



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 6, 2026 – 06:00 AM UTC

PDB ID : 2QCJ / pdb_00002qcj
Title : Native Structure of Lyp
Authors : Sun, J.P.; Yu, X.; Zhang, Z.Y.
Deposited on : 2007-06-19
Resolution : 3.00 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

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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

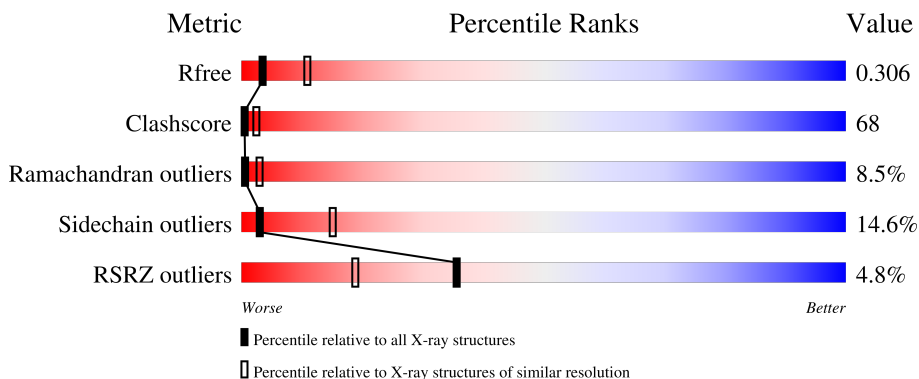
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

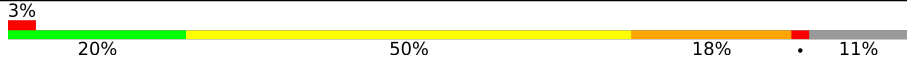
The reported resolution of this entry is 3.00 Å.

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}	180053	2672 (3.00-3.00)
Clashscore	190562	2977 (3.00-3.00)
Ramachandran outliers	187476	2877 (3.00-3.00)
Sidechain outliers	187428	2880 (3.00-3.00)
RSRZ outliers	180081	2671 (3.00-3.00)

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 ≥ 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	313	
1	B	313	

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 4631 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 22.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	280	Total	C	N	O	S	0	0	0
			2304	1488	374	427	15			
1	B	280	Total	C	N	O	S	0	0	0
			2304	1488	374	427	15			

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-18	HIS	-	expression tag	UNP Q9Y2R2
A	-17	HIS	-	expression tag	UNP Q9Y2R2
A	-16	HIS	-	expression tag	UNP Q9Y2R2
A	-15	HIS	-	expression tag	UNP Q9Y2R2
A	-14	HIS	-	expression tag	UNP Q9Y2R2
A	-13	HIS	-	expression tag	UNP Q9Y2R2
A	-12	SER	-	expression tag	UNP Q9Y2R2
A	-11	SER	-	expression tag	UNP Q9Y2R2
A	-10	GLY	-	expression tag	UNP Q9Y2R2
A	-9	LEU	-	expression tag	UNP Q9Y2R2
A	-8	VAL	-	expression tag	UNP Q9Y2R2
A	-7	PRO	-	expression tag	UNP Q9Y2R2
A	-6	ARG	-	expression tag	UNP Q9Y2R2
A	-5	GLY	-	expression tag	UNP Q9Y2R2
A	-4	SER	-	expression tag	UNP Q9Y2R2
A	-3	HIS	-	expression tag	UNP Q9Y2R2
A	-2	MET	-	expression tag	UNP Q9Y2R2
A	-1	ALA	-	expression tag	UNP Q9Y2R2
A	0	SER	-	expression tag	UNP Q9Y2R2
B	-18	HIS	-	expression tag	UNP Q9Y2R2
B	-17	HIS	-	expression tag	UNP Q9Y2R2
B	-16	HIS	-	expression tag	UNP Q9Y2R2
B	-15	HIS	-	expression tag	UNP Q9Y2R2
B	-14	HIS	-	expression tag	UNP Q9Y2R2
B	-13	HIS	-	expression tag	UNP Q9Y2R2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-12	SER	-	expression tag	UNP Q9Y2R2
B	-11	SER	-	expression tag	UNP Q9Y2R2
B	-10	GLY	-	expression tag	UNP Q9Y2R2
B	-9	LEU	-	expression tag	UNP Q9Y2R2
B	-8	VAL	-	expression tag	UNP Q9Y2R2
B	-7	PRO	-	expression tag	UNP Q9Y2R2
B	-6	ARG	-	expression tag	UNP Q9Y2R2
B	-5	GLY	-	expression tag	UNP Q9Y2R2
B	-4	SER	-	expression tag	UNP Q9Y2R2
B	-3	HIS	-	expression tag	UNP Q9Y2R2
B	-2	MET	-	expression tag	UNP Q9Y2R2
B	-1	ALA	-	expression tag	UNP Q9Y2R2
B	0	SER	-	expression tag	UNP Q9Y2R2

