



Full wwPDB X-ray Structure Validation Report ⓘ

May 22, 2023 – 01:25 pm BST

PDB ID : 7ZW4
Title : Crystal structure of Talin R7R8 domains with Caskin-2 LD-peptide
Authors : Celie, P.H.N.; Joosten, R.P.; Sonnenberg, A.; Perrakis, A.
Deposited on : 2022-05-18
Resolution : 2.72 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 : 4.02b-467
Xtrriage (Phenix) : 1.13
EDS : 2.33
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) : 2.33

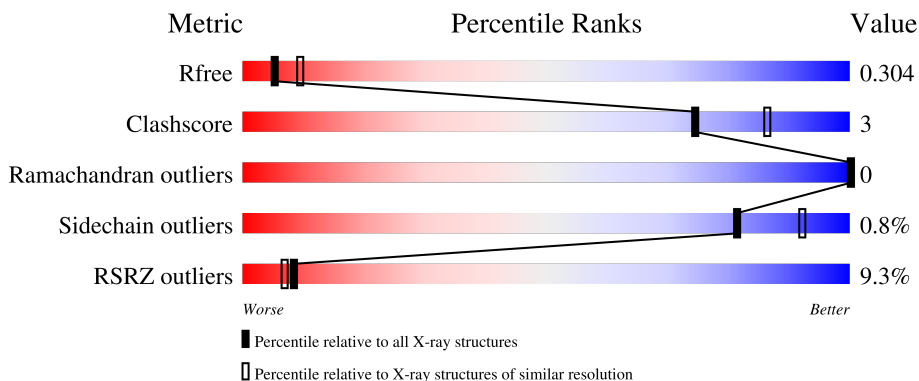
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.72 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	3359 (2.74-2.70)
Clashscore	141614	3686 (2.74-2.70)
Ramachandran outliers	138981	3622 (2.74-2.70)
Sidechain outliers	138945	3623 (2.74-2.70)
RSRZ outliers	127900	3276 (2.74-2.70)

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	303	 6% 91% 8%
2	B	25	 44% 88% 12%

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 2382 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 Talin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	298	2189	1342	394	438	15	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1357	GLY	-	expression tag	UNP P26039

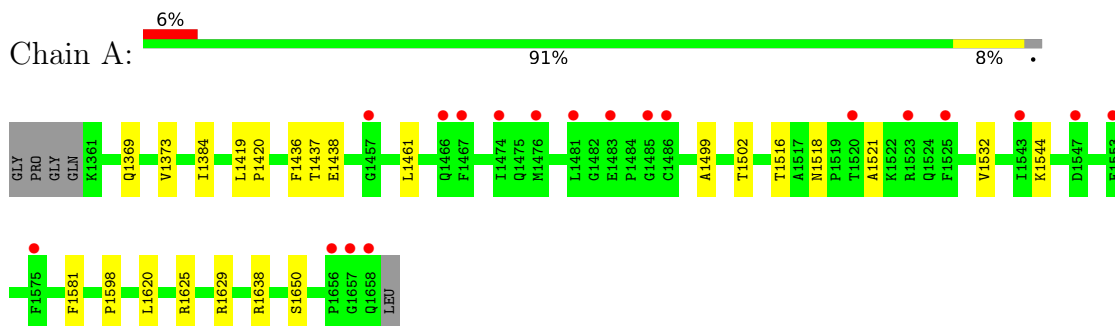
- Molecule 2 is a protein called Caskin-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	25	193	119	29	43	2	0	0	0

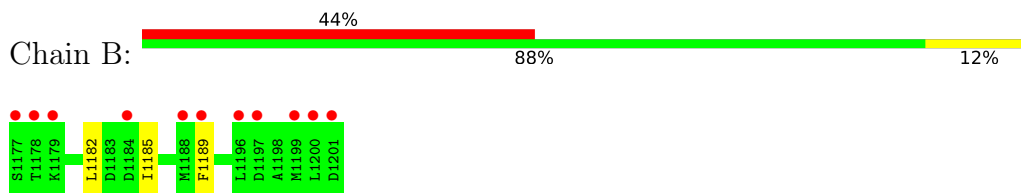
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: Talin-1



