



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 25, 2024 – 03:48 PM EST

PDB ID : 1ZCK
Title : native structure prl-1 (ptp4a1)
Authors : Sun, J.P.; Wang, W.Q.; Yang, H.; Liu, S.; Liang, F.; Fedorov, A.A.; Almo, S.C.; Zhang, Z.Y.
Deposited on : 2005-04-12
Resolution : 1.90 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 : 4.02b-467
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 1.21
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.004 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.40

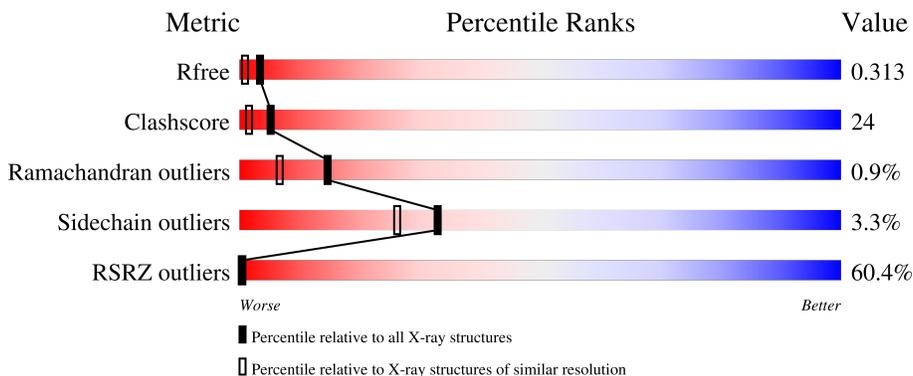
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.90 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}	164625	7293 (1.90-1.90)
Clashscore	180529	8090 (1.90-1.90)
Ramachandran outliers	177936	8022 (1.90-1.90)
Sidechain outliers	177891	8022 (1.90-1.90)
RSRZ outliers	164620	7292 (1.90-1.90)

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	154	 45% 64% 32% ..
1	B	154	 62% 65% 31% ..
1	C	154	 67% 60% 36% ..

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	ACY	B	415	-	X	X	-
2	ACY	C	416	-	X	X	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3952 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.

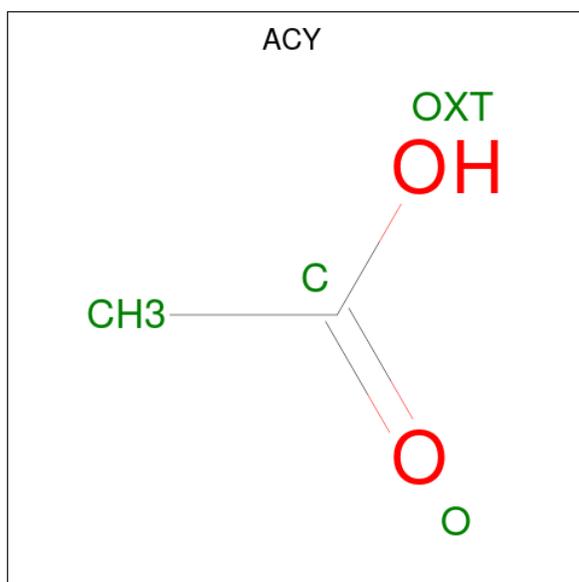
- Molecule 1 is a protein called protein tyrosine phosphatase 4a1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	150	1210	782	210	211	4	3	0	0	0
1	B	152	1226	790	212	217	4	3	0	0	0
1	C	152	1226	790	212	217	4	3	0	0	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	17	MSE	MET	modified residue	UNP Q78EG7
A	124	MSE	MET	modified residue	UNP Q78EG7
A	156	MSE	MET	modified residue	UNP Q78EG7
B	17	MSE	MET	modified residue	UNP Q78EG7
B	124	MSE	MET	modified residue	UNP Q78EG7
B	156	MSE	MET	modified residue	UNP Q78EG7
C	17	MSE	MET	modified residue	UNP Q78EG7
C	124	MSE	MET	modified residue	UNP Q78EG7
C	156	MSE	MET	insertion	UNP Q78EG7

- Molecule 2 is ACETIC ACID (three-letter code: ACY) (formula: C₂H₄O₂).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	C	1	Total C O 4 2 2	0	0

