W O R L D W I D E
wwPDB
P R O T E I N D A T A B A N K

wwPDB X-ray Structure Validation Summary Report (i)

May 21, 2020 – 09:18 pm BST

PDB ID	:	2QM0
Title	:	Crystal structure of BES protein from Bacillus cereus
Authors	:	Kim, Y.; Maltseva, N.; Zawadzka, A.; Holzle, D.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)
Deposited on	:	2007-07-13
Resolution	:	1.84 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.84 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		0.240
Clashscore		8
Ramachandran outliers		0.2%
Sidechain outliers		3.0%
RSRZ outliers		7.4%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
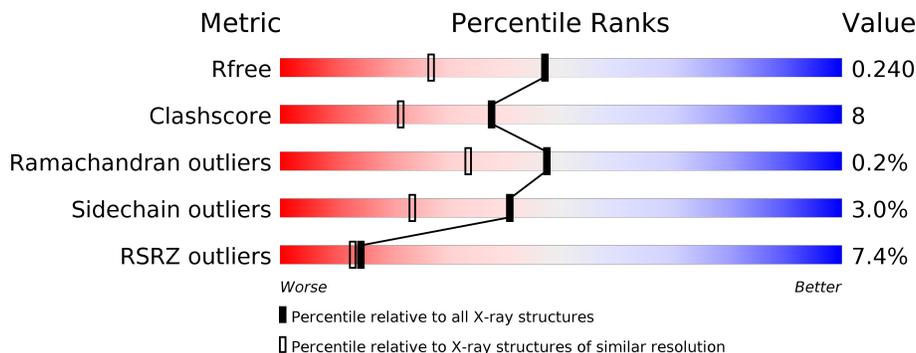
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	4003 (1.86-1.82)
Clashscore	141614	4233 (1.86-1.82)
Ramachandran outliers	138981	4185 (1.86-1.82)
Sidechain outliers	138945	4186 (1.86-1.82)
RSRZ outliers	127900	3957 (1.86-1.82)

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 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	275	7% ▮ 78% 16% · 5%
1	B	275	7% ▮ 87% 9% · ·

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
1	SVY	A	157	X	-	-	-
1	SVY	B	157[A]	X	-	-	-
1	SVY	B	157[B]	X	-	-	-
2	SO4	B	605	-	-	X	-

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4991 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 BES.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
1	A	262	Total	C	N	O	P	S	Se	0	9	0
			2196	1419	363	410	1	1	2			
1	B	268	Total	C	N	O	P	S	Se	0	8	0
			2232	1441	367	419	2	1	2			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	CLONING ARTIFACT	UNP Q81A57
A	-1	ASN	-	CLONING ARTIFACT	UNP Q81A57
A	0	ALA	-	CLONING ARTIFACT	UNP Q81A57
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q81A57
A	20	MSE	MET	MODIFIED RESIDUE	UNP Q81A57
A	157	SVY	SER	MODIFIED RESIDUE	UNP Q81A57
A	223	MSE	MET	MODIFIED RESIDUE	UNP Q81A57
B	-2	SER	-	CLONING ARTIFACT	UNP Q81A57
B	-1	ASN	-	CLONING ARTIFACT	UNP Q81A57
B	0	ALA	-	CLONING ARTIFACT	UNP Q81A57
B	1	MSE	MET	MODIFIED RESIDUE	UNP Q81A57
B	20	MSE	MET	MODIFIED RESIDUE	UNP Q81A57
B	157	SVY	SER	MODIFIED RESIDUE	UNP Q81A57
B	223	MSE	MET	MODIFIED RESIDUE	UNP Q81A57

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	276	Total 276	O 276	0	0
3	B	262	Total 262	O 262	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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: BES

Chain A:

