



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 22, 2023 – 09:21 AM EDT

PDB ID : 3GJO  
Title : Crystal structure of human EB1 in complex with microtubule Tip localization signal peptide of MACF  
Authors : Honnappa, S.; Steinmetz, M.O.  
Deposited on : 2009-03-09  
Resolution : 2.50 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 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  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

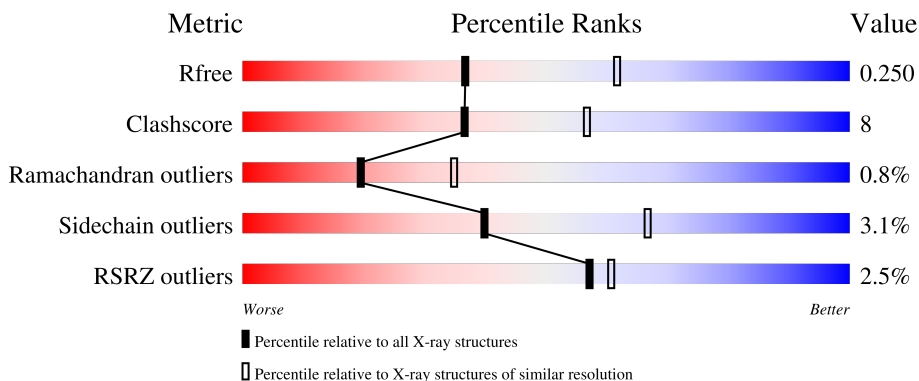
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 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)
RSRZ outliers	127900	4559 (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 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  $\leq 5\%$ . 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	72	 74% 18% 8%
1	B	72	 74% 15% 11%
1	C	72	 81% 10% 10%
1	D	72	 57% 18% 25% 3%
2	E	30	 33% 63%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	30	 17% 10% 70%
2	G	30	 13% 7% 7% 73%
2	H	30	 7% 10% 83%

## 2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 2309 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 Microtubule-associated protein RP/EB family member 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	66	Total 538	C 340	N 86	O 110	S 2	0	0	0
1	B	64	Total 521	C 331	N 83	O 105	S 2	0	0	0
1	C	65	Total 530	C 336	N 85	O 107	S 2	0	0	0
1	D	54	Total 446	C 284	N 73	O 87	S 2	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	189	GLY	-	expression tag	UNP Q15691
A	190	SER	-	expression tag	UNP Q15691
B	189	GLY	-	expression tag	UNP Q15691
B	190	SER	-	expression tag	UNP Q15691
C	189	GLY	-	expression tag	UNP Q15691
C	190	SER	-	expression tag	UNP Q15691
D	189	GLY	-	expression tag	UNP Q15691
D	190	SER	-	expression tag	UNP Q15691

- Molecule 2 is a protein called Dystonin.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	E	11	Total 89	C 57	N 18	O 14	0	0	0
2	F	9	Total 69	C 45	N 12	O 12	0	0	0
2	G	8	Total 60	C 39	N 10	O 11	0	0	0
2	H	5	Total 37	C 24	N 6	O 7	0	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	6	Total O 6 6	0	0
3	B	3	Total O 3 3	0	0
3	C	3	Total O 3 3	0	0
3	D	5	Total O 5 5	0	0
3	E	2	Total O 2 2	0	0

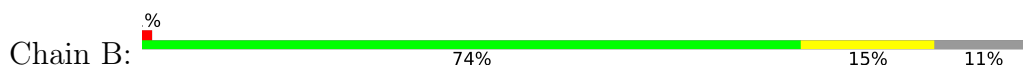
### 3 Residue-property plots [i](#)

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.