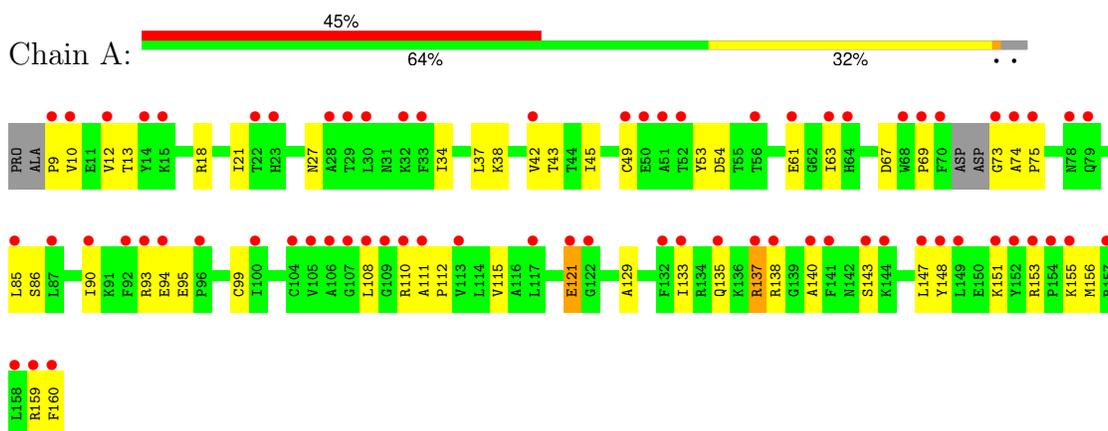
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	82	Total O 82 82	0	0
3	B	99	Total O 99 99	0	0
3	C	97	Total O 97 97	0	0

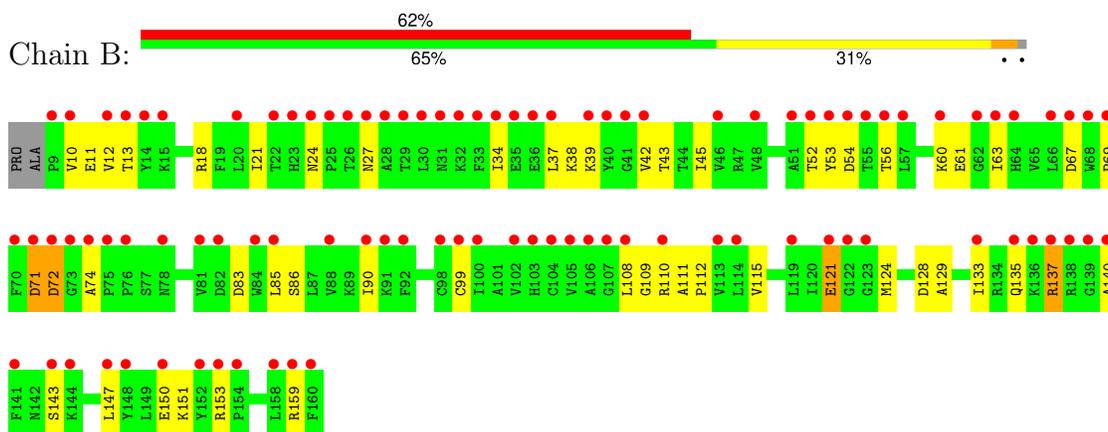
3 Residue-property plots [i](#)

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. 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. 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.

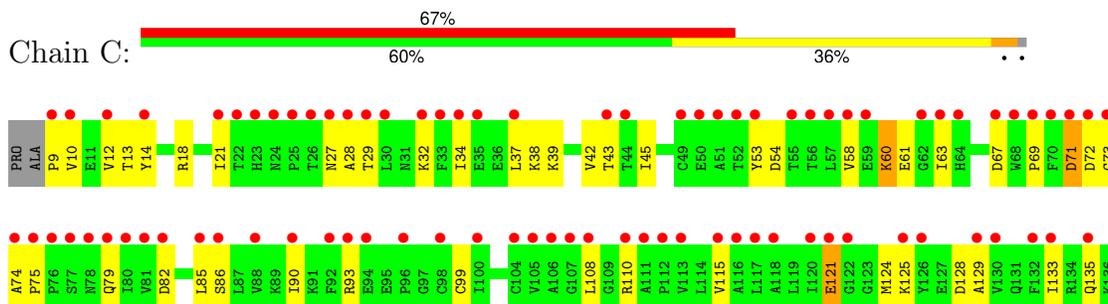
- Molecule 1: protein tyrosine phosphatase 4a1



- Molecule 1: protein tyrosine phosphatase 4a1



- Molecule 1: protein tyrosine phosphatase 4a1



R137	R138	G139	A140	F141	M142	S143	K144	Q145	L146	L147	Y148	L149	E150	K151	Y152	R153	P154	K155	M156	R157	L158	R159	F160
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	71.29Å 105.57Å 181.35Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 1.90 30.00 – 1.92	Depositor EDS
% Data completeness (in resolution range)	(Not available) (30.00-1.90) 92.4 (30.00-1.92)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.91 (at 1.92Å)	Xtrriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.245 , 0.257 0.307 , 0.313	Depositor DCC
R_{free} test set	4920 reflections (10.05%)	wwPDB-VP
Wilson B-factor (Å ²)	26.8	Xtrriage
Anisotropy	0.299	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 37.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	3952	wwPDB-VP
Average B, all atoms (Å ²)	32.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.07% of the height of the origin peak. No significant pseudotranslation is detected.*