- Molecule 1: BES

Chain B:
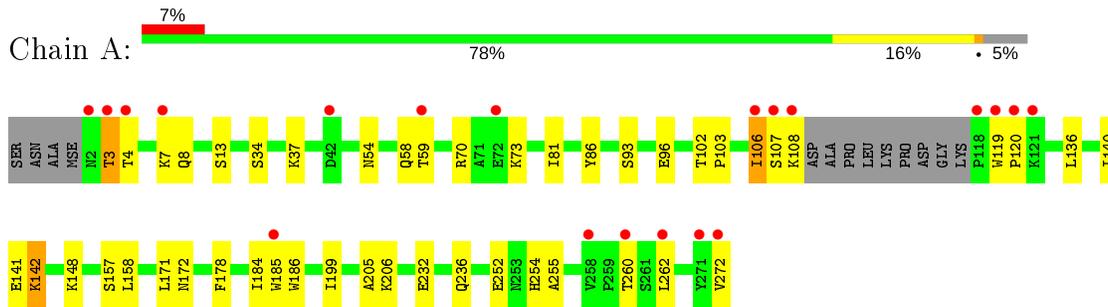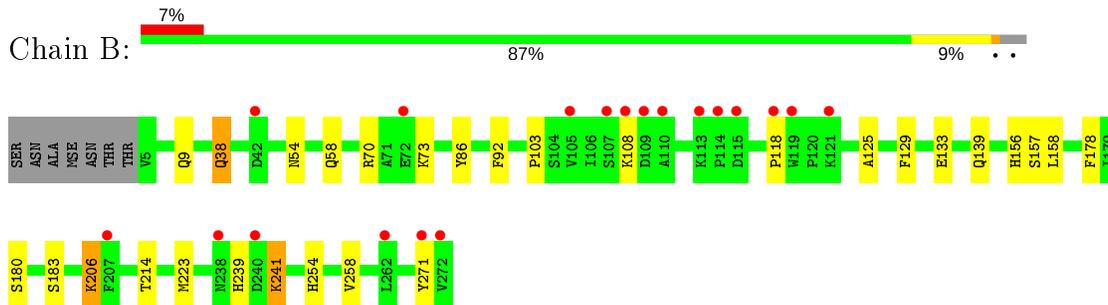
4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants	48.85Å 135.06Å 200.87Å	Depositor
a, b, c, α, β, γ	90.00° 90.00° 90.00°	
Resolution (Å)	45.94 – 1.84	Depositor
	45.94 – 1.84	EDS
% Data completeness	95.5 (45.94-1.84)	Depositor
(in resolution range)	95.5 (45.94-1.84)	EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
$<I/\sigma(I)>$ [1]	1.82 (at 1.84Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.184 , 0.228	Depositor
	0.200 , 0.240	DCC
R_{free} test set	5567 reflections (10.07%)	wwPDB-VP
Wilson B-factor (Å^2)	31.5	Xtriage
Anisotropy	0.517	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 49.6	EDS
L-test for twinning[2]	$<|L|> = 0.46$, $<L^2> = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4991	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.52% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, SVY

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	0.85	0/2233	0.68	0/3010
1	B	0.75	0/2254	0.66	0/3039
All	All	0.80	0/4487	0.67	0/6049

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	1	0
1	B	2	0
All	All	3	0

There are no bond length outliers.

There are no bond angle outliers.

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	157	SVY	CA
1	B	157[A]	SVY	CA
1	B	157[B]	SVY	CA

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	2196	0	2160	38	0
1	B	2232	0	2206	33	0
2	A	10	0	0	0	0
2	B	15	0	0	2	0
3	A	276	0	0	10	0
3	B	262	0	0	7	0
All	All	4991	0	4366	68	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 8.

The worst 5 of 68 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:223:MSE:HB2	2:B:605:SO4:O3	1.27	1.27
1:A:70[B]:ARG:HG2	1:A:73:LYS:HB2	1.27	1.17
1:B:38:GLN:HE21	1:B:38:GLN:H	1.06	1.01
1:B:54:ASN:HD21	1:B:86:TYR:H	1.12	0.98
1:B:223:MSE:CB	2:B:605:SO4:O3	2.11	0.97

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	266/275 (97%)	256 (96%)	9 (3%)	1 (0%)	34	20
1	B	272/275 (99%)	263 (97%)	9 (3%)	0	100	100
All	All	538/550 (98%)	519 (96%)	18 (3%)	1 (0%)	47	33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	107	SER

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	238/236 (101%)	229 (96%)	9 (4%)	33	15
1	B	240/236 (102%)	235 (98%)	5 (2%)	53	38
All	All	478/472 (101%)	464 (97%)	14 (3%)	41	25

