WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Aug 15, 2023 – 10:25 PM EDT

PDB ID	:	1XMP
Title	:	Crystal Structure of PurE (BA0288) from Bacillus anthracis at 1.8 Resolution
Authors	:	Boyle, M.P.; Kalliomaa, A.K.; Levdikov, V.; Blagova, E.; Fogg, M.J.; Brannigan, J.A.; Wilkinson, A.J.; Wilson, K.S.
Deposited on	:	2004-10-04
Resolution	:	1.80 Å(reported)

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
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)	:	2.35

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 1.80 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	Percentile Ranks	Value
Rfree		0.212
Clashscore		5
Ramachandran outliers		0
Sidechain outliers		1.8%
RSRZ outliers		2.0%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
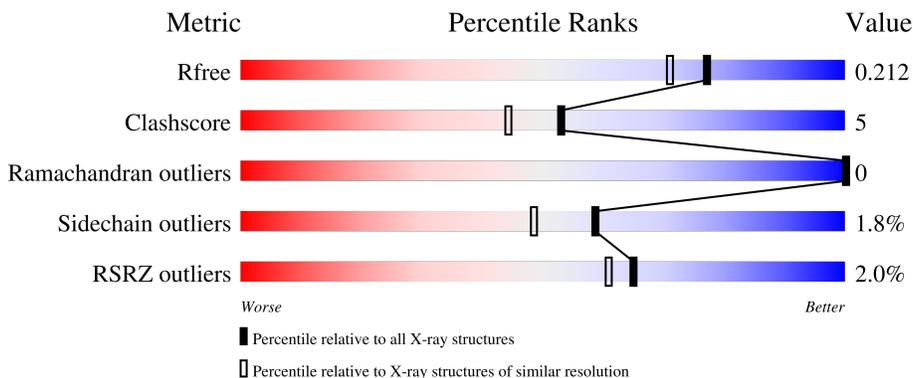
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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 <=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	170	% 79% 12% · 9%
1	B	170	2% 84% 8% 8%
1	C	170	% 88% 9% ·
1	D	170	81% 10% 9%
1	E	170	3% 78% 12% · 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	170	2% ▌ 85% 9% · 6%
1	G	170	2% ▌ 89% 7% ·
1	H	170	5% ▌ 85% 9% · 6%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 10667 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 phosphoribosylaminoimidazole carboxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	155	Total	C	N	O	S	0	0	0
			1151	727	198	220	6			
1	B	157	Total	C	N	O	S	0	0	0
			1161	732	200	223	6			
1	C	164	Total	C	N	O	S	0	0	0
			1231	774	219	231	7			
1	D	155	Total	C	N	O	S	0	0	0
			1151	727	198	220	6			
1	E	155	Total	C	N	O	S	0	0	0
			1151	727	198	220	6			
1	F	160	Total	C	N	O	S	0	0	0
			1186	748	203	229	6			
1	G	164	Total	C	N	O	S	0	0	0
			1231	774	219	231	7			
1	H	160	Total	C	N	O	S	0	0	0
			1186	748	203	229	6			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-8	GLY	-	cloning artifact	UNP Q81ZH8
A	-7	SER	-	cloning artifact	UNP Q81ZH8
A	-6	SER	-	cloning artifact	UNP Q81ZH8
A	-5	HIS	-	expression tag	UNP Q81ZH8
A	-4	HIS	-	expression tag	UNP Q81ZH8
A	-3	HIS	-	expression tag	UNP Q81ZH8
A	-2	HIS	-	expression tag	UNP Q81ZH8
A	-1	HIS	-	expression tag	UNP Q81ZH8
A	0	HIS	-	expression tag	UNP Q81ZH8
B	-8	GLY	-	cloning artifact	UNP Q81ZH8
B	-7	SER	-	cloning artifact	UNP Q81ZH8
B	-6	SER	-	cloning artifact	UNP Q81ZH8
