Full wwPDB X-ray Structure Validation Report ⓘ

Nov 21, 2023 – 07:13 AM JST

PDB ID	:	7F4B
Title	:	The crystal structure of the immature apo-enzyme of homoserine dehydrogenase from the hyperthermophilic archaeon Sulfurisphaera tokodaii.
Authors	:	Kurihara, E.; Kubota, T.; Watanabe, K.; Ogata, K.; Kaneko, R.; Oshima, T.; Yoshimune, K.; Goto, M.
Deposited on	:	2021-06-18
Resolution	:	2.05 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

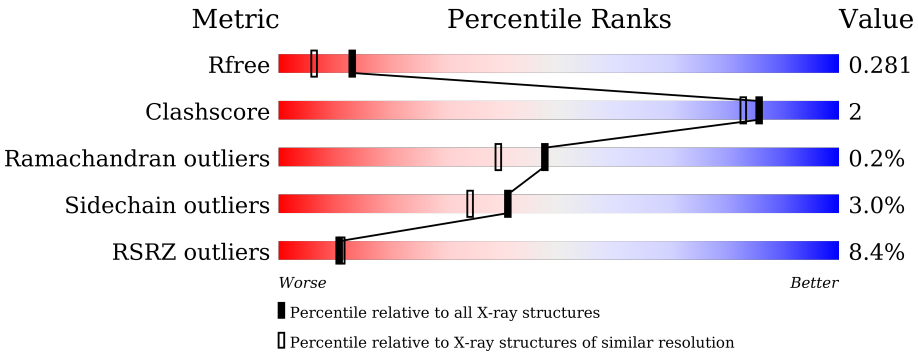
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.05 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	1692 (2.04-2.04)
Clashscore	141614	1773 (2.04-2.04)
Ramachandran outliers	138981	1752 (2.04-2.04)
Sidechain outliers	138945	1752 (2.04-2.04)
RSRZ outliers	127900	1672 (2.04-2.04)

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	304	5% 94% 5%
1	B	304	11% 87% 8%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4623 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 homoserine dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	304	Total	C	N	O	S	0	0	0
			2290	1466	384	434	6			
1	B	292	Total	C	N	O	S	0	0	0
			2167	1391	357	413	6			

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Mg	0	0
			2	2		

- Molecule 3 is water.

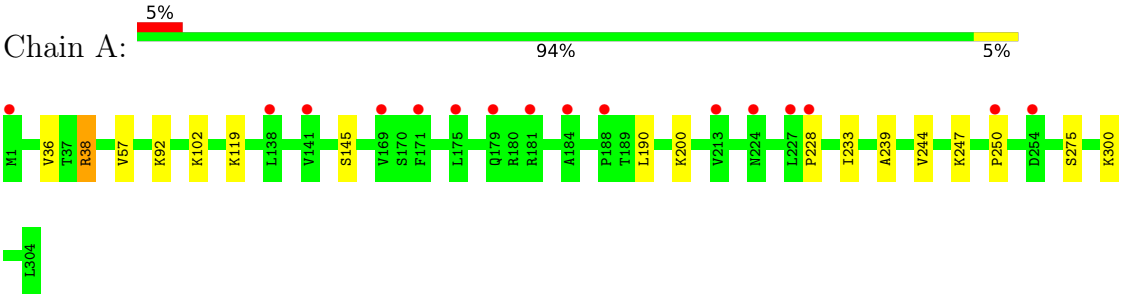
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	80	Total	O	0	0
			80	80		
3	B	84	Total	O	0	0
			84	84		

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.

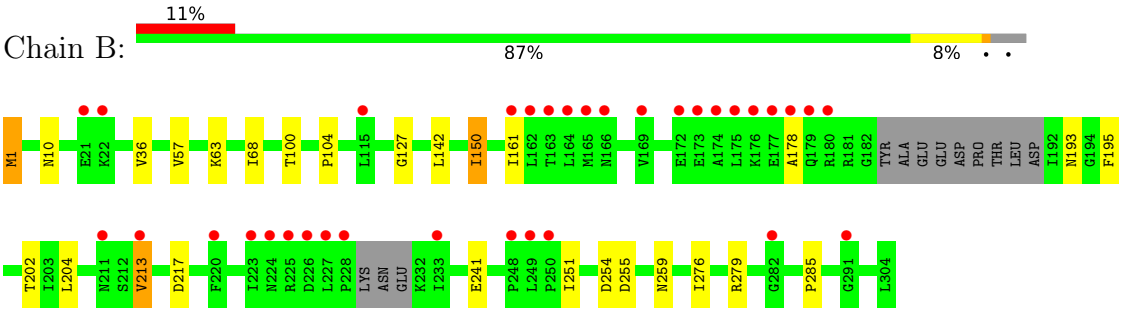
• Molecule 1: homoserine dehydrogenase

Chain A:



• Molecule 1: homoserine dehydrogenase

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	57.10Å 78.31Å 65.55Å 90.00° 106.17° 90.00°	Depositor
Resolution (Å)	48.59 – 2.05 36.61 – 2.05	Depositor EDS
% Data completeness (in resolution range)	99.6 (48.59-2.05) 99.6 (36.61-2.05)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	2.18 (at 2.05Å)	Xtriage
Refinement program	REFMAC 5.8.0131	Depositor
R, R_{free}	0.231 , 0.281 0.231 , 0.281	Depositor DCC
R_{free} test set	1682 reflections (4.84%)	wwPDB-VP
Wilson B-factor (Å^2)	38.5	Xtriage
Anisotropy	0.104	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 60.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	4623	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.21% 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

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

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	0.50	0/2327	0.71	0/3157
1	B	0.49	0/2199	0.69	0/2982
All	All	0.50	0/4526	0.70	0/6139

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

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	2290	0	2305	9	0
1	B	2167	0	2172	11	0
2	A	2	0	0	0	0
3	A	80	0	0	1	0
3	B	84	0	0	0	0
All	All	4623	0	4477	20	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 2.

