



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 24, 2022 – 12:19 pm GMT

PDB ID : 5MV7  
Title : Structure of human Myosin 7b C-terminal MyTH4-FERM MF2 domain  
Authors : Planelles-Herrero, V.J.; Sourigues, Y.; Titus, M.A.; Houdusse, A.  
Deposited on : 2017-01-15  
Resolution : 2.44 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](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  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.27  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0267  
CCP4 : 7.1.010 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.27

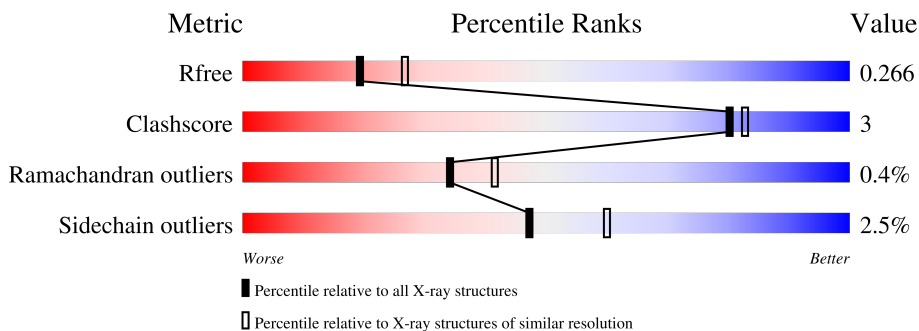
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*


The reported resolution of this entry is 2.44 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	1564 (2.46-2.42)
Clashscore	141614	1631 (2.46-2.42)
Ramachandran outliers	138981	1617 (2.46-2.42)
Sidechain outliers	138945	1617 (2.46-2.42)

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  $\geq 3$ , 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	522	 86% 8% 6%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	TAM	A	2202	-	-	X	-

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4202 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.

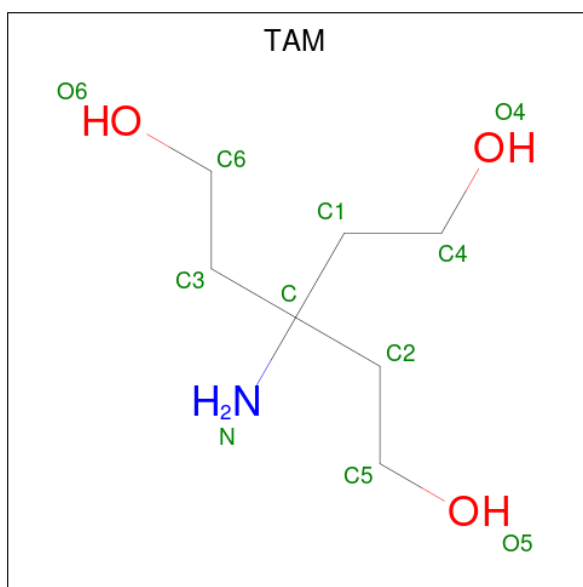
- Molecule 1 is a protein called Unconventional myosin-VIIb.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	493	4026	2590	708	710	18	0	0	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	HIS	-	expression tag	UNP Q6PIF6
A	2	HIS	-	expression tag	UNP Q6PIF6
A	3	HIS	-	expression tag	UNP Q6PIF6
A	4	HIS	-	expression tag	UNP Q6PIF6
A	5	HIS	-	expression tag	UNP Q6PIF6
A	6	HIS	-	expression tag	UNP Q6PIF6
A	7	SER	-	expression tag	UNP Q6PIF6
A	8	SER	-	expression tag	UNP Q6PIF6
A	9	GLY	-	expression tag	UNP Q6PIF6
A	10	HIS	-	expression tag	UNP Q6PIF6
A	1626	ALA	MET	conflict	UNP Q6PIF6

- Molecule 2 is TRIS(HYDROXYETHYL)AMINOMETHANE (three-letter code: TAM) (formula: C<sub>7</sub>H<sub>17</sub>NO<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	11	7	1	3	0	0
2	A	1	11	7	1	3	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
3	A	1	5	4	1	0	0
3	A	1	5	4	1	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		

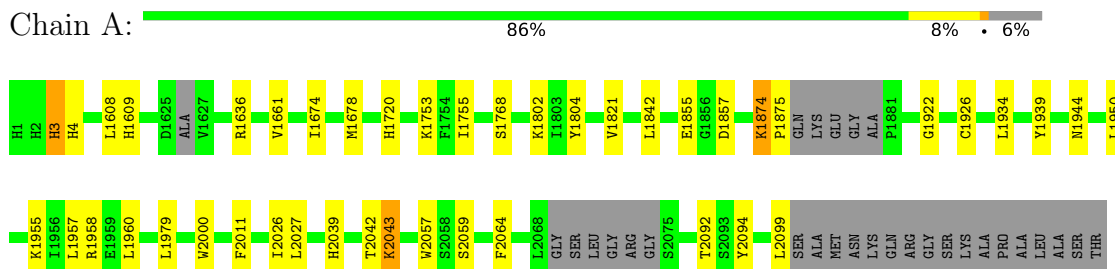
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	139	Total	O	0	0
			139	139		

### 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. 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: Unconventional myosin-VIIIb



