



# Full wwPDB X-ray Structure Validation Report ⓘ

May 29, 2020 – 09:34 am BST

PDB ID : 6BXZ  
Title : Crystal Structure of Pig Protocadherin-15 EC10-MAD12  
Authors : De-la-Torre, P.; Araya-Secchi, R.; Choudhary, D.; Sotomayor, M.  
Deposited on : 2017-12-19  
Resolution : 2.09 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 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

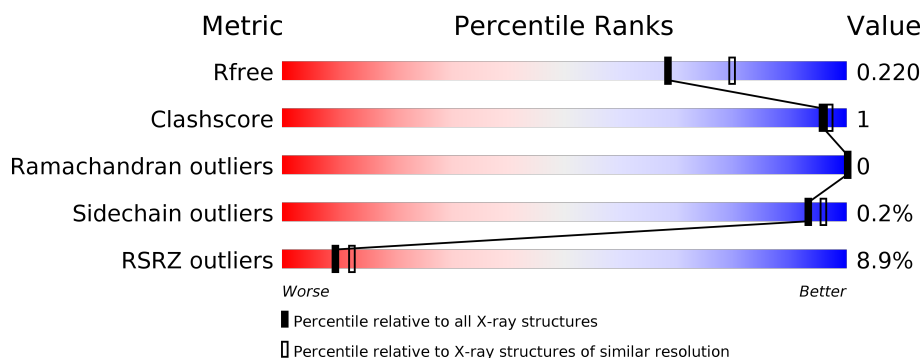
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.09 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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	356	<div> <div>11%</div> <div> <div></div> <div>87%</div> <div>•</div> <div>11%</div> </div> </div>
1	C	356	<div> <div>6%</div> <div> <div></div> <div>91%</div> <div>•</div> <div>6%</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5441 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 Protocadherin related 15.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	334	Total	C	N	O	S	0	3	0
			2641	1684	440	510	7			
1	A	317	Total	C	N	O	S	0	5	0
			2534	1616	426	485	7			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	1006	MET	-	initiating methionine	UNP F1SD06
C	1354	LEU	-	expression tag	UNP F1SD06
C	1355	GLU	-	expression tag	UNP F1SD06
C	1356	HIS	-	expression tag	UNP F1SD06
C	1357	HIS	-	expression tag	UNP F1SD06
C	1358	HIS	-	expression tag	UNP F1SD06
C	1359	HIS	-	expression tag	UNP F1SD06
C	1360	HIS	-	expression tag	UNP F1SD06
C	1361	HIS	-	expression tag	UNP F1SD06
A	1006	MET	-	initiating methionine	UNP F1SD06
A	1354	LEU	-	expression tag	UNP F1SD06
A	1355	GLU	-	expression tag	UNP F1SD06
A	1356	HIS	-	expression tag	UNP F1SD06
A	1357	HIS	-	expression tag	UNP F1SD06
A	1358	HIS	-	expression tag	UNP F1SD06
A	1359	HIS	-	expression tag	UNP F1SD06
A	1360	HIS	-	expression tag	UNP F1SD06
A	1361	HIS	-	expression tag	UNP F1SD06

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	3	Total	Ca	0	0
			3	3		

*Continued on next page...*

*Continued from previous page...*

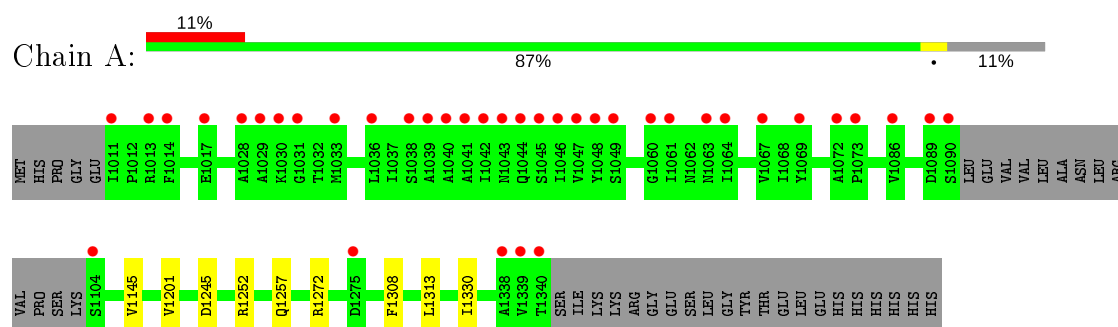
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	C	3	Total 3	Ca 3	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	106	Total 106	O 106	0	0
3	A	154	Total 154	O 154	0	0



- Molecule 1: Protocadherin related 15



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	92.28Å 49.25Å 108.25Å 90.00° 92.15° 90.00°	Depositor
Resolution (Å)	108.18 – 2.09 40.56 – 2.09	Depositor EDS
% Data completeness (in resolution range)	96.4 (108.18-2.09) 96.4 (40.56-2.09)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.35 (at 2.10Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.187 , 0.217 0.195 , 0.220	Depositor DCC
$R_{free}$ test set	2786 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.7	Xtriage
Anisotropy	0.171	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 44.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.023 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	5441	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.62% 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: CA

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	0.61	0/2595	0.79	3/3512 (0.1%)
1	C	0.56	0/2698	0.81	1/3659 (0.0%)
All	All	0.58	0/5293	0.80	4/7171 (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	1245	ASP	CB-CG-OD1	6.24	123.92	118.30
1	A	1252	ARG	NE-CZ-NH1	5.32	122.96	120.30
1	A	1272	ARG	NE-CZ-NH2	-5.11	117.74	120.30
1	C	1298	ARG	NE-CZ-NH2	-5.10	117.75	120.30

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	2534	0	2557	3	0
1	C	2641	0	2662	8	0
2	A	3	0	0	0	0
2	C	3	0	0	0	0
3	A	154	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	106	0	0	0	0
All	All	5441	0	5219	11	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 1.