All (20) 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:38:ARG:H	1:A:38:ARG:HD2	1.62	0.64
1:B:213:VAL:HG23	1:B:217:ASP:HB2	1.81	0.63
1:A:38:ARG:H	1:A:38:ARG:CD	2.13	0.62
1:A:102:LYS:HD3	1:A:200:LYS:HE3	1.84	0.60
1:B:259:ASN:ND2	1:B:279:ARG:HH12	2.03	0.56
1:B:100:THR:HG22	1:B:127:GLY:HA3	1.96	0.47
1:B:202:THR:HG23	1:B:213:VAL:HG22	1.96	0.47
1:A:102:LYS:HE2	3:A:629:HOH:O	2.14	0.46
1:A:36:VAL:HG11	1:A:57:VAL:HG23	1.98	0.45
1:B:36:VAL:HG11	1:B:57:VAL:HG23	1.99	0.45
1:B:193:ASN:HD21	1:B:195:PHE:HB2	1.81	0.45
1:B:142:LEU:HD11	1:B:276:ILE:HD11	1.99	0.44
1:B:150:ILE:HD12	1:B:204:LEU:HB3	2.00	0.44
1:A:300:LYS:HA	1:A:300:LYS:HD2	1.81	0.43
1:B:161:ILE:HG12	1:B:178:ALA:HB2	2.01	0.42
1:A:239:ALA:HB2	1:A:244:VAL:HG13	2.02	0.42
1:A:233:ILE:HD13	1:A:250:PRO:HA	2.02	0.41
1:B:1:MET:HG2	1:B:68:ILE:CD1	2.51	0.41
1:A:92:LYS:HA	1:A:119:LYS:HD3	2.02	0.41
1:B:10:ASN:HB3	1:B:285:PRO:HB3	2.03	0.41

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	302/304 (99%)	288 (95%)	13 (4%)	1 (0%)	41 31
1	B	286/304 (94%)	269 (94%)	17 (6%)	0	100 100
All	All	588/608 (97%)	557 (95%)	30 (5%)	1 (0%)	47 39

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	228	PRO

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	241/256 (94%)	236 (98%)	5 (2%)	53 48
1	B	226/256 (88%)	217 (96%)	9 (4%)	31 24
All	All	467/512 (91%)	453 (97%)	14 (3%)	41 34

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	38	ARG
1	A	145	SER
1	A	190	LEU
1	A	247	LYS
1	A	275	SER
1	B	1	MET
1	B	63	LYS
1	B	104	PRO
1	B	150	ILE
1	B	213	VAL
1	B	241	GLU
1	B	251	ILE
1	B	254	ASP
1	B	255	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	110	ASN
1	B	193	ASN
1	B	224	ASN
1	B	265	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	287	ASN

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

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

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.

No monomer is involved in short contacts.

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, 95^{th} 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	304/304 (100%)	0.41	16 (5%) 26 28	28, 42, 64, 89	0
1	B	292/304 (96%)	0.72	34 (11%) 4 4	30, 43, 92, 120	0
All	All	596/608 (98%)	0.56	50 (8%) 11 11	28, 43, 77, 120	0

All (50) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	226	ASP	8.0
1	B	169	VAL	7.4
1	B	161	ILE	7.4
1	B	228	PRO	6.5
1	B	223	ILE	6.2
1	B	233	ILE	5.9
1	B	163	THR	5.5
1	B	174	ALA	5.1
1	B	162	LEU	4.6
1	B	175	LEU	4.6
1	B	179	GLN	4.4
1	B	172	GLU	4.2
1	A	254	ASP	4.1
1	B	173	GLU	4.0
1	B	165	MET	3.6
1	A	175	LEU	3.5
1	B	249	LEU	3.4
1	B	177	GLU	3.2
1	B	225	ARG	3.2
1	B	178	ALA	3.1
1	A	171	PHE	3.1
1	B	227	LEU	3.1
1	B	180	ARG	3.1
1	A	228	PRO	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	184	ALA	2.9
1	B	164	LEU	2.9
1	B	250	PRO	2.9
1	A	227	LEU	2.8
1	B	115	LEU	2.7
1	A	250	PRO	2.7
1	B	211	ASN	2.6
1	A	141	VAL	2.5
1	B	248	PRO	2.5
1	A	224	ASN	2.5
1	B	176	LYS	2.5
1	A	169	VAL	2.4
1	B	220	PHE	2.4
1	A	213	VAL	2.3
1	A	179	GLN	2.2
1	B	166	ASN	2.2
1	B	282	GLY	2.2
1	B	22	LYS	2.2
1	A	181	ARG	2.2
1	B	213	VAL	2.1
1	B	291	GLY	2.1
1	B	21	GLU	2.1
1	B	224	ASN	2.1
1	A	1	MET	2.1
1	A	138	LEU	2.0
1	A	188	PRO	2.0

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, 95^{th} 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	MG	A	502	1/1	0.83	0.08	50,50,50,50	0
2	MG	A	501	1/1	0.93	0.05	41,41,41,41	0

6.5 Other polymers (i)

There are no such residues in this entry.