



# wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 30, 2026 – 12:09 PM EDT

PDB ID : 9YBZ / pdb\_00009ybz  
Title : CRYSTAL STRUCTURE OF THE A149T VARIANT OF SERINE HYDROXYMETHYLTRANSFERASE 8 FROM SOYBEAN CULTIVAR ESSEX IN COMPLEX WITH PLP  
Authors : Beamer, L.J.; Samarakoon, V.; Owuocha, L.F.  
Deposited on : 2025-09-17  
Resolution : 2.28 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](mailto: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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

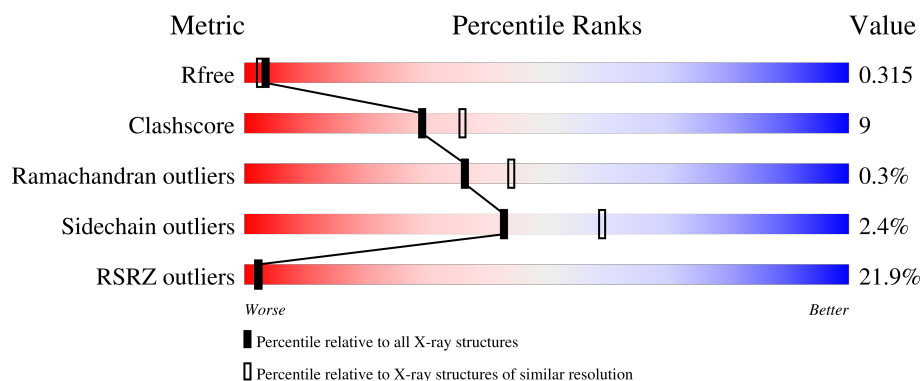
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.28 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))
$R_{free}$	180053	9078 (2.30-2.26)
Clashscore	190562	9802 (2.30-2.26)
Ramachandran outliers	187476	9690 (2.30-2.26)
Sidechain outliers	187428	9691 (2.30-2.26)
RSRZ outliers	180081	9085 (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  $\geq 3$ , 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	A	496	<div> <div>15%</div> <div>74%</div> <div>18%</div> <div>7%</div> </div>
1	B	496	<div> <div>10%</div> <div>75%</div> <div>17%</div> <div>7%</div> </div>
1	C	496	<div> <div>12%</div> <div>75%</div> <div>17%</div> <div>7%</div> </div>
1	D	496	<div> <div>13%</div> <div>75%</div> <div>18%</div> <div>7%</div> </div>
1	E	496	<div> <div>14%</div> <div>79%</div> <div>14%</div> <div>7%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	496	<div><div></div><div>56%</div><div></div><div>74%</div><div></div><div>16%</div><div></div><div>9%</div></div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 20214 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 Serine hydroxymethyltransferase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	462	Total	C	N	O	P	S	0	0	0
			3414	2159	588	649	1	17			
1	B	459	Total	C	N	O	P	S	0	0	0
			3409	2163	585	644	1	16			
1	C	460	Total	C	N	O	P	S	0	0	0
			3434	2180	584	652	1	17			
1	D	463	Total	C	N	O	P	S	0	0	0
			3419	2172	580	649	1	17			
1	E	461	Total	C	N	O	P	S	0	0	0
			3367	2137	578	635	1	16			
1	F	450	Total	C	N	O	P	S	0	0	0
			2972	1847	527	583	1	14			

There are 156 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-24	MET	-	expression tag	UNP K4FZF8
A	-23	GLY	-	expression tag	UNP K4FZF8
A	-22	MET	-	expression tag	UNP K4FZF8
A	-21	HIS	-	expression tag	UNP K4FZF8
A	-20	HIS	-	expression tag	UNP K4FZF8
A	-19	HIS	-	expression tag	UNP K4FZF8
A	-18	HIS	-	expression tag	UNP K4FZF8
A	-17	HIS	-	expression tag	UNP K4FZF8
A	-16	HIS	-	expression tag	UNP K4FZF8
A	-15	SER	-	expression tag	UNP K4FZF8
A	-14	SER	-	expression tag	UNP K4FZF8
A	-13	GLY	-	expression tag	UNP K4FZF8
A	-12	VAL	-	expression tag	UNP K4FZF8
A	-11	ASP	-	expression tag	UNP K4FZF8
A	-10	LEU	-	expression tag	UNP K4FZF8
A	-9	GLY	-	expression tag	UNP K4FZF8
A	-8	THR	-	expression tag	UNP K4FZF8

