



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 13, 2023 – 12:42 pm GMT

PDB ID : 7R4B
Title : The Bacillus pumilus chorismate mutase
Authors : Lund, B.A.
Deposited on : 2022-02-08
Resolution : 1.10 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

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
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.32.1
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.32.1

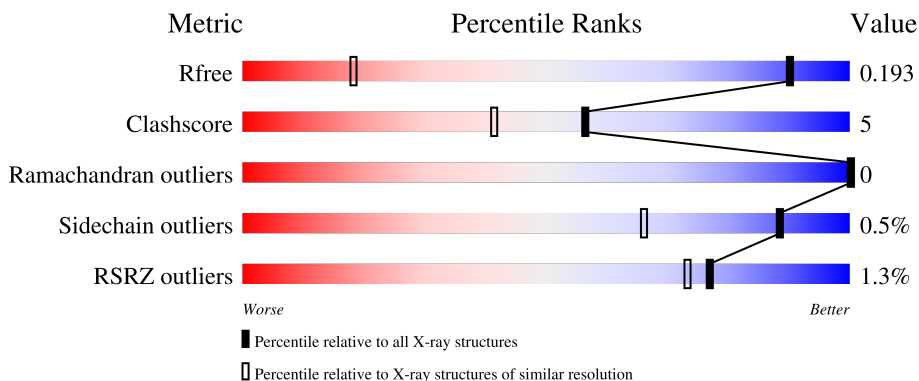
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.10 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.



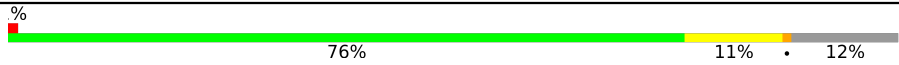
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1619 (1.14-1.06)
Clashscore	141614	1671 (1.14-1.06)
Ramachandran outliers	138981	1615 (1.14-1.06)
Sidechain outliers	138945	1613 (1.14-1.06)
RSRZ outliers	127900	1588 (1.14-1.06)

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 ≥ 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	133	 84% 6% 10%
1	B	133	 2% 81% 7% 12%
1	C	133	 2% 81% 8% 11%
1	D	133	 2% 78% 11% 11%
1	E	133	 % 82% 8% 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	133	 <p>A horizontal bar chart representing the quality of the chain. The bar is divided into three segments: a green segment representing 76%, a yellow segment representing 11%, and a grey segment representing 12%. A small red square is at the beginning of the bar, and a small black dot is at the end of the grey segment. The percentage values are printed below the bar.</p>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 12720 atoms, of which 5997 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 Chorismate mutase AroH.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	120	2010	625	1012	173	193	7	0	9	0
1	B	117	1958	611	986	168	186	7	0	10	0
1	C	119	2022	627	1022	174	191	8	0	11	0
1	D	118	1959	611	985	168	187	8	0	7	0
1	E	120	2010	626	1013	171	193	7	0	9	0
1	F	117	1937	603	979	165	182	8	0	6	0