B	-5	HIS	-	expression tag	UNP Q81ZH8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-4	HIS	-	expression tag	UNP Q81ZH8
B	-3	HIS	-	expression tag	UNP Q81ZH8
B	-2	HIS	-	expression tag	UNP Q81ZH8
B	-1	HIS	-	expression tag	UNP Q81ZH8
B	0	HIS	-	expression tag	UNP Q81ZH8
C	-8	GLY	-	cloning artifact	UNP Q81ZH8
C	-7	SER	-	cloning artifact	UNP Q81ZH8
C	-6	SER	-	cloning artifact	UNP Q81ZH8
C	-5	HIS	-	expression tag	UNP Q81ZH8
C	-4	HIS	-	expression tag	UNP Q81ZH8
C	-3	HIS	-	expression tag	UNP Q81ZH8
C	-2	HIS	-	expression tag	UNP Q81ZH8
C	-1	HIS	-	expression tag	UNP Q81ZH8
C	0	HIS	-	expression tag	UNP Q81ZH8
D	-8	GLY	-	cloning artifact	UNP Q81ZH8
D	-7	SER	-	cloning artifact	UNP Q81ZH8
D	-6	SER	-	cloning artifact	UNP Q81ZH8
D	-5	HIS	-	expression tag	UNP Q81ZH8
D	-4	HIS	-	expression tag	UNP Q81ZH8
D	-3	HIS	-	expression tag	UNP Q81ZH8
D	-2	HIS	-	expression tag	UNP Q81ZH8
D	-1	HIS	-	expression tag	UNP Q81ZH8
D	0	HIS	-	expression tag	UNP Q81ZH8
E	-8	GLY	-	cloning artifact	UNP Q81ZH8
E	-7	SER	-	cloning artifact	UNP Q81ZH8
E	-6	SER	-	cloning artifact	UNP Q81ZH8
E	-5	HIS	-	expression tag	UNP Q81ZH8
E	-4	HIS	-	expression tag	UNP Q81ZH8
E	-3	HIS	-	expression tag	UNP Q81ZH8
E	-2	HIS	-	expression tag	UNP Q81ZH8
E	-1	HIS	-	expression tag	UNP Q81ZH8
E	0	HIS	-	expression tag	UNP Q81ZH8
F	-8	GLY	-	cloning artifact	UNP Q81ZH8
F	-7	SER	-	cloning artifact	UNP Q81ZH8
F	-6	SER	-	cloning artifact	UNP Q81ZH8
F	-5	HIS	-	expression tag	UNP Q81ZH8
F	-4	HIS	-	expression tag	UNP Q81ZH8
F	-3	HIS	-	expression tag	UNP Q81ZH8
F	-2	HIS	-	expression tag	UNP Q81ZH8
F	-1	HIS	-	expression tag	UNP Q81ZH8
F	0	HIS	-	expression tag	UNP Q81ZH8
G	-8	GLY	-	cloning artifact	UNP Q81ZH8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	-7	SER	-	cloning artifact	UNP Q81ZH8
G	-6	SER	-	cloning artifact	UNP Q81ZH8
G	-5	HIS	-	expression tag	UNP Q81ZH8
G	-4	HIS	-	expression tag	UNP Q81ZH8
G	-3	HIS	-	expression tag	UNP Q81ZH8
G	-2	HIS	-	expression tag	UNP Q81ZH8
G	-1	HIS	-	expression tag	UNP Q81ZH8
G	0	HIS	-	expression tag	UNP Q81ZH8
H	-8	GLY	-	cloning artifact	UNP Q81ZH8
H	-7	SER	-	cloning artifact	UNP Q81ZH8
H	-6	SER	-	cloning artifact	UNP Q81ZH8
H	-5	HIS	-	expression tag	UNP Q81ZH8
H	-4	HIS	-	expression tag	UNP Q81ZH8
H	-3	HIS	-	expression tag	UNP Q81ZH8
H	-2	HIS	-	expression tag	UNP Q81ZH8
H	-1	HIS	-	expression tag	UNP Q81ZH8
H	0	HIS	-	expression tag	UNP Q81ZH8