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	GLU	-	expression tag	UNP K4FZF8
A	-6	ASN	-	expression tag	UNP K4FZF8
A	-5	LEU	-	expression tag	UNP K4FZF8
A	-4	TYR	-	expression tag	UNP K4FZF8
A	-3	PHE	-	expression tag	UNP K4FZF8
A	-2	GLN	-	expression tag	UNP K4FZF8
A	-1	SER	-	expression tag	UNP K4FZF8
A	0	ASN	-	expression tag	UNP K4FZF8
A	149	THR	ALA	engineered mutation	UNP K4FZF8
B	-24	MET	-	expression tag	UNP K4FZF8
B	-23	GLY	-	expression tag	UNP K4FZF8
B	-22	MET	-	expression tag	UNP K4FZF8
B	-21	HIS	-	expression tag	UNP K4FZF8
B	-20	HIS	-	expression tag	UNP K4FZF8
B	-19	HIS	-	expression tag	UNP K4FZF8
B	-18	HIS	-	expression tag	UNP K4FZF8
B	-17	HIS	-	expression tag	UNP K4FZF8
B	-16	HIS	-	expression tag	UNP K4FZF8
B	-15	SER	-	expression tag	UNP K4FZF8
B	-14	SER	-	expression tag	UNP K4FZF8
B	-13	GLY	-	expression tag	UNP K4FZF8
B	-12	VAL	-	expression tag	UNP K4FZF8
B	-11	ASP	-	expression tag	UNP K4FZF8
B	-10	LEU	-	expression tag	UNP K4FZF8
B	-9	GLY	-	expression tag	UNP K4FZF8
B	-8	THR	-	expression tag	UNP K4FZF8
B	-7	GLU	-	expression tag	UNP K4FZF8
B	-6	ASN	-	expression tag	UNP K4FZF8
B	-5	LEU	-	expression tag	UNP K4FZF8
B	-4	TYR	-	expression tag	UNP K4FZF8
B	-3	PHE	-	expression tag	UNP K4FZF8
B	-2	GLN	-	expression tag	UNP K4FZF8
B	-1	SER	-	expression tag	UNP K4FZF8
B	0	ASN	-	expression tag	UNP K4FZF8
B	149	THR	ALA	engineered mutation	UNP K4FZF8
C	-24	MET	-	expression tag	UNP K4FZF8
C	-23	GLY	-	expression tag	UNP K4FZF8
C	-22	MET	-	expression tag	UNP K4FZF8
C	-21	HIS	-	expression tag	UNP K4FZF8
C	-20	HIS	-	expression tag	UNP K4FZF8
C	-19	HIS	-	expression tag	UNP K4FZF8
C	-18	HIS	-	expression tag	UNP K4FZF8

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	-17	HIS	-	expression tag	UNP K4FZF8
C	-16	HIS	-	expression tag	UNP K4FZF8
C	-15	SER	-	expression tag	UNP K4FZF8
C	-14	SER	-	expression tag	UNP K4FZF8
C	-13	GLY	-	expression tag	UNP K4FZF8
C	-12	VAL	-	expression tag	UNP K4FZF8
C	-11	ASP	-	expression tag	UNP K4FZF8
C	-10	LEU	-	expression tag	UNP K4FZF8
C	-9	GLY	-	expression tag	UNP K4FZF8
C	-8	THR	-	expression tag	UNP K4FZF8
C	-7	GLU	-	expression tag	UNP K4FZF8
C	-6	ASN	-	expression tag	UNP K4FZF8
C	-5	LEU	-	expression tag	UNP K4FZF8
C	-4	TYR	-	expression tag	UNP K4FZF8
C	-3	PHE	-	expression tag	UNP K4FZF8
C	-2	GLN	-	expression tag	UNP K4FZF8
C	-1	SER	-	expression tag	UNP K4FZF8
C	0	ASN	-	expression tag	UNP K4FZF8
C	149	THR	ALA	engineered mutation	UNP K4FZF8
D	-24	MET	-	expression tag	UNP K4FZF8
D	-23	GLY	-	expression tag	UNP K4FZF8
D	-22	MET	-	expression tag	UNP K4FZF8
D	-21	HIS	-	expression tag	UNP K4FZF8
D	-20	HIS	-	expression tag	UNP K4FZF8
D	-19	HIS	-	expression tag	UNP K4FZF8
D	-18	HIS	-	expression tag	UNP K4FZF8
D	-17	HIS	-	expression tag	UNP K4FZF8
D	-16	HIS	-	expression tag	UNP K4FZF8
D	-15	SER	-	expression tag	UNP K4FZF8
D	-14	SER	-	expression tag	UNP K4FZF8
D	-13	GLY	-	expression tag	UNP K4FZF8
D	-12	VAL	-	expression tag	UNP K4FZF8
D	-11	ASP	-	expression tag	UNP K4FZF8
D	-10	LEU	-	expression tag	UNP K4FZF8
D	-9	GLY	-	expression tag	UNP K4FZF8
D	-8	THR	-	expression tag	UNP K4FZF8
D	-7	GLU	-	expression tag	UNP K4FZF8
D	-6	ASN	-	expression tag	UNP K4FZF8
D	-5	LEU	-	expression tag	UNP K4FZF8
D	-4	TYR	-	expression tag	UNP K4FZF8
D	-3	PHE	-	expression tag	UNP K4FZF8
D	-2	GLN	-	expression tag	UNP K4FZF8