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	128	HIS	-	expression tag	UNP A8FEK3
A	129	HIS	-	expression tag	UNP A8FEK3
A	130	HIS	-	expression tag	UNP A8FEK3
A	131	HIS	-	expression tag	UNP A8FEK3
A	132	HIS	-	expression tag	UNP A8FEK3
A	133	HIS	-	expression tag	UNP A8FEK3
B	128	HIS	-	expression tag	UNP A8FEK3
B	129	HIS	-	expression tag	UNP A8FEK3
B	130	HIS	-	expression tag	UNP A8FEK3
B	131	HIS	-	expression tag	UNP A8FEK3
B	132	HIS	-	expression tag	UNP A8FEK3
B	133	HIS	-	expression tag	UNP A8FEK3
C	128	HIS	-	expression tag	UNP A8FEK3
C	129	HIS	-	expression tag	UNP A8FEK3
C	130	HIS	-	expression tag	UNP A8FEK3
C	131	HIS	-	expression tag	UNP A8FEK3
C	132	HIS	-	expression tag	UNP A8FEK3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	133	HIS	-	expression tag	UNP A8FEK3
D	128	HIS	-	expression tag	UNP A8FEK3
D	129	HIS	-	expression tag	UNP A8FEK3
D	130	HIS	-	expression tag	UNP A8FEK3
D	131	HIS	-	expression tag	UNP A8FEK3
D	132	HIS	-	expression tag	UNP A8FEK3
D	133	HIS	-	expression tag	UNP A8FEK3
E	128	HIS	-	expression tag	UNP A8FEK3
E	129	HIS	-	expression tag	UNP A8FEK3
E	130	HIS	-	expression tag	UNP A8FEK3
E	131	HIS	-	expression tag	UNP A8FEK3
E	132	HIS	-	expression tag	UNP A8FEK3
E	133	HIS	-	expression tag	UNP A8FEK3
F	128	HIS	-	expression tag	UNP A8FEK3
F	129	HIS	-	expression tag	UNP A8FEK3
F	130	HIS	-	expression tag	UNP A8FEK3
F	131	HIS	-	expression tag	UNP A8FEK3
F	132	HIS	-	expression tag	UNP A8FEK3
F	133	HIS	-	expression tag	UNP A8FEK3

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



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


- Molecule 3 is water.

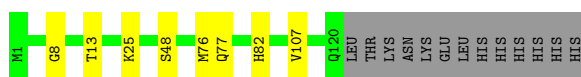
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	149	Total O 149 149	0	0
3	B	139	Total O 139 139	0	0
3	C	148	Total O 148 148	0	0
3	D	125	Total O 125 125	0	0
3	E	127	Total O 127 127	0	0
3	F	132	Total O 132 132	0	0

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.

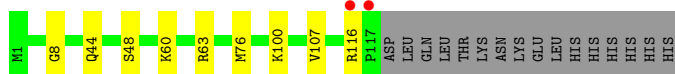
- Molecule 1: Chorismate mutase AroH

Chain A: 




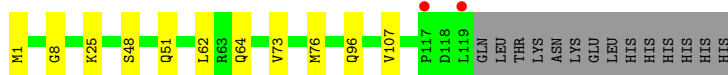
- Molecule 1: Chorismate mutase AroH

Chain B: 




- Molecule 1: Chorismate mutase AroH

Chain C: 




- Molecule 1: Chorismate mutase AroH

Chain D: 




- Molecule 1: Chorismate mutase AroH

Chain E: 



- Molecule 1: Chorismate mutase AroH

Chain F:  %



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	49.90Å 94.86Å 77.31Å 90.00° 107.49° 90.00°	Depositor
Resolution (Å)	24.58 – 1.10 47.43 – 1.10	Depositor EDS
% Data completeness (in resolution range)	78.2 (24.58-1.10) 78.1 (47.43-1.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.07 (at 1.10Å)	Xtrriage
Refinement program	PHENIX 1.20	Depositor
R, R_{free}	0.182 , 0.197 0.178 , 0.193	Depositor DCC
R_{free} test set	1903 reflections (0.88%)	wwPDB-VP
Wilson B-factor (Å ²)	11.4	Xtrriage
Anisotropy	0.114	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.43 , 39.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.019 for h,-k,-h-l	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	12720	wwPDB-VP
Average B, all atoms (Å ²)	17.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.20% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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

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	0.39	0/1047	0.67	0/1419
1	B	0.42	0/1037	0.67	0/1405
1	C	0.40	0/1064	0.69	0/1438
1	D	0.38	0/1014	0.64	0/1373
1	E	0.37	0/1051	0.63	0/1423
1	F	0.40	0/1000	0.67	0/1353
All	All	0.39	0/6213	0.66	0/8411