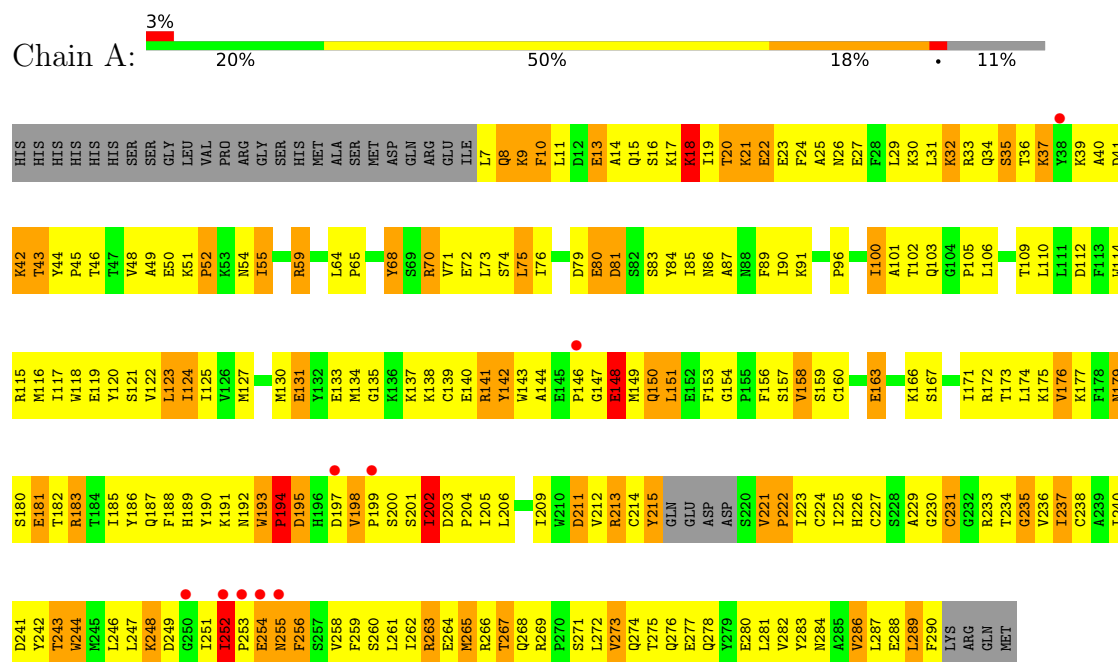
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	12	Total O 12 12	0	0
2	B	11	Total O 11 11	0	0

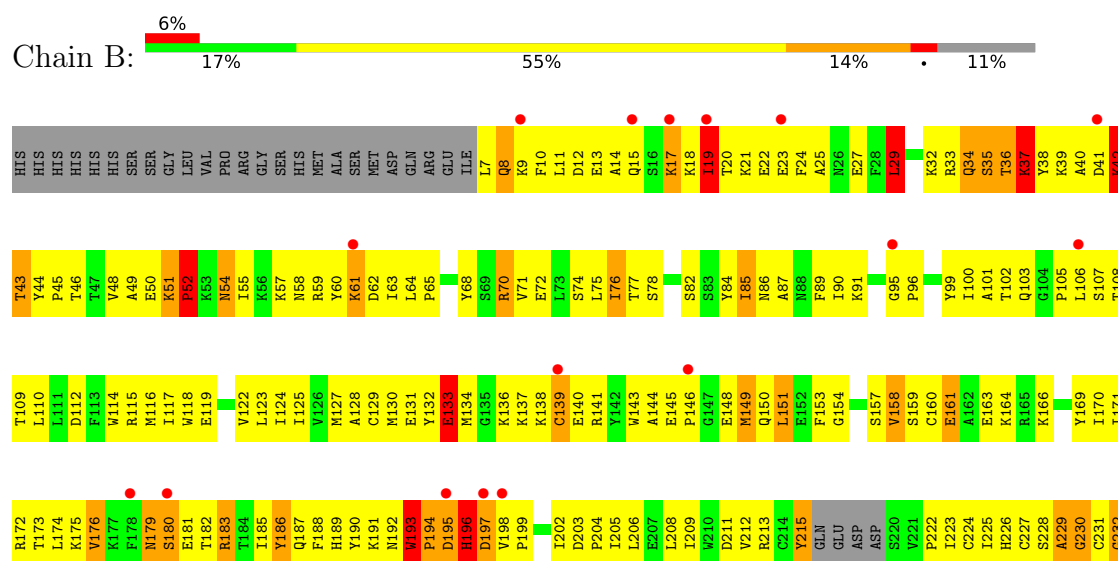
3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: Tyrosine-protein phosphatase non-receptor type 22



