



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 09:04 am BST

PDB ID : 4MFV  
Title : Crystal structure of human CTNNBL1(residues 33 563)  
Authors : Ahn, J.W.; Kim, S.; Kim, K.J.  
Deposited on : 2013-08-28  
Resolution : 2.92 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](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 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

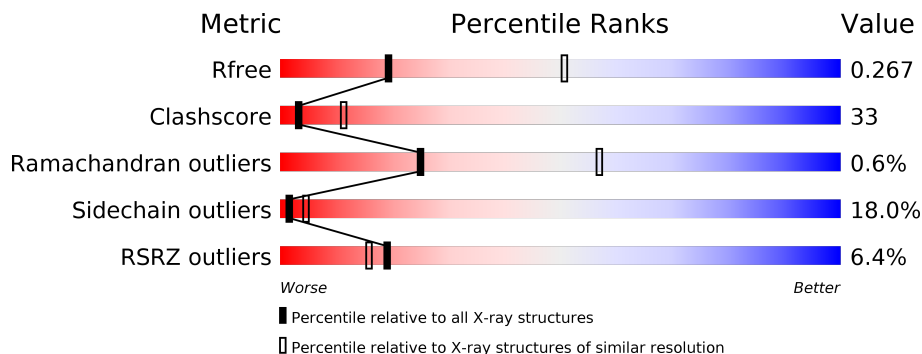
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.92 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	2307 (2.94-2.90)
Clashscore	141614	2531 (2.94-2.90)
Ramachandran outliers	138981	2462 (2.94-2.90)
Sidechain outliers	138945	2464 (2.94-2.90)
RSRZ outliers	127900	2248 (2.94-2.90)

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 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	534	
1	B	534	

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 7866 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 Beta-catenin-like protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	489	3933	2466	686	754	27	0	0	0
1	B	489	3933	2466	686	754	27	0	0	0

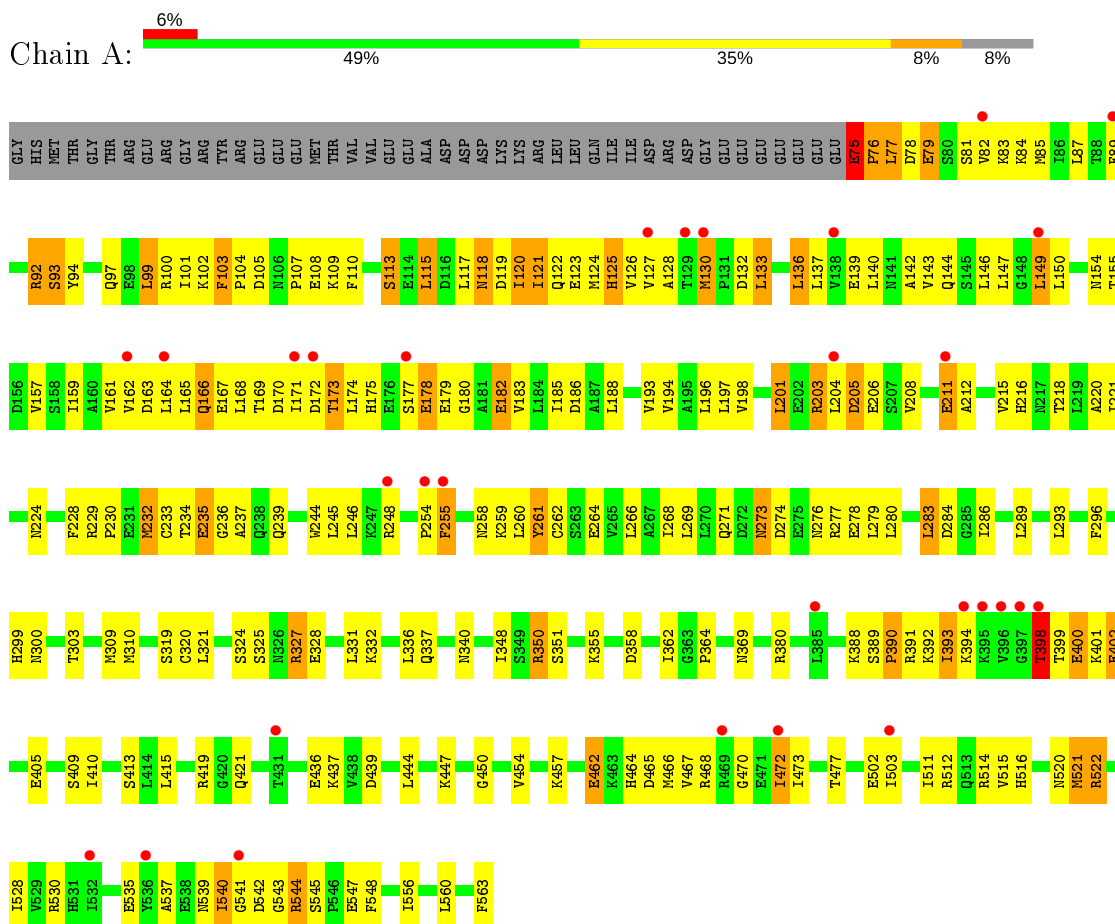
There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	30	GLY	-	EXPRESSION TAG	UNP Q8WYA6
A	31	HIS	-	EXPRESSION TAG	UNP Q8WYA6
A	32	MET	-	EXPRESSION TAG	UNP Q8WYA6
B	30	GLY	-	EXPRESSION TAG	UNP Q8WYA6
B	31	HIS	-	EXPRESSION TAG	UNP Q8WYA6
B	32	MET	-	EXPRESSION TAG	UNP Q8WYA6

