



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 28, 2018 – 11:20 AM EST

PDB ID : 1GFY  
Title : RESIDUE 259 IS A KEY DETERMINANT OF SUBSTRATE SPECIFICITY  
OF PROTEIN-TYROSINE PHOSPHATASE 1B AND ALPHA  
Authors : Iversen, L.F.  
Deposited on : 2000-06-26  
Resolution : 2.13 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030736

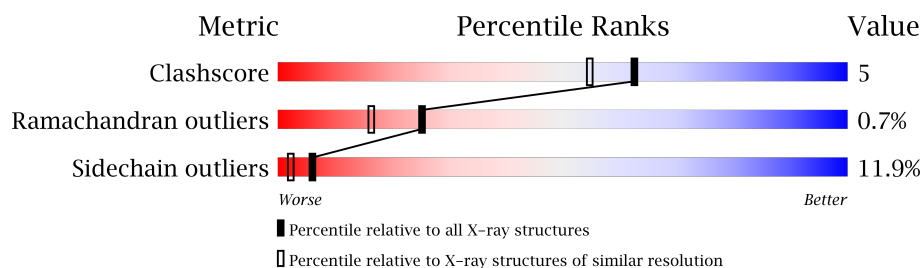
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.13 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))
Clashscore	112137	2047 (2.16-2.12)
Ramachandran outliers	110173	2020 (2.16-2.12)
Sidechain outliers	110143	2019 (2.16-2.12)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	298	 67% 23% 8% ..

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 2644 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.

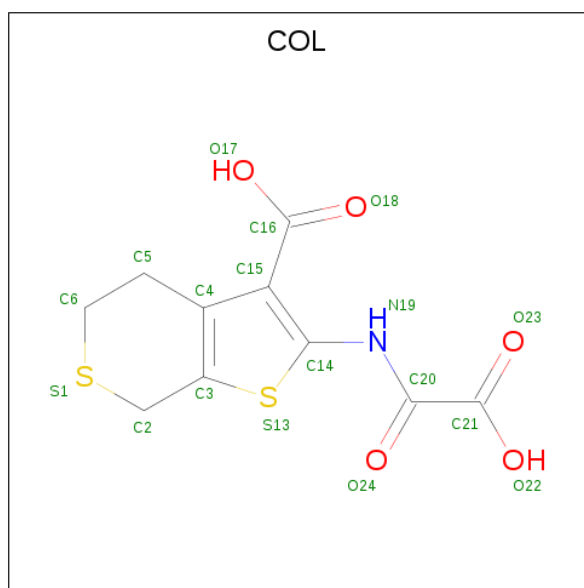
- Molecule 1 is a protein called PROTEIN (PROTEIN-TYROSINE PHOSPHATASE 1B).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	296	2417	1530	416	455	16	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	47	VAL	ARG	ENGINEERED	UNP P18031
A	48	ASN	ASP	ENGINEERED	UNP P18031
A	151	THR	SER	CONFLICT	UNP P18031
A	252	ASP	GLU	CONFLICT	UNP P18031
A	258	CYS	MET	ENGINEERED	UNP P18031
A	259	GLN	GLY	ENGINEERED	UNP P18031

- Molecule 2 is 2-(OXALYL-AMINO)-4,7-DIHYDRO-5H-THIENO[2,3-C]THIOPYRAN-3-CARBOXYLIC ACID (three-letter code: COL) (formula: C<sub>10</sub>H<sub>9</sub>NO<sub>5</sub>S<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			18	10	1	5	2		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	209	Total	O	0	0
			209	209		

### 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 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.

Note EDS was not executed.

- Molecule 1: PROTEIN (PROTEIN-TYROSINE PHOSPHATASE 1B)



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	88.19 Å 88.19 Å 103.73 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	6.00 – 2.13	Depositor
% Data completeness (in resolution range)	100.0 (6.00-2.13)	Depositor
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R, $R_{free}$	0.191 , 0.250	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2644	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

## 5 Model quality

### 5.1 Standard geometry

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

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	1.38	15/2472 (0.6%)	1.91	62/3334 (1.9%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	5

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	205	SER	CB-OG	14.35	1.60	1.42
1	A	285	SER	CB-OG	8.63	1.53	1.42
1	A	21	GLN	CG-CD	7.28	1.67	1.51
1	A	186	GLU	CG-CD	7.26	1.62	1.51
1	A	242	SER	CB-OG	7.17	1.51	1.42
1	A	201	SER	CB-OG	6.02	1.50	1.42
1	A	15	SER	CB-OG	5.98	1.50	1.42
1	A	50	SER	CB-OG	5.89	1.50	1.42
1	A	131	LYS	CD-CE	5.72	1.65	1.51
1	A	80	SER	CB-OG	5.63	1.49	1.42
1	A	104	SER	CB-OG	5.51	1.49	1.42
1	A	252	ASP	CB-CG	5.49	1.63	1.51
1	A	216	SER	CA-CB	5.41	1.61	1.52
1	A	25	HIS	CG-CD2	5.05	1.44	1.35
1	A	161	GLU	CD-OE1	-5.01	1.20	1.25

