



# Full wwPDB X-ray Structure Validation Report ⓘ

May 13, 2020 – 12:55 am BST

PDB ID : 5C5S  
Title : Crystal Structure of human Myosin 9b RhoGAP domain at 2.2 angstrom  
Authors : Yi, F.S.; Ren, J.Q.; Feng, W.  
Deposited on : 2015-06-22  
Resolution : 2.20 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  
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

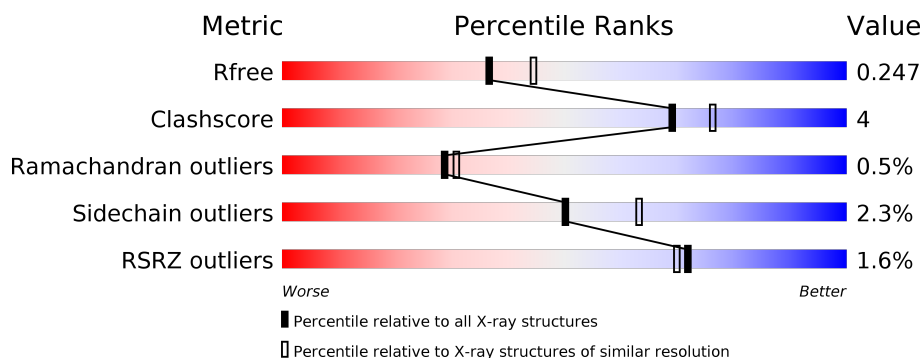
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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  $\leq 5\%$ . 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	232	<div> <div></div> <div> <div></div> <div>84%</div> <div>6%</div> <div>9%</div> </div> </div>
1	B	232	<div> <div>2%</div> <div> <div></div> <div>77%</div> <div>8%</div> <div>13%</div> </div> </div>
1	C	232	<div> <div>2%</div> <div> <div></div> <div>80%</div> <div>6%</div> <div>13%</div> </div> </div>
1	D	232	<div> <div> <div></div> <div>75%</div> <div>12%</div> <div>13%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6824 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-IXb.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	210	Total	C	N	O	S	0	2	0
			1653	1058	283	300	12			
1	B	202	Total	C	N	O	S	0	2	0
			1602	1026	272	293	11			
1	C	202	Total	C	N	O	S	0	4	0
			1610	1031	278	290	11			
1	D	203	Total	C	N	O	S	0	0	0
			1606	1027	274	294	11			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	104	GLY	-	expression tag	UNP Q13459
A	105	PRO	-	expression tag	UNP Q13459
A	106	GLY	-	expression tag	UNP Q13459
A	107	SER	-	expression tag	UNP Q13459
A	108	GLU	-	expression tag	UNP Q13459
A	109	PHE	-	expression tag	UNP Q13459
B	104	GLY	-	expression tag	UNP Q13459
B	105	PRO	-	expression tag	UNP Q13459
B	106	GLY	-	expression tag	UNP Q13459
B	107	SER	-	expression tag	UNP Q13459
B	108	GLU	-	expression tag	UNP Q13459
B	109	PHE	-	expression tag	UNP Q13459
C	104	GLY	-	expression tag	UNP Q13459
C	105	PRO	-	expression tag	UNP Q13459
C	106	GLY	-	expression tag	UNP Q13459
C	107	SER	-	expression tag	UNP Q13459
C	108	GLU	-	expression tag	UNP Q13459
C	109	PHE	-	expression tag	UNP Q13459
D	104	GLY	-	expression tag	UNP Q13459
D	105	PRO	-	expression tag	UNP Q13459
D	106	GLY	-	expression tag	UNP Q13459

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	107	SER	-	expression tag	UNP Q13459
D	108	GLU	-	expression tag	UNP Q13459
D	109	PHE	-	expression tag	UNP Q13459

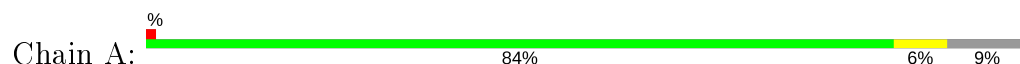
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	134	Total 134	O 134	0	0
2	B	75	Total 75	O 75	0	0
2	C	77	Total 77	O 77	0	0
2	D	67	Total 67	O 67	0	0

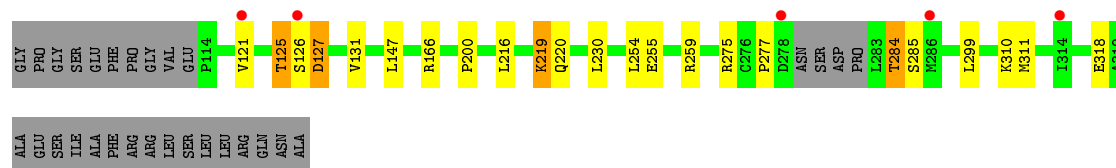
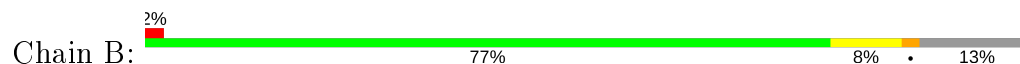
### 3 Residue-property plots [i](#)

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 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.

