



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 10:11 AM GMT

PDB ID : 3L4F  
Title : Crystal Structure of betaPIX Coiled-Coil Domain and Shank PDZ Complex  
Authors : Im, Y.J.; Kang, G.B.; Lee, J.H.; Song, H.E.; Park, K.R.; Kim, E.; Song, W.K.;  
Park, D.; Eom, S.H.  
Deposited on : 2009-12-19  
Resolution : 2.80 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  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

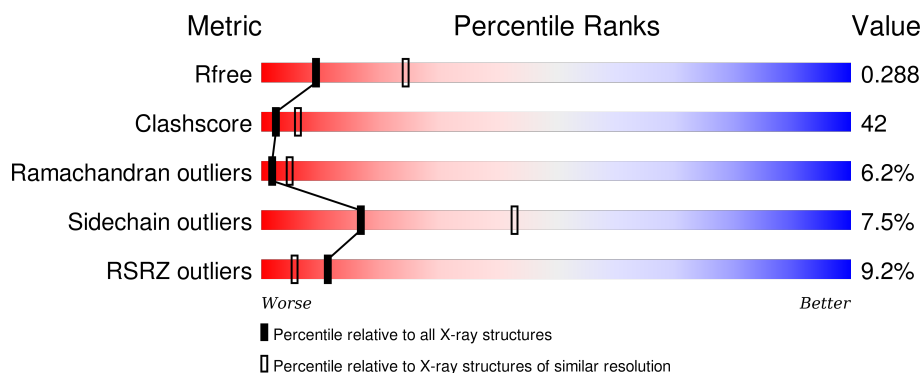
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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))
$R_{free}$	91344	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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. 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	61	<div> <div>2%</div> <div> <div></div> <div>43%</div> <div>48%</div> <div>10%</div> </div> </div>
1	B	61	<div> <div>7%</div> <div> <div></div> <div>46%</div> <div>44%</div> <div>10%</div> </div> </div>
1	C	61	<div> <div>5%</div> <div> <div></div> <div>41%</div> <div>36%</div> <div>10%</div> <div>13%</div> </div> </div>
2	D	132	<div> <div>14%</div> <div> <div></div> <div>23%</div> <div>50%</div> <div>8%</div> <div>19%</div> </div> </div>

## 2 Entry composition

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 Rho guanine nucleotide exchange factor 7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	61	Total	C	N	O	S	0	0	0
			511	315	91	102	3			
1	B	61	Total	C	N	O	S	0	0	0
			511	315	91	102	3			
1	C	53	Total	C	N	O	S	0	0	0
			444	273	81	87	3			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	586	MET	-	EXPRESSION TAG	UNP O55043
B	586	MET	-	EXPRESSION TAG	UNP O55043
C	586	MET	-	EXPRESSION TAG	UNP O55043

- Molecule 2 is a protein called SH3 and multiple ankyrin repeat domains protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	107	Total	C	N	O	S	0	0	0
			823	523	145	151	4			

There are 19 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	634	MET	-	EXPRESSION TAG	UNP Q9WV48
D	635	GLY	-	EXPRESSION TAG	UNP Q9WV48
D	636	SER	-	EXPRESSION TAG	UNP Q9WV48
D	637	SER	-	EXPRESSION TAG	UNP Q9WV48
D	638	HIS	-	EXPRESSION TAG	UNP Q9WV48
D	639	HIS	-	EXPRESSION TAG	UNP Q9WV48
D	640	HIS	-	EXPRESSION TAG	UNP Q9WV48
D	641	HIS	-	EXPRESSION TAG	UNP Q9WV48
D	642	HIS	-	EXPRESSION TAG	UNP Q9WV48

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	643	HIS	-	EXPRESSION TAG	UNP Q9WV48
D	644	SER	-	EXPRESSION TAG	UNP Q9WV48
D	645	GLN	-	EXPRESSION TAG	UNP Q9WV48
D	646	ASP	-	EXPRESSION TAG	UNP Q9WV48
D	647	PRO	-	EXPRESSION TAG	UNP Q9WV48
D	648	LEU	-	EXPRESSION TAG	UNP Q9WV48
D	649	VAL	-	EXPRESSION TAG	UNP Q9WV48
D	650	PRO	-	EXPRESSION TAG	UNP Q9WV48
D	651	ARG	-	EXPRESSION TAG	UNP Q9WV48
D	652	GLY	-	EXPRESSION TAG	UNP Q9WV48

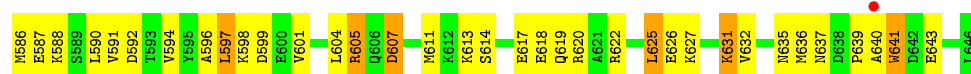
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	3	Total O 3 3	0	0
3	B	7	Total O 7 7	0	0
3	C	2	Total O 2 2	0	0
3	D	8	Total O 8 8	0	0

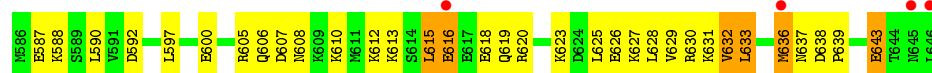
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

