expression tag	UNP Q8WV41
A	198	GLY	-	expression tag	UNP Q8WV41
A	199	VAL	-	expression tag	UNP Q8WV41
A	200	ASP	-	expression tag	UNP Q8WV41
A	201	LEU	-	expression tag	UNP Q8WV41
A	202	GLY	-	expression tag	UNP Q8WV41
A	203	THR	-	expression tag	UNP Q8WV41
A	204	GLU	-	expression tag	UNP Q8WV41
A	205	ASN	-	expression tag	UNP Q8WV41
A	206	LEU	-	expression tag	UNP Q8WV41
A	207	TYR	-	expression tag	UNP Q8WV41
A	208	PHE	-	expression tag	UNP Q8WV41
A	209	GLN	-	expression tag	UNP Q8WV41
A	210	SER	-	expression tag	UNP Q8WV41
A	211	MET	-	expression tag	UNP Q8WV41
B	189	MET	-	expression tag	UNP Q8WV41
B	190	HIS	-	expression tag	UNP Q8WV41

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	191	HIS	-	expression tag	UNP Q8WV41
B	192	HIS	-	expression tag	UNP Q8WV41
B	193	HIS	-	expression tag	UNP Q8WV41
B	194	HIS	-	expression tag	UNP Q8WV41
B	195	HIS	-	expression tag	UNP Q8WV41
B	196	SER	-	expression tag	UNP Q8WV41
B	197	SER	-	expression tag	UNP Q8WV41
B	198	GLY	-	expression tag	UNP Q8WV41
B	199	VAL	-	expression tag	UNP Q8WV41
B	200	ASP	-	expression tag	UNP Q8WV41
B	201	LEU	-	expression tag	UNP Q8WV41
B	202	GLY	-	expression tag	UNP Q8WV41
B	203	THR	-	expression tag	UNP Q8WV41
B	204	GLU	-	expression tag	UNP Q8WV41
B	205	ASN	-	expression tag	UNP Q8WV41
B	206	LEU	-	expression tag	UNP Q8WV41
B	207	TYR	-	expression tag	UNP Q8WV41
B	208	PHE	-	expression tag	UNP Q8WV41
B	209	GLN	-	expression tag	UNP Q8WV41
B	210	SER	-	expression tag	UNP Q8WV41
B	211	MET	-	expression tag	UNP Q8WV41

- Molecule 2 is SULFATE ION (CCD ID: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		

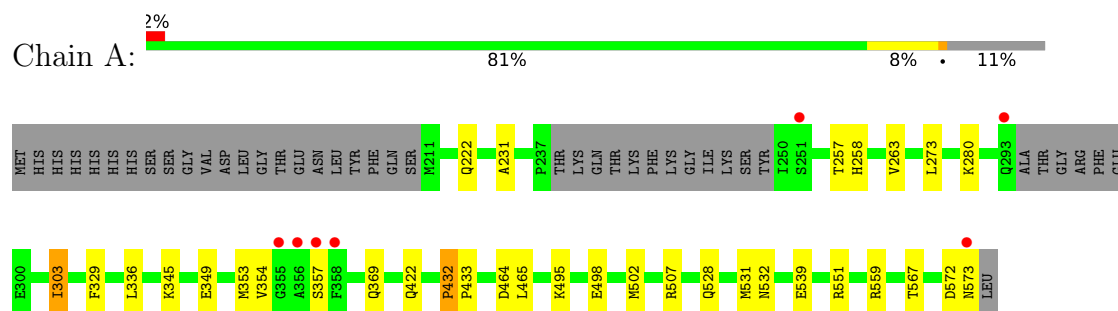
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	28	Total	O	0	0
			28	28		
3	B	34	Total	O	0	0
			34	34		

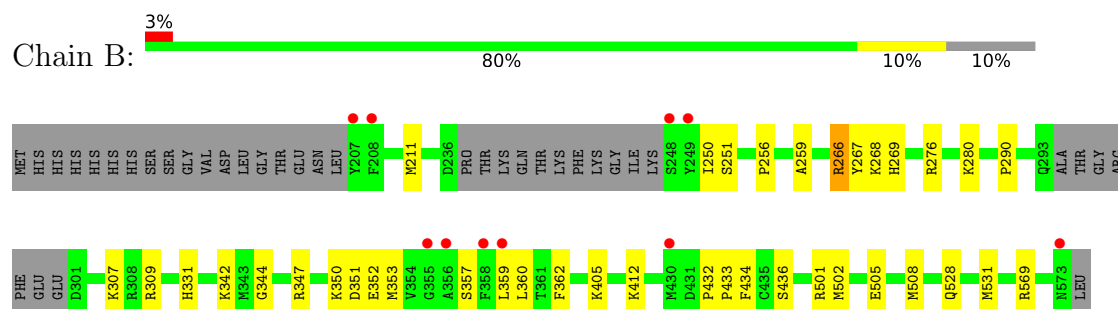
3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: SORTING NEXIN-33



