wwPDB X-ray Structure Validation Summary Report (i)

Oct 6, 2023 – 04:36 AM EDT

PDB ID	:	8SSD
Title	:	Methionine synthase, C-terminal fragment, Cobalamin and Reactivation Domains from Thermus thermophilus HB8
Authors	:	Yamada, K.; Mendoza, J.; Koutmos, M.
Deposited on	:	2023-05-08
Resolution	:	2.40 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 (1)) were used in the production of this report:

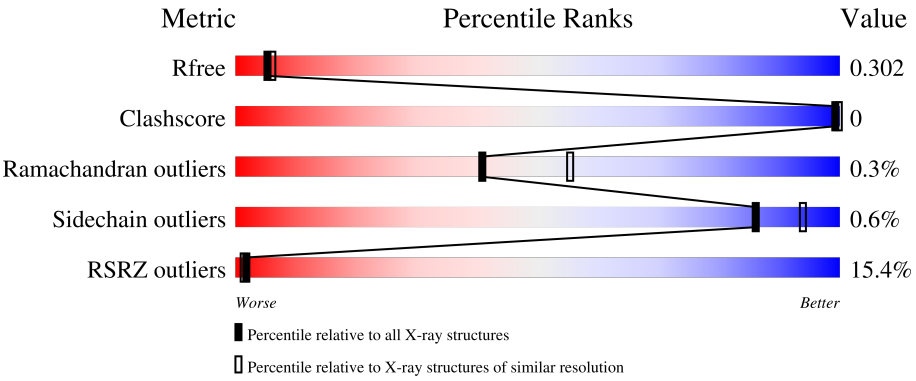
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	523	15% 94%
1	B	523	12% 94%
1	C	523	17% 94%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 12452 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 Methionine synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	506	Total	C	N	O	S	0	0	0
			4003	2566	703	718	16			
1	C	505	Total	C	N	O	S	0	0	0
			3987	2558	699	715	15			
1	B	502	Total	C	N	O	S	0	1	0
			3977	2551	696	714	16			

- Molecule 2 is water.

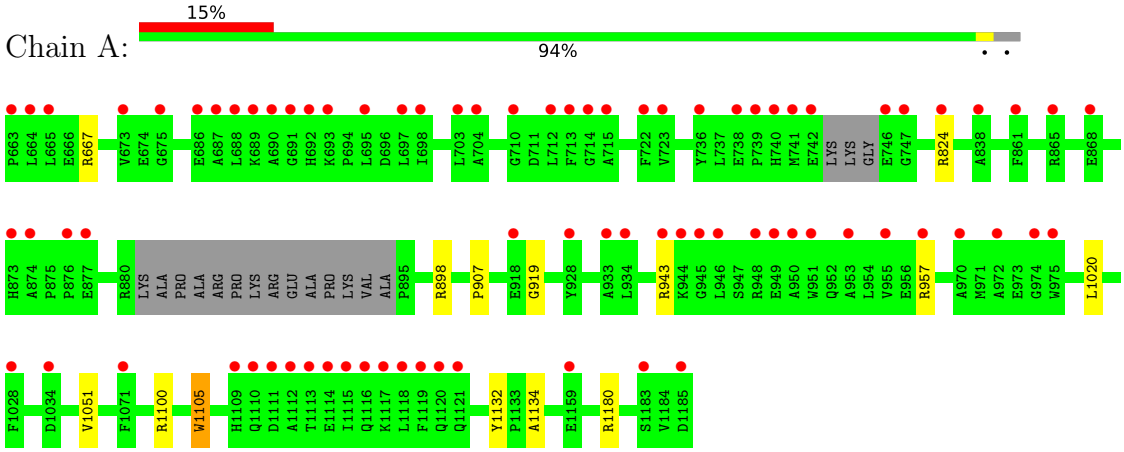
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	176	Total	O	0	0
			176	176		
2	C	150	Total	O	0	0
			150	150		
2	B	159	Total	O	0	0
			159	159		

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.

