



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 13, 2025 – 10:05 AM EST

PDB ID : 9DHF  
Title : The Retinoblastoma Protein with Mutation E554K  
Authors : Ruiz-Rivera, A.; Castro, A.; Burke, J.R.  
Deposited on : 2024-09-03  
Resolution : 2.26 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 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
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

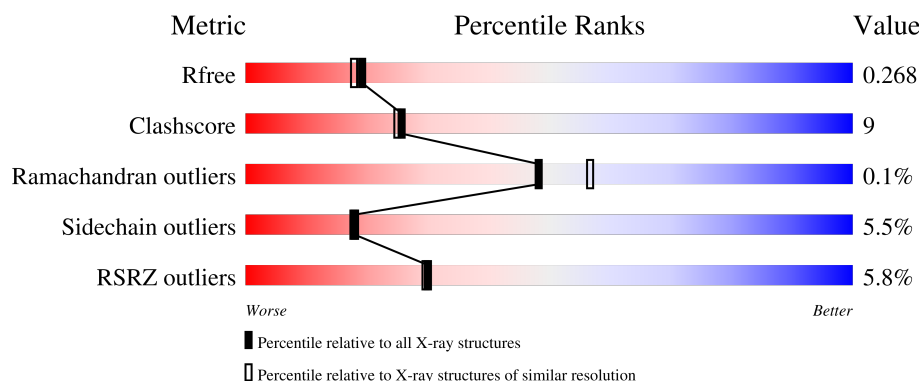
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

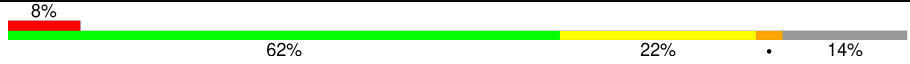

The reported resolution of this entry is 2.26 Å.

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	1763 (2.26-2.26)
Clashscore	180529	1919 (2.26-2.26)
Ramachandran outliers	177936	1884 (2.26-2.26)
Sidechain outliers	177891	1885 (2.26-2.26)
RSRZ outliers	164620	1763 (2.26-2.26)

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  $\geq 3$ , 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	390	
1	B	390	

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 11545 atoms, of which 5754 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 Retinoblastoma-associated protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	337	Total	C	H	N	O	S	0	0	0
			5623	1799	2840	466	498	20			
1	B	349	Total	C	H	N	O	S	0	0	0
			5784	1850	2914	480	519	21			

There are 68 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	377	GLY	-	expression tag	UNP P06400
A	378	GLU	-	expression tag	UNP P06400
A	379	PHE	-	expression tag	UNP P06400
A	554	LYS	GLU	conflict	UNP P06400
A	608	GLU	SER	conflict	UNP P06400
A	612	ALA	SER	conflict	UNP P06400
A	?	-	LYS	deletion	UNP P06400
A	?	-	GLY	deletion	UNP P06400
A	?	-	SER	deletion	UNP P06400
A	?	-	THR	deletion	UNP P06400
A	?	-	THR	deletion	UNP P06400
A	?	-	ARG	deletion	UNP P06400
A	?	-	VAL	deletion	UNP P06400
A	?	-	ASN	deletion	UNP P06400
A	?	-	SER	deletion	UNP P06400
A	?	-	THR	deletion	UNP P06400
A	?	-	ALA	deletion	UNP P06400
A	?	-	ASN	deletion	UNP P06400
A	?	-	ALA	deletion	UNP P06400
A	?	-	GLU	deletion	UNP P06400
A	?	-	THR	deletion	UNP P06400
A	?	-	GLN	deletion	UNP P06400
A	?	-	ALA	deletion	UNP P06400
A	?	-	THR	deletion	UNP P06400
A	?	-	SER	deletion	UNP P06400

