



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 28, 2020 – 08:28 pm BST

PDB ID : 1SJ7  
Title : Crystal Structure of Talin Rod 482-655  
Authors : Papagrigoriou, E.; Gingras, A.R.; Barsukov, I.L.; Critchley, D.R.; Emsley, J.  
Deposited on : 2004-03-03  
Resolution : 2.50 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](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

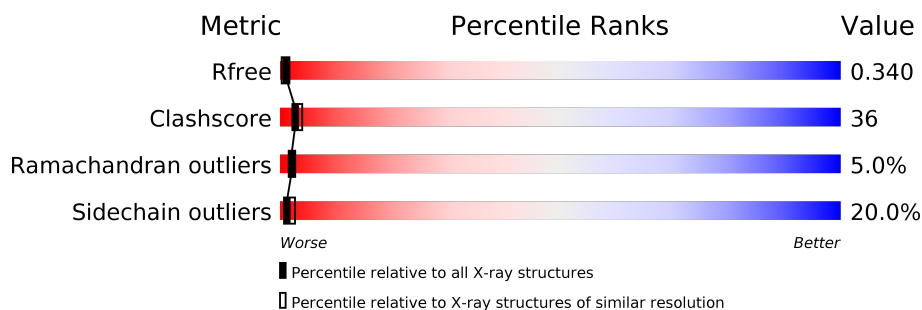
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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))
$R_{free}$	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	174	<div> <div style="width: 46%; background-color: green;"></div> <div style="width: 37%; background-color: yellow;"></div> <div style="width: 11%; background-color: orange;"></div> <div style="width: 5%; background-color: red;"></div> <div style="width: 1%; background-color: grey;"></div> </div> <div>46% 37% 11% . .</div>
1	B	174	<div> <div style="width: 43%; background-color: green;"></div> <div style="width: 40%; background-color: yellow;"></div> <div style="width: 14%; background-color: orange;"></div> <div style="width: 3%; background-color: red;"></div> <div style="width: 1%; background-color: grey;"></div> </div> <div>43% 40% 14% . .</div>
1	C	174	<div> <div style="width: 46%; background-color: green;"></div> <div style="width: 33%; background-color: yellow;"></div> <div style="width: 14%; background-color: orange;"></div> <div style="width: 5%; background-color: red;"></div> <div style="width: 1%; background-color: grey;"></div> </div> <div>46% 33% 14% . 5%</div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3731 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
1	A	167	Total	C	N	O	S	0	0	0
			1195	727	214	252	2			
1	B	170	Total	C	N	O	S	0	0	0
			1213	739	217	255	2			
1	C	166	Total	C	N	O	S	0	0	0
			1186	720	213	251	2			

- Molecule 2 is PLATINUM (II) ION (three-letter code: PT) (formula: Pt).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	3	Total	Pt	0	0
			3	3		
2	A	2	Total	Pt	0	0
			2	2		
2	C	2	Total	Pt	0	0
			2	2		

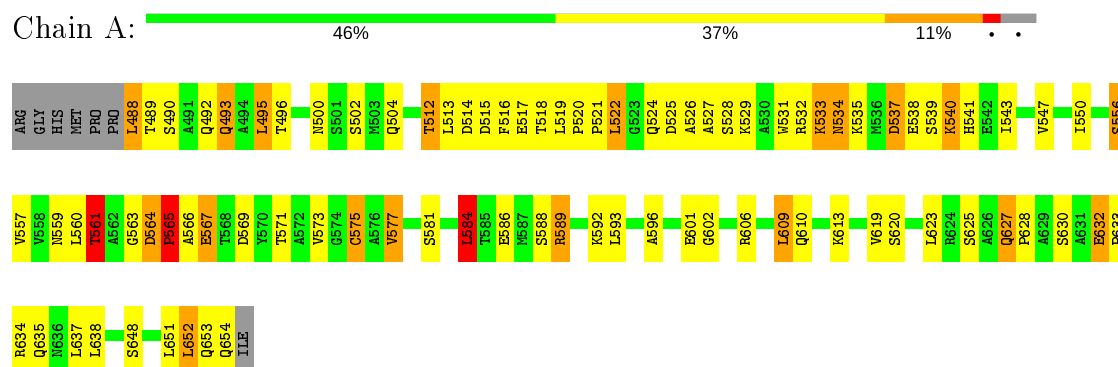
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	49	Total	O	0	0
			49	49		
3	B	42	Total	O	0	0
			42	42		
3	C	39	Total	O	0	0
			39	39		

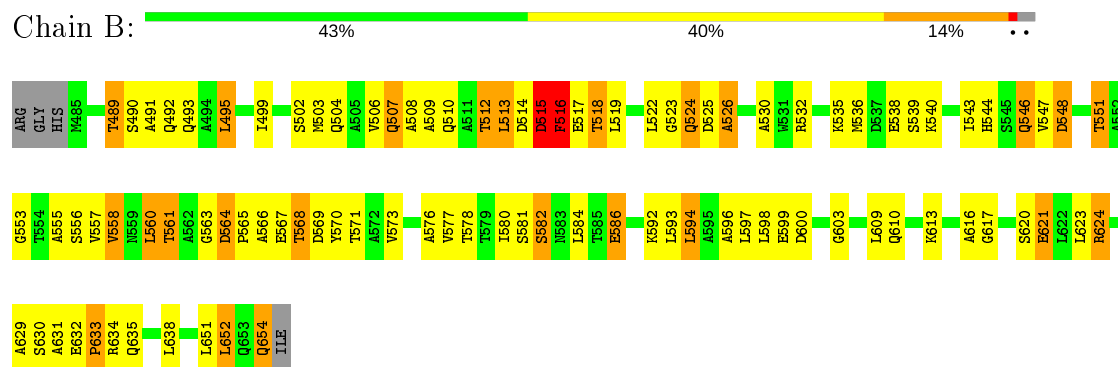
### 3 Residue-property plots [i](#)

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. 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. 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.

