



# Full wwPDB X-ray Structure Validation Report ⓘ

Jul 24, 2023 – 08:13 PM JST

PDB ID : 7YKF  
Title : Crystal structure of MAGI2 PDZ0-GK/pEphexin4 complex  
Authors : Zhang, M.; Lin, L.; Zhu, J.  
Deposited on : 2022-07-22  
Resolution : 2.28 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.34  
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) : 2.34

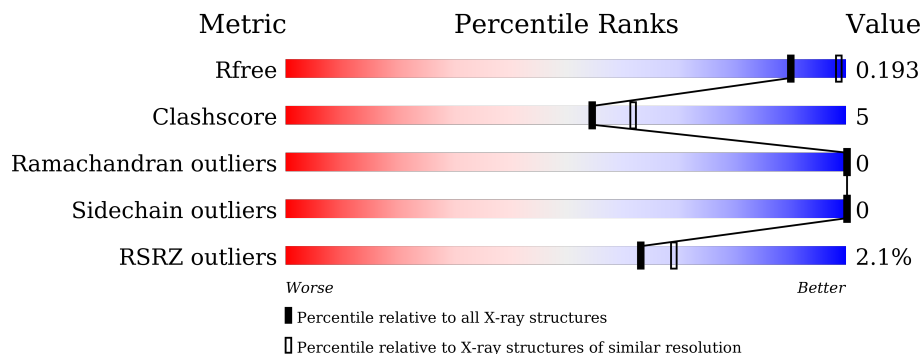
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.28 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	6980 (2.30-2.26)
Clashscore	141614	7711 (2.30-2.26)
Ramachandran outliers	138981	7597 (2.30-2.26)
Sidechain outliers	138945	7598 (2.30-2.26)
RSRZ outliers	127900	6849 (2.30-2.26)

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	234	 2% 69% 9% 22%
1	C	234	 % 69% 9% 22%
2	B	12	 50% 33% 8% 8%
2	D	12	 8% 58% 42%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 3021 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 Membrane-associated guanylate kinase, WW and PDZ domain-containing protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	183	1389	882	235	268	4	0	0	0
1	C	183	1391	884	233	270	4	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	GLY	-	expression tag	UNP Q9WVQ1
A	-3	PRO	-	expression tag	UNP Q9WVQ1
A	-2	GLY	-	expression tag	UNP Q9WVQ1
A	-1	SER	-	expression tag	UNP Q9WVQ1
C	-4	GLY	-	expression tag	UNP Q9WVQ1
C	-3	PRO	-	expression tag	UNP Q9WVQ1
C	-2	GLY	-	expression tag	UNP Q9WVQ1
C	-1	SER	-	expression tag	UNP Q9WVQ1

- Molecule 2 is a protein called Ephexin4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	N	O	P				S
2	B	11	95	54	20	19	1	1	0	0	0
2	D	12	104	60	22	20	1	1	0	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	18	Total	O	0	0
			18	18		

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
3	B	4	Total O 4 4	0	0
3	C	19	Total O 19 19	0	0
3	D	1	Total O 1 1	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	40.21Å 55.62Å 61.07Å 65.31° 76.62° 71.56°	Depositor
Resolution (Å)	29.67 – 2.28 29.67 – 2.28	Depositor EDS
% Data completeness (in resolution range)	96.7 (29.67-2.28) 96.9 (29.67-2.28)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.75 (at 2.29Å)	Xtrriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, $R_{free}$	0.186 , 0.236 0.188 , 0.193	Depositor DCC
$R_{free}$ test set	1012 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	38.0	Xtrriage
Anisotropy	0.783	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 41.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	3021	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	57.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.69% 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: SEP

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	0.49	1/1420 (0.1%)	0.68	1/1934 (0.1%)
1	C	0.46	1/1422 (0.1%)	0.78	4/1936 (0.2%)
2	B	0.35	0/84	0.74	0/109
2	D	0.43	0/93	0.63	0/120
All	All	0.47	2/3019 (0.1%)	0.73	5/4099 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	129	VAL	CB-CG2	6.24	1.66	1.52
1	A	71	ARG	CZ-NH1	5.38	1.40	1.33

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	100	ARG	CB-CG-CD	-12.20	79.88	111.60
1	C	100	ARG	CA-CB-CG	10.10	135.62	113.40
1	A	121	ARG	NE-CZ-NH2	-7.65	116.47	120.30
1	C	100	ARG	CG-CD-NE	6.68	125.83	111.80
1	C	129	VAL	CG1-CB-CG2	-5.94	101.40	110.90

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	1389	0	1328	15	0
1	C	1391	0	1332	12	0
2	B	95	0	88	3	1
2	D	104	0	101	2	1
3	A	18	0	0	2	0
3	B	4	0	0	1	0
3	C	19	0	0	1	0
3	D	1	0	0	0	0
All	All	3021	0	2849	30	1

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 5.