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	ALA	deletion	UNP P06400
A	?	-	PHE	deletion	UNP P06400
A	?	-	GLN	deletion	UNP P06400
A	?	-	THR	deletion	UNP P06400
A	?	-	GLN	deletion	UNP P06400
A	?	-	LYS	deletion	UNP P06400
A	?	-	PRO	deletion	UNP P06400
A	?	-	LEU	deletion	UNP P06400
A	780	ALA	SER	conflict	UNP P06400
B	377	GLY	-	expression tag	UNP P06400
B	378	GLU	-	expression tag	UNP P06400
B	379	PHE	-	expression tag	UNP P06400
B	554	LYS	GLU	conflict	UNP P06400
B	608	GLU	SER	conflict	UNP P06400
B	639	ALA	SER	conflict	UNP P06400
B	?	-	LYS	deletion	UNP P06400
B	?	-	GLY	deletion	UNP P06400
B	?	-	SER	deletion	UNP P06400
B	?	-	THR	deletion	UNP P06400
B	?	-	THR	deletion	UNP P06400
B	?	-	ARG	deletion	UNP P06400
B	?	-	VAL	deletion	UNP P06400
B	?	-	ASN	deletion	UNP P06400
B	?	-	SER	deletion	UNP P06400
B	?	-	THR	deletion	UNP P06400
B	?	-	ALA	deletion	UNP P06400
B	?	-	ASN	deletion	UNP P06400
B	?	-	ALA	deletion	UNP P06400
B	?	-	GLU	deletion	UNP P06400
B	?	-	THR	deletion	UNP P06400
B	?	-	GLN	deletion	UNP P06400
B	?	-	ALA	deletion	UNP P06400
B	?	-	THR	deletion	UNP P06400
B	?	-	SER	deletion	UNP P06400
B	?	-	ALA	deletion	UNP P06400
B	?	-	PHE	deletion	UNP P06400
B	?	-	GLN	deletion	UNP P06400
B	?	-	THR	deletion	UNP P06400
B	?	-	GLN	deletion	UNP P06400
B	?	-	LYS	deletion	UNP P06400
B	?	-	PRO	deletion	UNP P06400
B	?	-	LEU	deletion	UNP P06400

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	780	ALA	SER	conflict	UNP P06400

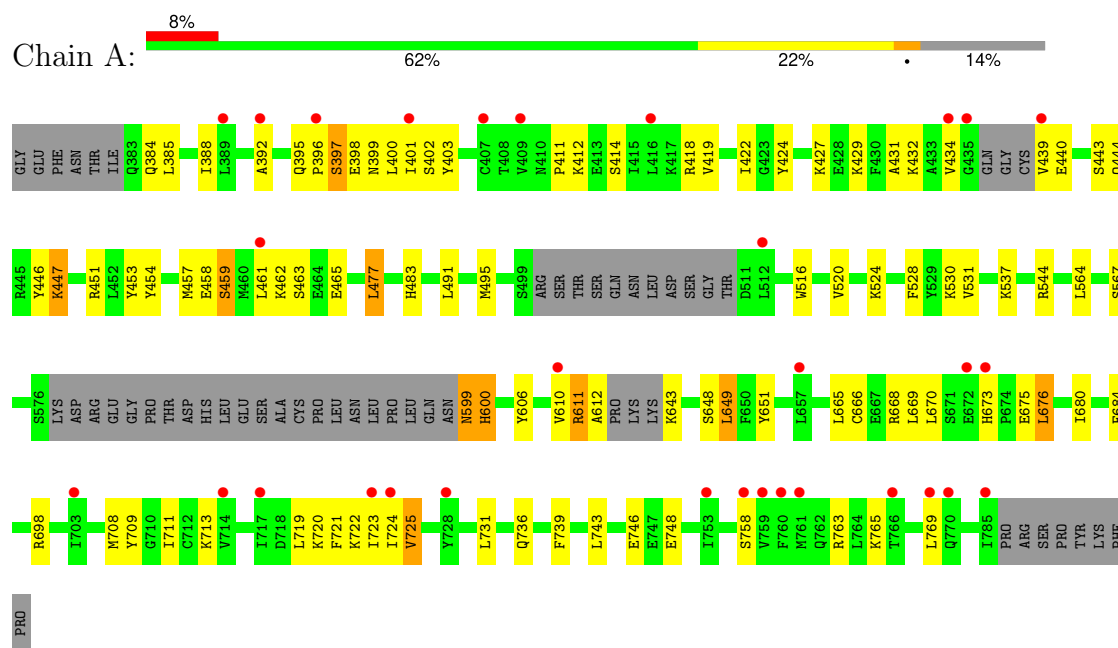
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	30	Total 30	O 30	0	0
2	B	108	Total 108	O 108	0	0