All (62) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	238	ARG	NE-CZ-NH2	-21.36	109.62	120.30
1	A	56	ARG	NE-CZ-NH2	-20.50	110.05	120.30
1	A	56	ARG	NE-CZ-NH1	20.23	130.42	120.30
1	A	238	ARG	NE-CZ-NH1	15.85	128.22	120.30
1	A	268	ARG	NE-CZ-NH2	-14.74	112.93	120.30
1	A	33	ARG	NE-CZ-NH1	12.07	126.34	120.30
1	A	45	ARG	NE-CZ-NH1	11.23	125.91	120.30
1	A	45	ARG	NE-CZ-NH2	-10.48	115.06	120.30
1	A	257	ARG	CD-NE-CZ	8.57	135.60	123.60
1	A	234	LEU	CB-CG-CD1	8.52	125.48	111.00
1	A	169	ARG	NE-CZ-NH2	-7.87	116.37	120.30
1	A	81	TYR	CB-CG-CD1	-7.60	116.44	121.00
1	A	66	TYR	CB-CG-CD1	7.58	125.55	121.00
1	A	49	VAL	CA-CB-CG1	7.50	122.16	110.90
1	A	205	SER	CB-CA-C	7.22	123.81	110.10
1	A	268	ARG	NE-CZ-NH1	7.21	123.90	120.30
1	A	79	ARG	NE-CZ-NH1	7.08	123.84	120.30
1	A	66	TYR	CB-CG-CD2	-6.98	116.81	121.00
1	A	79	ARG	NE-CZ-NH2	-6.97	116.81	120.30
1	A	131	LYS	CD-CE-NZ	6.84	127.42	111.70
1	A	162	ASN	CB-CA-C	6.72	123.84	110.40
1	A	25	HIS	CA-CB-CG	-6.69	102.22	113.60
1	A	252	ASP	CB-CG-OD2	6.67	124.30	118.30
1	A	151	THR	CB-CA-C	-6.59	93.79	111.60
1	A	243	SER	CB-CA-C	-6.47	97.81	110.10
1	A	18	ALA	CB-CA-C	-6.28	100.69	110.10
1	A	289	ASP	CB-CA-C	6.27	122.94	110.40
1	A	237	LYS	CA-CB-CG	6.25	127.16	113.40
1	A	253	MET	CG-SD-CE	-6.21	90.26	100.20
1	A	199	ARG	NE-CZ-NH2	-6.11	117.24	120.30
1	A	272	LEU	CB-CG-CD1	6.11	121.39	111.00
1	A	254	ARG	NE-CZ-NH2	6.11	123.35	120.30
1	A	52	PHE	CB-CG-CD1	-6.09	116.54	120.80
1	A	160	LEU	CB-CG-CD2	6.07	121.31	111.00
1	A	118	SER	CA-CB-OG	-6.03	94.91	111.20
1	A	244	VAL	CA-CB-CG1	5.87	119.70	110.90
1	A	130	GLU	CB-CA-C	-5.83	98.75	110.40
1	A	91	THR	CA-CB-OG1	-5.79	96.85	109.00
1	A	249	VAL	CA-CB-CG2	5.77	119.56	110.90
1	A	21	GLN	CB-CG-CD	5.75	126.54	111.60
1	A	120	LYS	CG-CD-CE	5.68	128.94	111.90
1	A	140	LEU	CB-CG-CD1	5.66	120.62	111.00
1	A	33	ARG	NE-CZ-NH2	-5.53	117.54	120.30

Continued on next page...



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	148	ASP	N-CA-CB	-5.51	100.68	110.60
1	A	292	LYS	CD-CE-NZ	-5.42	99.25	111.70
1	A	9	GLN	N-CA-CB	-5.39	100.90	110.60
1	A	113	VAL	CA-CB-CG1	5.36	118.94	110.90
1	A	237	LYS	N-CA-CB	5.34	120.21	110.60
1	A	72	ILE	CA-CB-CG1	5.33	121.13	111.00
1	A	132	GLU	CB-CA-C	-5.33	99.74	110.40
1	A	84	THR	OG1-CB-CG2	5.32	122.23	110.00
1	A	121	CYS	CA-CB-SG	5.13	123.24	114.00
1	A	29	ASP	N-CA-CB	-5.10	101.42	110.60
1	A	80	SER	N-CA-CB	-5.09	102.87	110.50
1	A	131	LYS	CG-CD-CE	5.09	127.16	111.90
1	A	156	ARG	NE-CZ-NH2	5.08	122.84	120.30
1	A	71	LEU	CB-CG-CD2	5.05	119.58	111.00
1	A	168	THR	OG1-CB-CG2	5.04	121.60	110.00
1	A	88	LEU	CB-CG-CD1	5.04	119.57	111.00
1	A	292	LYS	CA-CB-CG	-5.04	102.32	113.40
1	A	65	ASP	CB-CG-OD1	5.02	122.82	118.30
1	A	237	LYS	CD-CE-NZ	5.00	123.21	111.70

