



# Full wwPDB X-ray Structure Validation Report ⓘ

Jul 28, 2016 – 05:39 PM EDT

PDB ID : 5KC6  
Title : Crystal structure of Cbln1 (Val55-Gly58 deletion mutant)  
Authors : Elegheert, J.; Clay, J.E.; Aricescu, A.R.  
Deposited on : 2016-06-05  
Resolution : 2.80 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

<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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027939  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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-20027939

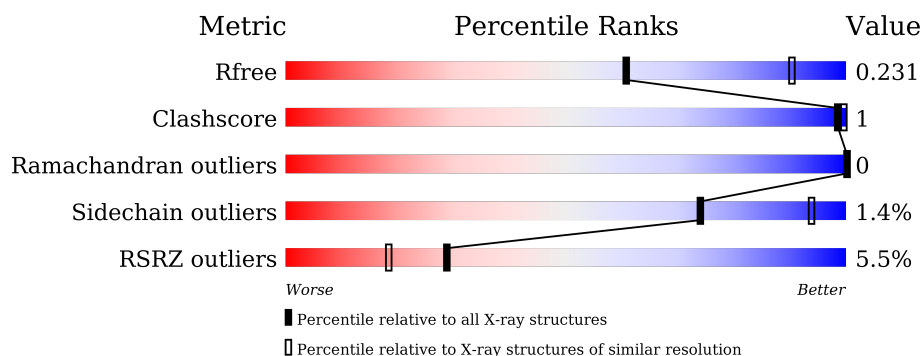
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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))
$R_{free}$	91344	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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  $\geq 3$ , 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	178	<div> <div>4%</div> <div> <div></div> <div>75%</div> <div>..</div> <div>23%</div> </div> </div>
1	B	178	<div> <div>4%</div> <div> <div></div> <div>78%</div> <div>..</div> <div>19%</div> </div> </div>
1	C	178	<div> <div>4%</div> <div> <div></div> <div>74%</div> <div>..</div> <div>24%</div> </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	C	301	-	-	-	X

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3368 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 Cerebellin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	137	Total	C	N	O	S	0	0	0
			1084	692	188	199	5			
1	B	144	Total	C	N	O	S	0	1	0
			1165	743	209	208	5			
1	C	136	Total	C	N	O	S	0	0	0
			1077	688	187	197	5			

There are 48 discrepancies between the modelled and reference sequences:

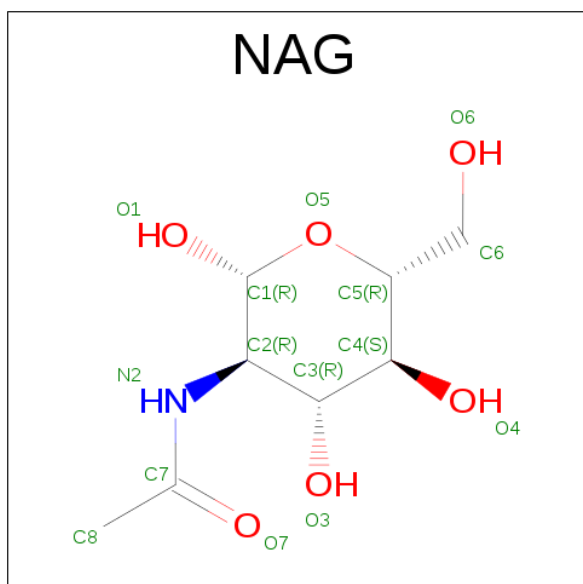
Chain	Residue	Modelled	Actual	Comment	Reference
A	25	GLU	-	expression tag	UNP P23435
A	26	THR	-	expression tag	UNP P23435
A	27	GLY	-	expression tag	UNP P23435
A	?	-	VAL	deletion	UNP P23435
A	?	-	ARG	deletion	UNP P23435
A	?	-	SER	deletion	UNP P23435
A	?	-	GLY	deletion	UNP P23435
A	194	GLY	-	expression tag	UNP P23435
A	195	THR	-	expression tag	UNP P23435
A	196	LYS	-	expression tag	UNP P23435
A	197	HIS	-	expression tag	UNP P23435
A	198	HIS	-	expression tag	UNP P23435
A	199	HIS	-	expression tag	UNP P23435
A	200	HIS	-	expression tag	UNP P23435
A	201	HIS	-	expression tag	UNP P23435
A	202	HIS	-	expression tag	UNP P23435
B	25	GLU	-	expression tag	UNP P23435
B	26	THR	-	expression tag	UNP P23435
B	27	GLY	-	expression tag	UNP P23435
B	?	-	VAL	deletion	UNP P23435
B	?	-	ARG	deletion	UNP P23435
B	?	-	SER	deletion	UNP P23435
B	?	-	GLY	deletion	UNP P23435