All (30) 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:22:GLU:OE2	3:A:301:HOH:O	1.94	0.85
1:A:37:GLU:O	3:A:301:HOH:O	1.95	0.84
1:A:103:LEU:HD22	1:A:121:ARG:HE	1.55	0.72
2:B:233:ARG:O	2:B:238:ARG:NH1	2.26	0.69
2:B:236:SEP:O1P	3:B:301:HOH:O	2.15	0.64
1:C:2:TRP:HD1	1:C:94:ILE:HG12	1.63	0.64
1:C:121:ARG:NH1	1:C:165:GLU:OE2	2.33	0.62
1:A:60:VAL:HG11	1:A:77:ILE:HD13	1.82	0.61
1:C:14:ASN:HB2	1:C:15:PRO:HD2	1.86	0.57
1:A:150:THR:HG23	1:A:153:GLU:H	1.71	0.56
1:A:103:LEU:HD22	1:A:121:ARG:NE	2.20	0.54
1:C:28:GLU:HG3	1:C:181:ALA:HB2	1.90	0.54
1:A:73:VAL:O	1:A:77:ILE:HG12	2.08	0.53
2:D:232:LEU:HD12	2:D:233:ARG:H	1.73	0.53
1:A:168:THR:HG22	1:A:173:TYR:CD1	2.44	0.52
1:C:2:TRP:CD1	1:C:94:ILE:HG12	2.44	0.49
1:A:52:VAL:HG22	1:A:55:GLU:HG3	1.97	0.47
1:A:68:LEU:HB3	1:A:72:ASP:HB2	1.97	0.47
1:C:68:LEU:HB3	1:C:72:ASP:HB2	1.97	0.47
1:A:69:THR:O	1:A:73:VAL:HG23	2.15	0.46
1:A:167:GLY:HA3	2:B:240:ALA:O	2.17	0.44
1:A:10:VAL:HG22	1:A:85:ARG:HG2	2.00	0.44
1:A:168:THR:HG22	1:A:173:TYR:HD1	1.84	0.43
1:C:33:PRO:HD3	3:C:304:HOH:O	2.18	0.43

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:167:GLY:HA3	2:D:240:ALA:O	2.19	0.42
1:A:133:THR:HA	1:A:149:ILE:O	2.19	0.42
1:C:133:THR:HA	1:C:149:ILE:O	2.20	0.41
1:C:69:THR:O	1:C:73:VAL:HG23	2.20	0.41
1:C:59:GLU:HB2	1:C:87:LYS:HB2	2.03	0.41
1:C:65:VAL:HB	1:C:73:VAL:HG22	2.03	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:231:ASN:ND2	2:D:231:ASN:OD1[1_545]	2.06	0.14

## 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	181/234 (77%)	180 (99%)	1 (1%)	0	100	100
1	C	181/234 (77%)	181 (100%)	0	0	100	100
2	B	8/12 (67%)	8 (100%)	0	0	100	100
2	D	9/12 (75%)	9 (100%)	0	0	100	100
All	All	379/492 (77%)	378 (100%)	1 (0%)	0	100	100

There are no Ramachandran outliers to report.

### 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	146/202 (72%)	146 (100%)	0	100	100
1	C	147/202 (73%)	147 (100%)	0	100	100
2	B	8/9 (89%)	8 (100%)	0	100	100
2	D	9/9 (100%)	9 (100%)	0	100	100
All	All	310/422 (74%)	310 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues 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
2	SEP	D	236	2	8,9,10	1.54	1 (12%)	8,12,14	1.52	2 (25%)
2	SEP	B	236	2	8,9,10	1.50	1 (12%)	8,12,14	1.29	1 (12%)

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	SEP	D	236	2	-	0/5/8/10	-
2	SEP	B	236	2	-	0/5/8/10	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	236	SEP	P-O1P	3.32	1.61	1.50
2	B	236	SEP	P-O1P	3.09	1.60	1.50

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	236	SEP	P-OG-CB	-3.20	109.49	118.30
2	B	236	SEP	P-OG-CB	-2.78	110.64	118.30
2	D	236	SEP	OG-CB-CA	2.58	110.65	108.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	236	SEP	1	0

## 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	183/234 (78%)	0.03	5 (2%) 54 60	37, 52, 84, 104	0
1	C	183/234 (78%)	0.14	2 (1%) 80 84	39, 54, 82, 99	0
2	B	10/12 (83%)	0.21	0 100 100	53, 66, 72, 84	0
2	D	11/12 (91%)	0.58	1 (9%) 9 11	52, 64, 80, 87	0
All	All	387/492 (78%)	0.10	8 (2%) 63 69	37, 54, 84, 104	0

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	0	SER	4.9
1	C	0	SER	3.5
2	D	242	LYS	3.2
1	A	171	ASP	3.1
1	A	182	GLU	2.8
1	C	110	GLY	2.3
1	A	169	TYR	2.1
1	A	109	LYS	2.0

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	SEP	D	236	10/11	0.97	0.10	48,50,60,61	0
2	SEP	B	236	10/11	0.99	0.12	47,59,64,65	0

### 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.