



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 29, 2026 – 01:45 pm BST

PDB ID : 31HE / pdb\_000031he  
Title : Xray crystal structure of PtsM receptor binding protein  
Authors : Wilson, J.S.; Bullough, P.A.; Fagan, R.P.  
Deposited on : 2026-06-04  
Resolution : 1.42 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 : 4-5-2 with Phenix2.0  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.015 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.50

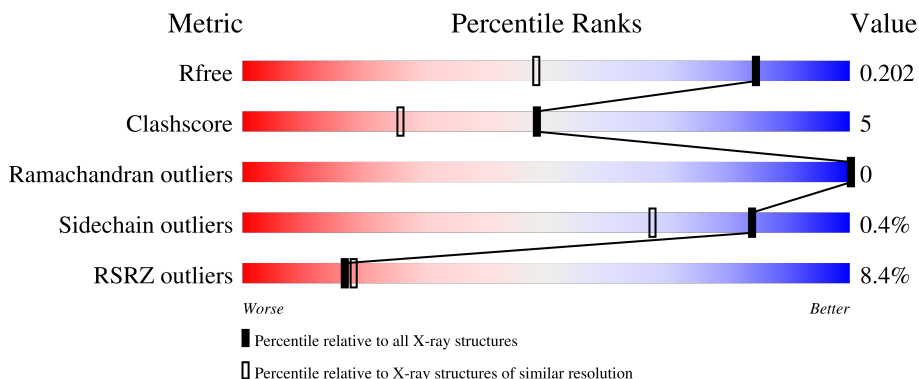
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.42 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	4041 (1.44-1.40)
Clashscore	190562	4154 (1.44-1.40)
Ramachandran outliers	187476	4083 (1.44-1.40)
Sidechain outliers	187428	4082 (1.44-1.40)
RSRZ outliers	180081	4039 (1.44-1.40)

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 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	AAA	514	 2% 28% 69%
1	BBB	514	 3% 28% 68%

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2862 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 Tail fiber protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	AAA	158	1308	851	208	247	2	0	6	0
1	BBB	162	1310	849	208	251	2	0	2	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	509	HIS	-	expression tag	UNP A0AB74R428
AAA	510	HIS	-	expression tag	UNP A0AB74R428
AAA	511	HIS	-	expression tag	UNP A0AB74R428
AAA	512	HIS	-	expression tag	UNP A0AB74R428
AAA	513	HIS	-	expression tag	UNP A0AB74R428
AAA	514	HIS	-	expression tag	UNP A0AB74R428
BBB	509	HIS	-	expression tag	UNP A0AB74R428
BBB	510	HIS	-	expression tag	UNP A0AB74R428
BBB	511	HIS	-	expression tag	UNP A0AB74R428
BBB	512	HIS	-	expression tag	UNP A0AB74R428
BBB	513	HIS	-	expression tag	UNP A0AB74R428
BBB	514	HIS	-	expression tag	UNP A0AB74R428

- Molecule 2 is SODIUM ION (CCD ID: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	AAA	1	Total	Na	0	0
			1	1		

- Molecule 3 is 1,2-ETHANEDIOL (CCD ID: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).

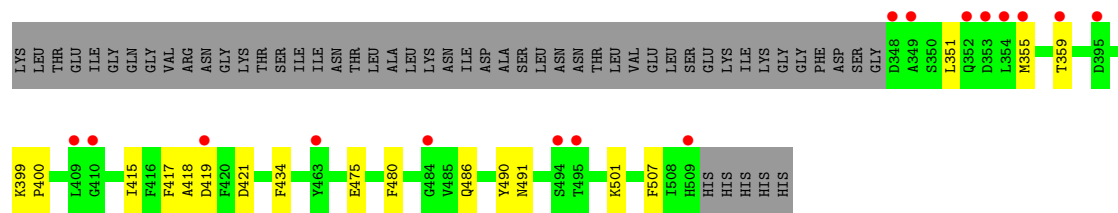


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	BBB	1	Total	C O	0	0
			4	2 2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	AAA	123	Total	O	0	0
			123	123		
4	BBB	116	Total	O	0	0
			116	116		





