Full wwPDB X-ray Structure Validation Report (i)

Aug 16, 2021 – 04:02 PM EDT

PDB ID	:	3DYJ
Title	:	Crystal Structure of the R11R12 Domains of Talin
Authors	:	Gingras, A.R.; Joyce, M.G.; Critchley, D.R.; Emsley, J.
Deposited on	:	2008-07-28
Resolution	:	1.85 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

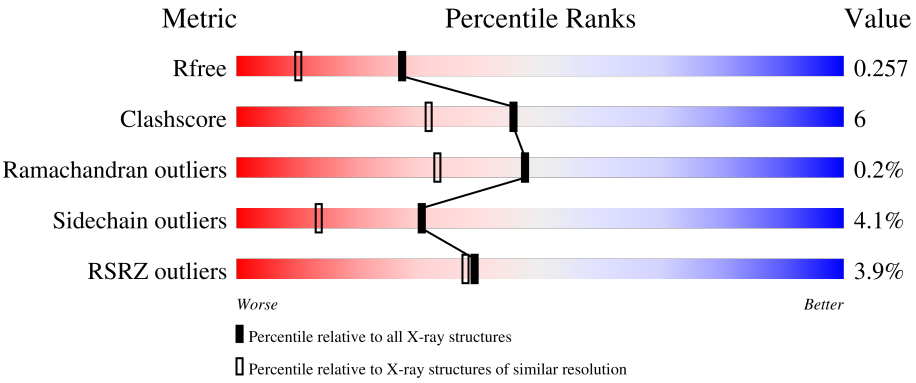
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.85 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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 <=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	332	4% 86% 9% • 5%
1	B	332	3% 80% 14% • 5%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5010 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	317	Total	C	N	O	S	Se	0	0	0
			2339	1451	416	461	5	6			
1	B	317	Total	C	N	O	S	Se	0	0	0
			2339	1451	416	461	5	6			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1962	GLY	-	expression tag	UNP P26039
A	1963	ILE	-	expression tag	UNP P26039
A	1964	ASP	-	expression tag	UNP P26039
A	1965	PRO	-	expression tag	UNP P26039
A	1966	PHE	-	expression tag	UNP P26039
A	1967	THR	-	expression tag	UNP P26039
A	1968	GLY	-	expression tag	UNP P26039
A	1969	ILE	-	expression tag	UNP P26039
A	1970	ASP	-	expression tag	UNP P26039
A	1971	PRO	-	expression tag	UNP P26039
A	1972	PHE	-	expression tag	UNP P26039
A	1973	THR	-	expression tag	UNP P26039
B	1962	GLY	-	expression tag	UNP P26039
B	1963	ILE	-	expression tag	UNP P26039
B	1964	ASP	-	expression tag	UNP P26039
B	1965	PRO	-	expression tag	UNP P26039
B	1966	PHE	-	expression tag	UNP P26039
B	1967	THR	-	expression tag	UNP P26039
B	1968	GLY	-	expression tag	UNP P26039
B	1969	ILE	-	expression tag	UNP P26039
B	1970	ASP	-	expression tag	UNP P26039
B	1971	PRO	-	expression tag	UNP P26039
B	1972	PHE	-	expression tag	UNP P26039
B	1973	THR	-	expression tag	UNP P26039

- Molecule 2 is water.

