



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 03:58 PM UTC

PDB ID : 4K1N / pdb_00004k1n
Title : Crystal structure of full-length mouse alphaE-catenin
Authors : Ishiyama, N.; Ikura, M.
Deposited on : 2013-04-05
Resolution : 6.50 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
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

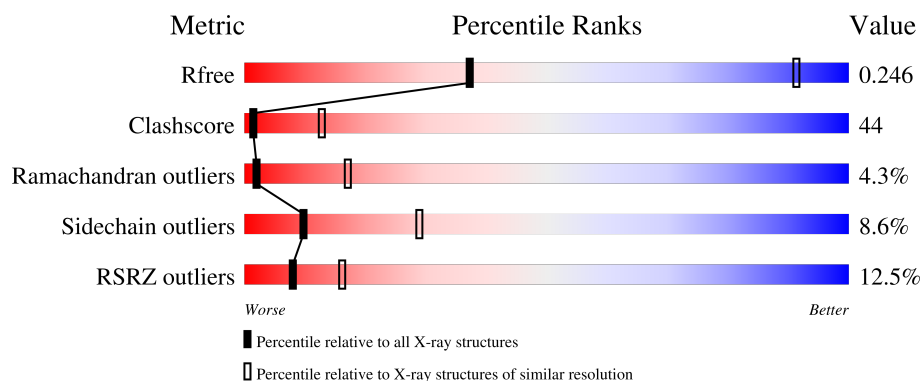
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 6.50 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	1162 (9.00-4.00)
Clashscore	190562	1020 (8.98-4.02)
Ramachandran outliers	187476	1050 (9.00-4.00)
Sidechain outliers	187428	1014 (9.00-4.00)
RSRZ outliers	180081	1155 (9.00-4.00)

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	912	6% 24% 28% • 43%
1	B	912	8% 24% 29% • 43%

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 8020 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 Catenin alpha-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	520	Total	C	N	O	S	0	0	0
			4010	2486	713	789	22			
1	B	520	Total	C	N	O	S	0	0	0
			4010	2486	713	789	22			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	GLY	-	expression tag	UNP P26231
A	-4	SER	-	expression tag	UNP P26231
A	-3	HIS	-	expression tag	UNP P26231
A	-2	MET	-	expression tag	UNP P26231
A	-1	ALA	-	expression tag	UNP P26231
A	0	SER	-	expression tag	UNP P26231
B	-5	GLY	-	expression tag	UNP P26231
B	-4	SER	-	expression tag	UNP P26231
B	-3	HIS	-	expression tag	UNP P26231
B	-2	MET	-	expression tag	UNP P26231
B	-1	ALA	-	expression tag	UNP P26231
B	0	SER	-	expression tag	UNP P26231

LEU ILE ALA ALA GLY GLN SER ALA ARG ALA ILE ILE MET ALA ALA GLN LEU LEU PRO GLN GLN GLN GLN LYS LYS LYS ILE ILE ILE LYS LYS LYS LEU ASP ALA GLU ALA VAL VAL LYS LYS TRP ASP ASP SER SER SER GLY ASN ASP ILE ILE ILE VAL VAL ALA LYS LYS MET MET MET ILE

MET	GLN	TYR	ILE	MET
ASP	LYS	SER	LEU	GLU
SER	GLN	GLY	CYS	THR
ILE	MET	ALA	LEU	THR
	ALA	SER	ASN	ARG
	LEU	LEU	ILE	GLY
	ASN	CYS	CYS	LYS
	LEU	SER	SER	GLY
	PRO	LYS	VAL	PRO
	ALA	VAL	LYS	LEU
	SER	VAL	ALA	LYS
	TRP	GLU	ASN	THR
	LYS	VAL	GLN	SER
	MET	ASN	GLN	ASP
	LYS	GLY	ILE	VAL
	ALA	PRO	GLY	SER
	GLU	LYS	GLY	ILE
	LYS	VAL	LEU	ALA
	PRO	VAL	VAL	LYS
	VAL	SER	VAL	ILE
	LYS	SER	GLY	ALA
	ARG	ASP	VAL	ALA
	GLU	GLU	ASP	GLY
	THR	SER	SER	ALA
	GLN	LEU	LEU	LYS
	LYS	ILE	LEU	GLY
	THR	GLN	ALA	ARG
	THR	ILE	ALA	THR
	LYS	LYS	LYS	ILE
	ARG	ARG	ASN	ALA
	ALA	ALA	LEU	ASP
	SER	SER	MET	HIS
	GLN	CYS	ASN	CYS
	LYS	LYS	ALA	PRO
	LYS	HIS	VAL	ASP
	VAL	VAL	VAL	SER
	ASN	THR	THR	ALA
	PRO	VAL	LYS	CYS
	VAL	VAL	GLN	LYS
	GLN	ALA	ALA	GLN
	ALA	SER	SER	ASP
	LEU	TYR	VAL	LEU
	SER	VAL	VAL	ALA
	THR	ALA	ALA	TYR
	GLU	SER	SER	LEU
	PHE	LYS	THR	GLN
	ALA	LYS	LYS	ARG

