



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 29, 2016 – 05:14 PM EST

PDB ID : 1VF6  
Title : 2.1 Angstrom crystal structure of the PALS-1-L27N and PATJ L27 heterodimer complex  
Authors : Li, Y.; Lavie, A.; Margolis, B.; Karnak, D.  
Deposited on : 2004-04-09  
Resolution : 2.10 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	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20028320
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)	:	rb-20028320

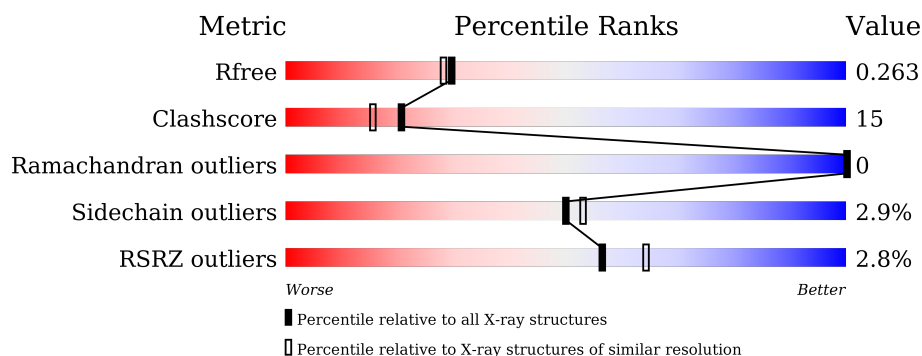
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	3939 (2.10-2.10)
Clashscore	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)
RSRZ outliers	91569	3948 (2.10-2.10)

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	83	<div> <div> <div>6%</div> <div>41%</div> <div>28%</div> <div>30%</div> </div> </div>
1	B	83	<div> <div> <div>53%</div> <div>19%</div> <div>28%</div> </div> </div>
2	C	72	<div> <div> <div>58%</div> <div>13%</div> <div>29%</div> </div> </div>
2	D	72	<div> <div> <div>6%</div> <div>42%</div> <div>24%</div> <div>33%</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 1867 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 PALS1-associated tight junction protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	58	Total	C	N	O	S	0	0	0
			476	304	81	89	2			
1	B	60	Total	C	N	O	S	0	0	0
			493	315	84	92	2			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	8	MET	-	INITIATING METHIONINE	UNP Q8NI35
B	8	MET	-	INITIATING METHIONINE	UNP Q8NI35

- Molecule 2 is a protein called MAGUK p55 subfamily member 5.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	51	Total	C	N	O	0	0	0
			414	258	72	84			
2	D	48	Total	C	N	O	0	0	0
			390	244	68	78			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	109	MET	-	EXPRESSION TAG	UNP Q9JLB2
C	110	GLY	-	EXPRESSION TAG	UNP Q9JLB2
C	111	SER	-	EXPRESSION TAG	UNP Q9JLB2
C	112	SER	-	EXPRESSION TAG	UNP Q9JLB2
C	113	HIS	-	EXPRESSION TAG	UNP Q9JLB2
C	114	HIS	-	EXPRESSION TAG	UNP Q9JLB2
C	115	HIS	-	EXPRESSION TAG	UNP Q9JLB2
C	116	HIS	-	EXPRESSION TAG	UNP Q9JLB2
C	117	HIS	-	EXPRESSION TAG	UNP Q9JLB2
C	118	HIS	-	EXPRESSION TAG	UNP Q9JLB2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	119	SER	-	EXPRESSION TAG	UNP Q9JLB2
C	120	GLN	-	EXPRESSION TAG	UNP Q9JLB2
C	121	ASP	-	EXPRESSION TAG	UNP Q9JLB2
C	122	PRO	-	EXPRESSION TAG	UNP Q9JLB2
D	109	MET	-	EXPRESSION TAG	UNP Q9JLB2
D	110	GLY	-	EXPRESSION TAG	UNP Q9JLB2
D	111	SER	-	EXPRESSION TAG	UNP Q9JLB2
D	112	SER	-	EXPRESSION TAG	UNP Q9JLB2
D	113	HIS	-	EXPRESSION TAG	UNP Q9JLB2
D	114	HIS	-	EXPRESSION TAG	UNP Q9JLB2
D	115	HIS	-	EXPRESSION TAG	UNP Q9JLB2
D	116	HIS	-	EXPRESSION TAG	UNP Q9JLB2
D	117	HIS	-	EXPRESSION TAG	UNP Q9JLB2
D	118	HIS	-	EXPRESSION TAG	UNP Q9JLB2
D	119	SER	-	EXPRESSION TAG	UNP Q9JLB2
D	120	GLN	-	EXPRESSION TAG	UNP Q9JLB2
D	121	ASP	-	EXPRESSION TAG	UNP Q9JLB2
D	122	PRO	-	EXPRESSION TAG	UNP Q9JLB2

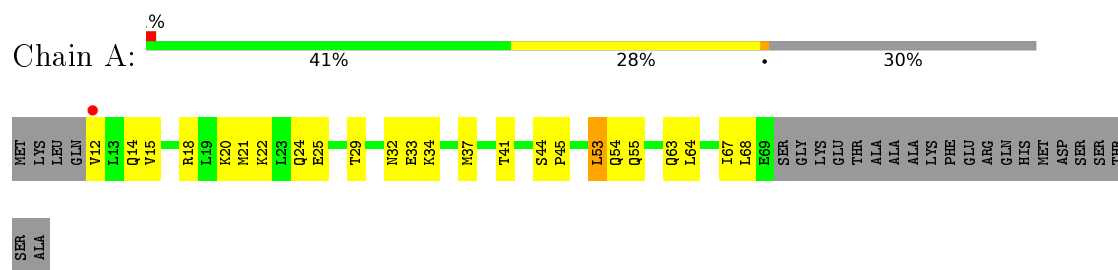
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	31	Total O 31 31	0	0
3	B	21	Total O 21 21	0	0
3	C	30	Total O 30 30	0	0
3	D	12	Total O 12 12	0	0

