



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 10:31 pm BST

PDB ID : 5V6E  
Title : Crystal structure of Myosin VI in complex with GH2 domain of GIPC1  
Authors : Shang, G.; Zhang, X.  
Deposited on : 2017-03-16  
Resolution : 3.51 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

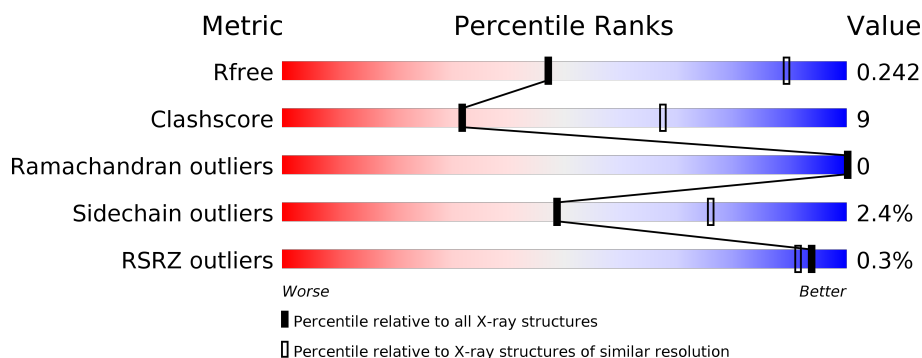
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.51 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	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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	80	
1	C	80	
1	E	80	
1	G	80	
1	I	80	
2	B	49	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	49	<div><div></div><div>84%8%8%</div></div>
2	F	49	<div><div></div><div>71%18%10%</div></div>
2	H	49	<div><div></div><div>82%10%8%</div></div>
2	J	49	<div><div></div><div>78%14%8%</div></div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4858 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 PDZ domain-containing protein GIPC1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	74	Total	C	N	O	S	0	0	0
			589	374	93	120	2			
1	C	74	Total	C	N	O	S	0	0	0
			581	369	92	118	2			
1	E	74	Total	C	N	O	S	0	0	0
			589	374	93	120	2			
1	G	74	Total	C	N	O	S	0	0	0
			589	374	93	120	2			
1	I	74	Total	C	N	O	S	0	0	0
			589	374	93	120	2			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	254	GLY	-	expression tag	UNP Q9Z0G0
A	255	PRO	-	expression tag	UNP Q9Z0G0
A	256	HIS	-	expression tag	UNP Q9Z0G0
A	257	MET	-	expression tag	UNP Q9Z0G0
C	254	GLY	-	expression tag	UNP Q9Z0G0
C	255	PRO	-	expression tag	UNP Q9Z0G0
C	256	HIS	-	expression tag	UNP Q9Z0G0
C	257	MET	-	expression tag	UNP Q9Z0G0
E	254	GLY	-	expression tag	UNP Q9Z0G0
E	255	PRO	-	expression tag	UNP Q9Z0G0
E	256	HIS	-	expression tag	UNP Q9Z0G0
E	257	MET	-	expression tag	UNP Q9Z0G0
G	254	GLY	-	expression tag	UNP Q9Z0G0
G	255	PRO	-	expression tag	UNP Q9Z0G0
G	256	HIS	-	expression tag	UNP Q9Z0G0
G	257	MET	-	expression tag	UNP Q9Z0G0
I	254	GLY	-	expression tag	UNP Q9Z0G0
I	255	PRO	-	expression tag	UNP Q9Z0G0
I	256	HIS	-	expression tag	UNP Q9Z0G0

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
I	257	MET	-	expression tag	UNP Q9Z0G0