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	-1	SER	-	expression tag	UNP K4FZF8
D	0	ASN	-	expression tag	UNP K4FZF8
D	149	THR	ALA	engineered mutation	UNP K4FZF8
E	-24	MET	-	expression tag	UNP K4FZF8
E	-23	GLY	-	expression tag	UNP K4FZF8
E	-22	MET	-	expression tag	UNP K4FZF8
E	-21	HIS	-	expression tag	UNP K4FZF8
E	-20	HIS	-	expression tag	UNP K4FZF8
E	-19	HIS	-	expression tag	UNP K4FZF8
E	-18	HIS	-	expression tag	UNP K4FZF8
E	-17	HIS	-	expression tag	UNP K4FZF8
E	-16	HIS	-	expression tag	UNP K4FZF8
E	-15	SER	-	expression tag	UNP K4FZF8
E	-14	SER	-	expression tag	UNP K4FZF8
E	-13	GLY	-	expression tag	UNP K4FZF8
E	-12	VAL	-	expression tag	UNP K4FZF8
E	-11	ASP	-	expression tag	UNP K4FZF8
E	-10	LEU	-	expression tag	UNP K4FZF8
E	-9	GLY	-	expression tag	UNP K4FZF8
E	-8	THR	-	expression tag	UNP K4FZF8
E	-7	GLU	-	expression tag	UNP K4FZF8
E	-6	ASN	-	expression tag	UNP K4FZF8
E	-5	LEU	-	expression tag	UNP K4FZF8
E	-4	TYR	-	expression tag	UNP K4FZF8
E	-3	PHE	-	expression tag	UNP K4FZF8
E	-2	GLN	-	expression tag	UNP K4FZF8
E	-1	SER	-	expression tag	UNP K4FZF8
E	0	ASN	-	expression tag	UNP K4FZF8
E	149	THR	ALA	engineered mutation	UNP K4FZF8
F	-24	MET	-	expression tag	UNP K4FZF8
F	-23	GLY	-	expression tag	UNP K4FZF8
F	-22	MET	-	expression tag	UNP K4FZF8
F	-21	HIS	-	expression tag	UNP K4FZF8
F	-20	HIS	-	expression tag	UNP K4FZF8
F	-19	HIS	-	expression tag	UNP K4FZF8
F	-18	HIS	-	expression tag	UNP K4FZF8
F	-17	HIS	-	expression tag	UNP K4FZF8
F	-16	HIS	-	expression tag	UNP K4FZF8
F	-15	SER	-	expression tag	UNP K4FZF8
F	-14	SER	-	expression tag	UNP K4FZF8
F	-13	GLY	-	expression tag	UNP K4FZF8
F	-12	VAL	-	expression tag	UNP K4FZF8

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	-11	ASP	-	expression tag	UNP K4FZF8
F	-10	LEU	-	expression tag	UNP K4FZF8
F	-9	GLY	-	expression tag	UNP K4FZF8
F	-8	THR	-	expression tag	UNP K4FZF8
F	-7	GLU	-	expression tag	UNP K4FZF8
F	-6	ASN	-	expression tag	UNP K4FZF8
F	-5	LEU	-	expression tag	UNP K4FZF8
F	-4	TYR	-	expression tag	UNP K4FZF8
F	-3	PHE	-	expression tag	UNP K4FZF8
F	-2	GLN	-	expression tag	UNP K4FZF8
F	-1	SER	-	expression tag	UNP K4FZF8
F	0	ASN	-	expression tag	UNP K4FZF8
F	149	THR	ALA	engineered mutation	UNP K4FZF8

- Molecule 2 is water.