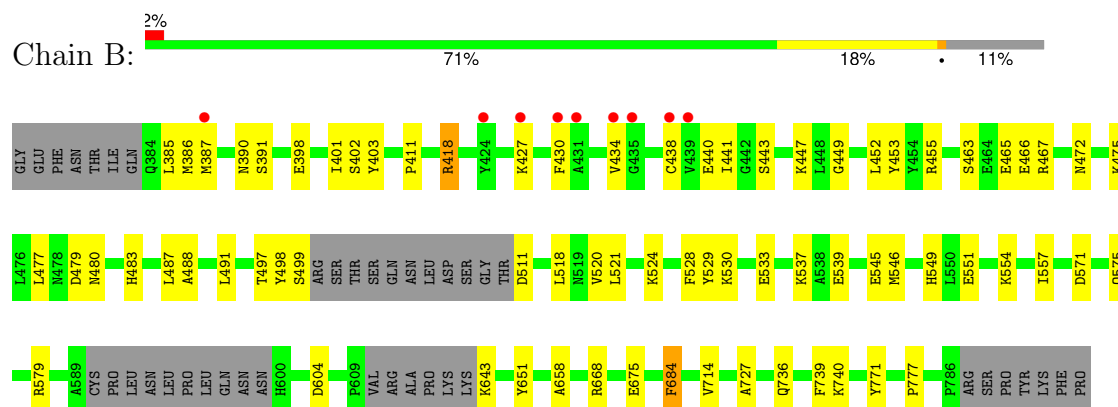
### 3 Residue-property plots [i](#)

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 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: Retinoblastoma-associated protein



#### • Molecule 1: Retinoblastoma-associated protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	250.36Å 250.36Å 35.31Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.31 – 2.26 47.31 – 2.26	Depositor EDS
% Data completeness (in resolution range)	99.9 (47.31-2.26) 99.9 (47.31-2.26)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.28 (at 2.27Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, $R_{free}$	0.227 , 0.265 0.231 , 0.268	Depositor DCC
$R_{free}$ test set	36649 reflections (5.18%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	40.5	Xtriage
Anisotropy	0.286	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 44.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.039 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	11545	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.54% 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	0.53	0/2838	0.67	0/3821
1	B	0.88	1/2929 (0.0%)	0.87	2/3946 (0.1%)
All	All	0.73	1/5767 (0.0%)	0.78	2/7767 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	675	GLU	CG-CD	5.59	1.60	1.51

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	571	ASP	CB-CG-OD1	-5.21	113.61	118.30
1	B	530	LYS	CD-CE-NZ	-5.08	100.02	111.70

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	2783	2840	2840	62	2
1	B	2870	2914	2914	45	2
2	A	30	0	0	5	0
2	B	108	0	0	11	0
All	All	5791	5754	5754	107	2



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 9.

The worst 5 of 107 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:533:GLU:HG3	2:B:882:HOH:O	1.78	0.83
1:B:736:GLN:NE2	2:B:806:HOH:O	2.18	0.75
1:A:599:ASN:N	2:A:802:HOH:O	2.22	0.72
1:A:395:GLN:HB2	1:A:396:PRO:HD2	1.72	0.72
1:A:461:LEU:HD21	1:A:477:LEU:HD11	1.73	0.70

All (2) 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:A:444:GLN:NE2	1:B:440:GLU:OE1[6_454]	2.07	0.13
1:A:444:GLN:HE21	1:B:440:GLU:OE1[6_454]	1.56	0.04

## 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	327/390 (84%)	314 (96%)	12 (4%)	1 (0%)	37	41
1	B	341/390 (87%)	332 (97%)	9 (3%)	0	100	100
All	All	668/780 (86%)	646 (97%)	21 (3%)	1 (0%)	48	57

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	397	SER

### 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	314/362 (87%)	289 (92%)	25 (8%)	10	8
1	B	324/362 (90%)	314 (97%)	10 (3%)	35	43
All	All	638/724 (88%)	603 (94%)	35 (6%)	18	18

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	463	SER
1	B	475	LYS
1	B	651	TYR
1	A	610	VAL
1	A	600	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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, 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	337/390 (86%)	0.93	31 (9%) 16 15	54, 81, 110, 121	0
1	B	349/390 (89%)	0.03	9 (2%) 57 58	23, 45, 79, 105	0
All	All	686/780 (87%)	0.47	40 (5%) 30 29	23, 66, 106, 121	0

The worst 5 of 40 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	438	CYS	4.6
1	A	435	GLY	4.3
1	A	769	LEU	3.8
1	A	439	VAL	3.7
1	A	728	TYR	3.5

### 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](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.