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	194	GLY	-	expression tag	UNP P23435
B	195	THR	-	expression tag	UNP P23435
B	196	LYS	-	expression tag	UNP P23435
B	197	HIS	-	expression tag	UNP P23435
B	198	HIS	-	expression tag	UNP P23435
B	199	HIS	-	expression tag	UNP P23435
B	200	HIS	-	expression tag	UNP P23435
B	201	HIS	-	expression tag	UNP P23435
B	202	HIS	-	expression tag	UNP P23435
C	25	GLU	-	expression tag	UNP P23435
C	26	THR	-	expression tag	UNP P23435
C	27	GLY	-	expression tag	UNP P23435
C	?	-	VAL	deletion	UNP P23435
C	?	-	ARG	deletion	UNP P23435
C	?	-	SER	deletion	UNP P23435
C	?	-	GLY	deletion	UNP P23435
C	194	GLY	-	expression tag	UNP P23435
C	195	THR	-	expression tag	UNP P23435
C	196	LYS	-	expression tag	UNP P23435
C	197	HIS	-	expression tag	UNP P23435
C	198	HIS	-	expression tag	UNP P23435
C	199	HIS	-	expression tag	UNP P23435
C	200	HIS	-	expression tag	UNP P23435
C	201	HIS	-	expression tag	UNP P23435
C	202	HIS	-	expression tag	UNP P23435

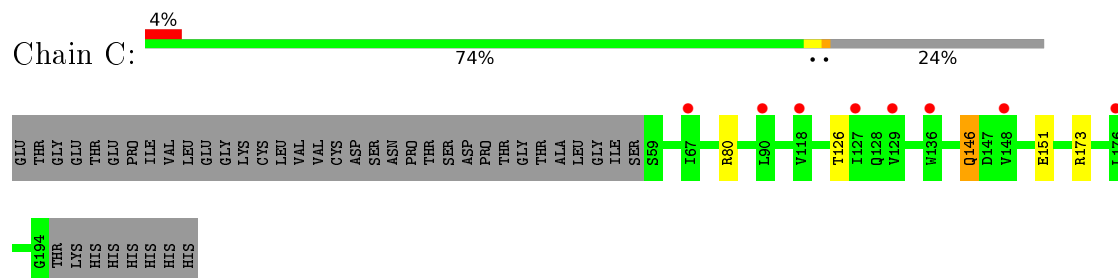
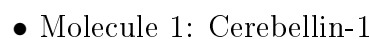
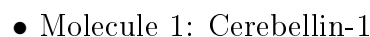
- Molecule 2 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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



- Molecule 1: Cerebellin-1



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	79.58Å 170.45Å 116.43Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	68.77 – 2.80 85.22 – 2.80	Depositor EDS
% Data completeness (in resolution range)	73.2 (68.77-2.80) 94.6 (85.22-2.80)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.24 (at 2.82Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1772)	Depositor
R, $R_{free}$	0.188 , 0.227 0.194 , 0.231	Depositor DCC
$R_{free}$ test set	962 reflections (5.12%)	DCC
Wilson B-factor (Å <sup>2</sup> )	74.7	Xtriage
Anisotropy	1.479	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 38.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	3368	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.85% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 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	0/1108	0.42	0/1496
1	B	0.24	0/1196	0.43	0/1615
1	C	0.24	0/1101	0.43	0/1486
All	All	0.24	0/3405	0.43	0/4597

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	1084	0	1067	2	0
1	B	1165	0	1130	2	1
1	C	1077	0	1060	2	1
2	A	14	0	13	0	0
2	B	14	0	13	0	0
2	C	14	0	13	0	0
All	All	3368	0	3296	6	1

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 1.

