



wwPDB X-ray Structure Validation Summary Report

Nov 13, 2023 – 06:07 pm GMT

PDB ID : 7Z56
Title : Crystal Structure of the Ring Nuclease 0455 from *Sulfolobus islandicus* (Sis0455) in its apo form
Authors : Molina, R.; Martin-Garcia, R.; Lopez-Mendez, B.; Jensen, A.L.G.; Marchena-Hurtado, J.; Stella, S.; Montoya, G.
Deposited on : 2022-03-08
Resolution : 2.27 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
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

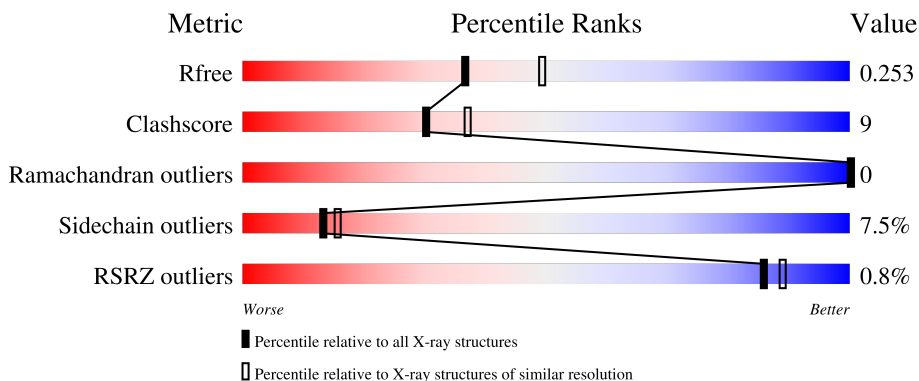
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.27 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	6980 (2.30-2.26)
Clashscore	141614	7711 (2.30-2.26)
Ramachandran outliers	138981	7597 (2.30-2.26)
Sidechain outliers	138945	7598 (2.30-2.26)
RSRZ outliers	127900	6849 (2.30-2.26)

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	AAA	207	
1	BBB	207	

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 2881 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 CRISPR-associated protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	AAA	178	1404	894	236	267	7	0	0	0
1	BBB	183	1447	922	241	277	7	0	0	0

There are 58 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	179	GLY	-	expression tag	UNP F0NGX6
AAA	180	SER	-	expression tag	UNP F0NGX6
AAA	181	GLU	-	expression tag	UNP F0NGX6
AAA	182	PHE	-	expression tag	UNP F0NGX6
AAA	183	GLU	-	expression tag	UNP F0NGX6
AAA	184	LEU	-	expression tag	UNP F0NGX6
AAA	185	GLU	-	expression tag	UNP F0NGX6
AAA	186	ASN	-	expression tag	UNP F0NGX6
AAA	187	LEU	-	expression tag	UNP F0NGX6
AAA	188	TYR	-	expression tag	UNP F0NGX6
AAA	189	PHE	-	expression tag	UNP F0NGX6
AAA	190	GLN	-	expression tag	UNP F0NGX6
AAA	191	GLY	-	expression tag	UNP F0NGX6
AAA	192	GLU	-	expression tag	UNP F0NGX6
AAA	193	LEU	-	expression tag	UNP F0NGX6
AAA	194	ARG	-	expression tag	UNP F0NGX6
AAA	195	ARG	-	expression tag	UNP F0NGX6
AAA	196	GLN	-	expression tag	UNP F0NGX6
AAA	197	ALA	-	expression tag	UNP F0NGX6
AAA	198	SER	-	expression tag	UNP F0NGX6
AAA	199	ALA	-	expression tag	UNP F0NGX6
AAA	200	LEU	-	expression tag	UNP F0NGX6
AAA	201	GLU	-	expression tag	UNP F0NGX6
AAA	202	HIS	-	expression tag	UNP F0NGX6
AAA	203	HIS	-	expression tag	UNP F0NGX6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	204	HIS	-	expression tag	UNP F0NGX6
AAA	205	HIS	-	expression tag	UNP F0NGX6
AAA	206	HIS	-	expression tag	UNP F0NGX6
AAA	207	HIS	-	expression tag	UNP F0NGX6
BBB	179	GLY	-	expression tag	UNP F0NGX6
BBB	180	SER	-	expression tag	UNP F0NGX6
BBB	181	GLU	-	expression tag	UNP F0NGX6
BBB	182	PHE	-	expression tag	UNP F0NGX6
BBB	183	GLU	-	expression tag	UNP F0NGX6
BBB	184	LEU	-	expression tag	UNP F0NGX6
BBB	185	GLU	-	expression tag	UNP F0NGX6
BBB	186	ASN	-	expression tag	UNP F0NGX6
BBB	187	LEU	-	expression tag	UNP F0NGX6
BBB	188	TYR	-	expression tag	UNP F0NGX6
BBB	189	PHE	-	expression tag	UNP F0NGX6
BBB	190	GLN	-	expression tag	UNP F0NGX6
BBB	191	GLY	-	expression tag	UNP F0NGX6
BBB	192	GLU	-	expression tag	UNP F0NGX6
BBB	193	LEU	-	expression tag	UNP F0NGX6
BBB	194	ARG	-	expression tag	UNP F0NGX6
BBB	195	ARG	-	expression tag	UNP F0NGX6
BBB	196	GLN	-	expression tag	UNP F0NGX6
BBB	197	ALA	-	expression tag	UNP F0NGX6
BBB	198	SER	-	expression tag	UNP F0NGX6
BBB	199	ALA	-	expression tag	UNP F0NGX6
BBB	200	LEU	-	expression tag	UNP F0NGX6
BBB	201	GLU	-	expression tag	UNP F0NGX6
BBB	202	HIS	-	expression tag	UNP F0NGX6
BBB	203	HIS	-	expression tag	UNP F0NGX6
BBB	204	HIS	-	expression tag	UNP F0NGX6
BBB	205	HIS	-	expression tag	UNP F0NGX6
BBB	206	HIS	-	expression tag	UNP F0NGX6
BBB	207	HIS	-	expression tag	UNP F0NGX6

