



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 23, 2026 – 04:05 AM UTC

PDB ID : 3D7K / pdb\_00003d7k  
Title : Crystal structure of benzaldehyde lyase in complex with the inhibitor MBP  
Authors : Brandt, G.S.  
Deposited on : 2008-05-21  
Resolution : 2.49 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 : **FAILED**  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Buster-report : **FAILED**  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

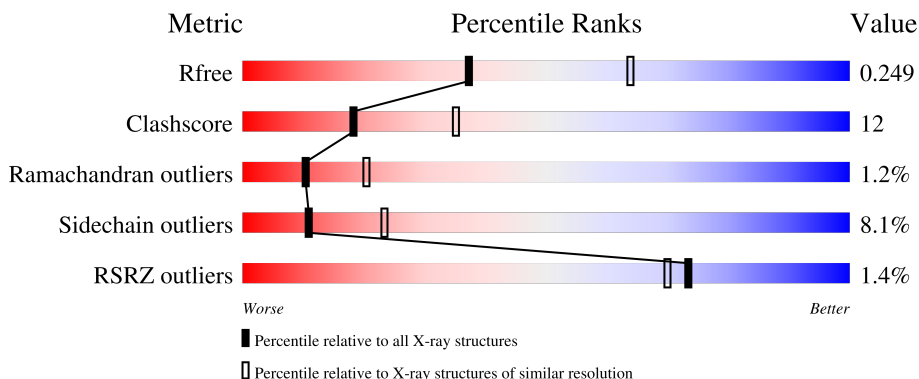
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

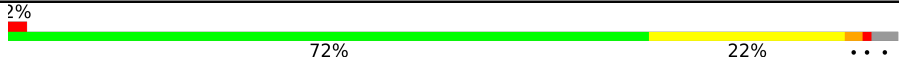
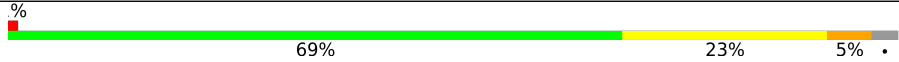
The reported resolution of this entry is 2.49 Å.

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	5829 (2.50-2.50)
Clashscore	190562	6492 (2.50-2.50)
Ramachandran outliers	187476	6378 (2.50-2.50)
Sidechain outliers	187428	6380 (2.50-2.50)
RSRZ outliers	180081	5833 (2.50-2.50)

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	570	
1	B	570	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8413 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 Benzaldehyde lyase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	554	4079	2576	723	764	16	0	0	0
1	B	554	4079	2576	723	764	16	0	0	0

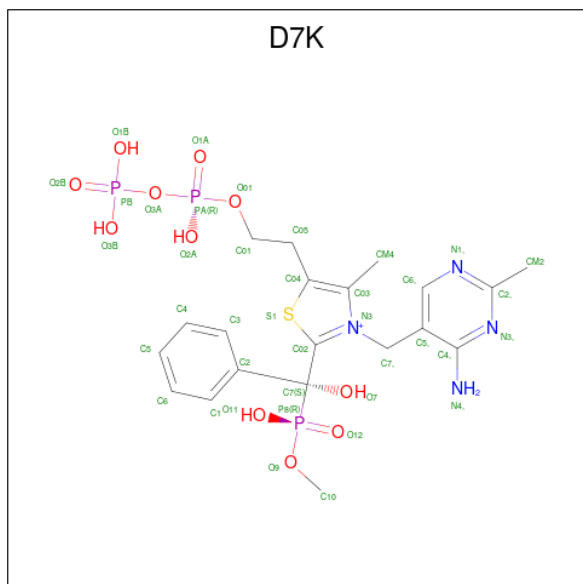
There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	563	GLY	-	expression tag	UNP Q9F4L3
A	564	SER	-	expression tag	UNP Q9F4L3
A	565	HIS	-	expression tag	UNP Q9F4L3
A	566	HIS	-	expression tag	UNP Q9F4L3
A	567	HIS	-	expression tag	UNP Q9F4L3
A	568	HIS	-	expression tag	UNP Q9F4L3
A	569	HIS	-	expression tag	UNP Q9F4L3
A	570	HIS	-	expression tag	UNP Q9F4L3
B	563	GLY	-	expression tag	UNP Q9F4L3
B	564	SER	-	expression tag	UNP Q9F4L3
B	565	HIS	-	expression tag	UNP Q9F4L3
B	566	HIS	-	expression tag	UNP Q9F4L3
B	567	HIS	-	expression tag	UNP Q9F4L3
B	568	HIS	-	expression tag	UNP Q9F4L3
B	569	HIS	-	expression tag	UNP Q9F4L3
B	570	HIS	-	expression tag	UNP Q9F4L3

- Molecule 2 is CALCIUM ION (CCD ID: CA) (formula: Ca).

