



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 28, 2025 – 04:24 PM EDT

PDB ID : 9OQB / pdb\_00009oqb  
Title : Crystal structure of E. coli ApaH in complex with Up4AGG  
Authors : Nuthanakanti, A.; Serganov, A.  
Deposited on : 2025-05-20  
Resolution : 2.04 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 Phenix2.0rc1  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 2.0rc1  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.006 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.45.1

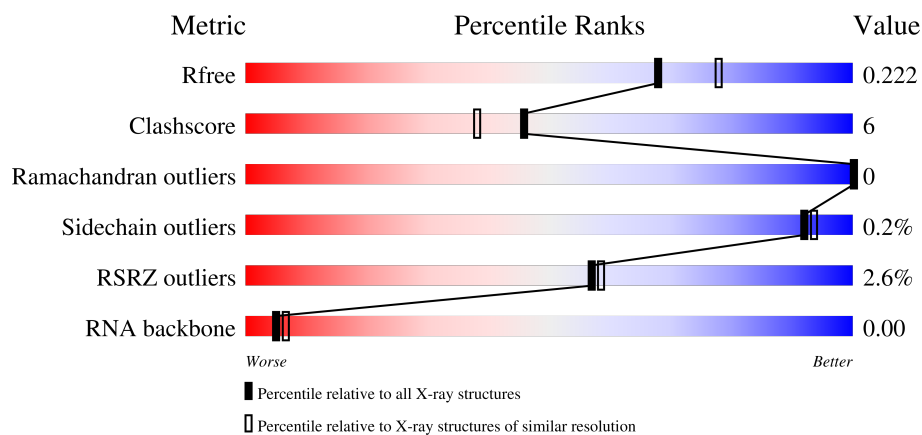
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.04 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	2096 (2.04-2.04)
Clashscore	180529	2229 (2.04-2.04)
Ramachandran outliers	177936	2217 (2.04-2.04)
Sidechain outliers	177891	2217 (2.04-2.04)
RSRZ outliers	164620	2096 (2.04-2.04)
RNA backbone	3690	1036 (2.40-1.68)

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	286	<div> <div>2%</div> <div> <div></div> <div>84%</div> <div>10%</div> <div>6%</div> </div> </div>
1	B	286	<div> <div>2%</div> <div> <div></div> <div>81%</div> <div>13%</div> <div>6%</div> </div> </div>
2	C	3	<div> <div>33%</div> <div> <div></div> <div>67%</div> <div>33%</div> </div> </div>
2	D	3	<div> <div>33%</div> <div> <div></div> <div>33%</div> <div>33%</div> </div> </div>

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 4615 atoms, of which 25 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 Bis(5'-nucleosyl)-tetraphosphatase, symmetrical.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	269	Total	C	N	O	S	0	0	0
			2094	1347	353	383	11			
1	B	269	Total	C	N	O	S	0	2	0
			2112	1359	356	385	12			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	281	HIS	-	expression tag	UNP C4ZPX5
A	282	HIS	-	expression tag	UNP C4ZPX5
A	283	HIS	-	expression tag	UNP C4ZPX5
A	284	HIS	-	expression tag	UNP C4ZPX5
A	285	HIS	-	expression tag	UNP C4ZPX5
A	286	HIS	-	expression tag	UNP C4ZPX5
B	281	HIS	-	expression tag	UNP C4ZPX5
B	282	HIS	-	expression tag	UNP C4ZPX5
B	283	HIS	-	expression tag	UNP C4ZPX5
B	284	HIS	-	expression tag	UNP C4ZPX5
B	285	HIS	-	expression tag	UNP C4ZPX5
B	286	HIS	-	expression tag	UNP C4ZPX5

- Molecule 2 is a RNA chain called RNA Up4AGG.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	2	Total	C	N	O	P	0	0	0
			39	10	5	19	5			
2	D	2	Total	C	N	O	P	0	0	0
			39	10	5	19	5			