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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ACY

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	0.36	0/1233	0.61	0/1659
1	B	0.35	0/1250	0.61	0/1684
1	C	0.37	0/1250	0.63	0/1684
All	All	0.36	0/3733	0.61	0/5027

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	1210	0	1243	50	2
1	B	1226	0	1252	63	2
1	C	1226	0	1252	77	2
2	A	4	0	3	1	0
2	B	4	0	4	7	0
2	C	4	0	3	3	0
3	A	82	0	0	10	0
3	B	99	0	0	9	1
3	C	97	0	0	26	0
All	All	3952	0	3757	179	4

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 24.

The worst 5 of 179 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)
2:B:415:ACY:CH3	2:B:415:ACY:C	1.78	1.57
2:B:415:ACY:C	2:B:415:ACY:O	1.67	1.43
1:B:72:ASP:HB3	1:B:110:ARG:NH2	1.74	1.03
1:A:43:THR:HG21	1:A:95:GLU:OE1	1.59	1.01
1:A:155:LYS:HE3	3:A:485:HOH:O	1.65	0.94

All (4) 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)
1:B:83:ASP:OD2	1:C:79:GLN:NE2[7_455]	1.85	0.35
1:A:153:ARG:O	1:C:153:ARG:NH2[3_555]	1.92	0.28
1:A:160:PHE:OXT	1:B:159:ARG:NH1[3_555]	2.07	0.13
3:B:428:HOH:O	3:B:430:HOH:O[3_555]	2.17	0.03

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	146/154 (95%)	138 (94%)	7 (5%)	1 (1%)	19 11
1	B	150/154 (97%)	140 (93%)	9 (6%)	1 (1%)	19 11
1	C	150/154 (97%)	139 (93%)	9 (6%)	2 (1%)	10 3
All	All	446/462 (96%)	417 (94%)	25 (6%)	4 (1%)	14 7

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	71	ASP
1	A	140	ALA
1	B	140	ALA
1	C	140	ALA

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	131/131 (100%)	127 (97%)	4 (3%)	35	29
1	B	133/131 (102%)	128 (96%)	5 (4%)	28	21
1	C	133/131 (102%)	129 (97%)	4 (3%)	36	30
All	All	397/393 (101%)	384 (97%)	13 (3%)	33	26

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	121	GLU
1	B	137	ARG
1	C	137	ARG
1	C	85	LEU
1	C	121	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	27	ASN
1	B	27	ASN
1	C	27	ASN

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](#)

3 ligands 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	ACY	B	415	-	3,3,3	7.85	3 (100%)	3,3,3	4.91	2 (66%)
2	ACY	A	414	-	3,3,3	1.10	0	3,3,3	3.62	2 (66%)
2	ACY	C	416	-	3,3,3	0.89	0	3,3,3	4.33	3 (100%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	415	ACY	O-C	10.43	1.67	1.22
2	B	415	ACY	CH3-C	7.42	1.78	1.49
2	B	415	ACY	OXT-C	4.59	1.52	1.30

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	415	ACY	O-C-CH3	-7.39	92.23	122.53
2	C	416	ACY	OXT-C-CH3	-5.88	90.42	115.05
2	A	414	ACY	OXT-C-CH3	-4.69	95.41	115.05
2	C	416	ACY	O-C-CH3	4.12	139.42	122.53
2	A	414	ACY	O-C-CH3	4.08	139.25	122.53

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	415	ACY	7	0
2	A	414	ACY	1	0
2	C	416	ACY	3	0

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](#)

Warning: The R factor obtained from EDS is 0.311, which does not match the depositor's R factor of 0.245. Please interpret the results in this section carefully.

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	147/154 (95%)	2.11	70 (47%) 0 0	19, 33, 46, 58	0
1	B	149/154 (96%)	2.43	96 (64%) 0 0	17, 30, 50, 58	0
1	C	149/154 (96%)	2.60	103 (69%) 0 0	18, 31, 50, 66	0
All	All	445/462 (96%)	2.38	269 (60%) 0 0	17, 32, 50, 66	0

The worst 5 of 269 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	108	LEU	6.6
1	A	148	TYR	6.5
1	C	26	THR	6.5
1	C	140	ALA	6.3
1	C	74	ALA	6.1

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 monosaccharides in this entry.

6.4 Ligands [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, 95th 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 ²)	Q<0.9
2	ACY	B	415	4/4	0.33	0.32	59,61,61,62	0
2	ACY	A	414	4/4	0.45	0.20	54,55,55,56	0
2	ACY	C	416	4/4	0.57	0.22	38,45,46,50	0

6.5 Other polymers [i](#)

There are no such residues in this entry.