wwPDB X-ray Structure Validation Summary Report (i)

Apr 29, 2024 – 10:41 am BST

PDB ID	:	2CJB
Title	:	Crystal structure of Methanosarcina barkeri seryl-tRNA synthetase complexed with serine
Authors	:	Bilokapic, S.; Maier, T.; Ahel, D.; Gruic-Sovulj, I.; Soll, D.; Weygand-Durasevic, I.; Ban, N.
Deposited on	:	2006-03-30
Resolution	:	2.70 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 (i) 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 (i)) were used in the production of this report:

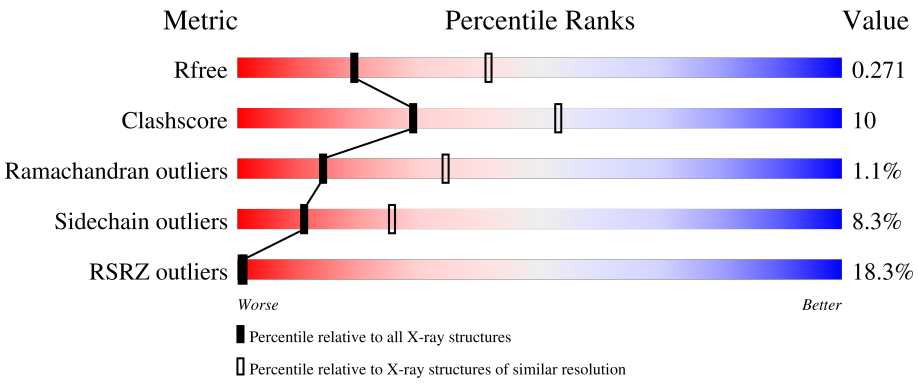
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	522	20% 67% 21% • 9%
1	B	522	14% 71% 22% • 5%

2 Entry composition (i)

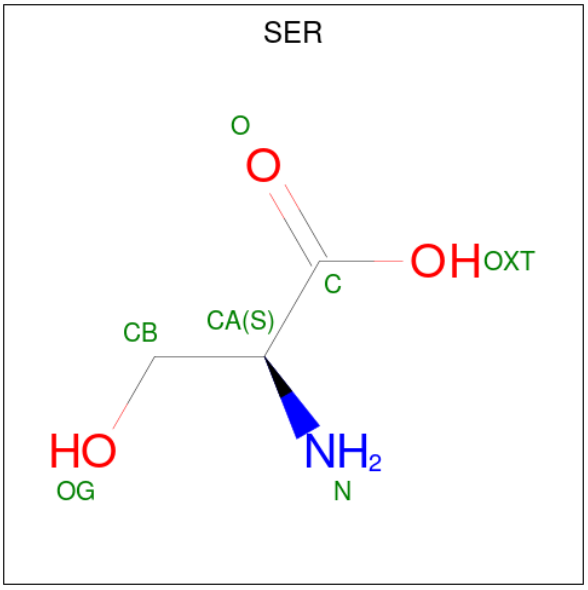
There are 5 unique types of molecules in this entry. The entry contains 8045 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 SERYL-TRNA SYNTHETASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	474	3886	2495	660	710	21	0	1	0
1	B	495	4042	2596	688	737	21	0	1	0

- Molecule 2 is SERINE (three-letter code: SER) (formula: $C_3H_7NO_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	7	3	1	3	0	0
2	B	1	7	3	1	3	0	0

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
3	A	1	1	1	0	0
			Total	Zn		
3	B	1	1	1	0	0

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
4	A	1	1	1	0	0
			Total	Cl		
4	B	1	1	1	0	0

- Molecule 5 is water.