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	153	Total 153	O 153	0	0
2	B	178	Total 178	O 178	0	0
2	C	187	Total 187	O 187	0	0
2	D	166	Total 166	O 166	0	0
2	E	115	Total 115	O 115	0	0
2	F	127	Total 127	O 127	0	0
2	G	154	Total 154	O 154	0	0
2	H	139	Total 139	O 139	0	0

3 Residue-property plots ⓘ

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: phosphoribosylaminoimidazole carboxylase

Chain A:

- Molecule 1: phosphoribosylaminoimidazole carboxylase

Chain B:

- Molecule 1: phosphoribosylaminoimidazole carboxylase

Chain C:

- Molecule 1: phosphoribosylaminoimidazole carboxylase

Chain D:

- Molecule 1: phosphoribosylaminoimidazole carboxylase

Chain E:

• Molecule 1: phosphoribosylaminoimidazole carboxylase

Chain F:

• Molecule 1: phosphoribosylaminoimidazole carboxylase

Chain G:

• Molecule 1: phosphoribosylaminoimidazole carboxylase

Chain H:
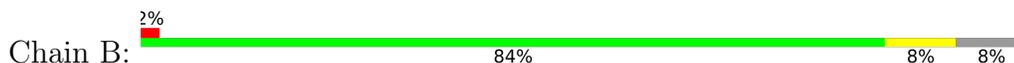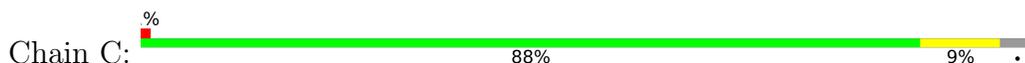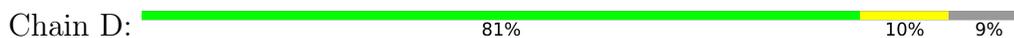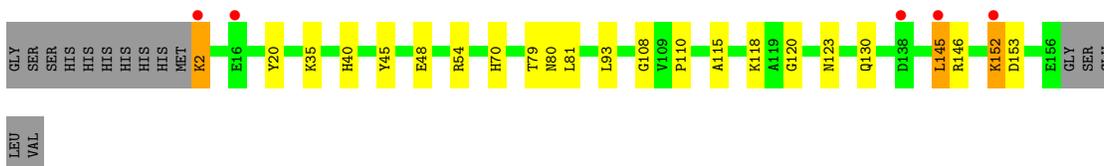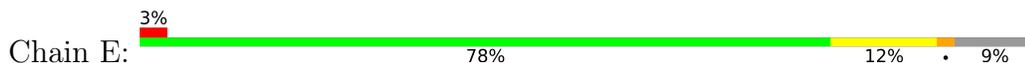
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	168.26Å 76.48Å 102.67Å 90.00° 96.68° 90.00°	Depositor
Resolution (Å)	31.14 – 1.80 31.14 – 1.80	Depositor EDS
% Data completeness (in resolution range)	50.0 (31.14-1.80) 99.8 (31.14-1.80)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.45 (at 1.80Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.167 , 0.205 0.179 , 0.212	Depositor DCC
R_{free} test set	5996 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	17.0	Xtriage
Anisotropy	0.028	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 61.7	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	10667	wwPDB-VP
Average B, all atoms (Å^2)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.01% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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	0.58	0/1169	0.67	0/1584
1	B	0.59	0/1179	0.64	0/1597
1	C	0.56	0/1255	0.68	1/1700 (0.1%)
1	D	0.56	0/1169	0.66	0/1584
1	E	0.56	0/1169	0.62	0/1584
1	F	0.52	0/1204	0.63	1/1630 (0.1%)
1	G	0.57	0/1255	0.65	0/1700
1	H	0.56	0/1204	0.62	0/1630
All	All	0.56	0/9604	0.65	2/13009 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	F	97	ASP	CB-CG-OD1	5.55	123.30	118.30
1	C	54	ARG	NE-CZ-NH1	5.52	123.06	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-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 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	1151	0	1175	26	0
1	B	1161	0	1183	14	0
1	C	1231	0	1239	16	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	1151	0	1175	16	0
1	E	1151	0	1175	24	0
1	F	1186	0	1209	18	0
1	G	1231	0	1239	13	0
1	H	1186	0	1209	16	0
2	A	153	0	0	5	0
2	B	178	0	0	2	0
2	C	187	0	0	3	0
2	D	166	0	0	1	0
2	E	115	0	0	3	0
2	F	127	0	0	1	0
2	G	154	0	0	1	0
2	H	139	0	0	1	0
All	All	10667	0	9604	102	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 102 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:A:8:ILE:HD11	1:A:49:TYR:CZ	1.96	1.00
1:A:55:GLU:HG3	2:A:201:HOH:O	1.75	0.85
1:E:2:LYS:HB2	2:E:231:HOH:O	1.84	0.75
1:B:35:LYS:HD2	2:B:225:HOH:O	1.84	0.75
1:H:20:TYR:CE1	1:H:118:LYS:HG2	2.22	0.74