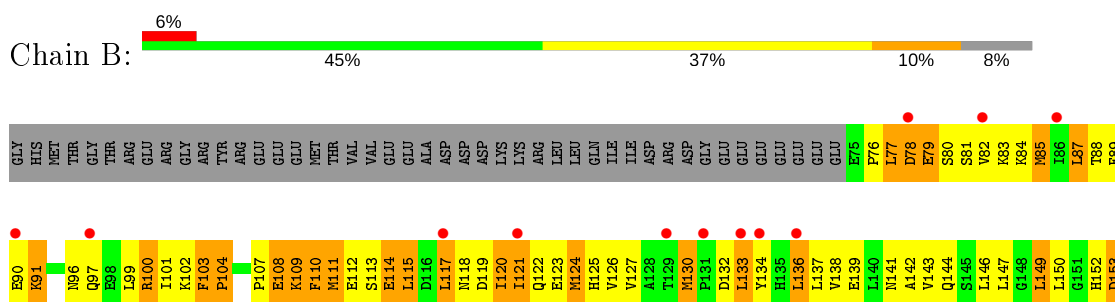
### 3 Residue-property plots

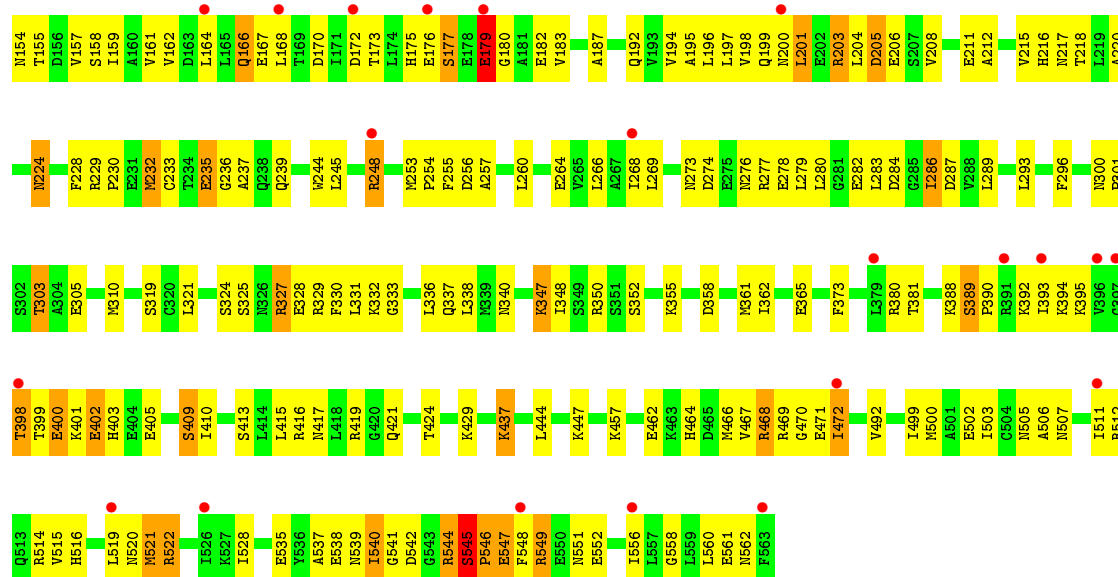
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Beta-catenin-like protein 1