All (11) 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:C:1321[A]:GLN:NE2	1:C:1321[A]:GLN:HA	2.06	0.70
1:A:1313:LEU:HD11	1:A:1330:ILE:HG21	1.86	0.57
1:C:1321[B]:GLN:HE21	1:C:1326:GLU:HG2	1.74	0.52
1:C:1321[A]:GLN:HE21	1:C:1321[A]:GLN:HA	1.75	0.50
1:C:1295:GLN:HA	1:C:1295:GLN:HE21	1.82	0.45
1:C:1145:VAL:HG21	1:C:1201:VAL:HG21	2.00	0.44
1:C:1257:GLN:HG3	1:C:1308:PHE:CD1	2.53	0.44
1:A:1145:VAL:HG21	1:A:1201:VAL:HG21	1.99	0.43
1:C:1324:TYR:HB3	1:C:1328:GLY:HA3	2.01	0.43
1:C:1321[A]:GLN:CA	1:C:1321[A]:GLN:NE2	2.78	0.43
1:A:1257:GLN:HG3	1:A:1308:PHE:CD1	2.56	0.41

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	318/356 (89%)	314 (99%)	4 (1%)	0	100	100
1	C	335/356 (94%)	333 (99%)	2 (1%)	0	100	100
All	All	653/712 (92%)	647 (99%)	6 (1%)	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	279/309 (90%)	279 (100%)	0	100	100
1	C	292/309 (94%)	290 (99%)	2 (1%)	84	88
All	All	571/618 (92%)	569 (100%)	2 (0%)	93	94

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	1321[A]	GLN
1	C	1321[B]	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	1257	GLN
1	C	1273	HIS
1	C	1295	GLN
1	A	1124	GLN
1	A	1257	GLN
1	A	1273	HIS
1	A	1321	GLN

### 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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 6 are monoatomic - leaving 0 for Mogul analysis.

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	317/356 (89%)	0.79	38 (11%) 4 5	21, 35, 92, 120	0
1	C	334/356 (93%)	0.46	20 (5%) 21 27	23, 43, 70, 79	0
All	All	651/712 (91%)	0.62	58 (8%) 9 12	21, 40, 81, 120	0

All (58) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1042	ILE	10.0
1	A	1041	ALA	7.8
1	A	1043	ASN	6.7
1	A	1013	ARG	5.9
1	A	1045	SER	5.8
1	A	1046	ILE	5.4
1	A	1104	SER	5.2
1	A	1011	ILE	4.8
1	A	1040	ALA	4.5
1	A	1340	THR	4.2
1	C	1342	ILE	3.7
1	A	1069	TYR	3.6
1	A	1090	SER	3.6
1	A	1089	ASP	3.5
1	A	1044	GLN	3.5
1	A	1039	ALA	3.4
1	C	1072	ALA	3.4
1	A	1061	ILE	3.4
1	A	1064	ILE	3.3
1	C	1073	PRO	3.2
1	A	1033	MET	3.0
1	C	1045	SER	3.0
1	A	1038	SER	2.9
1	A	1036	LEU	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	1030	LYS	2.8
1	C	1098	LEU	2.8
1	A	1031	GLY	2.8
1	C	1056	GLU	2.7
1	A	1063	ASN	2.7
1	A	1030	LYS	2.7
1	A	1275	ASP	2.6
1	C	1097	ASN	2.6
1	A	1086	VAL	2.6
1	A	1049	SER	2.5
1	A	1338	ALA	2.5
1	A	1029	ALA	2.4
1	A	1017	GLU	2.4
1	C	1042	ILE	2.4
1	C	1028	ALA	2.4
1	C	1078	THR	2.4
1	A	1060	GLY	2.4
1	C	1341	SER	2.3
1	C	1275	ASP	2.2
1	A	1073	PRO	2.2
1	C	1065	THR	2.2
1	C	1074	LEU	2.2
1	A	1014	PHE	2.2
1	C	1095	LEU	2.2
1	A	1072	ALA	2.2
1	C	1027	LEU	2.1
1	A	1028	ALA	2.1
1	A	1047	VAL	2.1
1	C	1041	ALA	2.1
1	A	1339	VAL	2.1
1	A	1048	TYR	2.1
1	C	1166	PRO	2.0
1	C	1010	GLU	2.0
1	A	1067	VAL	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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. 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( $\text{\AA}^2$ )	Q<0.9
2	CA	C	1401	1/1	0.94	0.04	61,61,61,61	0
2	CA	A	1402	1/1	0.97	0.06	43,43,43,43	0
2	CA	C	1402	1/1	0.97	0.04	56,56,56,56	0
2	CA	A	1401	1/1	0.97	0.03	51,51,51,51	0
2	CA	A	1403	1/1	0.98	0.06	36,36,36,36	0
2	CA	C	1403	1/1	0.99	0.06	49,49,49,49	0

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

There are no such residues in this entry.