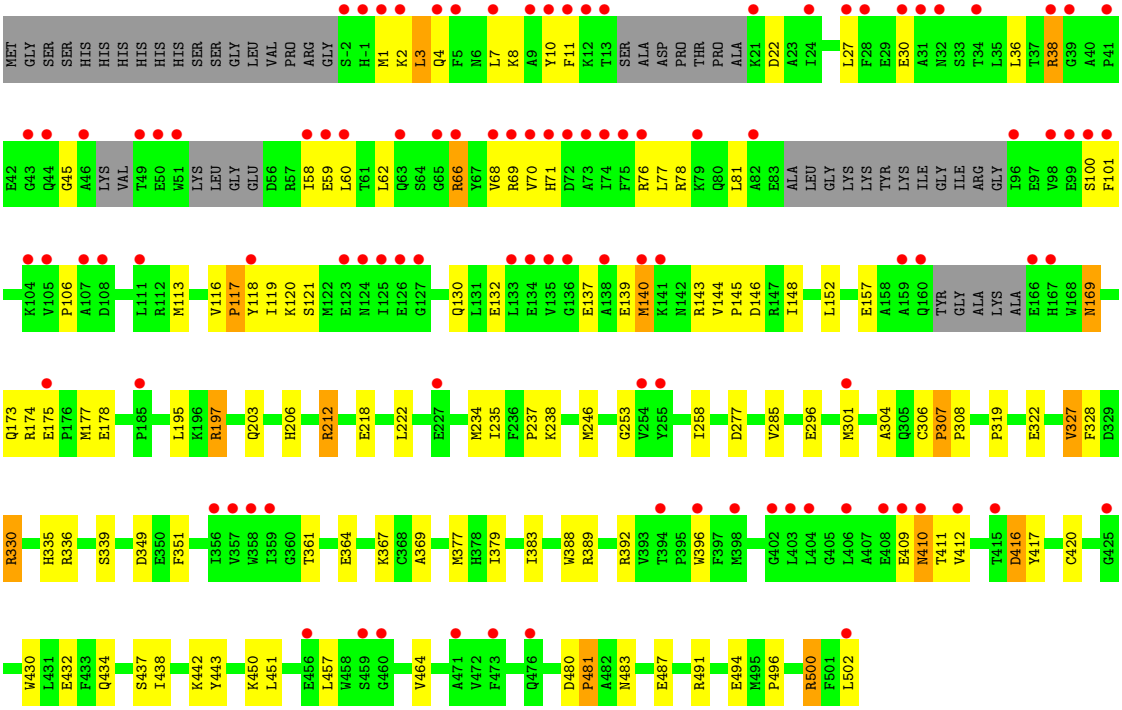
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	57	57	57	0	0
			Total	O		
5	B	42	42	42	0	0

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.

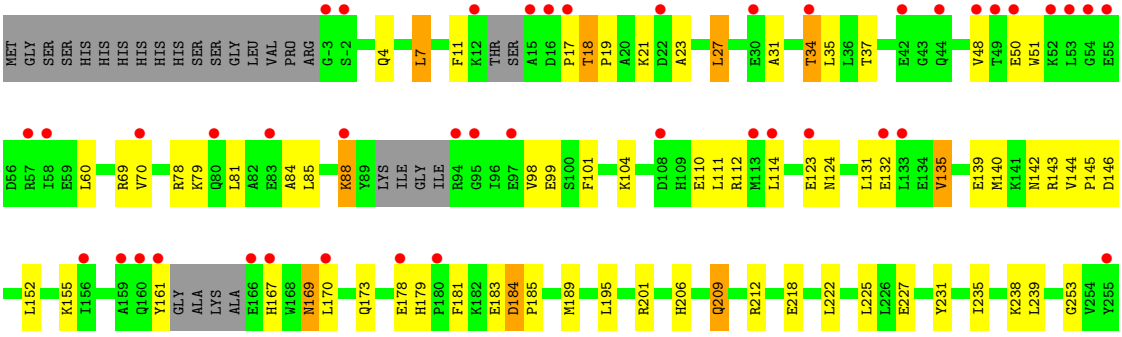
• Molecule 1: SERYL-TRNA SYNTHETASE

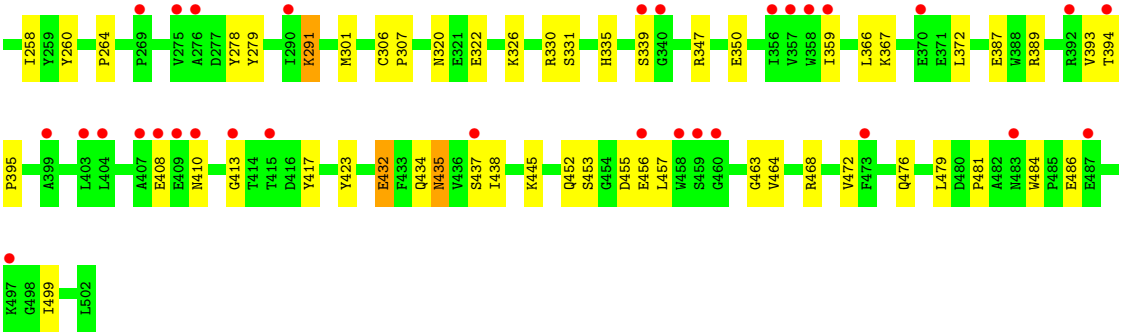
Chain A:



• Molecule 1: SERYL-TRNA SYNTHETASE

Chain B:




I258
Y259
Y260
P264
P269
V275
A276
D277
Y278
Y279
I290
K291
M301
C306
P307
N320
E321
E322
K326
R330
S331
H335
S339
G340
R347
E350
I356
V357
W358
I359
L366
K367
E370
E371
L372
E387
W388
R389
R392
V393
T394
P395
A399
L403
L404
A407
E408
E409
N410
G413
T414
T415
D416
Y417
Y423
E432
F433
Q434
N435
V436
S437
I438
K445
Q452
S453
G454
D455
E456
L457
W458
S459
G460
G463
V464
R468
V472
F473
Q476
L479
D480
P481
A482
M483
W484
P485
E486
E487
K497
G498
I499
L502

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	96.93Å 96.93Å 270.50Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.94 – 2.70 19.94 – 2.70	Depositor EDS
% Data completeness (in resolution range)	98.7 (19.94-2.70) 98.7 (19.94-2.70)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.99 (at 2.71Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.217 , 0.272 0.225 , 0.271	Depositor DCC
R_{free} test set	2036 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	69.1	Xtriage
Anisotropy	0.316	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.24 , 77.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.017 for -h,-k,l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	8045	wwPDB-VP
Average B, all atoms (Å^2)	102.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.03% 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: CL, ZN

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	0.52	0/3988	0.64	0/5390
1	B	0.49	0/4150	0.64	2/5610 (0.0%)
All	All	0.51	0/8138	0.64	2/11000 (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	B	184	ASP	CB-CG-OD1	5.82	123.54	118.30
1	B	184	ASP	CB-CG-OD2	-5.15	113.66	118.30

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	A	3886	0	3795	82	0
1	B	4042	0	3962	82	0
2	A	7	0	4	1	0
2	B	7	0	4	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	1	0	0	0	0
5	A	57	0	0	4	0
5	B	42	0	0	3	0
All	All	8045	0	7765	150	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 10.

The worst 5 of 150 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:1:MET:HG2	1:A:2:LYS:H	1.27	1.00
1:A:500:ARG:HH11	1:A:500:ARG:HG3	1.22	0.98
1:A:330:ARG:HH11	1:A:330:ARG:HG3	1.27	0.96
1:A:417:TYR:HB2	1:A:434:GLN:HB3	1.54	0.90
1:A:197:ARG:HD2	5:A:2038:HOH:O	1.73	0.89

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	463/522 (89%)	421 (91%)	35 (8%)	7 (2%)	10 26
1	B	488/522 (94%)	452 (93%)	33 (7%)	3 (1%)	25 50
All	All	951/1044 (91%)	873 (92%)	68 (7%)	10 (1%)	14 34

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	45	GLY

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	38	ARG
1	A	4	GLN
1	A	253	GLY
1	B	18	THR

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	410/444 (92%)	376 (92%)	34 (8%)	11	25
1	B	424/444 (96%)	389 (92%)	35 (8%)	11	25
All	All	834/888 (94%)	765 (92%)	69 (8%)	11	25

5 of 69 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	239	LEU
1	B	347	ARG
1	B	437	SER
1	A	330	ARG
1	A	327	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	343	HIS
1	B	435	ASN
1	B	6	ASN
1	B	109	HIS
1	B	124	ASN

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)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 for Mogul analysis.

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	SER	B	1505	3	5,6,6	0.95	0	5,7,7	1.36	2 (40%)
2	SER	A	1505	3	5,6,6	1.01	0	5,7,7	1.34	1 (20%)

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
2	SER	B	1505	3	-	0/6/6/6	-
2	SER	A	1505	3	-	0/6/6/6	-

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1505	SER	OXT-C-O	-2.37	118.72	124.09
2	B	1505	SER	OXT-C-O	-2.12	119.29	124.09
2	B	1505	SER	OXT-C-CA	2.01	120.24	113.38

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1505	SER	1	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	474/522 (90%)	1.18	103 (21%)	0	0	85, 96, 152, 163	0
1	B	495/522 (94%)	0.83	74 (14%)	2	1	81, 96, 110, 138	0
All	All	969/1044 (92%)	1.00	177 (18%)	1	1	81, 96, 143, 163	0

The worst 5 of 177 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	11	PHE	13.1
1	A	-2	SER	11.4
1	A	43	GLY	8.0
1	A	71	HIS	7.9
1	A	28	PHE	7.9

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, 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
4	CL	A	1504	1/1	0.92	0.14	56,56,56,56	0
2	SER	A	1505	7/7	0.94	0.51	91,92,93,93	0
3	ZN	A	1503	1/1	0.95	0.22	83,83,83,83	0
2	SER	B	1505	7/7	0.95	0.40	73,74,74,74	0
3	ZN	B	1503	1/1	0.98	0.18	81,81,81,81	0
4	CL	B	1504	1/1	0.98	0.12	58,58,58,58	0

6.5 Other polymers (i)

There are no such residues in this entry.