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	161.59Å 67.37Å 51.89Å 90.00° 99.77° 90.00°	Depositor
Resolution (Å)	46.79 – 1.42 46.79 – 1.42	Depositor EDS
% Data completeness (in resolution range)	99.8 (46.79-1.42) 99.8 (46.79-1.42)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.32 (at 1.42Å)	Xtrriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.185 , 0.202 0.185 , 0.202	Depositor DCC
$R_{free}$ test set	5081 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	20.0	Xtrriage
Anisotropy	0.109	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 25.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	2862	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.30% 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](#)

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

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	AAA	1.10	1/1364 (0.1%)	1.22	3/1854 (0.2%)
1	BBB	1.14	1/1353 (0.1%)	1.20	3/1839 (0.2%)
All	All	1.12	2/2717 (0.1%)	1.21	6/3693 (0.2%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	BBB	351	LEU	C-O	5.84	1.30	1.24
1	AAA	381	SER	CA-CB	-5.19	1.45	1.53

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	AAA	480	PHE	CA-CB-CG	6.64	120.44	113.80
1	BBB	480	PHE	CA-CB-CG	6.33	120.13	113.80
1	BBB	399	LYS	CB-CA-C	6.22	120.79	110.97
1	AAA	399	LYS	CB-CA-C	5.85	120.21	110.97
1	BBB	421	ASP	CA-CB-CG	5.56	118.16	112.60

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	AAA	1308	0	1251	18	0
1	BBB	1310	0	1236	20	0
2	AAA	1	0	0	0	0
3	BBB	4	0	6	1	0
4	AAA	123	0	0	1	0
4	BBB	116	0	0	2	0
All	All	2862	0	2493	28	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.

The worst 5 of 28 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:AAA:374[B]:VAL:HG23	1:AAA:505[B]:ILE:HD13	1.34	1.05
1:BBB:419:ASP:OD1	1:BBB:434:PHE:HD1	1.37	1.05
1:AAA:417:PHE:CZ	1:BBB:415[A]:ILE:HD11	2.05	0.90
1:AAA:467:ASN:OD1	4:AAA:701:HOH:O	1.90	0.89
1:BBB:419:ASP:OD1	1:BBB:434:PHE:CD1	2.25	0.88

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	AAA	162/514 (32%)	159 (98%)	3 (2%)	0	100	100
1	BBB	162/514 (32%)	159 (98%)	3 (2%)	0	100	100
All	All	324/1028 (32%)	318 (98%)	6 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	AAA	143/458 (31%)	142 (99%)	1 (1%)	76	52
1	BBB	140/458 (31%)	140 (100%)	0	100	100
All	All	283/916 (31%)	282 (100%)	1 (0%)	84	67

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	509	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](#)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 for Mogul analysis.

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	$\# Z  > 2$
3	EDO	BBB	701	-	3,3,3	0.30	0	2,2,2	0.75	0

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
3	EDO	BBB	701	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	BBB	701	EDO	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	BBB	701	EDO	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 [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	AAA	158/514 (30%)	0.38	11 (6%) 22 25	10, 20, 34, 47	6 (3%)
1	BBB	162/514 (31%)	0.53	16 (9%) 13 14	10, 20, 34, 57	2 (1%)
All	All	320/1028 (31%)	0.45	27 (8%) 17 18	10, 20, 34, 57	8 (2%)

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	419	ASP	6.0
1	BBB	419	ASP	5.8
1	BBB	348	ASP	5.8
1	BBB	509	HIS	4.5
1	AAA	483	THR	4.0

### 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](#)

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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
3	EDO	BBB	701	4/4	0.85	0.14	31,39,45,49	0
2	NA	AAA	601	1/1	0.98	0.12	27,27,27,27	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.