- Molecule 2: Caskin-2



4 Data and refinement statistics

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, α , β , γ	46.66Å 61.35Å 98.18Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.09 – 2.72 49.09 – 2.72	Depositor EDS
% Data completeness (in resolution range)	99.8 (49.09-2.72) 99.8 (49.09-2.72)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.33 (at 2.73Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.242 , 0.269 0.280 , 0.304	Depositor DCC
R_{free} test set	424 reflections (5.30%)	wwPDB-VP
Wilson B-factor (Å ²)	78.3	Xtrriage
Anisotropy	0.344	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 50.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	2382	wwPDB-VP
Average B, all atoms (Å ²)	86.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.75% 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	0.31	0/2218	0.49	0/3009
2	B	0.28	0/194	0.45	0/260
All	All	0.31	0/2412	0.49	0/3269

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 [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	2189	0	2185	15	0
2	B	193	0	183	3	0
All	All	2382	0	2368	15	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 (15) 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:1502:THR:HG23	1:A:1532:VAL:HG12	1.87	0.56
1:A:1384:ILE:HD11	1:A:1438:GLU:O	2.11	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1499:ALA:HB1	2:B:1185:ILE:HD12	1.94	0.49
1:A:1502:THR:HG23	1:A:1532:VAL:CG1	2.44	0.47
1:A:1461:LEU:HD21	1:A:1516:THR:HB	1.97	0.47
1:A:1598:PRO:HB2	1:A:1650:SER:OG	2.15	0.47
1:A:1502:THR:HG22	2:B:1189:PHE:CE1	2.50	0.46
1:A:1373:VAL:HG12	1:A:1437:THR:HG21	1.97	0.46
1:A:1620:LEU:HD21	1:A:1629:ARG:HB2	1.98	0.45
1:A:1419:LEU:N	1:A:1420:PRO:HD2	2.33	0.44
1:A:1369:GLN:O	1:A:1373:VAL:HG23	2.19	0.42
1:A:1518:ASN:HB3	1:A:1521:ALA:HB3	2.01	0.42
1:A:1461:LEU:HD22	1:A:1581:PHE:CD1	2.55	0.41
1:A:1544:LYS:HA	2:B:1182:LEU:HD23	2.03	0.41
1:A:1638:ARG:NE	1:A:1638:ARG:HA	2.37	0.40

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	296/303 (98%)	293 (99%)	3 (1%)	0	100	100
2	B	23/25 (92%)	23 (100%)	0	0	100	100
All	All	319/328 (97%)	316 (99%)	3 (1%)	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	234/237 (99%)	232 (99%)	2 (1%)	78	91
2	B	22/22 (100%)	22 (100%)	0	100	100
All	All	256/259 (99%)	254 (99%)	2 (1%)	81	92

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

Mol	Chain	Res	Type
1	A	1436	PHE
1	A	1625	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 monosaccharides 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

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	298/303 (98%)	0.54	19 (6%) 19 18	47, 84, 136, 160	0
2	B	25/25 (100%)	1.96	11 (44%) 0 0	66, 76, 90, 91	0
All	All	323/328 (98%)	0.65	30 (9%) 8 7	47, 82, 134, 160	0

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	1177	SER	8.0
1	A	1486	CYS	5.8
1	A	1481	LEU	5.6
1	A	1485	GLY	4.7
2	B	1178	THR	4.3
2	B	1199	MET	4.2
1	A	1553	GLU	4.0
1	A	1520	THR	3.7
1	A	1474	ILE	3.4
1	A	1483	GLU	3.3
2	B	1201	ASP	3.3
2	B	1188	MET	3.1
2	B	1179	LYS	3.1
2	B	1200	LEU	3.0
2	B	1196	LEU	3.0
2	B	1189	PHE	2.9
1	A	1575	PHE	2.8
1	A	1466	GLN	2.7
1	A	1657	GLY	2.7
1	A	1547	ASP	2.7
1	A	1543	ILE	2.5
2	B	1184	ASP	2.5
1	A	1467	PHE	2.5
1	A	1523	ARG	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1658	GLN	2.4
1	A	1656	PRO	2.4
1	A	1476	MET	2.4
1	A	1457	GLY	2.3
1	A	1525	PHE	2.1
2	B	1197	ASP	2.1

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 monosaccharides 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.