- Molecule 2 is a protein called Unconventional myosin-VI.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	47	Total	C	N	O	S	0	0	0
			394	248	74	70	2			
2	D	45	Total	C	N	O	S	0	0	0
			385	243	72	68	2			
2	F	44	Total	C	N	O	S	0	0	0
			376	237	69	68	2			
2	H	45	Total	C	N	O	S	0	0	0
			381	240	71	68	2			
2	J	45	Total	C	N	O	S	0	0	0
			385	243	72	68	2			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1048	GLY	-	expression tag	UNP Q64331
B	1049	PRO	-	expression tag	UNP Q64331
B	1050	GLY	-	expression tag	UNP Q64331
B	1051	SER	-	expression tag	UNP Q64331
D	1048	GLY	-	expression tag	UNP Q64331
D	1049	PRO	-	expression tag	UNP Q64331
D	1050	GLY	-	expression tag	UNP Q64331
D	1051	SER	-	expression tag	UNP Q64331
F	1048	GLY	-	expression tag	UNP Q64331
F	1049	PRO	-	expression tag	UNP Q64331
F	1050	GLY	-	expression tag	UNP Q64331
F	1051	SER	-	expression tag	UNP Q64331
H	1048	GLY	-	expression tag	UNP Q64331
H	1049	PRO	-	expression tag	UNP Q64331
H	1050	GLY	-	expression tag	UNP Q64331
H	1051	SER	-	expression tag	UNP Q64331
J	1048	GLY	-	expression tag	UNP Q64331
J	1049	PRO	-	expression tag	UNP Q64331
J	1050	GLY	-	expression tag	UNP Q64331
J	1051	SER	-	expression tag	UNP Q64331

### 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.

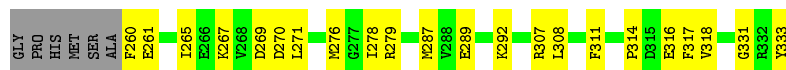
- Molecule 1: PDZ domain-containing protein GIPC1

Chain A: 



- Molecule 1: PDZ domain-containing protein GIPC1

Chain C: 




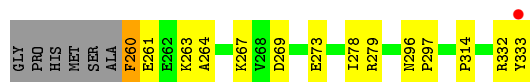
- Molecule 1: PDZ domain-containing protein GIPC1

Chain E: 



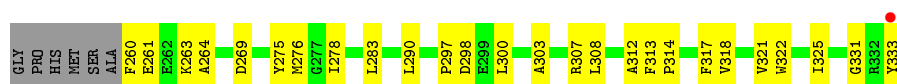
- Molecule 1: PDZ domain-containing protein GIPC1

Chain G: 




- Molecule 1: PDZ domain-containing protein GIPC1

Chain I: 




- Molecule 2: Unconventional myosin-VI

Chain B:  84% 10% . .



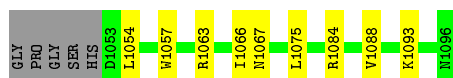
- Molecule 2: Unconventional myosin-VI

Chain D:  84% 8% 8%




- Molecule 2: Unconventional myosin-VI

Chain F:  71% 18% 10%




- Molecule 2: Unconventional myosin-VI

Chain H:  82% 10% 8%



- Molecule 2: Unconventional myosin-VI

Chain J:  78% 14% 8%



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	114.98Å 164.10Å 171.75Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.08 – 3.51 47.08 – 3.51	Depositor EDS
% Data completeness (in resolution range)	90.2 (47.08-3.51) 90.2 (47.08-3.51)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.95 (at 3.48Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.186 , 0.240 0.192 , 0.242	Depositor DCC
$R_{free}$ test set	1875 reflections (10.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	56.4	Xtriage
Anisotropy	0.191	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 47.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.015 for -h,-l,-k	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	4858	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.54% 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	0.37	0/599	0.52	0/806
1	C	0.40	0/591	0.51	0/797
1	E	0.36	0/599	0.50	0/806
1	G	0.28	0/599	0.42	0/806
1	I	0.32	0/599	0.45	0/806
2	B	0.41	0/403	0.54	0/541
2	D	0.30	0/394	0.40	0/529
2	F	0.40	0/384	0.46	0/514
2	H	0.30	0/390	0.41	0/525
2	J	0.29	0/394	0.42	0/529
All	All	0.35	0/4952	0.47	0/6659

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	589	0	559	13	0
1	C	581	0	544	14	0
1	E	589	0	559	14	0
1	G	589	0	559	13	0
1	I	589	0	559	14	0
2	B	394	0	382	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	385	0	377	5	0
2	F	376	0	370	9	0
2	H	381	0	366	7	0
2	J	385	0	377	6	0
All	All	4858	0	4652	85	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 9.