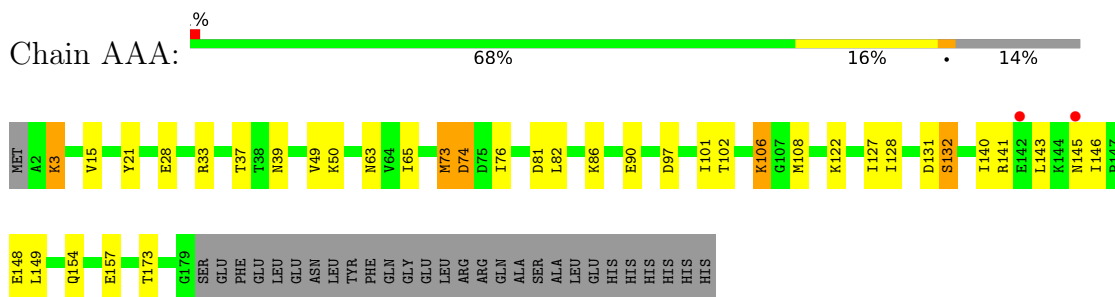
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	AAA	17	Total O 17 17	0	0
2	BBB	13	Total O 13 13	0	0

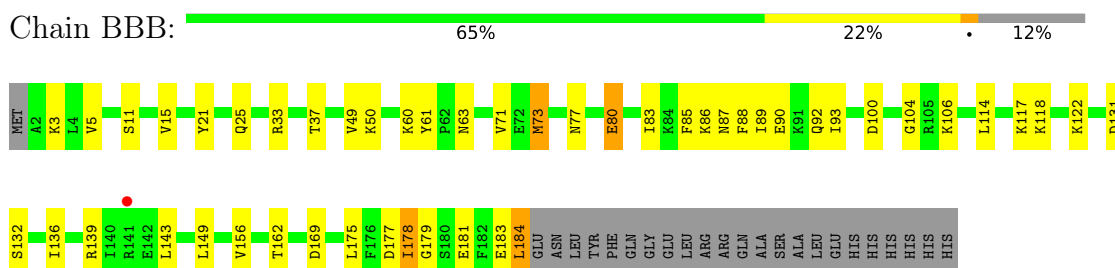
3 Residue-property plots [i](#)

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: CRISPR-associated protein



- Molecule 1: CRISPR-associated protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	42.42Å 51.00Å 80.35Å 90.00° 98.70° 90.00°	Depositor
Resolution (Å)	79.43 – 2.27 79.43 – 2.27	Depositor EDS
% Data completeness (in resolution range)	98.8 (79.43-2.27) 98.8 (79.43-2.27)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.34 (at 2.27Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.181 , 0.253 0.185 , 0.253	Depositor DCC
R_{free} test set	776 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å ²)	46.8	Xtrriage
Anisotropy	0.313	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 48.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	2881	wwPDB-VP
Average B, all atoms (Å ²)	57.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.84% 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](#)

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	0.79	0/1422	0.96	0/1917
1	BBB	0.77	1/1466 (0.1%)	0.97	1/1976 (0.1%)
All	All	0.78	1/2888 (0.0%)	0.97	1/3893 (0.0%)

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	BBB	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	BBB	184	LEU	C-O	7.20	1.37	1.23

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	BBB	33	ARG	NE-CZ-NH2	-5.37	117.61	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	BBB	92	GLN	Peptide

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	1404	0	1454	27	0
1	BBB	1447	0	1491	33	0
2	AAA	17	0	0	1	0
2	BBB	13	0	0	2	0
All	All	2881	0	2945	54	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 9.

The worst 5 of 54 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:BBB:143:LEU:HD21	1:BBB:162:THR:HG21	1.63	0.80
1:BBB:5:VAL:HG21	1:BBB:93:ILE:HD11	1.68	0.74
1:BBB:143:LEU:HD21	1:BBB:162:THR:CG2	2.21	0.71
1:AAA:76:ILE:HD12	1:AAA:108:MET:HG2	1.76	0.68
1:BBB:60:LYS:HG2	1:BBB:61:TYR:CE2	2.30	0.66

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	AAA	176/207 (85%)	173 (98%)	3 (2%)	0	100	100
1	BBB	181/207 (87%)	172 (95%)	9 (5%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	357/414 (86%)	345 (97%)	12 (3%)	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	AAA	157/183 (86%)	145 (92%)	12 (8%)	13	15
1	BBB	162/183 (88%)	150 (93%)	12 (7%)	13	16
All	All	319/366 (87%)	295 (92%)	24 (8%)	13	16

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	BBB	80	GLU
1	BBB	122	LYS
1	BBB	118	LYS
1	BBB	131	ASP
1	AAA	122	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](#)

There are no ligands in this entry.

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	AAA	178/207 (85%)	-0.07	2 (1%) 80 84	32, 50, 78, 127	0
1	BBB	183/207 (88%)	0.01	1 (0%) 91 93	34, 58, 90, 117	0
All	All	361/414 (87%)	-0.03	3 (0%) 86 89	32, 53, 87, 127	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	142	GLU	4.1
1	BBB	141	ARG	2.6
1	AAA	145	ASN	2.4

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.