



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 11:15 pm BST

PDB ID : 4OHL  
Title : LEOPARD Syndrome-Associated SHP2/T468M mutant  
Authors : Yu, Z.H.; Zhang, R.Y.; Walls, C.D.; Chen, L.; Zhang, S.; Wu, L.; Wang, L.;  
Liu, S.; Zhang, Z.Y.  
Deposited on : 2014-01-17  
Resolution : 2.40 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

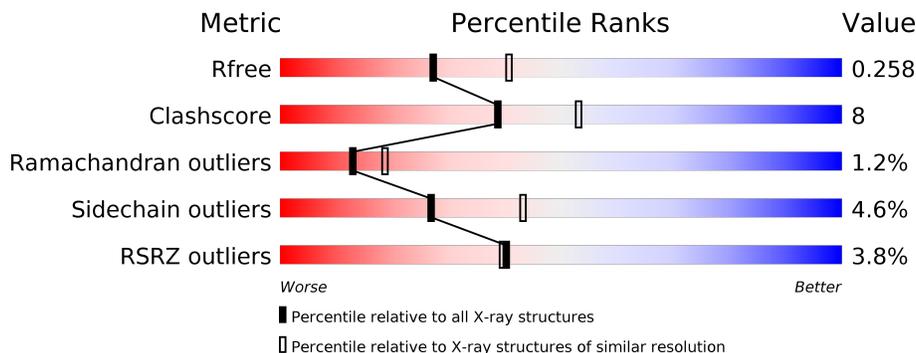
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	536	
1	B	536	

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 8154 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 Tyrosine-protein phosphatase non-receptor type 11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	484	3912	2465	698	730	19	0	0	0
1	B	485	3912	2464	699	730	19	0	0	1

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	GLN	DELETION	UNP Q06124
A	?	-	ALA	DELETION	UNP Q06124
A	?	-	LEU	DELETION	UNP Q06124
A	?	-	LEU	DELETION	UNP Q06124
A	468	MET	THR	ENGINEERED MUTATION	UNP Q06124
A	529	LEU	-	EXPRESSION TAG	UNP Q06124
A	530	GLU	-	EXPRESSION TAG	UNP Q06124
A	531	HIS	-	EXPRESSION TAG	UNP Q06124
A	532	HIS	-	EXPRESSION TAG	UNP Q06124
A	533	HIS	-	EXPRESSION TAG	UNP Q06124
A	534	HIS	-	EXPRESSION TAG	UNP Q06124
A	535	HIS	-	EXPRESSION TAG	UNP Q06124
A	536	HIS	-	EXPRESSION TAG	UNP Q06124
B	?	-	GLN	DELETION	UNP Q06124
B	?	-	ALA	DELETION	UNP Q06124
B	?	-	LEU	DELETION	UNP Q06124
B	?	-	LEU	DELETION	UNP Q06124
B	468	MET	THR	ENGINEERED MUTATION	UNP Q06124
B	529	LEU	-	EXPRESSION TAG	UNP Q06124
B	530	GLU	-	EXPRESSION TAG	UNP Q06124
B	531	HIS	-	EXPRESSION TAG	UNP Q06124
B	532	HIS	-	EXPRESSION TAG	UNP Q06124
B	533	HIS	-	EXPRESSION TAG	UNP Q06124
B	534	HIS	-	EXPRESSION TAG	UNP Q06124
B	535	HIS	-	EXPRESSION TAG	UNP Q06124

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	536	HIS	-	EXPRESSION TAG	UNP Q06124

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	183	Total O 183 183	0	0
2	B	147	Total O 147 147	0	0



Q526  
ARG  
ARG  
LEU  
GLU  
HIS  
HIS  
HIS  
HIS  
HIS  
HIS

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	45.34Å 209.26Å 56.08Å 90.00° 95.89° 90.00°	Depositor
Resolution (Å)	33.01 – 2.40 33.01 – 2.38	Depositor EDS
% Data completeness (in resolution range)	94.5 (33.01-2.40) 93.1 (33.01-2.38)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.61 (at 2.36Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, $R_{free}$	0.191 , 0.259 0.189 , 0.258	Depositor DCC
$R_{free}$ test set	1889 reflections (4.88%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	34.8	Xtrriage
Anisotropy	0.491	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 64.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	8154	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	41.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.73% 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	0.44	0/3993	0.58	0/5385
1	B	0.42	0/3993	0.58	1/5385 (0.0%)
All	All	0.43	0/7986	0.58	1/10770 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	501	ARG	NE-CZ-NH1	5.37	122.99	120.30

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	3912	0	3847	57	0
1	B	3912	0	3846	60	0
2	A	183	0	0	10	0
2	B	147	0	0	9	0
All	All	8154	0	7693	117	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 8.

The worst 5 of 117 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:205:THR:HG23	1:B:206:LEU:HD22	1.70	0.73
1:B:94:ASP:OD1	1:B:94:ASP:N	2.25	0.68
1:B:196:HIS:HA	1:B:199:LYS:HE2	1.74	0.68
1:B:390:GLU:HG2	1:B:399:ARG:HG2	1.77	0.66
1:A:399:ARG:NE	2:A:685:HOH:O	2.29	0.65

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	474/536 (88%)	445 (94%)	21 (4%)	8 (2%)	9	11
1	B	475/536 (89%)	447 (94%)	25 (5%)	3 (1%)	25	36
All	All	949/1072 (88%)	892 (94%)	46 (5%)	11 (1%)	13	19

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	206	LEU
1	B	92	ASN
1	A	116	HIS
1	A	36	SER
1	A	38	PRO

### 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	428/478 (90%)	404 (94%)	24 (6%)	21	34
1	B	428/478 (90%)	413 (96%)	15 (4%)	36	55
All	All	856/956 (90%)	817 (95%)	39 (5%)	27	43

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	362	ARG
1	A	480	ARG
1	B	480	ARG
1	A	391	SER
1	A	413	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	114	HIS
1	A	256	GLN
1	B	211	GLN
1	B	214	GLN
1	B	287	HIS

### 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	484/536 (90%)	0.01	14 (2%)	51 50	20, 39, 61, 82	0
1	B	485/536 (90%)	0.05	23 (4%)	31 30	19, 43, 64, 87	0
All	All	969/1072 (90%)	0.03	37 (3%)	40 39	19, 40, 63, 87	0

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	93	GLY	7.0
1	A	92	ASN	6.0
1	B	116	HIS	5.4
1	B	93	GLY	5.2
1	A	116	HIS	4.6

### 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.