- Molecule 3 is GLYCEROL (CCD ID: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	0
			14	3	8	3		

- Molecule 4 is SULFATE ION (CCD ID: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		

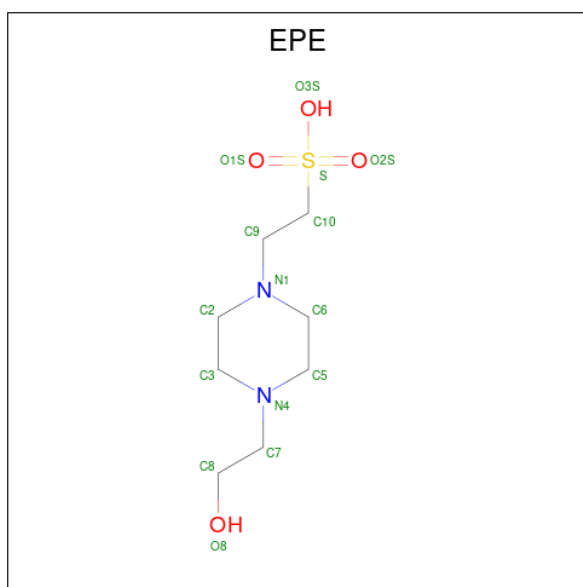
- Molecule 5 is MANGANESE (II) ION (CCD ID: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	2	Total	Mn	0	1
			2	2		
5	B	2	Total	Mn	0	1
			2	2		

- Molecule 6 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Mg	0	1
			1	1		
6	B	1	Total	Mg	0	1
			1	1		

- Molecule 7 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (CCD ID: EPE) (formula: C<sub>8</sub>H<sub>18</sub>N<sub>2</sub>O<sub>4</sub>S).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
7	B	1	Total	C	H	N	O	S	0	0
			32	8	17	2	4	1		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	136	Total	O	0	0
			136	136		
8	B	127	Total	O	0	0
			127	127		

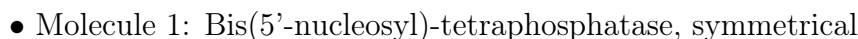
Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	C	4	Total 4	O 4	0	0
8	D	2	Total 2	O 2	0	0



- Molecule 1: Bis(5'-nucleosyl)-tetrphosphatase, symmetrical



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	167.18Å 55.03Å 120.18Å 90.00° 129.77° 90.00°	Depositor
Resolution (Å)	28.45 – 2.04 28.45 – 2.04	Depositor EDS
% Data completeness (in resolution range)	95.1 (28.45-2.04) 95.8 (28.45-2.04)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.27 (at 2.03Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.201 , 0.221 0.202 , 0.222	Depositor DCC
$R_{free}$ test set	1958 reflections (3.61%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.6	Xtriage
Anisotropy	0.123	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 49.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	4615	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.67% 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: SO4, MG, GOL, EPE, MN, A1L89

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	0.25	0/2154	0.51	0/2938
1	B	0.23	0/2173	0.49	0/2964
2	C	1.43	0/3	0.76	0/3
2	D	0.23	0/3	0.55	0/3
All	All	0.24	0/4333	0.50	0/5908

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	2094	0	2025	19	0
1	B	2112	0	2032	28	0
2	C	39	0	0	0	0
2	D	39	0	0	4	0
3	A	6	8	8	3	0
4	A	5	0	0	0	0
4	B	5	0	0	0	0
5	A	2	0	0	0	0
5	B	2	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	A	1	0	0	0	0
6	B	1	0	0	0	0
7	B	15	17	17	0	0
8	A	136	0	0	1	0
8	B	127	0	0	1	0
8	C	4	0	0	1	0
8	D	2	0	0	0	0
All	All	4590	25	4082	48	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 6.