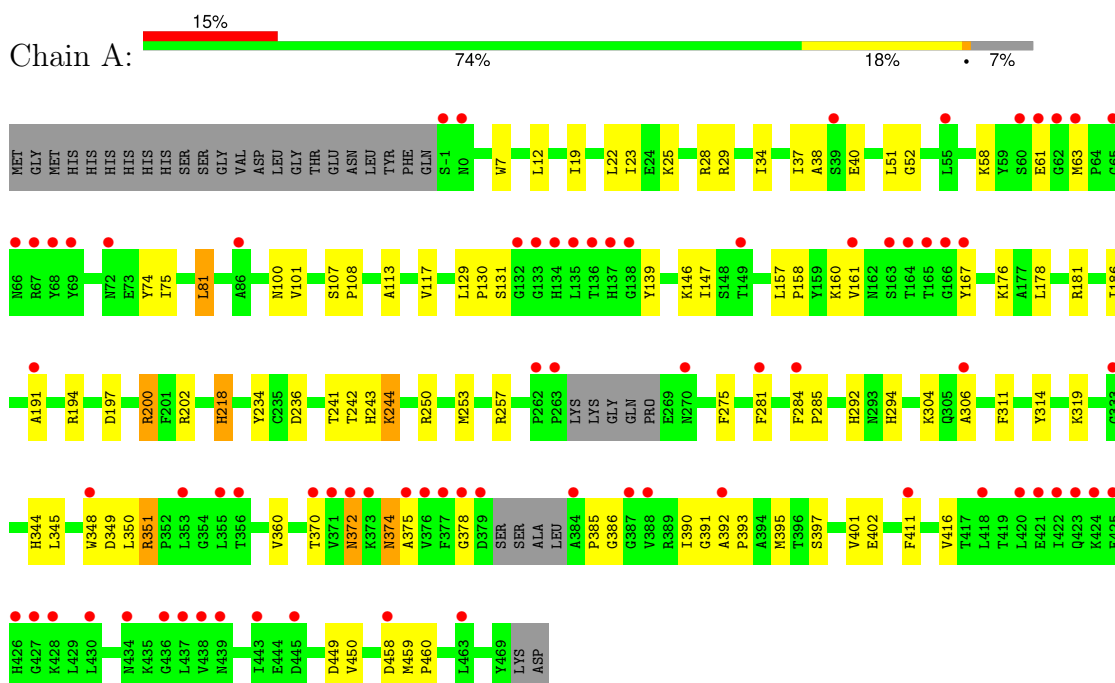
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	34	Total O 34 34	0	0
2	B	44	Total O 44 44	0	0
2	C	32	Total O 32 32	0	0
2	D	51	Total O 51 51	0	0
2	E	32	Total O 32 32	0	0
2	F	6	Total O 6 6	0	0



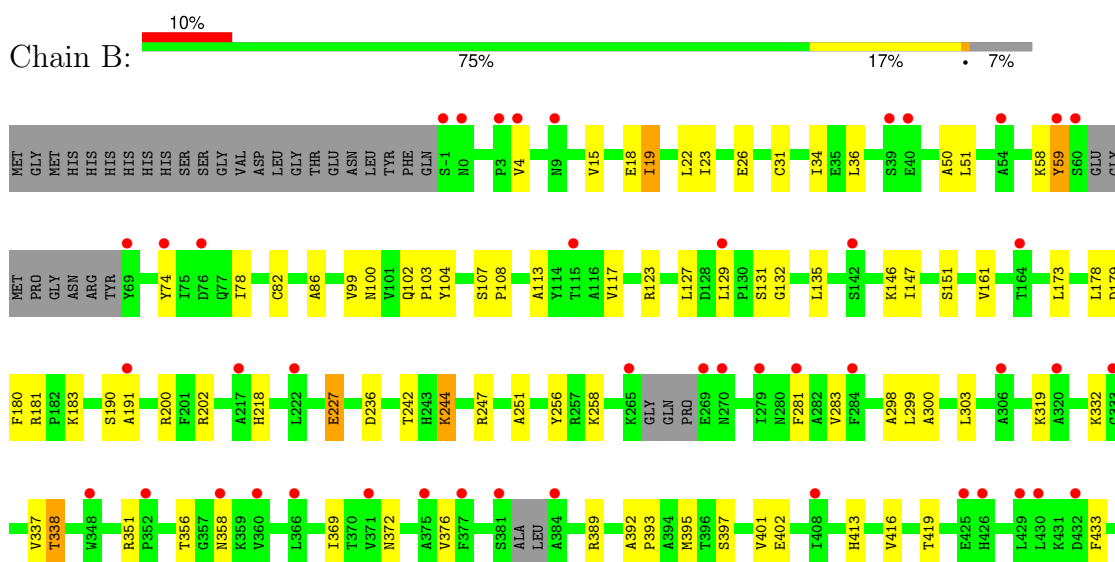
### 3 Residue-property plots [i](#)

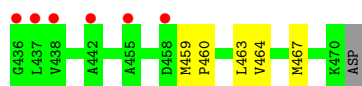
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Serine hydroxymethyltransferase