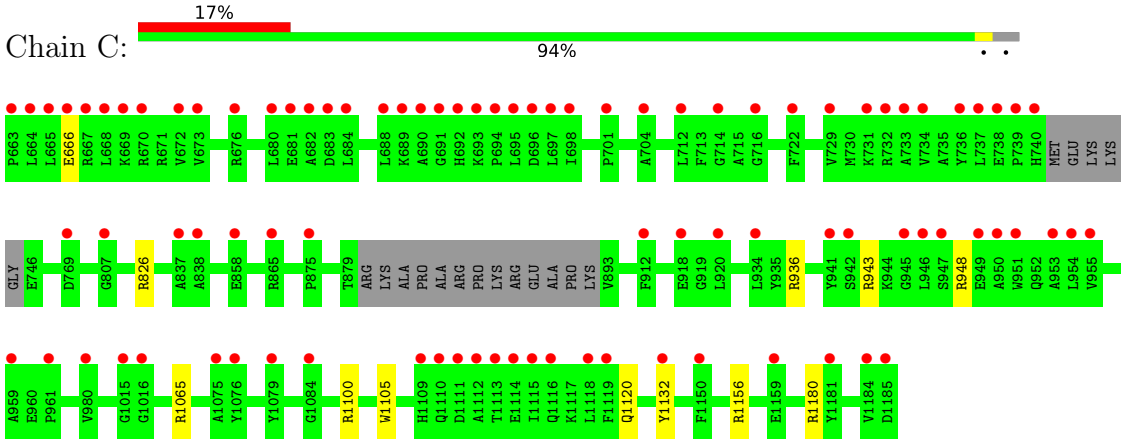
• Molecule 1: Methionine synthase

Chain A:



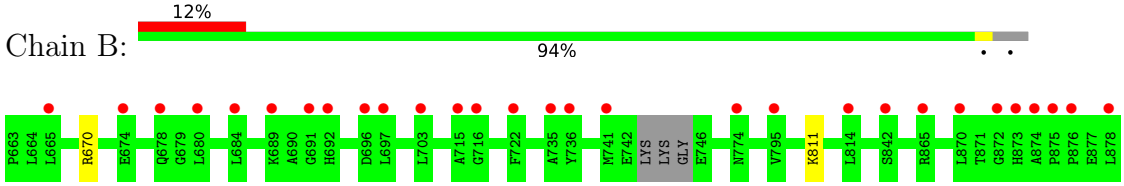
• Molecule 1: Methionine synthase

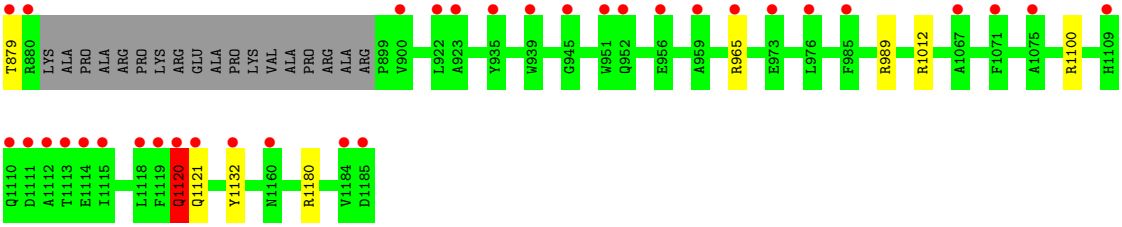
Chain C:



• Molecule 1: Methionine synthase

Chain B:




T879
R880
LYS
ALA
PRO
ALA
ARG
PRO
LYS
ARG
GLU
ALA
PRO
LYS
VAL
ALA
PRO
ARG
ALA
ARG
P899
V900
L922
A923
Y935
W939
G945
W951
Q952
E956
A959
R965
E973
L976
F985
R989
R1012
A1067
F1071
A1075
R1100
H1109
Q1110
D1111
A1112
T1113
E1114
I1115
L1118
F1119
Q1120
Q1121
Y1132
N1160
R1180
V1184
D1185