• Molecule 1: SORTING NEXIN-33



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	68.65Å 216.21Å 58.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	65.43 – 2.65 65.43 – 2.65	Depositor EDS
% Data completeness (in resolution range)	99.7 (65.43-2.65) 99.7 (65.43-2.65)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.35 (at 2.65Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.210 , 0.270 0.216 , 0.272	Depositor DCC
R_{free} test set	1319 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å ²)	72.3	Xtriage
Anisotropy	0.372	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 56.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	5805	wwPDB-VP
Average B, all atoms (Å ²)	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.93% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

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	0.27	0/2904	0.68	1/3904 (0.0%)
1	B	0.28	0/2950	0.72	3/3963 (0.1%)
All	All	0.28	0/5854	0.70	4/7867 (0.1%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	432	PRO	N-CA-C	6.28	118.36	110.70
1	B	432	PRO	N-CA-C	5.80	117.78	110.70
1	B	432	PRO	CA-C-N	-5.22	114.40	120.04
1	B	432	PRO	C-N-CA	-5.22	114.40	120.04

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	2834	0	2758	18	0
1	B	2879	0	2807	27	0
2	A	15	0	0	0	0
2	B	15	0	0	5	0
3	A	28	0	0	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	34	0	0	1	0
All	All	5805	0	5565	45	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.

The worst 5 of 45 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:266:ARG:HE	1:B:267:TYR:H	1.43	0.66
2:B:602:SO4:O1	3:B:2034:HOH:O	2.14	0.64
1:B:211:MET:HB3	1:B:359:LEU:HD23	1.81	0.63
1:B:528:GLN:HA	1:B:531:MET:HE2	1.84	0.60
3:A:2010:HOH:O	1:B:405:LYS:NZ	2.34	0.57

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	339/386 (88%)	335 (99%)	3 (1%)	1 (0%)	36	52
1	B	343/386 (89%)	336 (98%)	7 (2%)	0	100	100
All	All	682/772 (88%)	671 (98%)	10 (2%)	1 (0%)	48	66

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	303	ILE

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	305/346 (88%)	300 (98%)	5 (2%)	55	74
1	B	311/346 (90%)	305 (98%)	6 (2%)	50	71
All	All	616/692 (89%)	605 (98%)	11 (2%)	51	72

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	276	ARG
1	B	307	LYS
1	B	569	ARG
1	B	412	LYS
1	A	567	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	368	HIS
1	B	422	GLN
1	B	528	GLN
1	B	500	GLN
1	B	510	GLN

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](#)

6 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	SO4	A	603	-	4,4,4	0.24	0	6,6,6	0.07	0
2	SO4	A	601	-	4,4,4	0.24	0	6,6,6	0.09	0
2	SO4	B	602	-	4,4,4	0.19	0	6,6,6	0.35	0
2	SO4	B	601	-	4,4,4	0.24	0	6,6,6	0.08	0
2	SO4	A	602	-	4,4,4	0.23	0	6,6,6	0.08	0
2	SO4	B	603	-	4,4,4	0.24	0	6,6,6	0.07	0

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.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	602	SO4	3	0
2	B	603	SO4	2	0

5.7 Other polymers [i](#)

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 [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	345/386 (89%)	-0.03	7 (2%) 65 59	52, 78, 120, 153	0
1	B	349/386 (90%)	-0.01	10 (2%) 53 46	49, 71, 125, 151	0
All	All	694/772 (89%)	-0.02	17 (2%) 59 53	49, 75, 121, 153	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	207	TYR	4.6
1	B	248	SER	4.4
1	A	573	ASN	4.4
1	B	573	ASN	3.8
1	B	249	TYR	3.5

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 [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. 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	B-factors($\text{\AA}^2$)	Q<0.9
2	SO4	A	603	5/5	0.55	0.11	164,168,168,172	0
2	SO4	A	602	5/5	0.69	0.09	144,146,148,150	0
2	SO4	A	601	5/5	0.79	0.09	108,113,119,122	0
2	SO4	B	603	5/5	0.82	0.09	148,150,152,152	0
2	SO4	B	602	5/5	0.88	0.10	75,76,80,107	0
2	SO4	B	601	5/5	0.88	0.07	107,111,112,114	0

6.5 Other polymers [i](#)

There are no such residues in this entry.