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	105	ARG	Sidechain
1	A	112	ARG	Sidechain
1	A	153	TYR	Sidechain
1	A	254	ARG	Sidechain
1	A	33	ARG	Sidechain

## 5.2 Too-close contacts ⓘ

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	2417	0	2374	23	0
2	A	18	0	8	1	0
3	A	209	0	0	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	2644	0	2382	24	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 5.

All (24) 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:7:PHE:CZ	1:A:272:LEU:HD13	2.35	0.61
1:A:227:LEU:CD1	1:A:249:VAL:HG22	2.33	0.59
1:A:227:LEU:HD12	1:A:249:VAL:HG22	1.91	0.52
1:A:68:ASN:HB2	3:A:310:HOH:O	2.10	0.52
1:A:45:ARG:H	1:A:85:GLN:HE22	1.58	0.51
1:A:7:PHE:CE1	1:A:272:LEU:HD13	2.47	0.50
1:A:280:PHE:CD1	1:A:285:SER:HA	2.47	0.49
1:A:7:PHE:CZ	1:A:272:LEU:CD1	2.97	0.48
1:A:17:ALA:O	1:A:21:GLN:CG	2.62	0.48
1:A:103:LYS:NZ	3:A:342:HOH:O	2.48	0.45
1:A:73:LYS:NZ	3:A:349:HOH:O	2.50	0.45
1:A:165:THR:O	1:A:166:GLN:CB	2.64	0.45
1:A:237:LYS:NZ	3:A:340:HOH:O	2.48	0.45
1:A:238:ARG:HG2	1:A:238:ARG:H	1.65	0.45
1:A:219:ILE:HD13	1:A:219:ILE:HG21	1.79	0.44
1:A:31:PRO:HG3	1:A:33:ARG:HH21	1.84	0.43
1:A:59:LEU:HB3	1:A:61:GLN:HG2	2.00	0.43
2:A:301:COL:S13	2:A:301:COL:O24	2.76	0.43
1:A:145:ILE:O	1:A:146:SER:HB3	2.18	0.43
1:A:280:PHE:CE1	1:A:285:SER:HA	2.54	0.42
1:A:36:LYS:NZ	1:A:47:VAL:O	2.52	0.42
1:A:17:ALA:O	1:A:21:GLN:HG3	2.19	0.42
1:A:165:THR:O	1:A:166:GLN:HB2	2.20	0.41
1:A:45:ARG:H	1:A:85:GLN:NE2	2.19	0.40

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	294/298 (99%)	281 (96%)	11 (4%)	2 (1%)	25	16

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	63	ASP
1	A	261	ILE

### 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	270/272 (99%)	238 (88%)	32 (12%)	6	2

All (32) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	11	ASP
1	A	12	LYS
1	A	15	SER
1	A	59	LEU
1	A	71	LEU
1	A	72	ILE
1	A	88	LEU
1	A	105	ARG
1	A	128	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	131	LYS
1	A	140	LEU
1	A	158	LEU
1	A	160	LEU
1	A	162	ASN
1	A	163	LEU
1	A	166	GLN
1	A	168	THR
1	A	172	LEU
1	A	175	HIS
1	A	186	GLU
1	A	215	CYS
1	A	233	LEU
1	A	234	LEU
1	A	237	LYS
1	A	244	VAL
1	A	248	LYS
1	A	251	LEU
1	A	252	ASP
1	A	259	GLN
1	A	272	LEU
1	A	289	ASP
1	A	294	LEU

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	85	GLN
1	A	111	ASN
1	A	123	GLN
1	A	139	ASN
1	A	175	HIS

### 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](#)

1 ligand is 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	COL	A	301	-	11,19,19	1.97	3 (27%)	4,27,27	2.49	2 (50%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	COL	A	301	-	-	0/2/19/19	0/1/2/2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	301	COL	C15-C14	-4.09	1.32	1.41
2	A	301	COL	C14-S13	-3.15	1.67	1.72
2	A	301	COL	C15-C4	3.14	1.47	1.42

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	301	COL	C6-C5-C4	-3.98	102.70	112.39
2	A	301	COL	C5-C4-C15	-2.89	117.42	130.78

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	301	COL	1	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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.