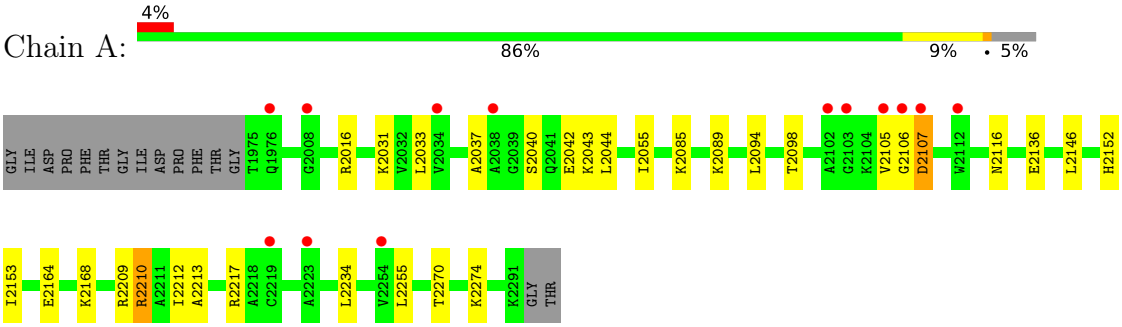
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	185	Total O 185 185	0	0
2	B	147	Total O 147 147	0	0

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.

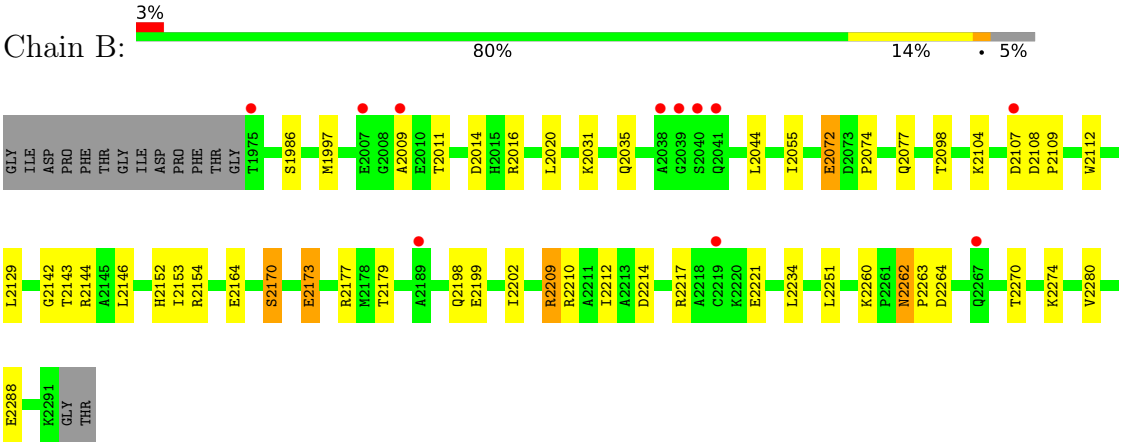
• Molecule 1: TALIN-1

Chain A:



• Molecule 1: TALIN-1

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	58.71Å 57.65Å 92.55Å 90.00° 102.96° 90.00°	Depositor
Resolution (Å)	19.57 – 1.85 19.57 – 1.85	Depositor EDS
% Data completeness (in resolution range)	98.3 (19.57-1.85) 98.3 (19.57-1.85)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	1.47 (at 1.85Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.214 , 0.260 0.212 , 0.257	Depositor DCC
R_{free} test set	1300 reflections (2.56%)	wwPDB-VP
Wilson B-factor (Å^2)	29.5	Xtriage
Anisotropy	0.041	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 42.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	5010	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 61.99 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.2059e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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	0.73	0/2357	0.71	1/3187 (0.0%)
1	B	0.70	0/2357	0.69	1/3187 (0.0%)
All	All	0.72	0/4714	0.70	2/6374 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	1997	MSE	CG-SE-CE	-6.70	84.17	98.90
1	A	2055	ILE	CG1-CB-CG2	-5.15	100.07	111.40

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	2339	0	2409	28	0
1	B	2339	0	2409	33	0
2	A	185	0	0	6	0
2	B	147	0	0	5	0
All	All	5010	0	4818	61	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 6.

