



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2026 – 09:52 AM UTC

PDB ID : 8K9C / pdb_00008k9c
Title : Structure of human Caprin-2 HR1 domain
Authors : Song, X.M.
Deposited on : 2023-07-31
Resolution : 2.32 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

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
Mogul	:	2022.3.0, CSD as543be (2022)
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

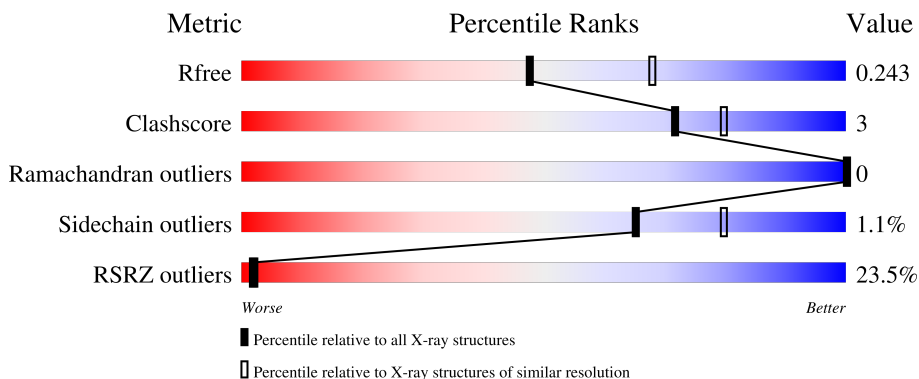
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.32 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	7754 (2.34-2.30)
Clashscore	190562	8383 (2.34-2.30)
Ramachandran outliers	187476	8303 (2.34-2.30)
Sidechain outliers	187428	8303 (2.34-2.30)
RSRZ outliers	180081	7760 (2.34-2.30)

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	254	24% 75% 6% 19%
1	B	254	13% 71% 9% 20%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3400 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 Caprin-2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	207	Total	C	N	O	S	Se	0	1	0
			1667	1061	283	319	2	2			
1	B	202	Total	C	N	O	S	Se	0	0	0
			1650	1057	279	310	2	2			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	98	GLY	-	expression tag	UNP Q6IMN6
A	99	ALA	-	expression tag	UNP Q6IMN6
A	100	MSE	-	expression tag	UNP Q6IMN6
A	101	ASP	-	expression tag	UNP Q6IMN6
B	98	GLY	-	expression tag	UNP Q6IMN6
B	99	ALA	-	expression tag	UNP Q6IMN6
B	100	MSE	-	expression tag	UNP Q6IMN6
B	101	ASP	-	expression tag	UNP Q6IMN6

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	41	Total	O	0	0
			41	41		
2	B	42	Total	O	0	0
			42	42		

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	139.72Å 65.32Å 112.20Å 90.00° 113.75° 90.00°	Depositor
Resolution (Å)	30.21 – 2.32 30.21 – 2.32	Depositor EDS
% Data completeness (in resolution range)	95.1 (30.21-2.32) 89.5 (30.21-2.32)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.02 (at 2.31Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, R_{free}	0.218 , 0.240 0.221 , 0.243	Depositor DCC
R_{free} test set	1969 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	53.3	Xtriage
Anisotropy	0.412	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 47.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	3400	wwPDB-VP
Average B, all atoms (Å ²)	79.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.40% 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

5.1 Standard geometry

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	0.22	0/1691	0.37	0/2267
1	B	0.24	0/1675	0.40	0/2243
All	All	0.23	0/3366	0.39	0/4510

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	1667	0	1675	12	0
1	B	1650	0	1697	12	0
2	A	41	0	0	1	0
2	B	42	0	0	1	0
All	All	3400	0	3372	22	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 3.

All (22) 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:217:ILE:HG13	1:A:286:LEU:HD21	1.47	0.92

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:282:PHE:O	1:A:286:LEU:HD13	1.83	0.77
1:A:213:LYS:HG3	1:A:286:LEU:HD23	1.69	0.74
1:A:235:ASP:OD1	2:A:401:HOH:O	2.07	0.71
1:A:217:ILE:HG13	1:A:286:LEU:CD2	2.27	0.61
1:A:128:CYS:SG	1:B:183:THR:HG21	2.43	0.59
1:A:251:LEU:O	1:A:255:ILE:HG13	2.04	0.57
1:B:235:ASP:HA	1:B:240:LEU:HB2	1.89	0.54
1:A:113:SER:OG	1:A:116:GLN:HG3	2.08	0.54
1:B:229:GLN:O	1:B:233:GLN:HG3	2.09	0.52
1:B:143:LEU:HD23	1:B:167:TYR:HA	1.93	0.50
1:A:149:ARG:O	1:A:154:GLU:HB3	2.13	0.47
1:A:114:PRO:HD3	1:B:198:ALA:HB1	1.97	0.47
1:B:269:LEU:HD22	1:B:273:ASP:HB3	1.96	0.47
1:B:289:SER:OG	2:B:401:HOH:O	2.21	0.47
1:B:182:LYS:HD2	1:B:182:LYS:HA	1.80	0.46
1:B:156:LEU:HB3	1:B:160:GLN:OE1	2.18	0.44
1:B:121:TYR:O	1:B:128:CYS:N	2.42	0.43
1:A:213:LYS:HG3	1:A:286:LEU:CD2	2.43	0.43
1:B:189:LEU:HD23	1:B:189:LEU:HA	1.88	0.43
1:B:145:ASP:O	1:B:149:ARG:HG3	2.19	0.42
1:A:288:GLY:O	1:A:299:LYS:HB2	2.21	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	204/254 (80%)	199 (98%)	5 (2%)	0	100	100
1	B	200/254 (79%)	194 (97%)	6 (3%)	0	100	100
All	All	404/508 (80%)	393 (97%)	11 (3%)	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	181/228 (79%)	180 (99%)	1 (1%)	78	88
1	B	184/228 (81%)	181 (98%)	3 (2%)	55	72
All	All	365/456 (80%)	361 (99%)	4 (1%)	65	80

