



wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 06:42 AM GMT

PDB ID : 3D7K
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 validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

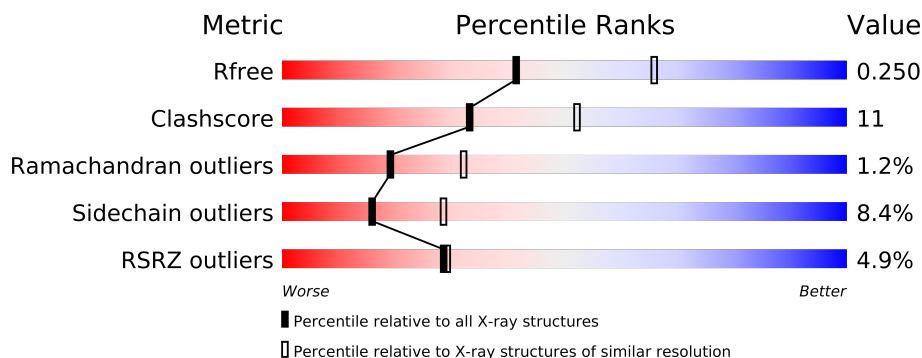
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

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}	66092	2784 (2.50-2.50)
Clashscore	79885	3562 (2.50-2.50)
Ramachandran outliers	78287	3480 (2.50-2.50)
Sidechain outliers	78261	3482 (2.50-2.50)
RSRZ outliers	66119	2785 (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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

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 hydrogen and 0 are deuterium.

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
1	A	554	Total	C	N	O	S	0	0	0
			4079	2576	723	764	16			
1	B	554	Total	C	N	O	S	0	0	0
			4079	2576	723	764	16			

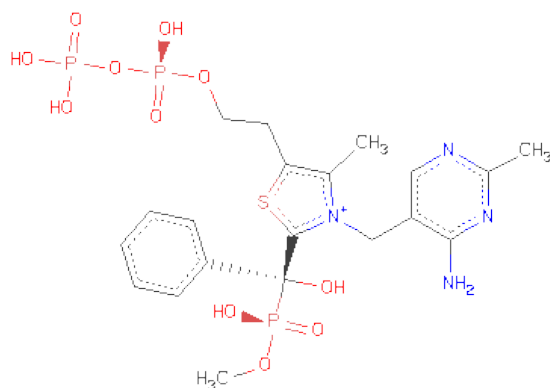
There are 16 discrepancies between the modelled and reference sequences:

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 (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Ca	0	0
			1	1		
2	A	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 (three-letter code: D7K) (formula: C₂₀H₂₈N₄O₁₁P₃S).



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

4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	151.94Å 151.94Å 98.22Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.00 – 2.49 46.01 – 2.49	Depositor EDS
% Data completeness (in resolution range)	97.2 (46.00-2.49) 97.2 (46.01-2.49)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.99 (at 2.48Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.194 , 0.250 0.194 , 0.250	Depositor DCC
R_{free} test set	4485 reflections (11.14%)	DCC
Wilson B-factor (Å ²)	49.6	Xtriage
Anisotropy	0.039	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 43.3	EDS
Estimated twinning fraction	0.025 for -h,-k,l	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 44738 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	8413	wwPDB-VP
Average B, all atoms (Å ²)	50.0	wwPDB-VP

Xtriage'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.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: D7K, CA

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	0.88	3/4162 (0.1%)	0.90	6/5677 (0.1%)
1	B	0.89	0/4162	0.89	7/5677 (0.1%)
All	All	0.88	3/8324 (0.0%)	0.90	13/11354 (0.1%)

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

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	164	ASP	CB-CG	-6.96	1.37	1.51
1	A	131	ARG	CG-CD	5.60	1.66	1.51
1	A	522	GLU	CG-CD	5.23	1.59	1.51

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	164	ASP	N-CA-CB	-8.19	95.86	110.60
1	A	62	ARG	NE-CZ-NH1	-7.07	116.76	120.30
1	A	392	ASP	CB-CG-OD1	6.60	124.24	118.30
1	A	393	GLY	N-CA-C	6.30	128.85	113.10
1	B	487	LEU	CA-CB-CG	6.11	129.35	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	106	ASP	Peptide

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4079	0	4069	80	0
1	B	4079	0	4069	99	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	183	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 11.

The worst 5 of 183 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(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:CG	1:B:342:ARG:HH11	1.81	0.91

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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%)	16	27
1	B	552/570 (97%)	513 (93%)	34 (6%)	5 (1%)	25	42
All	All	1104/1140 (97%)	1019 (92%)	72 (6%)	13 (1%)	19	32

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 ⓘ

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%)	387 (94%)	25 (6%)	26	46
1	B	412/426 (97%)	368 (89%)	44 (11%)	10	17
All	All	824/852 (97%)	755 (92%)	69 (8%)	16	28

5 of 69 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	96	LEU
1	B	220	ARG
1	B	494	VAL
1	B	106	ASP
1	B	176	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	26	HIS
1	B	113	GLN
1	B	530	GLN
1	B	40	HIS
1	B	87	ASN

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	D7K	A	572	2	40,41,41	1.64	5 (12%)	56,63,63	2.04	15 (26%)
3	D7K	B	572	2	40,41,41	1.84	7 (17%)	56,63,63	2.05	17 (30%)

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	D7K	A	572	2	-	0/35/41/41	0/3/3/3
3	D7K	B	572	2	-	0/35/41/41	0/3/3/3

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	572	D7K	C02-N3	6.53	1.47	1.36
3	A	572	D7K	C02-N3	6.38	1.47	1.36
3	B	572	D7K	C04-S1	-4.14	1.64	1.73
3	A	572	D7K	C04-S1	-3.59	1.65	1.73
3	A	572	D7K	P8-C7	-3.40	1.82	1.84

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	572	D7K	C02-N3-C03	6.48	115.42	109.76
3	A	572	D7K	C7,-N3-C02	-5.16	113.07	124.60
3	B	572	D7K	C7,-N3-C02	-4.96	113.53	124.60
3	A	572	D7K	CM2-C2,-N1,	4.92	122.89	117.02
3	B	572	D7K	C7,-N3-C03	4.26	133.09	124.34

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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, 95th 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 ²)	Q<0.9
1	A	554/570 (97%)	0.22	32 (5%) 22 23	26, 48, 89, 96	0
1	B	554/570 (97%)	0.22	22 (3%) 36 37	28, 47, 70, 80	0
All	All	1108/1140 (97%)	0.22	54 (4%) 28 29	26, 47, 80, 96	0

The worst 5 of 54 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	368	HIS	7.0
1	A	337	ALA	5.2
1	A	269	ALA	4.8
1	A	265	LYS	4.7
1	A	367	GLU	4.5

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th 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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	D7K	A	572	39/39	0.12	-0.73	31,43,53,54	1
3	D7K	B	572	39/39	0.10	-0.99	23,36,40,41	1
2	CA	A	571	1/1	0.09	-1.42	56,56,56,56	0
2	CA	B	571	1/1	0.03	-4.15	57,57,57,57	0

6.5 Other polymers

There are no such residues in this entry.