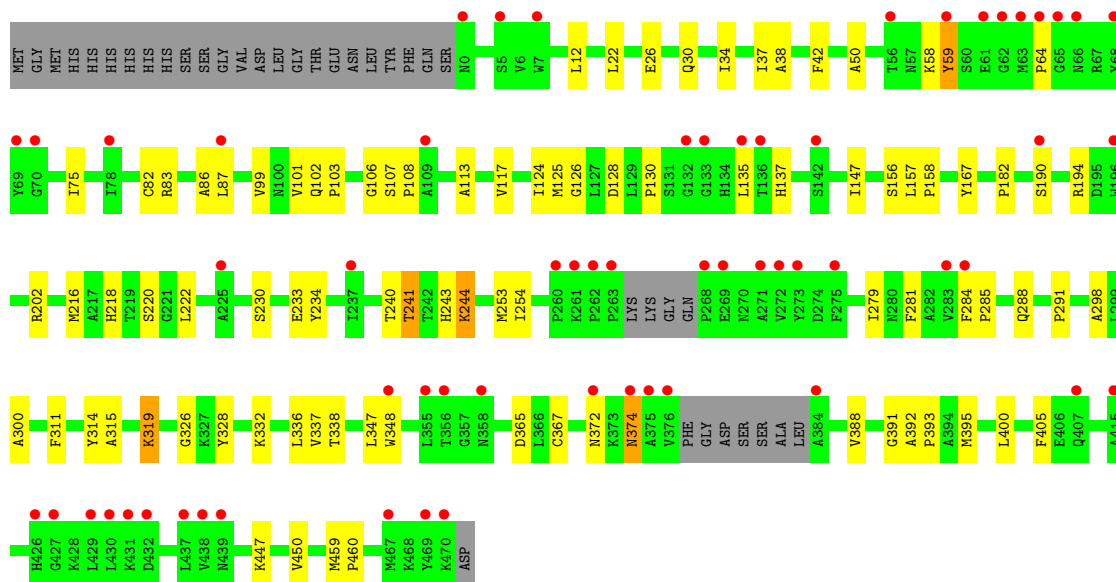
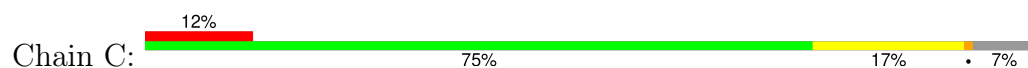


#### • Molecule 1: Serine hydroxymethyltransferase

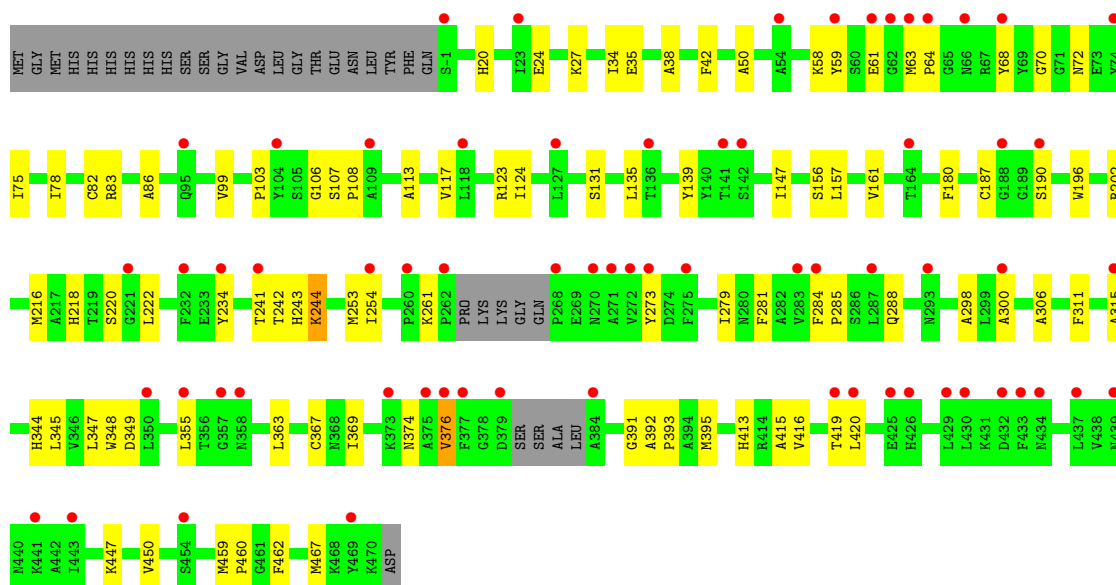
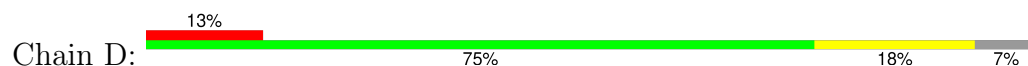