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

- Molecule 3 is 3-[(4-amino-2-methylpyrimidin-5-yl)methyl]-2-[(S)-hydroxy[(R)-hydroxy(methoxy)phosphoryl]phenylmethyl]-5-(2-[[[R)-hydroxy(phosphonoxy)phosphoryl]oxy]ethyl)-4-methyl-1,3-thiazol-3-ium (CCD ID: D7K) (formula: C<sub>20</sub>H<sub>28</sub>N<sub>4</sub>O<sub>11</sub>P<sub>3</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	N	O	P			S
3	A	1	Total	C	N	O	P	S	0	0
			39	20	4	11	3	1		
3	B	1	Total	C	N	O	P	S	0	0
			39	20	4	11	3	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	79	Total	O	0	0
			79	79		
4	B	96	Total	O	0	0
			96	96		





## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	151.94Å 151.94Å 98.22Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.00 – 2.49 46.00 – 2.49	Depositor EDS
% Data completeness (in resolution range)	97.2 (46.00-2.49) 97.2 (46.00-2.49)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.99 (at 2.48Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.194 , 0.250 0.193 , 0.249	Depositor DCC
$R_{free}$ test set	4485 reflections (9.75%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	49.6	Xtrriage
Anisotropy	0.039	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 41.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.025 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	8413	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	50.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.28% 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: CA, D7K

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	1.12	4/4162 (0.1%)	1.22	20/5677 (0.4%)
1	B	1.15	4/4162 (0.1%)	1.24	22/5677 (0.4%)
All	All	1.13	8/8324 (0.1%)	1.23	42/11354 (0.4%)

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	1

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	106	ASP	CA-C	7.56	1.62	1.52
1	A	394	ALA	N-CA	6.96	1.55	1.46
1	B	394	ALA	N-CA	6.47	1.54	1.46
1	B	494	VAL	CA-CB	6.12	1.61	1.54
1	A	164	ASP	CB-CG	-5.97	1.37	1.52

The worst 5 of 42 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	393	GLY	N-CA-C	10.49	128.85	115.21
1	B	490	GLY	CA-C-N	9.68	129.79	119.05
1	B	490	GLY	C-N-CA	9.68	129.79	119.05
1	A	105	ARG	CA-C-N	-9.18	108.98	124.31
1	A	105	ARG	C-N-CA	-9.18	108.98	124.31

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	106	ASP	Peptide

## 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	4079	0	4069	86	0
1	B	4079	0	4069	109	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	39	0	24	5	0
3	B	39	0	24	5	0
4	A	79	0	0	10	0
4	B	96	0	0	16	0
All	All	8413	0	8186	199	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 12.

The worst 5 of 199 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:164:ASP:HB2	4:B:648:HOH:O	1.18	1.27
1:B:342:ARG:HH11	1:B:342:ARG:HG3	1.08	1.12
1:B:366:SER:HB2	1:B:374:HIS:CD2	2.04	0.93
1:B:366:SER:HB2	1:B:374:HIS:HD2	1.34	0.92
1:B:342:ARG:HH11	1:B:342:ARG:CG	1.81	0.91

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	552/570 (97%)	506 (92%)	38 (7%)	8 (1%)	9	17
1	B	552/570 (97%)	513 (93%)	34 (6%)	5 (1%)	14	27
All	All	1104/1140 (97%)	1019 (92%)	72 (6%)	13 (1%)	10	20

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	265	LYS
1	B	337	ALA
1	A	364	SER
1	A	369	ALA
1	B	554	LEU

### 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	412/426 (97%)	389 (94%)	23 (6%)	19	39
1	B	412/426 (97%)	368 (89%)	44 (11%)	6	13
All	All	824/852 (97%)	757 (92%)	67 (8%)	11	23

5 of 67 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	487	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	492	ASN
1	B	544	VAL
1	B	16	LYS
1	B	3	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	118	GLN
1	B	297	GLN
1	B	530	GLN
1	B	294	HIS
1	B	331	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

Mogul failed to run properly - this section is therefore empty.

## 5.5 Carbohydrates [i](#)

Mogul failed to run properly - this section is therefore empty.

## 5.6 Ligand geometry [i](#)

Mogul failed to run properly - this section is therefore empty.

## 5.7 Other polymers [i](#)

Mogul failed to run properly - this section is therefore empty.

## 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	554/570 (97%)	0.05	10 (1%) 67 64	26, 48, 89, 96	0
1	B	554/570 (97%)	-0.06	5 (0%) 81 78	28, 47, 70, 80	0
All	All	1108/1140 (97%)	-0.01	15 (1%) 73 70	26, 47, 80, 96	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	555	ILE	7.4
1	B	555	ILE	4.5
1	A	338	ALA	3.3
1	A	368	HIS	3.1
1	A	265	LYS	3.1

### 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 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	D7K	A	572	39/39	0.97	0.06	31,43,53,54	1
3	D7K	B	572	39/39	0.97	0.06	23,36,40,41	1
2	CA	A	571	1/1	0.99	0.05	56,56,56,56	0
2	CA	B	571	1/1	0.99	0.08	57,57,57,57	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.