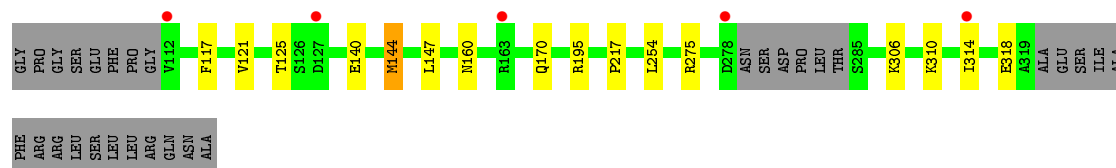
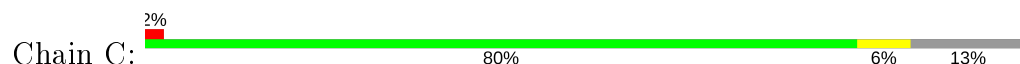
#### • Molecule 1: Unconventional myosin-IXb



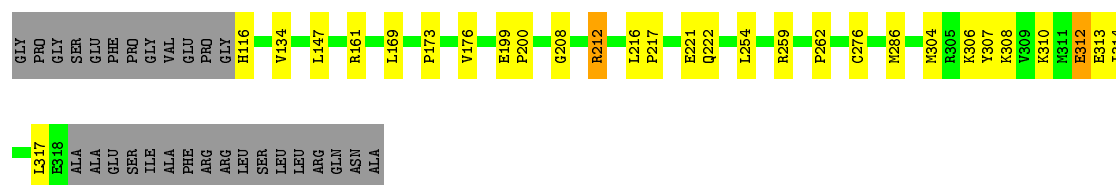
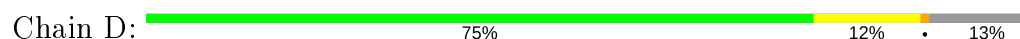
#### • Molecule 1: Unconventional myosin-IXb



#### • Molecule 1: Unconventional myosin-IXb



#### • Molecule 1: Unconventional myosin-IXb



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	77.08 Å 84.98 Å 134.14 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.57 – 2.20 43.48 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.7 (39.57-2.20) 93.6 (43.48-2.20)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.64 (at 2.20 Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.221 , 0.246 0.223 , 0.247	Depositor DCC
$R_{free}$ test set	2000 reflections (4.40%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.0	Xtriage
Anisotropy	0.631	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 36.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	6824	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	52.0	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 35.32 % 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 5.9397e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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	0.23	0/1691	0.40	0/2292
1	B	0.23	0/1637	0.42	0/2214
1	C	0.22	0/1645	0.40	0/2228
1	D	0.22	0/1636	0.40	0/2216
All	All	0.23	0/6609	0.40	0/8950

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	1653	0	1691	10	0
1	B	1602	0	1640	13	0
1	C	1610	0	1640	14	0
1	D	1606	0	1644	19	0
2	A	134	0	0	1	0
2	B	75	0	0	2	0
2	C	77	0	0	1	0
2	D	67	0	0	3	0
All	All	6824	0	6615	49	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 4.