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	96.18Å 96.18Å 356.04Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	39.85 – 2.40 39.85 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.9 (39.85-2.40) 100.0 (39.85-2.40)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.34 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.8.0411	Depositor
R, R_{free}	0.265 , 0.301 0.268 , 0.302	Depositor DCC
R_{free} test set	3830 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	40.3	Xtriage
Anisotropy	0.091	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 36.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.40$, $< L^2 > = 0.22$	Xtriage
Estimated twinning fraction	0.086 for -h,-k,l	Xtriage
Reported twinning fraction	0.831 for H, K, L 0.169 for -h,-k,l	Depositor
Outliers	0 of 75948 reflections	Xtriage
F_o,F_c correlation	0.86	EDS
Total number of atoms	12452	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 9.10% 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

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	0.45	0/4098	0.83	2/5537 (0.0%)
1	B	0.44	0/4074	0.85	4/5505 (0.1%)
1	C	0.43	0/4082	0.84	3/5519 (0.1%)
All	All	0.44	0/12254	0.84	9/16561 (0.1%)

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	0	6
1	B	0	4
1	C	0	5
All	All	0	15

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	1100	ARG	NE-CZ-NH1	6.04	123.32	120.30
1	A	1100	ARG	NE-CZ-NH1	5.96	123.28	120.30
1	C	1180	ARG	NE-CZ-NH2	5.88	123.24	120.30
1	B	1100	ARG	NE-CZ-NH1	5.79	123.20	120.30
1	B	1120	GLN	CB-CA-C	5.66	121.72	110.40

There are no chirality outliers.

5 of 15 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	667	ARG	Sidechain
1	A	824	ARG	Sidechain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	898	ARG	Sidechain
1	A	943	ARG	Sidechain
1	A	957	ARG	Sidechain

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	4003	0	4000	4	0
1	B	3977	0	3975	0	0
1	C	3987	0	3985	0	0
2	A	176	0	0	0	0
2	B	159	0	0	0	0
2	C	150	0	0	0	0
All	All	12452	0	11960	4	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 0.

All (4) 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:1020:LEU:CD1	1:A:1134:ALA:HB1	2.45	0.46
1:A:1020:LEU:HD13	1:A:1134:ALA:HB1	2.00	0.44
1:A:1051:VAL:HB	1:A:1105:TRP:HH2	1.82	0.43
1:A:907:PRO:O	1:A:1105:TRP:CZ3	2.72	0.42

There are no symmetry-related clashes.

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	A	500/523 (96%)	490 (98%)	9 (2%)	1 (0%)	47 62
1	B	497/523 (95%)	485 (98%)	9 (2%)	3 (1%)	25 36
1	C	499/523 (95%)	485 (97%)	14 (3%)	0	100 100
All	All	1496/1569 (95%)	1460 (98%)	32 (2%)	4 (0%)	41 55

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	879	THR
1	B	1120	GLN
1	A	919	GLY
1	B	1121	GLN

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	407/419 (97%)	406 (100%)	1 (0%)	93 97
1	B	405/419 (97%)	403 (100%)	2 (0%)	88 95
1	C	405/419 (97%)	401 (99%)	4 (1%)	76 88
All	All	1217/1257 (97%)	1210 (99%)	7 (1%)	86 94

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	1120	GLN
1	C	1156	ARG
1	B	1120	GLN
1	B	811	LYS
1	C	1105	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such

sidechains are listed below:

Mol	Chain	Res	Type
1	B	725	GLN
1	B	1077	GLN

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

There are no ligands in this entry.

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	506/523 (96%)	1.09	80 (15%)	2 1	20, 41, 84, 135	0
1	B	502/523 (95%)	1.01	63 (12%)	3 3	23, 42, 76, 100	0
1	C	505/523 (96%)	1.17	90 (17%)	1 1	20, 40, 93, 130	0
All	All	1513/1569 (96%)	1.09	233 (15%)	2 1	20, 41, 86, 135	0

The worst 5 of 233 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1115	ILE	14.4
1	C	1112	ALA	11.6
1	B	876	PRO	11.5
1	A	1119	PHE	9.7
1	C	945	GLY	8.4

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

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.