



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 22, 2022 – 01:45 am BST

PDB ID : 6R99  
Title : Crystal Structure of Human CLN protein 5 (Ceroid Lipofuscinosis Neuronal Protein 5)  
Authors : Luebben, A.V.; Sheldrick, G.M.  
Deposited on : 2019-04-03  
Resolution : 2.70 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](mailto: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 : 4.02b-467  
Mogul : **FAILED**  
Xtriage (Phenix) : 1.13  
EDS : 2.30  
buster-report : **FAILED**  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0267  
CCP4 : 7.1.010 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.30

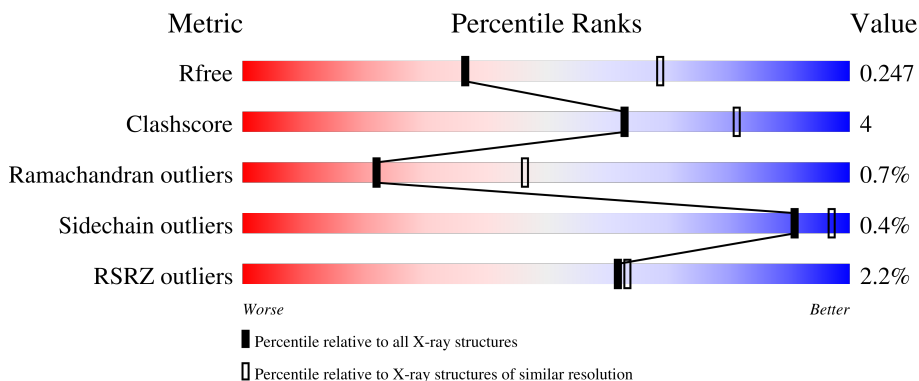
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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  $\geq 3$ , 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	407	 62% 8% 30%
2	B	2	 100%
3	C	3	 100%

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 2441 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 Ceroid-lipofuscinosis neuronal protein 5.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	285	2311	1522	363	413	6	7	0	0	0

- Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	2	28	16	2	10	0	0	0

- Molecule 3 is an oligosaccharide called beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	3	39	22	2	15	0	0	0

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	35	Total	O	0	0
			35	35		



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	58.42Å 58.42Å 179.05Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.05 – 2.70 48.69 – 2.70	Depositor EDS
% Data completeness (in resolution range)	100.0 (44.05-2.70) 100.0 (48.69-2.70)	Depositor EDS
$R_{merge}$	0.23	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.76 (at 2.69Å)	Xtrriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.221 , 0.247 0.221 , 0.247	Depositor DCC
$R_{free}$ test set	533 reflections (5.15%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	63.1	Xtrriage
Anisotropy	0.072	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.033 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	2441	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	57.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.13% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: 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	0.28	0/2383	0.46	0/3231

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	2311	0	2162	20	0
2	B	28	0	25	0	0
3	C	39	0	34	0	0
4	A	28	0	26	0	0
5	A	35	0	0	2	0
All	All	2441	0	2247	20	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.

All (20) 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:173:SER:OG	1:A:176:THR:HB	1.91	0.71
1:A:198:LEU:H	1:A:345:LYS:HZ1	1.41	0.67
1:A:253:GLU:OE1	5:A:601:HOH:O	2.15	0.63
1:A:176:THR:HG22	1:A:178:LYS:H	1.65	0.61
1:A:286:ARG:NH2	5:A:603:HOH:O	2.32	0.60
1:A:145:ARG:NH2	1:A:304:ASN:O	2.36	0.57
1:A:370:PHE:HB3	1:A:381:LEU:HB2	1.89	0.54
1:A:196:PRO:HB3	1:A:207:TRP:CE2	2.44	0.53
1:A:373:PHE:CZ	1:A:376:PHE:HA	2.47	0.49
1:A:263:VAL:HG12	1:A:274:TRP:HD1	1.78	0.49
1:A:169:ILE:HG13	1:A:284:VAL:HG22	1.95	0.47
1:A:192:ASN:OD1	1:A:197:HIS:NE2	2.38	0.47
1:A:108:ARG:NH1	1:A:254:THR:OG1	2.46	0.47
1:A:384:LYS:HG3	1:A:385:PHE:CB	2.45	0.47
1:A:198:LEU:H	1:A:345:LYS:NZ	2.14	0.46
1:A:198:LEU:N	1:A:345:LYS:HZ1	2.11	0.46
1:A:205:PRO:HG2	1:A:378:TYR:OH	2.16	0.45
1:A:102:TRP:CB	1:A:103:PRO:HD3	2.47	0.45
1:A:317:TYR:CZ	1:A:319:GLY:HA2	2.54	0.42
1:A:251:ASP:OD2	1:A:286:ARG:NH1	2.53	0.41

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	279/407 (69%)	264 (95%)	13 (5%)	2 (1%)	22 46

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	385	PHE
1	A	387	PHE



### 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	241/331 (73%)	240 (100%)	1 (0%)	<a href="#">91</a> <a href="#">97</a>

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	353	PHE

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

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

Mogul failed to run properly - this section is therefore empty.

### 5.5 Carbohydrates [i](#)

Mogul failed to run properly - this section is therefore empty.

### 5.6 Ligand geometry [i](#)

Mogul failed to run properly - this section is therefore empty.

### 5.7 Other polymers [i](#)

Mogul failed to run properly - this section is therefore empty.

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	278/407 (68%)	-0.00	6 (2%) 62 63	37, 55, 83, 112	0

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	161	HIS	4.3
1	A	399	ILE	3.8
1	A	200	PRO	3.7
1	A	298	PHE	2.4
1	A	297	GLU	2.3
1	A	206	PHE	2.1

### 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](#)

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, 95<sup>th</sup> 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(Å <sup>2</sup> )	Q<0.9
3	BMA	C	3	11/12	0.85	0.13	72,80,91,97	0
3	NAG	C	2	14/15	0.91	0.13	50,59,76,78	0
2	NAG	B	2	14/15	0.94	0.13	59,67,79,87	0
2	NAG	B	1	14/15	0.96	0.13	37,53,58,63	0
3	NAG	C	1	14/15	0.96	0.14	31,41,55,56	0

## 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, 95<sup>th</sup> 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(Å <sup>2</sup> )	Q<0.9
4	NAG	A	501	14/15	0.94	0.17	55,66,79,80	0
4	NAG	A	507	14/15	0.96	0.14	41,53,62,69	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.