The worst 5 of 48 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:197:LYS:HE3	2:D:1:A1L89:O2D	1.73	0.89
1:A:142:VAL:HG11	1:A:151:PHE:HB2	1.56	0.88
1:B:151[B]:PHE:O	1:B:155[B]:MET:HG3	1.81	0.80
1:A:159:MET:HE3	1:A:160:PRO:HA	1.74	0.68
1:A:145:SER:HB2	3:A:301:GOL:H2	1.78	0.66

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	267/286 (93%)	262 (98%)	5 (2%)	0	100	100
1	B	269/286 (94%)	265 (98%)	4 (2%)	0	100	100
All	All	536/572 (94%)	527 (98%)	9 (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	A	218/239 (91%)	217 (100%)	1 (0%)	86	88
1	B	219/239 (92%)	219 (100%)	0	100	100
All	All	437/478 (91%)	436 (100%)	1 (0%)	92	93

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

Mol	Chain	Res	Type
1	A	173	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	80	ASN
1	B	189	ASN
1	B	191	GLN
1	A	162	ASN
1	A	126	GLN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	C	0/3	-	-
2	D	0/3	-	-
All	All	0/6	-	-

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

## 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 10 ligands modelled in this entry, 6 are monoatomic - leaving 4 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	$\# Z  > 2$
4	SO4	B	302	-	4,4,4	0.23	0	6,6,6	0.10	0
7	EPE	B	301	-	15,15,15	0.66	0	19,20,20	1.81	5 (26%)
3	GOL	A	301	-	5,5,5	0.67	0	5,5,5	1.04	0
4	SO4	A	302	-	4,4,4	0.24	0	6,6,6	0.06	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
7	EPE	B	301	-	-	6/9/19/19	0/1/1/1
3	GOL	A	301	-	-	0/4/4/4	-

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	B	301	EPE	C5-N4-C3	3.58	116.55	108.84
7	B	301	EPE	C6-N1-C2	3.38	116.12	108.84
7	B	301	EPE	C10-C9-N1	-3.02	100.94	112.36
7	B	301	EPE	O3S-S-C10	2.75	111.38	106.00
7	B	301	EPE	O2S-S-C10	-2.20	103.40	106.73

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	B	301	EPE	C9-C10-S-O1S
7	B	301	EPE	C9-C10-S-O3S
7	B	301	EPE	N4-C7-C8-O8
7	B	301	EPE	C10-C9-N1-C2
7	B	301	EPE	C10-C9-N1-C6

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	301	GOL	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](#)

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	269/286 (94%)	0.10	6 (2%) 62 64	21, 33, 53, 70	0
1	B	269/286 (94%)	0.09	6 (2%) 62 64	17, 34, 49, 66	2 (0%)
2	C	1/3 (33%)	4.72	1 (100%) 0 0	38, 38, 38, 38	1 (100%)
2	D	1/3 (33%)	3.72	1 (100%) 0 0	43, 43, 43, 43	1 (100%)
All	All	540/578 (93%)	0.11	14 (2%) 57 58	17, 33, 52, 70	4 (0%)

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	2	G	4.7
2	D	2	G	3.7
1	B	150	PHE	3.1
1	A	204	PRO	3.0
1	B	269	ASN	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](#)

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
7	EPE	B	301	15/15	0.80	0.16	44,57,74,92	0
3	GOL	A	301	6/6	0.83	0.14	45,56,68,68	0
6	MG	A	304[B]	1/1	0.91	0.07	32,32,32,32	1
4	SO4	B	302	5/5	0.92	0.13	47,49,51,54	2
4	SO4	A	302	5/5	0.94	0.12	41,52,53,54	4
5	MN	A	303[A]	1/1	0.94	0.06	32,32,32,32	1
5	MN	A	305	1/1	0.98	0.04	27,27,27,27	1
5	MN	B	305	1/1	0.98	0.04	32,32,32,32	1
6	MG	B	304[B]	1/1	0.99	0.03	34,34,34,34	1
5	MN	B	303[A]	1/1	0.99	0.03	34,34,34,34	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.