- Molecule 1: Beta-catenin-like protein 1





## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.04Å 90.04Å 175.75Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.84 – 2.92 46.84 – 2.92	Depositor EDS
% Data completeness (in resolution range)	94.1 (46.84-2.92) 94.1 (46.84-2.92)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.33 (at 2.91Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.233 , 0.305 0.249 , 0.267	Depositor DCC
$R_{free}$ test set	1635 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	68.1	Xtriage
Anisotropy	0.174	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 39.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.048 for -h,-k,l 0.480 for h,-h-k,-l 0.054 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	7866	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	88.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.50% 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](#)

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	0.54	1/3985 (0.0%)	0.75	4/5358 (0.1%)
1	B	0.54	2/3985 (0.1%)	0.76	5/5358 (0.1%)
All	All	0.54	3/7970 (0.0%)	0.76	9/10716 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	3
1	B	0	4
All	All	0	7

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	104	PRO	N-CD	5.26	1.55	1.47
1	B	546	PRO	N-CD	5.16	1.55	1.47
1	A	76	PRO	N-CD	5.09	1.54	1.47

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	300	ASN	C-N-CD	5.99	140.98	128.40
1	B	327	ARG	NE-CZ-NH1	5.78	123.19	120.30
1	A	75	GLU	C-N-CD	5.75	140.47	128.40
1	B	103	PHE	C-N-CD	5.74	140.45	128.40
1	B	545	SER	C-N-CD	5.70	140.37	128.40

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	175	HIS	Peptide
1	A	254	PRO	Peptide
1	A	398	THR	Peptide
1	B	254	PRO	Peptide
1	B	256	ASP	Peptide

## 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	3933	0	3972	258	0
1	B	3933	0	3972	266	0
All	All	7866	0	7944	520	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 33.

The worst 5 of 520 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:393:ILE:CG2	1:A:399:THR:HG21	1.35	1.53
1:B:201:LEU:HD23	1:B:244:TRP:CE2	1.48	1.48
1:B:201:LEU:HD21	1:B:244:TRP:CH2	1.64	1.31
1:B:124:MET:O	1:B:127:VAL:HG23	1.34	1.27
1:B:201:LEU:HD23	1:B:244:TRP:CZ2	1.72	1.25

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	487/534 (91%)	452 (93%)	32 (7%)	3 (1%)	25	57
1	B	487/534 (91%)	444 (91%)	40 (8%)	3 (1%)	25	57
All	All	974/1068 (91%)	896 (92%)	72 (7%)	6 (1%)	25	57

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	540	ILE
1	B	540	ILE
1	B	76	PRO
1	B	390	PRO
1	A	103	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	442/482 (92%)	368 (83%)	74 (17%)	2	6
1	B	442/482 (92%)	357 (81%)	85 (19%)	1	4
All	All	884/964 (92%)	725 (82%)	159 (18%)	1	4

5 of 159 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	542	ASP
1	B	112	GLU
1	B	457	LYS
1	B	77	LEU
1	B	90	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	166	GLN
1	B	175	HIS
1	B	513	GLN
1	A	562	ASN
1	B	122	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 carbohydrates in this entry.

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

There are no ligands in this entry.

### 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, 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	489/534 (91%)	0.54	30 (6%) 21 18	24, 85, 128, 157	0
1	B	489/534 (91%)	0.53	33 (6%) 17 14	24, 87, 130, 158	0
All	All	978/1068 (91%)	0.53	63 (6%) 19 16	24, 86, 130, 158	0

The worst 5 of 63 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	397	GLY	11.0
1	B	398	THR	7.0
1	A	395	LYS	5.8
1	B	134	TYR	5.6
1	A	82	VAL	5.0

### 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.