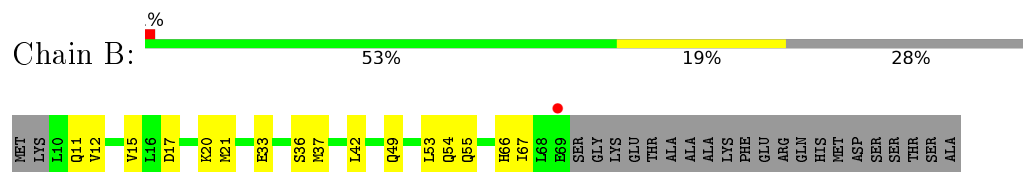
### 3 Residue-property plots [i](#)

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.

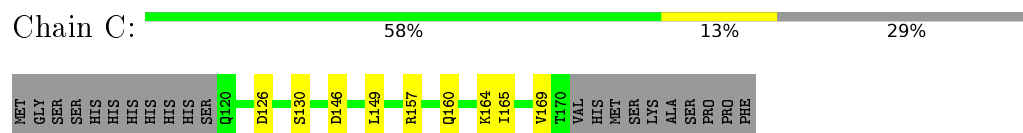
- Molecule 1: PALS1-associated tight junction protein



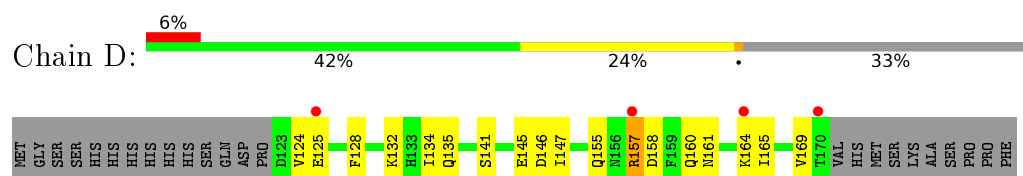
- Molecule 1: PALS1-associated tight junction protein



- Molecule 2: MAGUK p55 subfamily member 5



- Molecule 2: MAGUK p55 subfamily member 5



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	111.06Å 111.06Å 193.69Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 2.10 43.25 – 2.10	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-2.10) 94.3 (43.25-2.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.20 (at 2.10Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.242 , 0.262 0.242 , 0.263	Depositor DCC
$R_{free}$ test set	1235 reflections (4.83%)	DCC
Wilson B-factor (Å <sup>2</sup> )	42.2	Xtriage
Anisotropy	0.248	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 73.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	1867	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.96% 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	0.33	0/480	0.51	0/641
1	B	0.34	0/497	0.49	0/664
2	C	0.37	0/420	0.48	0/569
2	D	0.34	0/395	0.45	0/534
All	All	0.35	0/1792	0.49	0/2408

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	476	0	506	23	0
1	B	493	0	525	17	0
2	C	414	0	398	10	0
2	D	390	0	379	15	0
3	A	31	0	0	5	0
3	B	21	0	0	3	0
3	C	30	0	0	1	0
3	D	12	0	0	0	0
All	All	1867	0	1808	52	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 15.

The worst 5 of 52 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:68:LEU:HD22	2:D:169:VAL:HG21	1.61	0.82
1:A:20:LYS:O	1:A:24:GLN:HG2	1.84	0.77
1:A:15:VAL:HG13	2:C:146:ASP:HB3	1.68	0.76
1:A:53:LEU:HD12	1:B:67:ILE:HD12	1.70	0.74
1:B:11:GLN:H	1:B:11:GLN:NE2	1.86	0.72

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	56/83 (68%)	54 (96%)	2 (4%)	0	100	100
1	B	58/83 (70%)	58 (100%)	0	0	100	100
2	C	49/72 (68%)	49 (100%)	0	0	100	100
2	D	46/72 (64%)	46 (100%)	0	0	100	100
All	All	209/310 (67%)	207 (99%)	2 (1%)	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	56/76 (74%)	54 (96%)	2 (4%)	42	43
1	B	58/76 (76%)	56 (97%)	2 (3%)	44	45
2	C	49/68 (72%)	49 (100%)	0	100	100
2	D	46/68 (68%)	44 (96%)	2 (4%)	35	34
All	All	209/288 (73%)	203 (97%)	6 (3%)	50	53

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	17	ASP
2	D	158	ASP
1	B	55	GLN
1	A	64	LEU
2	D	157	ARG

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	A	63	GLN
1	A	65	ASN
1	B	55	GLN
1	A	59	GLN
1	B	54	GLN

### 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 ⓘ

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	58/83 (69%)	0.37	1 (1%) 73 78	36, 51, 67, 82	0
1	B	60/83 (72%)	0.34	1 (1%) 73 78	37, 46, 66, 78	0
2	C	51/72 (70%)	0.17	0 100 100	34, 44, 65, 74	0
2	D	48/72 (66%)	0.67	4 (8%) 14 19	42, 57, 74, 79	0
All	All	217/310 (70%)	0.38	6 (2%) 56 64	34, 49, 71, 82	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	164	LYS	2.6
1	B	69	GLU	2.4
2	D	125	GLU	2.3
2	D	170	THR	2.2
2	D	157	ARG	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.