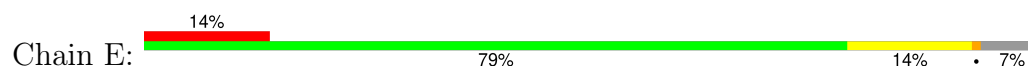
• Molecule 1: Serine hydroxymethyltransferase

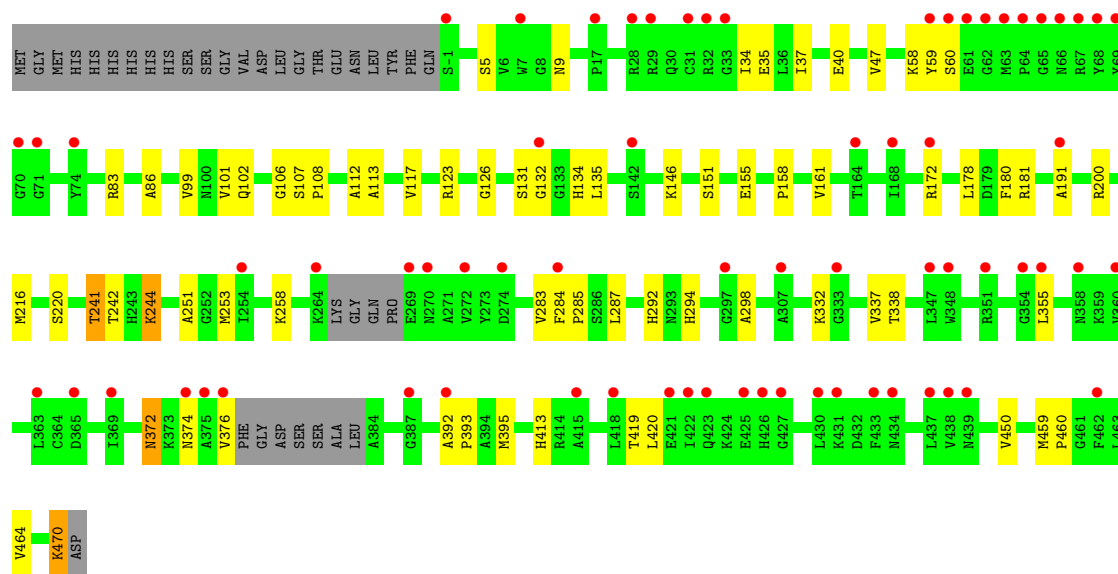


• Molecule 1: Serine hydroxymethyltransferase

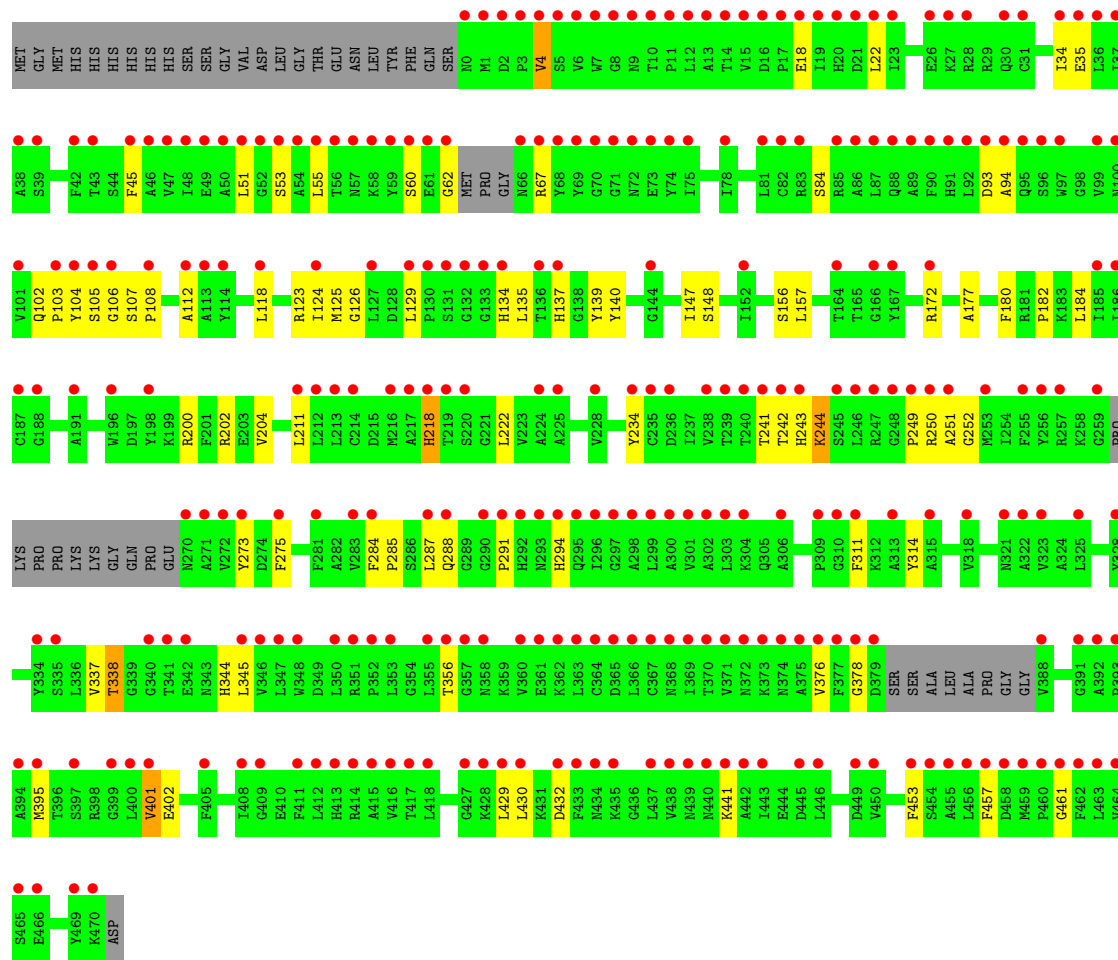
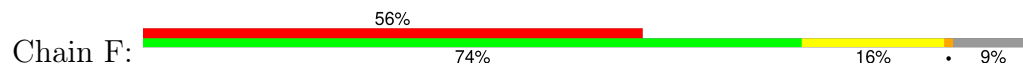