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	123.67Å 42.81Å 118.37Å 90.00° 97.69° 90.00°	Depositor
Resolution (Å)	35.17 – 2.44 35.17 – 2.44	Depositor EDS
% Data completeness (in resolution range)	96.9 (35.17-2.44) 97.5 (35.17-2.44)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.62 (at 2.45Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, $R_{free}$	0.209 , 0.258 0.215 , 0.266	Depositor DCC
$R_{free}$ test set	1132 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	52.0	Xtriage
Anisotropy	0.730	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	4202	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.59% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: TAM, SO4

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	0.50	0/4135	0.68	0/5602

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	4026	0	4020	21	0
2	A	22	0	34	6	0
3	A	15	0	0	0	0
4	A	139	0	0	0	0
All	All	4202	0	4054	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:1720:HIS:H	2:A:2202:TAM:HN2	1.20	0.88
1:A:1608:LEU:HD23	2:A:2202:TAM:H22	1.77	0.65
1:A:1922:GLY:O	1:A:2043:LYS:HE2	1.97	0.64
1:A:1934:LEU:HD21	1:A:2000:TRP:CD1	2.41	0.56
1:A:1720:HIS:HA	2:A:2202:TAM:H32	1.87	0.56
1:A:1720:HIS:CE1	2:A:2202:TAM:H41	2.46	0.50
1:A:1874:LYS:HG3	1:A:1875:PRO:HD3	1.94	0.50
1:A:1950:LEU:HD23	1:A:1979:LEU:HG	1.95	0.48
1:A:1674:ILE:HG23	1:A:1678:MET:CE	2.45	0.47
1:A:2057:TRP:HE1	1:A:2092:THR:HG23	1.78	0.47
2:A:2202:TAM:H12	2:A:2202:TAM:H61	1.70	0.45
1:A:2039:HIS:HB3	1:A:2042:THR:O	2.17	0.45
1:A:3:HIS:HD2	1:A:4:HIS:CD2	2.36	0.44
1:A:1955:LYS:HB3	1:A:1958:ARG:HD2	2.00	0.44
1:A:1821:VAL:HG13	1:A:1842:LEU:HD12	2.00	0.43
1:A:2011:PHE:HB3	1:A:2027:LEU:HD11	1.99	0.43
1:A:1802:LYS:HD3	1:A:1804:TYR:CZ	2.54	0.42
1:A:1720:HIS:ND1	2:A:2202:TAM:H32	2.36	0.41
1:A:1609:HIS:CD2	1:A:1753:LYS:HB3	2.56	0.41
1:A:1755:ILE:HG21	1:A:1768:SER:HB2	2.03	0.41
1:A:1939:TYR:HD2	1:A:1960:LEU:HD11	1.86	0.40
1:A:2059:SER:HB3	1:A:2064:PHE:HD1	1.86	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	485/522 (93%)	468 (96%)	15 (3%)	2 (0%)	34 41

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2043	LYS
1	A	1944	ASN

### 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	443/465 (95%)	432 (98%)	11 (2%)	47 60

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	3	HIS
1	A	1636	ARG
1	A	1661	VAL
1	A	1855	GLU
1	A	1857	ASP
1	A	1874	LYS
1	A	1926	CYS
1	A	1957	LEU
1	A	2026	ILE
1	A	2094	TYR
1	A	2099	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	1	HIS
1	A	3	HIS
1	A	4	HIS
1	A	1785	HIS
1	A	1786	GLN
1	A	1887	GLN
1	A	2039	HIS

### 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](#)

5 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	SO4	A	2205	-	4,4,4	0.13	0	6,6,6	0.10	0
2	TAM	A	2201	-	7,10,10	0.66	0	9,12,12	1.37	1 (11%)
3	SO4	A	2203	-	4,4,4	0.14	0	6,6,6	0.15	0
2	TAM	A	2202	-	7,10,10	0.72	0	9,12,12	1.08	1 (11%)
3	SO4	A	2204	-	4,4,4	0.14	0	6,6,6	0.06	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	TAM	A	2202	-	-	10/12/12/12	-
2	TAM	A	2201	-	-	4/12/12/12	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	2201	TAM	C2-C-C1	-3.37	104.56	110.50
2	A	2202	TAM	C3-C-C2	-2.13	106.75	110.50

There are no chirality outliers.

All (14) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	2201	TAM	C2-C-C1-C4
2	A	2201	TAM	C3-C-C1-C4
2	A	2201	TAM	N-C-C1-C4
2	A	2202	TAM	C2-C-C1-C4
2	A	2202	TAM	C3-C-C1-C4
2	A	2202	TAM	N-C-C1-C4
2	A	2202	TAM	C1-C-C2-C5
2	A	2202	TAM	C3-C-C2-C5
2	A	2202	TAM	N-C-C2-C5
2	A	2202	TAM	C1-C-C3-C6
2	A	2202	TAM	C2-C-C3-C6
2	A	2202	TAM	N-C-C3-C6
2	A	2202	TAM	C-C3-C6-O6
2	A	2201	TAM	C-C2-C5-O5

There are no ring outliers.

1 monomer is involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	2202	TAM	6	0

## 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](#)

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

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 6.3 Carbohydrates [i](#)

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

### 6.4 Ligands [i](#)

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

### 6.5 Other polymers [i](#)

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