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	192	LEU
1	B	161	LEU
1	B	171	LEU
1	B	191	LEU

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

Mol	Chain	Res	Type
1	A	135	ASN
1	A	219	GLN
1	A	221	GLN
1	A	310	ASN
1	B	135	ASN
1	B	195	GLN
1	B	221	GLN

5.3.3 RNA ⓘ

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 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

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, 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	205/254 (80%)	1.61	62 (30%)	1 1	29, 76, 137, 166	1 (0%)
1	B	200/254 (78%)	1.30	33 (16%)	4 5	47, 74, 114, 128	0
All	All	405/508 (79%)	1.46	95 (23%)	2 2	29, 75, 127, 166	1 (0%)

All (95) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	162	GLU	6.4
1	A	157	ASN	6.4
1	A	158	PRO	6.0
1	A	189	LEU	5.5
1	B	120	THR	5.4
1	A	161	LEU	5.4
1	A	177	ALA	5.2
1	A	191	LEU	4.9
1	B	123	GLU	4.8
1	A	160	GLN	4.8
1	A	128	CYS	4.8
1	A	190	ASP	4.5
1	A	150	LEU	4.5
1	B	320	PRO	4.4
1	A	193	LYS	4.4
1	A	123	GLU	4.3
1	A	289	SER	4.2
1	A	141	LEU	4.2
1	A	176	PHE	4.2
1	A	319	VAL	4.2
1	A	184	PHE	4.1
1	B	129	LEU	4.1
1	B	119	GLU	4.1
1	B	121	TYR	3.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	187	LEU	3.8
1	A	286	LEU	3.8
1	B	122	ILE	3.7
1	A	163	ALA	3.7
1	A	182	LYS	3.7
1	B	181	GLN	3.7
1	B	127	ILE	3.6
1	A	119	GLU	3.6
1	B	131	HIS	3.6
1	B	156	LEU	3.4
1	A	188	SER	3.3
1	A	244	VAL	3.3
1	A	303	ASP	3.3
1	A	185	SER	3.3
1	A	143	LEU	3.3
1	B	162	GLU	3.3
1	B	136	ILE	3.2
1	A	127	ILE	3.2
1	B	230	GLU	3.2
1	A	180	LEU	3.2
1	B	161	LEU	3.1
1	A	111	ALA	3.1
1	A	152	SER	3.1
1	B	134	ARG	3.0
1	A	159	ASP	3.0
1	B	158	PRO	3.0
1	A	172	HIS	3.0
1	B	241	ASN	2.9
1	A	183	THR	2.8
1	A	167	TYR	2.8
1	A	136	ILE	2.7
1	A	145	ASP	2.7
1	A	148	ASP	2.7
1	B	125	GLY	2.7
1	A	112	ALA	2.7
1	A	186	GLY	2.7
1	A	291	LYS	2.7
1	B	145	ASP	2.6
1	A	200	ARG	2.6
1	B	167	TYR	2.6
1	B	319	VAL	2.6
1	A	171	LEU	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	141	LEU	2.5
1	A	164	VAL	2.5
1	A	165	GLU	2.5
1	B	174	LEU	2.5
1	B	155	HIS	2.5
1	B	143	LEU	2.4
1	A	168	GLU	2.4
1	A	194	ALA	2.4
1	B	189	LEU	2.4
1	B	243	ALA	2.3
1	A	138	LYS	2.3
1	B	151	LYS	2.3
1	A	153	GLY	2.3
1	A	170	VAL	2.3
1	A	178	LYS	2.3
1	B	183	THR	2.3
1	B	210	GLU	2.3
1	A	126	LEU	2.2
1	A	130	LYS	2.2
1	A	299	LYS	2.2
1	B	248	SER	2.2
1	A	288	GLY	2.2
1	A	154	GLU	2.1
1	A	151	LYS	2.1
1	B	278	SER	2.1
1	A	207	LEU	2.1
1	A	114	PRO	2.0
1	A	139	LYS	2.0
1	A	149	ARG	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 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.