• Molecule 1: Serine hydroxymethyltransferase





• Molecule 1: Serine hydroxymethyltransferase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	174.32Å 174.32Å 184.21Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.54 – 2.28 48.54 – 2.28	Depositor EDS
% Data completeness (in resolution range)	98.8 (48.54-2.28) 98.9 (48.54-2.28)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.20 (at 2.29Å)	Xtriage
Refinement program	PHENIX 1.21.2_5419	Depositor
R, $R_{free}$	0.272 , 0.313 0.276 , 0.315	Depositor DCC
$R_{free}$ test set	7353 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	34.8	Xtriage
Anisotropy	0.637	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 41.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.010 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	20214	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 33.41 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 8.0390e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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](#)

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

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	0.38	0/3468	0.61	0/4720
1	B	0.37	0/3461	0.61	0/4706
1	C	0.34	0/3490	0.57	0/4746
1	D	0.36	0/3472	0.59	0/4722
1	E	0.38	0/3418	0.60	0/4657
1	F	0.34	0/3003	0.61	0/4106
All	All	0.36	0/20312	0.60	0/27657

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 [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	3414	0	3144	75	0
1	B	3409	0	3178	62	0
1	C	3434	0	3216	61	0
1	D	3419	0	3173	59	0
1	E	3367	0	3109	46	0
1	F	2972	0	2362	56	0
2	A	34	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	44	0	0	0	0
2	C	32	0	0	0	0
2	D	51	0	0	3	0
2	E	32	0	0	0	0
2	F	6	0	0	0	0
All	All	20214	0	18182	330	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 330 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:C:125:MET:HE2	1:C:182:PRO:HG3	1.40	1.04
1:A:243:HIS:HE2	1:B:59:TYR:HE1	1.13	0.93
1:C:106:GLY:HA3	1:C:241:THR:HG22	1.56	0.86
1:D:34:ILE:HA	1:D:395:MET:HE2	1.58	0.84
1:E:37:ILE:HG23	1:E:372:ASN:HD21	1.44	0.82

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	455/496 (92%)	435 (96%)	20 (4%)	0	100	100
1	B	450/496 (91%)	432 (96%)	18 (4%)	0	100	100
1	C	453/496 (91%)	433 (96%)	20 (4%)	0	100	100
1	D	456/496 (92%)	433 (95%)	23 (5%)	0	100	100
1	E	454/496 (92%)	436 (96%)	18 (4%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	441/496 (89%)	410 (93%)	22 (5%)	9 (2%)	6	4
All	All	2709/2976 (91%)	2579 (95%)	121 (4%)	9 (0%)	36	44

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	94	ALA
1	F	429	LEU
1	F	430	LEU
1	F	432	ASP
1	F	376	VAL