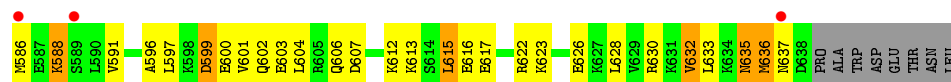
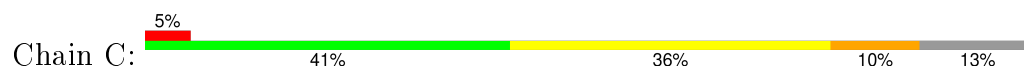
- Molecule 1: Rho guanine nucleotide exchange factor 7



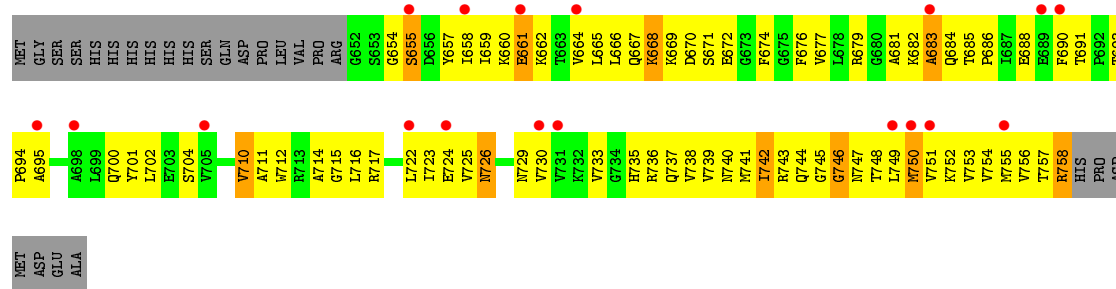
- Molecule 1: Rho guanine nucleotide exchange factor 7



- Molecule 1: Rho guanine nucleotide exchange factor 7



- Molecule 2: SH3 and multiple ankyrin repeat domains protein 1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32 1 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	47.67Å 47.67Å 263.16Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	41.28 – 2.80 41.28 – 2.80	Depositor EDS
% Data completeness (in resolution range)	96.5 (41.28-2.80) 95.8 (41.28-2.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.22 (at 2.81Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.272 , 0.307 0.276 , 0.288	Depositor DCC
$R_{free}$ test set	408 reflections (4.83%)	DCC
Wilson B-factor (Å <sup>2</sup> )	84.8	Xtriage
Anisotropy	0.313	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 106.6	EDS
Estimated twinning fraction	0.079 for -h,-k,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	1 of 8448 reflections (0.012%)	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	2309	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	91.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.71% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.39	0/514	0.64	0/683
1	B	0.38	0/514	0.67	0/683
1	C	0.38	0/444	0.64	0/586
2	D	0.37	0/836	0.61	0/1126
All	All	0.38	0/2308	0.63	0/3078

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	511	0	526	57	0
1	B	511	0	526	39	0
1	C	444	0	470	41	0
2	D	823	0	842	98	0
3	A	3	0	0	3	0
3	B	7	0	0	0	0
3	C	2	0	0	0	0
3	D	8	0	0	2	0
All	All	2309	0	2364	197	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 42.

The worst 5 of 197 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:B:627:LYS:O	1:B:630:ARG:HG2	1.33	1.27
2:D:660:LYS:HD3	2:D:661:GLU:H	1.04	1.13
2:D:661:GLU:HA	2:D:754:VAL:HG12	1.41	1.00
2:D:660:LYS:CD	2:D:661:GLU:H	1.81	0.93
1:C:588:LYS:H	1:C:588:LYS:HD3	1.33	0.91

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	59/61 (97%)	53 (90%)	5 (8%)	1 (2%)	11	36
1	B	59/61 (97%)	42 (71%)	13 (22%)	4 (7%)	1	4
1	C	51/61 (84%)	42 (82%)	6 (12%)	3 (6%)	2	5
2	D	105/132 (80%)	72 (69%)	24 (23%)	9 (9%)	1	2
All	All	274/315 (87%)	209 (76%)	48 (18%)	17 (6%)	2	5

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	587	GLU
1	B	636	MET
1	B	643	GLU
2	D	655	SER
2	D	659	ILE



### 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	58/58 (100%)	52 (90%)	6 (10%)	9	26
1	B	58/58 (100%)	52 (90%)	6 (10%)	9	26
1	C	51/58 (88%)	47 (92%)	4 (8%)	16	41
2	D	88/111 (79%)	85 (97%)	3 (3%)	44	78
All	All	255/285 (90%)	236 (92%)	19 (8%)	17	43

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	615	LEU
1	B	620	ARG
1	C	615	LEU
1	B	608	ASN
2	D	668	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	635	ASN
2	D	667	GLN
2	D	737	GLN
1	B	608	ASN
2	D	726	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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	61/61 (100%)	0.15	1 (1%) 74 66	46, 70, 99, 129	0
1	B	61/61 (100%)	0.46	4 (6%) 22 13	45, 74, 156, 170	0
1	C	53/61 (86%)	0.28	3 (5%) 27 17	47, 77, 140, 145	0
2	D	107/132 (81%)	0.87	18 (16%) 2 1	50, 100, 146, 170	0
All	All	282/315 (89%)	0.51	26 (9%) 11 5	45, 85, 146, 170	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	689	GLU	4.5
1	C	589	SER	4.4
2	D	683	ALA	4.2
2	D	751	VAL	4.0
2	D	750	MET	3.7

### 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.