All (6) 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:120:LYS:NZ	1:A:125:GLN:O	2.11	0.84
1:B:120:LYS:NZ	1:B:125:GLN:O	2.33	0.61
1:C:146:GLN:NE2	1:C:151:GLU:OE2	2.42	0.52
1:B:109:LYS:NZ	1:B:197:HIS:O	2.45	0.50
1:A:146:GLN:NE2	1:A:151:GLU:OE2	2.43	0.43
1:C:126:THR:O	1:C:173:ARG:NH1	2.53	0.41

All (1) 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)
1:B:163:LYS:O	1:C:173:ARG:NH2[8_555]	2.18	0.02

## 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	135/178 (76%)	132 (98%)	3 (2%)	0	100	100
1	B	143/178 (80%)	141 (99%)	2 (1%)	0	100	100
1	C	134/178 (75%)	131 (98%)	3 (2%)	0	100	100
All	All	412/534 (77%)	404 (98%)	8 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	117/153 (76%)	116 (99%)	1 (1%)	84	96
1	B	125/153 (82%)	123 (98%)	2 (2%)	70	93
1	C	116/153 (76%)	114 (98%)	2 (2%)	68	92
All	All	358/459 (78%)	353 (99%)	5 (1%)	74	94

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	146	GLN
1	B	65	SER
1	B	125	GLN
1	C	80	ARG
1	C	146	GLN

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 ⓘ

3 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	301	1	14,14,15	0.26	0	15,19,21	0.49	0
2	NAG	B	301	1	14,14,15	0.42	0	15,19,21	0.42	0
2	NAG	C	301	1	14,14,15	0.46	0	15,19,21	0.48	0

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	301	1	-	0/6/23/26	0/1/1/1
2	NAG	B	301	1	-	0/6/23/26	0/1/1/1
2	NAG	C	301	1	-	0/6/23/26	0/1/1/1

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.

No monomer is involved in short contacts.

## 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

### 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, 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	137/178 (76%)	0.53	7 (5%)	32 21	27, 54, 100, 133	0
1	B	144/178 (80%)	0.85	8 (5%)	28 18	27, 47, 119, 166	0
1	C	136/178 (76%)	0.67	8 (5%)	26 16	24, 49, 96, 134	0
All	All	417/534 (78%)	0.69	23 (5%)	29 18	24, 51, 115, 166	0

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	202	HIS	3.9
1	B	124	ARG	3.8
1	A	148	VAL	3.7
1	A	147	ASP	3.6
1	C	176	LEU	3.1
1	B	131	LEU	3.1
1	A	149	THR	2.9
1	A	118	VAL	2.7
1	B	199	HIS	2.7
1	B	147	ASP	2.6
1	C	148	VAL	2.6
1	B	138	VAL	2.4
1	C	136	TRP	2.4
1	C	90	LEU	2.3
1	A	84	ILE	2.2
1	A	138	VAL	2.2
1	C	67	ILE	2.1
1	C	129	VAL	2.1
1	C	118	VAL	2.1
1	A	131	LEU	2.1
1	B	159	ILE	2.1
1	B	123	ASN	2.0
1	C	127	ILE	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](#)

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, 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	NAG	C	301	14/15	0.78	0.31	2.54	78,90,108,112	0
2	NAG	A	301	14/15	0.88	0.16	-0.51	60,76,81,81	0
2	NAG	B	301	14/15	0.82	0.14	-1.35	75,89,100,102	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.