All (61) 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:2210:ARG:HH11	1:A:2210:ARG:HG3	1.04	1.15
1:A:2210:ARG:HG3	1:A:2210:ARG:NH1	1.82	0.90
1:B:2011:THR:HG23	1:B:2014:ASP:H	1.41	0.86
1:B:2044:LEU:HD11	1:B:2098:THR:HG21	1.63	0.81
1:A:2094:LEU:O	1:A:2098:THR:HG23	1.86	0.75
1:A:2085:LYS:HE3	1:A:2089:LYS:HZ1	1.54	0.73
1:A:2136:GLU:HG3	2:A:310:HOH:O	1.89	0.72
1:A:2085:LYS:HE3	1:A:2089:LYS:NZ	2.05	0.71
1:B:2016:ARG:O	1:B:2020:LEU:HG	1.92	0.69
1:B:2020:LEU:HD23	1:B:2129:LEU:HD22	1.73	0.69
1:B:2044:LEU:HD11	1:B:2098:THR:CG2	2.22	0.69
1:A:2217:ARG:HG3	2:A:66:HOH:O	1.96	0.66
1:B:2170:SER:HB2	1:B:2221:GLU:HB3	1.77	0.64
1:B:2262:ASN:HB2	1:B:2263:PRO:HD2	1.79	0.63
1:B:2143:THR:H	1:B:2198:GLN:HE22	1.46	0.62
1:A:2031:LYS:NZ	2:A:143:HOH:O	2.31	0.62
1:B:2031:LYS:HE2	1:B:2035:GLN:NE2	2.17	0.60
1:B:2199:GLU:O	1:B:2202:ILE:HG22	2.01	0.60
1:A:2040:SER:HB2	1:A:2043:LYS:HB2	1.83	0.59
1:A:2210:ARG:HH11	1:A:2210:ARG:CG	1.95	0.56
1:A:2217:ARG:HD2	2:A:276:HOH:O	2.06	0.55
1:A:2152:HIS:CD2	1:A:2209:ARG:HG3	2.42	0.54
1:B:2153:ILE:HG23	1:B:2212:ILE:HG21	1.89	0.53
1:B:2179:THR:HG22	1:B:2280:VAL:HG13	1.90	0.53
1:A:2210:ARG:HD3	2:A:186:HOH:O	2.10	0.52
1:B:2074:PRO:HA	1:B:2077:GLN:HE21	1.74	0.52
1:A:2085:LYS:CE	1:A:2089:LYS:NZ	2.73	0.52
1:B:2152:HIS:ND1	1:B:2209:ARG:HG3	2.25	0.51
1:B:2262:ASN:HD22	1:B:2264:ASP:H	1.59	0.49
1:B:2270:THR:HG22	1:B:2274:LYS:NZ	2.28	0.48
1:A:2085:LYS:HG2	1:A:2089:LYS:HZ3	1.79	0.48
1:B:2170:SER:HB2	1:B:2221:GLU:CB	2.44	0.48
1:A:2033:LEU:HD22	1:A:2098:THR:HG21	1.96	0.48
1:B:2107:ASP:CB	2:B:297:HOH:O	2.62	0.47
1:A:2153:ILE:HG23	1:A:2212:ILE:HG21	1.96	0.47
1:A:2210:ARG:NH1	1:A:2210:ARG:CG	2.62	0.47
1:A:2217:ARG:HG2	1:A:2217:ARG:HH21	1.80	0.47
1:A:2217:ARG:HB3	2:A:166:HOH:O	2.14	0.47
1:B:2107:ASP:HB2	2:B:297:HOH:O	2.16	0.46
1:B:2146:LEU:HD11	1:B:2251:LEU:HD22	1.97	0.46
1:B:2104:LYS:HD2	1:B:2108:ASP:OD1	2.16	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:2109:PRO:HA	1:B:2112:TRP:NE1	2.31	0.45
1:A:2213:ALA:O	1:A:2217:ARG:HG3	2.16	0.45
1:A:2085:LYS:CE	1:A:2089:LYS:HZ1	2.25	0.45
1:B:2199:GLU:HA	1:B:2202:ILE:HG22	1.98	0.45
1:B:2072:GLU:H	1:B:2072:GLU:CD	2.21	0.44
1:B:2154:ARG:HD2	2:B:338:HOH:O	2.17	0.44
1:A:2270:THR:HG22	1:A:2274:LYS:HE3	1.99	0.44
1:A:2146:LEU:HB3	1:A:2255:LEU:HD21	2.00	0.43
1:B:2011:THR:HG23	1:B:2014:ASP:N	2.22	0.43
1:B:2173:GLU:CD	1:B:2173:GLU:H	2.20	0.43
1:B:2142:GLY:H	1:B:2198:GLN:HE21	1.67	0.43
1:B:2009:ALA:HA	1:B:2260:LYS:HE2	2.01	0.42
1:A:2234:LEU:HD13	2:B:5:HOH:O	2.20	0.42
1:B:2107:ASP:HA	2:B:271:HOH:O	2.20	0.41
1:B:2020:LEU:CD2	1:B:2129:LEU:HD22	2.48	0.41
1:A:2037:ALA:HA	1:A:2044:LEU:HD13	2.03	0.40
1:B:2142:GLY:H	1:B:2198:GLN:NE2	2.19	0.40
1:B:2214:ASP:OD1	1:B:2217:ARG:NH2	2.54	0.40
1:A:2044:LEU:HD12	1:A:2044:LEU:HA	1.83	0.40
1:A:2105:VAL:O	1:A:2107:ASP:N	2.54	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	315/332 (95%)	311 (99%)	3 (1%)	1 (0%)	41 26
1	B	315/332 (95%)	314 (100%)	1 (0%)	0	100 100
All	All	630/664 (95%)	625 (99%)	4 (1%)	1 (0%)	47 33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2106	GLY