4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α , β , γ	144.68Å 144.68Å 136.76Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 6.50 50.00 – 6.50	Depositor EDS
% Data completeness (in resolution range)	94.6 (50.00-6.50) 94.4 (50.00-6.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.13	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.84 (at 6.67Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.224 , 0.247 0.224 , 0.246	Depositor DCC
R_{free} test set	716 reflections (10.90%)	wwPDB-VP
Wilson B-factor (Å ²)	439.2	Xtriage
Anisotropy	0.055	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 459.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.032 for -h,-k,l 0.051 for h,-h-k,-l 0.023 for -k,-h,-l	Xtriage
F_o, F_c correlation	0.86	EDS
Total number of atoms	8020	wwPDB-VP
Average B, all atoms (Å ²)	354.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.57% 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](#)

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	0.29	0/4048	0.80	5/5464 (0.1%)
1	B	0.29	0/4048	0.80	4/5464 (0.1%)
All	All	0.29	0/8096	0.80	9/10928 (0.1%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	368	ILE	N-CA-C	-6.01	106.64	111.81
1	B	368	ILE	N-CA-C	-5.95	106.69	111.81
1	B	302	ARG	CA-C-N	5.42	126.62	119.84
1	B	302	ARG	C-N-CA	5.42	126.62	119.84
1	A	302	ARG	CA-C-N	5.38	126.56	119.84

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	4010	0	4051	357	0
1	B	4010	0	4051	366	0
All	All	8020	0	8102	706	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 44.

The worst 5 of 706 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:369:ASP:O	1:A:373:LYS:HB3	1.65	0.97
1:A:104:MET:HG2	1:A:131:LEU:HD22	1.48	0.96
1:A:209:ALA:HB1	1:A:452:MET:HE2	1.44	0.96
1:B:86:LEU:HD22	1:B:87:LYS:H	1.31	0.95
1:B:104:MET:HG2	1:B:131:LEU:HD22	1.48	0.93

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	516/912 (57%)	444 (86%)	49 (10%)	23 (4%)	2	17
1	B	516/912 (57%)	447 (87%)	48 (9%)	21 (4%)	2	18
All	All	1032/1824 (57%)	891 (86%)	97 (9%)	44 (4%)	2	17

5 of 44 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	86	LEU
1	A	291	VAL
1	A	292	ASP
1	A	301	PHE
1	A	393	SER

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	431/766 (56%)	394 (91%)	37 (9%)	10	29
1	B	431/766 (56%)	394 (91%)	37 (9%)	10	29
All	All	862/1532 (56%)	788 (91%)	74 (9%)	10	29

5 of 74 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	404	LEU
1	B	598	LEU
1	B	410	ASN
1	B	508	ILE
1	A	440	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	354	ASN
1	B	426	HIS
1	B	466	ASN
1	A	354	ASN
1	A	342	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 ⓘ

There are no oligosaccharides 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, 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	520/912 (57%)	0.74	57 (10%)	10 18	355, 355, 355, 355	0
1	B	520/912 (57%)	0.84	73 (14%)	6 15	355, 355, 355, 355	0
All	All	1040/1824 (57%)	0.79	130 (12%)	8 16	355, 355, 355, 355	0

The worst 5 of 130 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	504	ASP	5.8
1	A	331	GLU	5.0
1	A	181	LYS	4.7
1	A	121	ARG	4.6
1	A	116	CYS	4.6

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.