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	153/170 (90%)	151 (99%)	2 (1%)	0	100	100
1	B	155/170 (91%)	153 (99%)	2 (1%)	0	100	100
1	C	162/170 (95%)	160 (99%)	2 (1%)	0	100	100
1	D	153/170 (90%)	150 (98%)	3 (2%)	0	100	100
1	E	153/170 (90%)	152 (99%)	1 (1%)	0	100	100
1	F	158/170 (93%)	154 (98%)	4 (2%)	0	100	100
1	G	162/170 (95%)	160 (99%)	2 (1%)	0	100	100
1	H	158/170 (93%)	156 (99%)	2 (1%)	0	100	100
All	All	1254/1360 (92%)	1236 (99%)	18 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	121/134 (90%)	119 (98%)	2 (2%)	60	51
1	B	122/134 (91%)	121 (99%)	1 (1%)	81	78
1	C	130/134 (97%)	128 (98%)	2 (2%)	65	56
1	D	121/134 (90%)	120 (99%)	1 (1%)	81	78
1	E	121/134 (90%)	116 (96%)	5 (4%)	30	16
1	F	125/134 (93%)	123 (98%)	2 (2%)	62	54
1	G	130/134 (97%)	128 (98%)	2 (2%)	65	56
1	H	125/134 (93%)	122 (98%)	3 (2%)	49	36
All	All	995/1072 (93%)	977 (98%)	18 (2%)	59	48

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	93	LEU
1	H	118	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	H	93	LEU
1	E	93	LEU
1	G	16	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	40	HIS
1	H	123	ASN
1	G	80	ASN
1	H	40	HIS
1	C	89	GLN

5.3.3 RNA ⓘ

There are no RNA molecules 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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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(Å²)	Q<0.9
1	A	155/170 (91%)	-0.14	1 (0%)	89	87	9, 15, 24, 29	0
1	B	157/170 (92%)	-0.28	4 (2%)	57	52	9, 14, 23, 39	0
1	C	164/170 (96%)	-0.16	1 (0%)	89	87	8, 14, 25, 37	0
1	D	155/170 (91%)	-0.20	0	100	100	10, 15, 22, 30	0
1	E	155/170 (91%)	0.15	5 (3%)	47	41	12, 20, 35, 39	0
1	F	160/170 (94%)	0.13	4 (2%)	57	52	12, 22, 31, 38	0
1	G	164/170 (96%)	-0.08	3 (1%)	68	64	9, 16, 30, 45	0
1	H	160/170 (94%)	0.07	8 (5%)	28	23	12, 19, 30, 50	0
All	All	1270/1360 (93%)	-0.06	26 (2%)	65	61	8, 17, 30, 50	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	161	VAL	7.1
1	B	158	SER	5.3
1	C	-7	SER	4.9
1	G	-7	SER	4.9
1	H	158	SER	4.2

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 monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.