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	184	ILE
1	A	199	ILE
1	B	178	PHE
1	A	178	PHE
1	B	38	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	8	GLN
1	B	9	GLN
1	B	38	GLN
1	A	172	ASN
1	B	17	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)

3 non-standard protein/DNA/RNA residues are 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	$\#|Z| > 2$
1	SVY	B	157[B]	1	14,15,16	1.43	3 (21%)	16,20,22	4.56	8 (50%)
1	SVY	A	157	1	14,15,16	1.23	2 (14%)	16,20,22	4.71	10 (62%)
1	SVY	B	157[A]	1	14,15,16	1.49	3 (21%)	16,20,22	5.27	10 (62%)

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
1	SVY	B	157[B]	1	1/1/4/5	4/15/18/20	-
1	SVY	A	157	1	1/1/4/5	3/15/18/20	-
1	SVY	B	157[A]	1	1/1/4/5	3/15/18/20	-

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	157[A]	SVY	P-O3P	2.84	1.58	1.48
1	B	157[B]	SVY	P-O3P	2.70	1.57	1.48
1	B	157[A]	SVY	P-O1P	2.67	1.61	1.57
1	B	157[B]	SVY	P-O1P	2.51	1.61	1.57
1	B	157[A]	SVY	P-OG	2.41	1.63	1.57

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	157[A]	SVY	OG-CB-CA	14.19	121.95	108.14
1	A	157	SVY	OG-CB-CA	13.09	120.89	108.14
1	B	157[B]	SVY	OG-CB-CA	10.62	118.48	108.14
1	B	157[B]	SVY	OG-P-O2P	-9.07	78.67	103.23

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	157[A]	SVY	OG-P-O2P	-8.43	80.39	103.23

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	B	157[B]	SVY	CA
1	A	157	SVY	CA
1	B	157[A]	SVY	CA

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	B	157[B]	SVY	C3'-C1'-O2P-P
1	A	157	SVY	N-CA-CB-OG
1	B	157[A]	SVY	C1-O1P-P-O3P
1	B	157[B]	SVY	C1'-O2P-P-O3P
1	A	157	SVY	CA-CB-OG-P

There are no ring outliers.

3 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	157[B]	SVY	5	0
1	A	157	SVY	3	0
1	B	157[A]	SVY	5	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

5 ligands are 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	SO4	B	605	-	4,4,4	0.56	0	6,6,6	0.86	0
2	SO4	B	601	-	4,4,4	0.16	0	6,6,6	0.25	0
2	SO4	B	603	-	4,4,4	0.23	0	6,6,6	0.35	0
2	SO4	A	602	-	4,4,4	0.19	0	6,6,6	0.44	0
2	SO4	A	604	-	4,4,4	0.37	0	6,6,6	0.44	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	605	SO4	2	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	259/275 (94%)	0.37	20 (7%)	13	12	23, 29, 42, 50	0
1	B	265/275 (96%)	0.44	19 (7%)	15	14	23, 34, 44, 49	0
All	All	524/550 (95%)	0.41	39 (7%)	14	13	23, 32, 44, 50	0

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	119	TRP	7.3
1	B	272	VAL	5.8
1	B	114	PRO	5.1
1	A	272	VAL	4.7
1	B	271	TYR	4.1

6.2 Non-standard residues in protein, DNA, RNA chains (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
1	SVY	B	157[B]	16/17	0.94	0.21	24,28,33,34	16
1	SVY	B	157[A]	16/17	0.94	0.21	23,28,34,34	16
1	SVY	A	157	16/17	0.97	0.13	23,28,37,39	0

6.3 Carbohydrates (i)

There are no carbohydrates 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	SO4	B	605	5/5	0.76	0.39	47,52,55,58	0
2	SO4	B	603	5/5	0.95	0.17	51,55,56,57	0
2	SO4	B	601	5/5	0.97	0.08	56,58,60,60	0
2	SO4	A	604	5/5	0.97	0.12	46,50,52,53	0
2	SO4	A	602	5/5	0.98	0.08	48,50,51,52	0

6.5 Other polymers (i)

There are no such residues in this entry.