All (49) 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:D:169:LEU:HD12	1:D:176:VAL:HG21	1.79	0.64
1:C:318:GLU:O	1:D:306:LYS:NZ	2.31	0.63
1:D:212:ARG:NH2	2:D:405:HOH:O	2.32	0.62
1:B:230:LEU:HD11	1:B:299:LEU:HD21	1.83	0.61
1:B:147:LEU:HB3	1:B:254:LEU:HD12	1.84	0.60
1:A:147:LEU:HB3	1:A:254:LEU:HD12	1.86	0.57
1:B:284:THR:O	1:B:284:THR:OG1	2.22	0.56
1:D:147:LEU:HB3	1:D:254:LEU:HD12	1.88	0.56
1:D:199:GLU:HG2	1:D:200:PRO:HD2	1.91	0.52
1:D:116:HIS:N	2:D:410:HOH:O	2.42	0.52
1:C:195[A]:ARG:HA	1:C:275[A]:ARG:HH12	1.74	0.52
1:C:310:LYS:HB2	1:D:314:ILE:HD11	1.91	0.51
1:D:286:MET:O	2:D:401:HOH:O	2.19	0.51
1:C:160:ASN:ND2	2:C:405:HOH:O	2.42	0.50
1:B:255[B]:GLU:OE2	1:B:259:ARG:NH1	2.46	0.49
1:D:308:LYS:O	1:D:312:GLU:HG2	2.12	0.49
1:C:310:LYS:HE2	1:D:310:LYS:HE3	1.95	0.48
1:D:116:HIS:HB3	1:D:134:VAL:HG21	1.96	0.47
1:B:259:ARG:NH2	2:B:411:HOH:O	2.47	0.46
1:A:169:LEU:HD23	1:A:176:VAL:HG21	1.96	0.46
1:C:310:LYS:NZ	1:D:313:GLU:OE2	2.35	0.46
1:B:219:LYS:HD2	1:B:220:GLN:HG2	1.97	0.46
1:B:200:PRO:HG3	1:B:275:ARG:HB2	1.97	0.46
1:D:216:LEU:O	1:D:222:GLN:NE2	2.28	0.46
1:A:212:ARG:HD3	1:A:212:ARG:HA	1.70	0.46
1:C:117:PHE:HB2	1:C:170:GLN:HE22	1.81	0.45
1:B:121:VAL:O	1:B:125:THR:HG22	2.16	0.45
1:D:134:VAL:HG22	1:D:173:PRO:HB3	1.98	0.45
1:D:208:GLY:O	1:D:212:ARG:HG2	2.17	0.44
1:D:216:LEU:HD13	1:D:221:GLU:HG3	1.99	0.44
1:D:161:ARG:HD2	1:D:161:ARG:HA	1.79	0.43
1:A:161[A]:ARG:NH1	1:A:164:GLU:OE1	2.43	0.43
1:C:306:LYS:O	1:C:310:LYS:HG3	2.19	0.43
1:B:126:SER:O	1:B:127:ASP:HB2	2.19	0.43
1:A:161[A]:ARG:HD2	1:A:161[A]:ARG:HA	1.83	0.43
1:B:216:LEU:HA	1:B:216:LEU:HD23	1.88	0.42
1:A:301:LYS:NZ	1:C:125[A]:THR:O	2.53	0.42
1:C:147:LEU:HB3	1:C:254:LEU:HD12	2.02	0.42
1:A:200:PRO:HD3	1:A:275:ARG:NH2	2.35	0.42
1:A:116:HIS:HB3	1:A:134:VAL:HG21	2.02	0.42

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:220:GLN:O	2:B:401:HOH:O	2.22	0.41
1:D:262:PRO:HG3	1:D:304:MET:HE1	2.01	0.41
1:A:199:GLU:OE2	2:A:401:HOH:O	2.22	0.41
1:C:140:GLU:O	1:C:144:MET:HB2	2.21	0.40
1:B:311:MET:HB2	1:B:311:MET:HE2	1.96	0.40
1:C:117:PHE:HB2	1:C:170:GLN:NE2	2.37	0.40
1:A:306:LYS:NZ	1:B:318:GLU:O	2.54	0.40
1:C:121:VAL:O	1:C:125[A]:THR:HG23	2.21	0.40
1:C:314:ILE:HG21	1:D:307:TYR:CD2	2.57	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	210/232 (90%)	204 (97%)	6 (3%)	0	100	100
1	B	200/232 (86%)	194 (97%)	4 (2%)	2 (1%)	15	14
1	C	202/232 (87%)	194 (96%)	7 (4%)	1 (0%)	29	31
1	D	201/232 (87%)	194 (96%)	6 (3%)	1 (0%)	29	31
All	All	813/928 (88%)	786 (97%)	23 (3%)	4 (0%)	29	31

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	127	ASP
1	C	217	PRO
1	B	277	PRO
1	D	217	PRO

### 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	181/200 (90%)	178 (98%)	3 (2%)	60	74
1	B	175/200 (88%)	168 (96%)	7 (4%)	31	40
1	C	173/200 (86%)	172 (99%)	1 (1%)	86	93
1	D	177/200 (88%)	172 (97%)	5 (3%)	43	56
All	All	706/800 (88%)	690 (98%)	16 (2%)	50	63

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	154	ARG
1	A	189	VAL
1	A	212	ARG
1	B	125	THR
1	B	131	VAL
1	B	166	ARG
1	B	219	LYS
1	B	284	THR
1	B	285	SER
1	B	310	LYS
1	C	144	MET
1	D	212	ARG
1	D	259	ARG
1	D	276	CYS
1	D	312	GLU
1	D	317	LEU

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

### 5.3.3 RNA ⓘ

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](#)

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	210/232 (90%)	-0.25	3 (1%) 75 73	23, 40, 75, 104	0
1	B	202/232 (87%)	-0.06	5 (2%) 57 55	27, 54, 99, 148	0
1	C	202/232 (87%)	-0.06	5 (2%) 57 55	27, 50, 97, 152	0
1	D	203/232 (87%)	-0.12	0 100 100	30, 50, 90, 108	0
All	All	817/928 (88%)	-0.12	13 (1%) 72 70	23, 49, 93, 152	0

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	127	ASP	6.2
1	C	112	VAL	5.8
1	C	278	ASP	5.0
1	A	109	PHE	2.8
1	B	278	ASP	2.8
1	C	314	ILE	2.8
1	A	110	PRO	2.6
1	B	314	ILE	2.6
1	B	126	SER	2.4
1	C	163	ARG	2.4
1	B	286	MET	2.4
1	B	121	VAL	2.3
1	A	286	MET	2.2

### 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 carbohydrates 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.