- Molecule 1: Tyrosine-protein phosphatase non-receptor type 22



	GLN	MET
R233		
T234		
G235		
V236		
I237		
C238		
A239		
L240		
D241		
Y242		
T243		
N244		
H245		
L246		
L247		
K248		
D249		
G250		
I251		
L252		
P253		
E254		
N255		
F256		
S257		
V258		
F259		
S260		
L261		
T262		
E263		
Z264		
N265		
R266		
T267		
Q268		
T269		
P270		
S271		
L272		
T273		
Q274		
T275		
Q276		
E277		
Q278		
Y279		
E280		
L281		
V282		
Y283		
N284		
A285		
V286		
L287		
Q288		
L289		
F290		
LYS		
ARG		

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	46.31Å 93.95Å 148.17Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.00 50.00 – 3.00	Depositor EDS
% Data completeness (in resolution range)	92.2 (50.00-3.00) 92.1 (50.00-3.00)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.14	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.18 (at 2.80Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.284 , 0.307 0.284 , 0.306	Depositor DCC
R_{free} test set	990 reflections (6.11%)	wwPDB-VP
Wilson B-factor (Å ²)	54.2	Xtriage
Anisotropy	0.227	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 108.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	4631	wwPDB-VP
Average B, all atoms (Å ²)	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 58.92 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.9045e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical 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	0.55	0/2360	1.16	19/3191 (0.6%)
1	B	0.59	0/2360	1.12	20/3191 (0.6%)
All	All	0.57	0/4720	1.14	39/6382 (0.6%)

There are no bond length outliers.

The worst 5 of 39 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	11	LEU	N-CA-C	-13.10	95.73	112.23
1	A	19	ILE	N-CA-C	10.85	131.90	109.34
1	A	252	ILE	CA-C-N	10.59	131.87	120.12
1	A	252	ILE	C-N-CA	10.59	131.87	120.12
1	A	43	THR	N-CA-C	-8.44	102.01	111.71

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	2304	0	2291	324	0
1	B	2304	0	2291	306	0
2	A	12	0	0	2	0
2	B	11	0	0	5	0
All	All	4631	0	4582	624	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 68.

The worst 5 of 624 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:10:PHE:HD2	1:A:13:GLU:HB2	1.09	1.13
1:B:46:THR:HG21	1:B:64:LEU:HB3	1.28	1.12
1:A:91:LYS:HG3	1:A:96:PRO:HA	1.30	1.10
1:A:46:THR:HG21	1:A:64:LEU:HB3	1.15	1.09
1:B:51:LYS:HG3	1:B:52:PRO:HD2	1.35	1.07

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	276/313 (88%)	191 (69%)	59 (21%)	26 (9%)	0	2
1	B	276/313 (88%)	205 (74%)	50 (18%)	21 (8%)	1	4
All	All	552/626 (88%)	396 (72%)	109 (20%)	47 (8%)	0	3

5 of 47 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	21	LYS
1	A	42	LYS
1	A	80	GLU
1	A	148	GLU
1	A	193	TRP

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	256/287 (89%)	218 (85%)	38 (15%)	3	15
1	B	256/287 (89%)	219 (86%)	37 (14%)	3	15
All	All	512/574 (89%)	437 (85%)	75 (15%)	3	15

5 of 75 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	151	LEU
1	B	252	ILE
1	B	161	GLU
1	B	197	ASP
1	A	194	PRO

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	278	GLN
1	B	276	GLN
1	B	54	ASN
1	B	268	GLN
1	B	15	GLN

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 oligosaccharides 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, 95th 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 ²)	Q<0.9
1	A	280/313 (89%)	0.44	9 (3%)	50 29	9, 52, 93, 104	0
1	B	280/313 (89%)	0.44	18 (6%)	25 13	5, 52, 95, 104	0
All	All	560/626 (89%)	0.44	27 (4%)	35 19	5, 52, 95, 104	0

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	252	ILE	4.4
1	A	146	PRO	3.4
1	B	41	ASP	3.4
1	B	198	VAL	3.3
1	B	146	PRO	3.0

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 oligosaccharides 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.