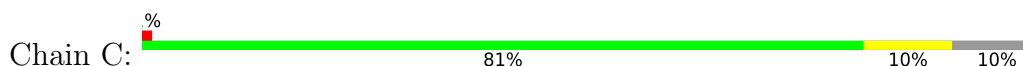
- Molecule 1: Microtubule-associated protein RP/EB family member 1



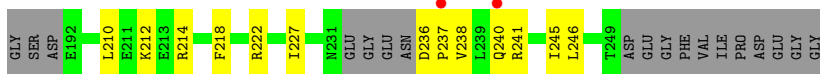
- Molecule 1: Microtubule-associated protein RP/EB family member 1



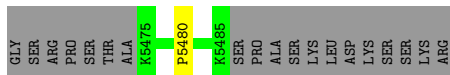
- Molecule 1: Microtubule-associated protein RP/EB family member 1



- Molecule 1: Microtubule-associated protein RP/EB family member 1



- Molecule 2: Dystonin



- Molecule 2: Dystonin



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	45.61Å 44.90Å 74.84Å 90.00° 98.57° 90.00°	Depositor
Resolution (Å)	25.00 – 2.50 37.00 – 2.50	Depositor EDS
% Data completeness (in resolution range)	96.2 (25.00-2.50) 96.2 (37.00-2.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.45 (at 2.51Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.213 , 0.255 0.212 , 0.250	Depositor DCC
$R_{free}$ test set	487 reflections (4.79%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	50.3	Xtrriage
Anisotropy	0.258	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 63.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	2309	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.58% 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](#)

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	0.54	0/544	0.60	0/735
1	B	0.48	0/526	0.54	0/709
1	C	0.51	0/536	0.54	0/724
1	D	0.49	0/449	0.59	0/604
2	E	0.44	0/91	0.52	0/121
2	F	0.47	0/71	0.60	0/96
2	G	0.32	0/62	0.60	0/84
2	H	0.56	0/37	0.75	0/49
All	All	0.50	0/2316	0.57	0/3122

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	538	0	525	12	0
1	B	521	0	512	8	0
1	C	530	0	521	8	0
1	D	446	0	450	10	0
2	E	89	0	103	1	0
2	F	69	0	77	4	0
2	G	60	0	65	4	0
2	H	37	0	42	2	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	6	0	0	0	0
3	B	3	0	0	0	0
3	C	3	0	0	0	0
3	D	5	0	0	0	0
3	E	2	0	0	0	0
All	All	2309	0	2295	36	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 8.

