WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 20, 2024 – 10:42 pm GMT

PDB ID	:	6ZZ5
Title	:	Structure of soluble SmhB of the tripartite alpha-pore forming toxin, Smh, from Serratia marcescens.
Authors	:	Churchill-Angus, A.M.; Baker, P.J.
Deposited on	:	2020-08-03
Resolution	:	1.84 Å(reported)

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

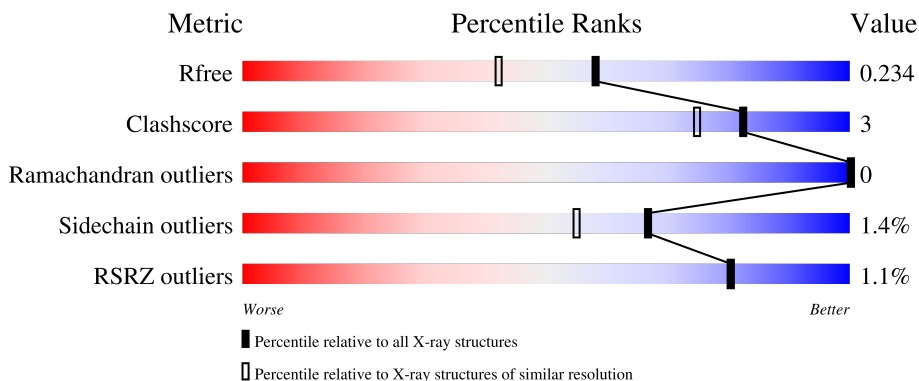|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.84 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		0.234
Clashscore		3
Ramachandran outliers		0
Sidechain outliers		1.4%
RSRZ outliers		1.1%

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 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	AAA	366	% 89% 8% •
1	BBB	366	2% 92% 5% •

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 11418 atoms, of which 5624 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 SmhB.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	356	Total	C	H	N	O	S	173	10	0
			5489	1690	2772	481	536	10			
1	BBB	360	Total	C	H	N	O	S	179	10	0
			5562	1713	2810	491	537	11			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	359	LEU	-	expression tag	UNP A0A1Q4NVM7
AAA	360	GLU	-	expression tag	UNP A0A1Q4NVM7
AAA	361	HIS	-	expression tag	UNP A0A1Q4NVM7
AAA	362	HIS	-	expression tag	UNP A0A1Q4NVM7
AAA	363	HIS	-	expression tag	UNP A0A1Q4NVM7
AAA	364	HIS	-	expression tag	UNP A0A1Q4NVM7
AAA	365	HIS	-	expression tag	UNP A0A1Q4NVM7
AAA	366	HIS	-	expression tag	UNP A0A1Q4NVM7
BBB	359	LEU	-	expression tag	UNP A0A1Q4NVM7
BBB	360	GLU	-	expression tag	UNP A0A1Q4NVM7
BBB	361	HIS	-	expression tag	UNP A0A1Q4NVM7
BBB	362	HIS	-	expression tag	UNP A0A1Q4NVM7
BBB	363	HIS	-	expression tag	UNP A0A1Q4NVM7
BBB	364	HIS	-	expression tag	UNP A0A1Q4NVM7
BBB	365	HIS	-	expression tag	UNP A0A1Q4NVM7
BBB	366	HIS	-	expression tag	UNP A0A1Q4NVM7

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

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	AAA	1	Total 5	O 4	S 1	0	0
2	AAA	1	Total 5	O 4	S 1	0	0
2	AAA	1	Total 5	O 4	S 1	0	0
2	AAA	1	Total 5	O 4	S 1	0	0
2	AAA	1	Total 5	O 4	S 1	0	0
2	BBB	1	Total 5	O 4	S 1	0	0
2	BBB	1	Total 5	O 4	S 1	0	0
2	BBB	1	Total 5	O 4	S 1	0	0

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).

EDO

O2

OH

C1

C2

HO

O1

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	AAA	1	Total 10	C 2	H 6	O 2	1	0
3	AAA	1	Total 10	C 2	H 6	O 2	1	0
3	AAA	1	Total 10	C 2	H 6	O 2	1	0
3	AAA	1	Total 10	C 2	H 6	O 2	1	0
3	BBB	1	Total 10	C 2	H 6	O 2	1	0
3	BBB	1	Total 10	C 2	H 6	O 2	1	0
3	BBB	1	Total 10	C 2	H 6	O 2	1	0