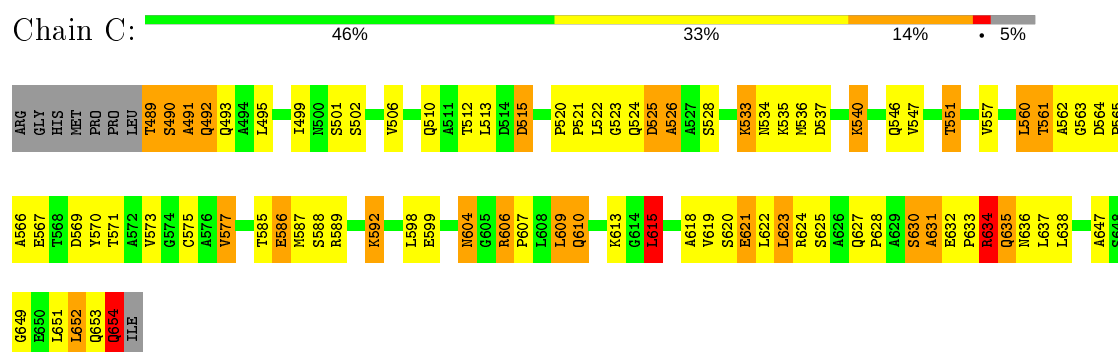
#### • Molecule 1: Talin 1



#### • Molecule 1: Talin 1



#### • Molecule 1: Talin 1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	95.66Å 95.66Å 115.22Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	28.16 – 2.50 47.30 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.9 (28.16-2.50) 99.8 (47.30-2.40)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.00 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.0	Depositor
R, $R_{free}$	0.241 , 0.286 0.322 , 0.340	Depositor DCC
$R_{free}$ test set	1122 reflections (4.63%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	52.2	Xtriage
Anisotropy	0.367	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 11.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.025 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	3731	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.61% 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 ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: PT

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	1.24	2/1206 (0.2%)	1.29	8/1640 (0.5%)
1	B	1.18	3/1226 (0.2%)	1.35	10/1670 (0.6%)
1	C	1.21	3/1197 (0.3%)	1.33	11/1628 (0.7%)
All	All	1.21	8/3629 (0.2%)	1.32	29/4938 (0.6%)

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	538	GLU	CD-OE2	9.37	1.35	1.25
1	A	575	CYS	CB-SG	6.85	1.93	1.82
1	C	621	GLU	CD-OE2	6.67	1.32	1.25
1	B	654	GLN	CG-CD	6.48	1.66	1.51
1	C	540	LYS	CD-CE	5.81	1.65	1.51

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	515	ASP	CB-CG-OD2	11.23	128.40	118.30
1	B	624	ARG	NE-CZ-NH1	10.94	125.77	120.30
1	B	624	ARG	NE-CZ-NH2	-10.59	115.01	120.30
1	A	514	ASP	CB-CG-OD2	8.33	125.80	118.30
1	B	548	ASP	CB-CG-OD2	8.20	125.68	118.30

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	1195	0	1180	92	0
1	B	1213	0	1193	104	0
1	C	1186	0	1167	74	1
2	A	2	0	0	0	0
2	B	3	0	0	0	0
2	C	2	0	0	0	0
3	A	49	0	0	3	0
3	B	42	0	0	3	1
3	C	39	0	0	6	1
All	All	3731	0	3540	258	2

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 36.

The worst 5 of 258 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:B:512:THR:CG2	1:B:592:LYS:HE2	1.55	1.37
1:A:632:GLU:CB	1:A:633:PRO:HA	1.62	1.25
1:B:516:PHE:CZ	1:B:599:GLU:HG3	1.75	1.20
1:B:512:THR:HG22	1:B:592:LYS:CE	1.70	1.19
1:B:516:PHE:HZ	1:B:599:GLU:HG3	1.00	1.14

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:73:HOH:O	3:C:10:HOH:O[4_456]	2.03	0.17
1:C:575:CYS:SG	1:C:575:CYS:SG[5_676]	2.16	0.04

## 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	165/174 (95%)	147 (89%)	12 (7%)	6 (4%)	3	4
1	B	168/174 (97%)	154 (92%)	6 (4%)	8 (5%)	2	2
1	C	164/174 (94%)	146 (89%)	7 (4%)	11 (7%)	1	1
All	All	497/522 (95%)	447 (90%)	25 (5%)	25 (5%)	2	2

5 of 25 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	565	PRO
1	A	566	ALA
1	A	632	GLU
1	B	516	PHE
1	B	524	GLN

### 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	123/131 (94%)	95 (77%)	28 (23%)	1	1
1	B	125/131 (95%)	99 (79%)	26 (21%)	1	2
1	C	122/131 (93%)	102 (84%)	20 (16%)	2	4
All	All	370/393 (94%)	296 (80%)	74 (20%)	1	2

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



Mol	Chain	Res	Type
1	B	512	THR
1	B	568	THR
1	C	615	LEU
1	B	513	LEU
1	B	551	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	546	GLN
1	B	636	ASN
1	C	639	GLN
1	B	635	GLN
1	B	639	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](#)

Of 7 ligands modelled in this entry, 7 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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.