All (36) 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)
2:F:5475:LYS:HB2	2:F:5476:PRO:HD3	1.50	0.94
2:F:5481:THR:HG22	2:F:5482:PRO:HD2	1.54	0.90
1:A:237:PRO:O	1:A:241:ARG:HG2	1.85	0.76
1:A:237:PRO:HA	1:A:240:GLN:HB3	1.69	0.72
2:F:5481:THR:CG2	2:F:5482:PRO:HD2	2.22	0.69
1:C:239:LEU:HD11	1:D:238:VAL:HG11	1.75	0.69
2:G:5481:THR:HG22	2:G:5482:PRO:HD2	1.75	0.68
1:A:192:GLU:HG2	1:A:193:ALA:N	2.09	0.67
1:C:225:GLU:OE1	2:G:5477:SER:HB2	1.99	0.63
1:A:192:GLU:HG2	1:A:193:ALA:H	1.64	0.62
2:F:5475:LYS:CB	2:F:5476:PRO:HD3	2.31	0.57
2:G:5481:THR:HG22	2:G:5482:PRO:CD	2.34	0.57
1:C:247:TYR:CZ	2:G:5478:LYS:HB2	2.39	0.57
1:B:193:ALA:O	1:B:197:MET:HG3	2.07	0.54
2:H:5478:LYS:O	2:H:5478:LYS:HG3	2.09	0.52
1:C:207:VAL:HG13	1:D:210:LEU:HD11	1.91	0.51
1:D:218:PHE:CE2	1:D:222:ARG:HD2	2.46	0.51
1:C:227:ILE:HD11	1:D:245:ILE:HD11	1.94	0.49
1:D:246:LEU:O	2:H:5480:PRO:HD2	2.14	0.47
1:B:201:ASN:HB3	1:C:209:ASP:OD2	2.15	0.47
1:A:192:GLU:CG	1:A:193:ALA:H	2.27	0.47
1:D:236:ASP:HA	1:D:237:PRO:HD3	1.71	0.47
1:A:214:ARG:HD2	1:B:213:GLU:OE2	2.16	0.46
1:B:198:GLN:O	1:B:202:VAL:HG23	2.16	0.46
1:A:250:ASP:HB2	1:B:216:PHE:HE1	1.81	0.45
1:A:254:VAL:O	2:E:5480:PRO:HA	2.16	0.45
1:C:227:ILE:CD1	1:D:245:ILE:HD11	2.47	0.44
1:C:227:ILE:HD11	1:D:241:ARG:HH11	1.82	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:237:PRO:HA	1:D:240:GLN:CD	2.38	0.43
1:A:198:GLN:HA	1:A:198:GLN:OE1	2.19	0.42
1:A:222:ARG:O	1:A:226:LEU:HG	2.19	0.42
1:B:255:ILE:HA	1:B:256:PRO:HD3	1.92	0.42
1:A:250:ASP:HB2	1:B:216:PHE:CE1	2.53	0.42
1:B:231:ASN:HD22	1:B:231:ASN:N	2.17	0.41
1:A:192:GLU:CG	1:A:193:ALA:N	2.79	0.40
1:D:227:ILE:HD13	1:D:227:ILE:HA	1.90	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	64/72 (89%)	64 (100%)	0	0	100	100
1	B	60/72 (83%)	56 (93%)	4 (7%)	0	100	100
1	C	63/72 (88%)	61 (97%)	2 (3%)	0	100	100
1	D	50/72 (69%)	50 (100%)	0	0	100	100
2	E	9/30 (30%)	9 (100%)	0	0	100	100
2	F	7/30 (23%)	5 (71%)	1 (14%)	1 (14%)	0	0
2	G	6/30 (20%)	6 (100%)	0	0	100	100
2	H	3/30 (10%)	2 (67%)	0	1 (33%)	0	0
All	All	262/408 (64%)	253 (97%)	7 (3%)	2 (1%)	19	35

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	H	5480	PRO
2	F	5482	PRO

### 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	60/63 (95%)	58 (97%)	2 (3%)	38	64
1	B	58/63 (92%)	57 (98%)	1 (2%)	60	82
1	C	59/63 (94%)	58 (98%)	1 (2%)	60	82
1	D	50/63 (79%)	48 (96%)	2 (4%)	31	56
2	E	11/27 (41%)	11 (100%)	0	100	100
2	F	9/27 (33%)	9 (100%)	0	100	100
2	G	8/27 (30%)	6 (75%)	2 (25%)	0	1
2	H	5/27 (18%)	5 (100%)	0	100	100
All	All	260/360 (72%)	252 (97%)	8 (3%)	40	67

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	243	VAL
1	A	257	ASP
1	B	226	LEU
1	C	249	THR
1	D	212	LYS
1	D	214	ARG
2	G	5478	LYS
2	G	5481	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	231	ASN
1	D	231	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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	66/72 (91%)	0.15	1 (1%) 73 75	29, 51, 89, 92	0
1	B	64/72 (88%)	0.27	1 (1%) 72 74	31, 50, 78, 84	0
1	C	65/72 (90%)	0.14	1 (1%) 73 75	36, 50, 75, 89	0
1	D	54/72 (75%)	0.16	2 (3%) 41 45	29, 49, 75, 90	0
2	E	11/30 (36%)	0.24	0 100 100	57, 65, 73, 76	0
2	F	9/30 (30%)	0.86	0 100 100	56, 64, 72, 73	0
2	G	8/30 (26%)	0.13	0 100 100	56, 60, 65, 70	0
2	H	5/30 (16%)	2.01	2 (40%) 0 0	77, 77, 78, 82	0
All	All	282/408 (69%)	0.24	7 (2%) 57 61	29, 54, 82, 92	0

All (7) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	193	ALA	3.5
1	C	254	VAL	3.2
2	H	5478	LYS	3.1
2	H	5480	PRO	2.6
1	D	237	PRO	2.5
1	B	255	ILE	2.3
1	D	240	GLN	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

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.