- Molecule 4 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	AAA	138	Total 138	O 138	0	0
4	BBB	119	Total 119	O 119	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: SmhB
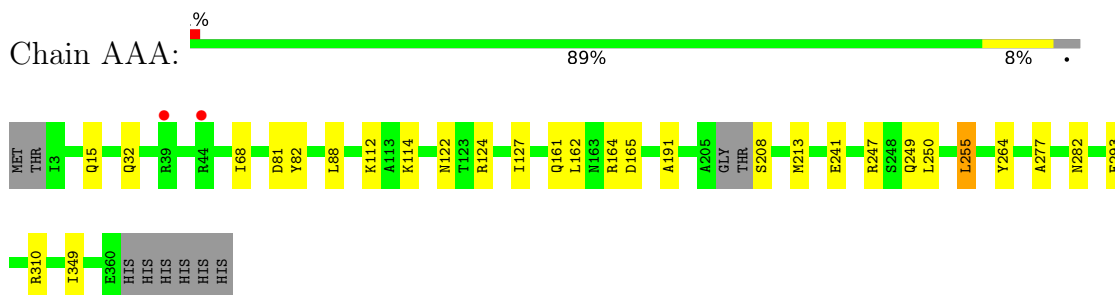
• Molecule 1: SmhB
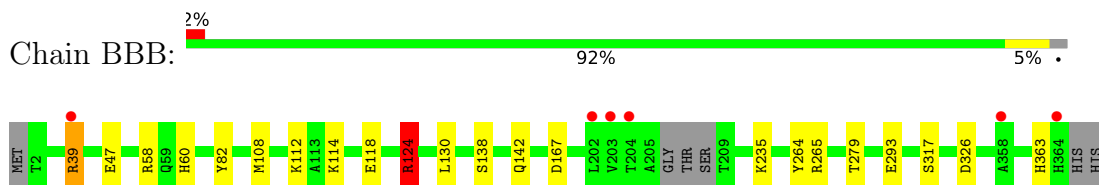
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	50.94Å 113.95Å 130.37Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	65.19 – 1.84 65.19 – 1.84	Depositor EDS
% Data completeness (in resolution range)	99.7 (65.19-1.84) 99.7 (65.19-1.84)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.37 (at 1.84Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.187 , 0.224 0.198 , 0.234	Depositor DCC
R_{free} test set	3268 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å^2)	23.4	Xtriage
Anisotropy	0.432	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.45 , 41.1	EDS
L-test for twinning[2]	$<|L|> = 0.51$, $<L^2> = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	11418	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.13% 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: EDO, SO4

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	AAA	0.84	1/2770 (0.0%)	0.92	1/3754 (0.0%)
1	BBB	0.87	1/2812 (0.0%)	0.91	2/3809 (0.1%)
All	All	0.86	2/5582 (0.0%)	0.91	3/7563 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	BBB	47	GLU	CD-OE2	-5.47	1.19	1.25
1	AAA	241	GLU	CD-OE2	-5.06	1.20	1.25

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	BBB	39	ARG	CG-CD-NE	-7.61	95.81	111.80
1	AAA	310	ARG	NE-CZ-NH1	6.08	123.34	120.30
1	BBB	124	ARG	NE-CZ-NH2	-5.00	117.80	120.30

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	AAA	2717	2772	2775	18	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	BBB	2752	2810	2811	12	4
2	AAA	25	0	0	0	0
2	BBB	15	0	0	0	0
3	AAA	16	24	24	1	0
3	BBB	12	18	18	1	0
4	AAA	138	0	0	2	0
4	BBB	119	0	0	1	0
All	All	5794	5624	5628	28	4

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 3.

The worst 5 of 28 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:AAA:247:ARG:NH1	4:AAA:501:HOH:O	2.01	0.92
1:AAA:165:ASP:OD2	1:AAA:247:ARG:NH1	2.16	0.79
1:AAA:124:ARG:HE	1:AAA:282:ASN:HD22	1.37	0.73
1:AAA:122:ASN:HB3	4:AAA:628:HOH:O	2.01	0.59
1:BBB:58:ARG:NH2	3:BBB:405:EDO:O1	2.36	0.58

All (4) 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)
1:BBB:39:ARG:CZ	1:BBB:326:ASP:OD1[3_654]	1.86	0.34
1:BBB:39:ARG:NH2	1:BBB:326:ASP:OD1[3_654]	1.90	0.30
1:BBB:39:ARG:NH1	1:BBB:326:ASP:OD1[3_654]	1.96	0.24
1:BBB:293:GLU:OE2	1:BBB:363:HIS:HD1[1_455]	1.55	0.05