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	998	1012	979	6	0
1	B	972	986	946	10	0
1	C	1000	1022	968	7	0
1	D	974	985	965	12	0
1	E	997	1013	967	13	0
1	F	958	979	948	11	0
2	E	4	0	6	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	149	0	0	3	2
3	B	139	0	0	7	2
3	C	148	0	0	1	4
3	D	125	0	0	7	1
3	E	127	0	0	7	3
3	F	132	0	0	4	0
All	All	6723	5997	5779	59	6

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 59 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:F:14:GLU:OE2	3:F:201:HOH:O	1.88	0.90
1:D:105:GLN:OE1	3:D:201:HOH:O	1.94	0.85
2:E:201:EDO:O2	3:E:301:HOH:O	1.94	0.85
1:D:23:LYS:HE2	3:D:222:HOH:O	1.76	0.84
1:F:7:ARG:NH2	3:F:202:HOH:O	2.11	0.83

The worst 5 of 6 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)
3:A:301:HOH:O	3:B:218:HOH:O[2_657]	1.83	0.37
3:C:338:HOH:O	3:E:422:HOH:O[1_455]	1.91	0.29
3:B:306:HOH:O	3:C:312:HOH:O[1_655]	2.02	0.18
3:A:348:HOH:O	3:D:314:HOH:O[2_656]	2.03	0.17
3:C:308:HOH:O	3:E:346:HOH:O[1_455]	2.05	0.15

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	127/133 (96%)	126 (99%)	1 (1%)	0	100	100
1	B	125/133 (94%)	124 (99%)	1 (1%)	0	100	100
1	C	127/133 (96%)	125 (98%)	2 (2%)	0	100	100
1	D	122/133 (92%)	121 (99%)	1 (1%)	0	100	100
1	E	127/133 (96%)	125 (98%)	2 (2%)	0	100	100
1	F	120/133 (90%)	118 (98%)	2 (2%)	0	100	100
All	All	748/798 (94%)	739 (99%)	9 (1%)	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	118/122 (97%)	118 (100%)	0	100	100
1	B	116/122 (95%)	116 (100%)	0	100	100
1	C	119/122 (98%)	119 (100%)	0	100	100
1	D	114/122 (93%)	114 (100%)	0	100	100
1	E	118/122 (97%)	116 (98%)	2 (2%)	60	23
1	F	112/122 (92%)	109 (97%)	3 (3%)	44	9
All	All	697/732 (95%)	692 (99%)	5 (1%)	88	57

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	76[A]	MET
1	E	76[B]	MET
1	F	1[A]	MET
1	F	1[B]	MET
1	F	86	LYS

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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

1 ligand is modelled in this entry.

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	$\# Z > 2$
2	EDO	E	201	-	3,3,3	0.46	0	2,2,2	0.12	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
2	EDO	E	201	-	-	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
2	E	201	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
2	E	201	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, 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	120/133 (90%)	-0.15	0 100 100	8, 13, 24, 37	0
1	B	117/133 (87%)	0.00	2 (1%) 70 66	8, 12, 25, 50	0
1	C	119/133 (89%)	-0.15	2 (1%) 70 66	8, 11, 24, 49	0
1	D	118/133 (88%)	0.15	3 (2%) 57 53	10, 16, 30, 62	0
1	E	120/133 (90%)	0.05	1 (0%) 86 84	10, 15, 29, 41	0
1	F	117/133 (87%)	-0.13	1 (0%) 84 82	9, 12, 22, 40	0
All	All	711/798 (89%)	-0.04	9 (1%) 77 73	8, 13, 26, 62	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	117	PRO	5.4
1	D	118	ASP	5.1
1	B	117	PRO	4.7
1	D	116	ARG	3.5
1	C	119	LEU	3.4

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](#)

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, 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	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	EDO	E	201	4/4	0.88	0.11	24,25,26,26	0

6.5 Other polymers [i](#)

There are no such residues in this entry.