The worst 5 of 85 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:273:GLU:OE1	1:A:279:ARG:NH1	2.20	0.73
1:G:273:GLU:OE1	1:G:279:ARG:NH1	2.22	0.72
1:E:311:PHE:HA	2:F:1088:VAL:HG11	1.72	0.71
1:A:270:ASP:OD1	1:A:279:ARG:NH2	2.24	0.70
2:D:1075:LEU:HD21	2:D:1079:ARG:HH11	1.60	0.66

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	72/80 (90%)	65 (90%)	7 (10%)	0	100	100
1	C	72/80 (90%)	67 (93%)	5 (7%)	0	100	100
1	E	72/80 (90%)	69 (96%)	3 (4%)	0	100	100
1	G	72/80 (90%)	70 (97%)	2 (3%)	0	100	100
1	I	72/80 (90%)	68 (94%)	4 (6%)	0	100	100
2	B	45/49 (92%)	41 (91%)	4 (9%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	43/49 (88%)	39 (91%)	4 (9%)	0	100	100
2	F	42/49 (86%)	41 (98%)	1 (2%)	0	100	100
2	H	43/49 (88%)	41 (95%)	2 (5%)	0	100	100
2	J	43/49 (88%)	41 (95%)	2 (5%)	0	100	100
All	All	576/645 (89%)	542 (94%)	34 (6%)	0	100	100

There are no Ramachandran outliers to report.

### 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	60/64 (94%)	56 (93%)	4 (7%)	16	48
1	C	58/64 (91%)	57 (98%)	1 (2%)	60	82
1	E	60/64 (94%)	58 (97%)	2 (3%)	38	68
1	G	60/64 (94%)	59 (98%)	1 (2%)	60	82
1	I	60/64 (94%)	57 (95%)	3 (5%)	24	58
2	B	41/43 (95%)	40 (98%)	1 (2%)	49	76
2	D	41/43 (95%)	41 (100%)	0	100	100
2	F	40/43 (93%)	40 (100%)	0	100	100
2	H	40/43 (93%)	40 (100%)	0	100	100
2	J	41/43 (95%)	41 (100%)	0	100	100
All	All	501/535 (94%)	489 (98%)	12 (2%)	49	76

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	269	ASP
1	E	260	PHE
1	I	269	ASP
2	B	1065	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	G	260	PHE

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

Mol	Chain	Res	Type
2	F	1067	ASN
2	H	1067	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 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	74/80 (92%)	-0.30	0	100 100	21, 40, 88, 121	0
1	C	74/80 (92%)	-0.53	0	100 100	25, 44, 90, 106	0
1	E	74/80 (92%)	-0.54	0	100 100	39, 53, 85, 112	0
1	G	74/80 (92%)	-0.08	1 (1%)	75 69	79, 98, 125, 133	0
1	I	74/80 (92%)	-0.17	1 (1%)	75 69	43, 82, 110, 142	0
2	B	47/49 (95%)	-0.52	0	100 100	11, 29, 72, 94	0
2	D	45/49 (91%)	-0.45	0	100 100	49, 67, 106, 130	0
2	F	44/49 (89%)	-0.23	0	100 100	32, 67, 96, 111	0
2	H	45/49 (91%)	-0.02	0	100 100	76, 94, 119, 162	0
2	J	45/49 (91%)	0.24	0	100 100	70, 98, 123, 131	0
All	All	596/645 (92%)	-0.28	2 (0%)	94 91	11, 70, 114, 162	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	I	333	TYR	2.3
1	G	333	TYR	2.3

### 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

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.