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	AAA	362/366 (99%)	359 (99%)	3 (1%)	0	100	100
1	BBB	366/366 (100%)	362 (99%)	4 (1%)	0	100	100
All	All	728/732 (100%)	721 (99%)	7 (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	AAA	299/298 (100%)	295 (99%)	4 (1%)	69	58
1	BBB	303/298 (102%)	299 (99%)	4 (1%)	69	58
All	All	602/596 (101%)	594 (99%)	8 (1%)	67	58

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	BBB	317	SER
1	BBB	264	TYR
1	BBB	124	ARG
1	AAA	264	TYR
1	BBB	235	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

15 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	#$|Z| > 2$
3	EDO	AAA	408	-	3,3,3	0.08	0	2,2,2	0.19	0
2	SO4	AAA	401	-	4,4,4	0.26	0	6,6,6	0.18	0
2	SO4	AAA	402	-	4,4,4	0.41	0	6,6,6	0.10	0
2	SO4	AAA	404	-	4,4,4	0.26	0	6,6,6	0.18	0
3	EDO	BBB	404	-	3,3,3	0.15	0	2,2,2	0.08	0
2	SO4	BBB	402	-	4,4,4	0.32	0	6,6,6	0.14	0
3	EDO	AAA	406	-	3,3,3	0.07	0	2,2,2	0.22	0
2	SO4	AAA	405	-	4,4,4	0.40	0	6,6,6	0.17	0
2	SO4	BBB	403	-	4,4,4	0.28	0	6,6,6	0.06	0
2	SO4	BBB	401	-	4,4,4	0.44	0	6,6,6	0.19	0
3	EDO	AAA	409	-	3,3,3	0.17	0	2,2,2	0.23	0
2	SO4	AAA	403	-	4,4,4	0.35	0	6,6,6	0.12	0
3	EDO	BBB	406	-	3,3,3	0.25	0	2,2,2	0.54	0
3	EDO	BBB	405	-	3,3,3	0.13	0	2,2,2	0.40	0
3	EDO	AAA	407	-	3,3,3	0.03	0	2,2,2	0.03	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
3	EDO	AAA	408	-	-	1/1/1/1	-
3	EDO	BBB	404	-	-	1/1/1/1	-
3	EDO	AAA	406	-	-	1/1/1/1	-
3	EDO	AAA	409	-	-	1/1/1/1	-
3	EDO	BBB	406	-	-	1/1/1/1	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	EDO	BBB	405	-	-	1/1/1/1	-
3	EDO	AAA	407	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	AAA	406	EDO	O1-C1-C2-O2
3	AAA	409	EDO	O1-C1-C2-O2
3	BBB	405	EDO	O1-C1-C2-O2
3	BBB	406	EDO	O1-C1-C2-O2
3	AAA	408	EDO	O1-C1-C2-O2

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	AAA	409	EDO	1	0
3	BBB	405	EDO	1	0

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	AAA	356/366 (97%)	0.14	2 (0%)	89	89	16, 23, 43, 65	0
1	BBB	360/366 (98%)	0.12	6 (1%)	70	69	17, 24, 42, 90	0
All	All	716/732 (97%)	0.13	8 (1%)	80	80	16, 24, 42, 90	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	203	VAL	7.6
1	BBB	204	THR	5.0
1	BBB	202	LEU	4.8
1	AAA	39	ARG	3.4
1	BBB	39	ARG	3.2

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 ⓘ

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(Å²)	Q<0.9
3	EDO	BBB	406	4/4	0.54	0.21	40,52,55,61	1
3	EDO	AAA	406	4/4	0.73	0.20	40,43,45,51	1
3	EDO	BBB	405	4/4	0.79	0.12	36,53,55,57	1
3	EDO	AAA	408	4/4	0.85	0.13	40,54,55,60	1
2	SO4	AAA	401	5/5	0.86	0.24	45,48,64,73	0
2	SO4	BBB	403	5/5	0.89	0.14	51,56,93,94	0
2	SO4	AAA	404	5/5	0.90	0.17	39,53,71,81	0
3	EDO	BBB	404	4/4	0.91	0.12	30,35,37,40	1
3	EDO	AAA	409	4/4	0.94	0.22	34,42,50,55	1
3	EDO	AAA	407	4/4	0.95	0.11	40,42,44,53	1
2	SO4	AAA	403	5/5	0.95	0.13	50,51,62,86	0
2	SO4	BBB	401	5/5	0.96	0.14	47,51,58,78	0
2	SO4	BBB	402	5/5	0.96	0.16	40,54,54,80	0
2	SO4	AAA	402	5/5	0.98	0.12	35,40,49,63	0
2	SO4	AAA	405	5/5	0.98	0.11	26,28,34,39	0

6.5 Other polymers ⓘ

There are no such residues in this entry.