### 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	A	327/407 (80%)	317 (97%)	10 (3%)	35	50
1	B	330/407 (81%)	321 (97%)	9 (3%)	39	55
1	C	337/407 (83%)	331 (98%)	6 (2%)	51	68
1	D	328/407 (81%)	325 (99%)	3 (1%)	70	82
1	E	318/407 (78%)	310 (98%)	8 (2%)	42	58
1	F	217/407 (53%)	209 (96%)	8 (4%)	30	43
All	All	1857/2442 (76%)	1813 (98%)	44 (2%)	43	59

5 of 44 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	40	GLU
1	E	470	LYS
1	E	60	SER
1	E	372	ASN
1	F	84	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	295	GLN
1	E	77	GLN
1	F	226	GLN
1	D	413	HIS
1	E	293	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

6 non-standard protein/DNA/RNA residues 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
1	LLP	E	244	1	23,24,25	1.18	2 (8%)	25,32,34	1.73	4 (16%)
1	LLP	F	244	1	23,24,25	1.17	1 (4%)	25,32,34	1.43	5 (20%)
1	LLP	D	244	1	23,24,25	1.17	3 (13%)	25,32,34	1.62	8 (32%)
1	LLP	A	244	1	23,24,25	1.14	1 (4%)	25,32,34	1.12	3 (12%)
1	LLP	B	244	1	23,24,25	1.16	2 (8%)	25,32,34	2.01	8 (32%)
1	LLP	C	244	1	23,24,25	1.08	1 (4%)	25,32,34	1.64	4 (16%)

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
1	LLP	E	244	1	-	11/16/17/19	0/1/1/1
1	LLP	F	244	1	-	10/16/17/19	0/1/1/1

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	LLP	D	244	1	-	11/16/17/19	0/1/1/1
1	LLP	A	244	1	-	11/16/17/19	0/1/1/1
1	LLP	B	244	1	-	8/16/17/19	0/1/1/1
1	LLP	C	244	1	-	13/16/17/19	0/1/1/1

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	244	LLP	P-OP4	4.37	1.74	1.60
1	E	244	LLP	P-OP4	3.94	1.72	1.60
1	D	244	LLP	P-OP4	3.70	1.72	1.60
1	A	244	LLP	P-OP4	3.66	1.71	1.60
1	C	244	LLP	P-OP4	3.54	1.71	1.60

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	244	LLP	C5-C4-C4'	4.68	128.69	121.47
1	B	244	LLP	C4-C3-C2	4.49	122.67	120.14
1	E	244	LLP	C5-C4-C4'	3.83	127.39	121.47
1	E	244	LLP	OP4-C5'-C5	3.66	116.22	109.36
1	C	244	LLP	OP3-P-OP2	3.61	121.35	107.80

There are no chirality outliers.

5 of 64 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	244	LLP	C4-C4'-NZ-CE
1	A	244	LLP	C5'-OP4-P-OP1
1	A	244	LLP	C5'-OP4-P-OP2
1	A	244	LLP	C5'-OP4-P-OP3
1	A	244	LLP	O-C-CA-CB

There are no ring outliers.

6 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	E	244	LLP	3	0
1	F	244	LLP	2	0
1	D	244	LLP	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	244	LLP	4	0
1	B	244	LLP	2	0
1	C	244	LLP	2	0

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	461/496 (92%)	1.13	75 (16%) 4 5	29, 46, 82, 118	0
1	B	458/496 (92%)	1.10	51 (11%) 10 11	26, 38, 71, 103	0
1	C	459/496 (92%)	1.17	61 (13%) 7 7	28, 44, 79, 108	0
1	D	462/496 (93%)	1.15	66 (14%) 6 6	28, 38, 76, 102	0
1	E	460/496 (92%)	1.11	69 (15%) 5 6	34, 48, 87, 110	0
1	F	449/496 (90%)	2.50	280 (62%) 0 0	53, 76, 109, 134	0
All	All	2749/2976 (92%)	1.36	602 (21%) 2 2	26, 47, 92, 134	0

The worst 5 of 602 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	74	TYR	9.6
1	F	460	PRO	8.5
1	F	54	ALA	8.3
1	F	240	THR	7.1
1	E	68	TYR	6.6

### 6.2 Non-standard residues in protein, DNA, RNA chains [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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	LLP	F	244	24/25	0.81	0.21	64,78,89,91	0
1	LLP	B	244	24/25	0.87	0.15	28,33,35,36	0
1	LLP	C	244	24/25	0.88	0.14	34,42,47,47	0
1	LLP	D	244	24/25	0.91	0.13	28,36,40,41	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
1	LLP	A	244	24/25	0.91	0.15	33,43,47,48	0
1	LLP	E	244	24/25	0.92	0.14	38,45,51,51	0

### 6.3 Carbohydrates [i](#)

There are no oligosaccharides 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.