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	246/251 (98%)	239 (97%)	7 (3%)	43 27
1	B	246/251 (98%)	233 (95%)	13 (5%)	22 8
All	All	492/502 (98%)	472 (96%)	20 (4%)	30 13

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2016	ARG
1	A	2042	GLU
1	A	2107	ASP
1	A	2116	ASN
1	A	2164	GLU
1	A	2168	LYS
1	A	2210	ARG
1	B	1986	SER
1	B	2055	ILE
1	B	2072	GLU
1	B	2144	ARG
1	B	2164	GLU
1	B	2170	SER
1	B	2173	GLU
1	B	2177	ARG
1	B	2209	ARG
1	B	2210	ARG
1	B	2234	LEU
1	B	2262	ASN
1	B	2288	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	2046	GLN
1	A	2152	HIS
1	A	2268	GLN
1	A	2272	HIS
1	B	2035	GLN
1	B	2041	GLN
1	B	2077	GLN
1	B	2198	GLN
1	B	2262	ASN

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 (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^{th} 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	311/332 (93%)	0.14	13 (4%)	36	34	18, 27, 42, 58	0
1	B	311/332 (93%)	0.25	11 (3%)	44	41	17, 31, 46, 58	0
All	All	622/664 (93%)	0.19	24 (3%)	39	38	17, 29, 46, 58	0

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	2038	ALA	6.9
1	A	2105	VAL	6.2
1	A	2112	TRP	5.0
1	A	2107	ASP	4.8
1	B	2007	GLU	3.3
1	B	2107	ASP	3.3
1	A	2106	GLY	3.3
1	A	2103	GLY	3.2
1	B	2189	ALA	2.8
1	B	2009	ALA	2.7
1	B	2267	GLN	2.6
1	A	2008	GLY	2.6
1	A	2034	VAL	2.6
1	A	2102	ALA	2.5
1	B	2039	GLY	2.5
1	B	2219	CYS	2.5
1	A	1976	GLN	2.4
1	B	2040	SER	2.3
1	A	2219	CYS	2.3
1	A	2038	ALA	2.2
1	A	2223	ALA	2.2
1	B	1975	THR	2.2
1	B	2041	GLN	2.2
1	A	2254	VAL	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 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.