



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 27, 2018 – 11:17 am GMT

PDB ID : 1F4D  
Title : CRYSTAL STRUCTURE OF E. COLI THYMIDYLATE SYNTHASE C146S,  
L143C COVALENTLY MODIFIED AT C143 WITH N-[TOSYL-D-PROLIN  
YL]AMINO-ETHANETHIOL  
Authors : Erlanson, D.A.; Braisted, A.C.; Raphael, D.R.; Randal, M.; Stroud, R.M.;  
Gordon, E.; Wells, J.A.  
Deposited on : 2000-06-07  
Resolution : 2.15 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30686

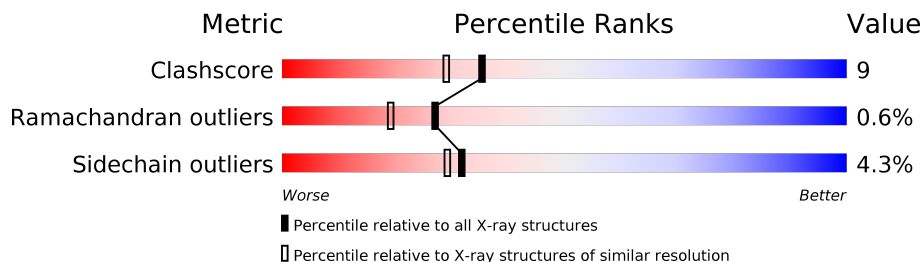
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.15 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))
Clashscore	122078	1390 (2.16-2.16)
Ramachandran outliers	120005	1368 (2.16-2.16)
Sidechain outliers	119972	1367 (2.16-2.16)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	264	
1	B	264	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	GOL	A	501	-	-	X	-
4	GOL	B	502	-	-	X	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 4682 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.

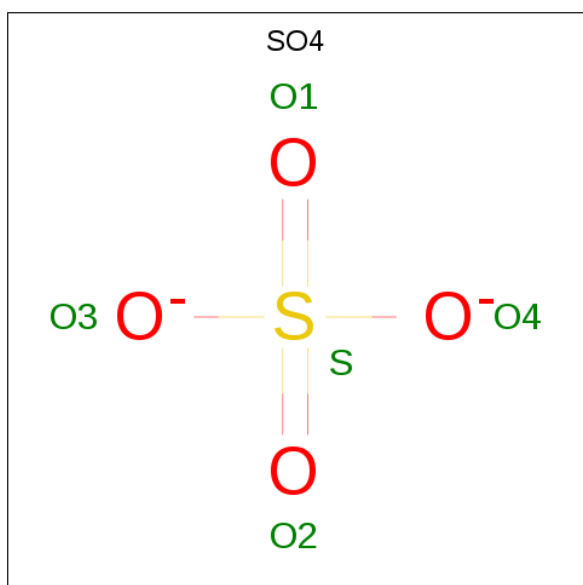
- Molecule 1 is a protein called THYMIDYLATE SYNTHASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	264	Total	C	N	O	S	0	1	0
			2156	1375	373	396	12			
1	B	264	Total	C	N	O	S	0	1	0
			2155	1374	372	397	12			

There are 6 discrepancies between the modelled and reference sequences:

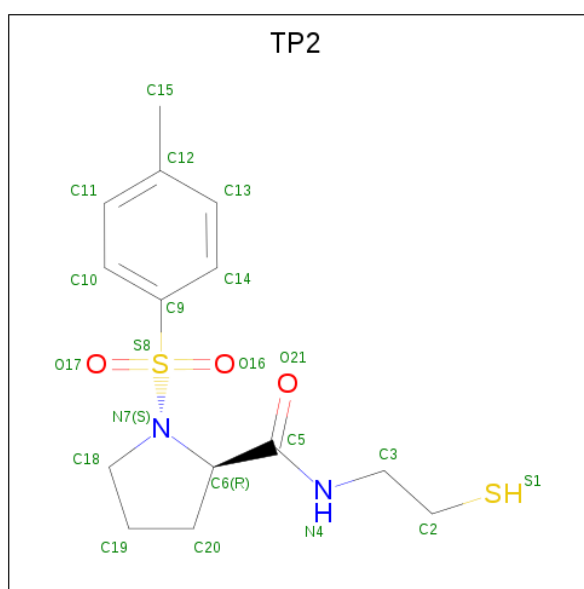
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	CXM	MET	ENGINEERED	UNP P0A884
A	143	CYS	LEU	ENGINEERED	UNP P0A884
A	146	SER	CYS	ENGINEERED	UNP P0A884
B	1	CXM	MET	ENGINEERED	UNP P0A884
B	143	CYS	LEU	ENGINEERED	UNP P0A884
B	146	SER	CYS	ENGINEERED	UNP P0A884

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is N-[TOSYL-D-PROLINYL]AMINO-ETHANETHIOL (three-letter code: TP2) (formula:  $C_{14}H_{20}N_2O_3S_2$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			21	14	2	3	2		
3	B	1	Total	C	N	O	S	0	0
			21	14	2	3	2		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is water.

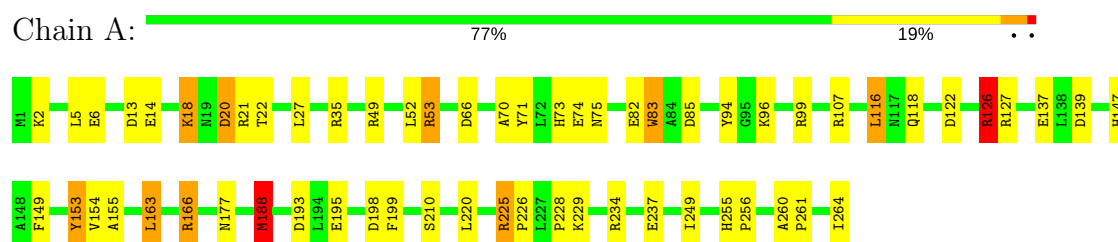
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	179	Total	O	0	0
			179	179		
5	B	118	Total	O	0	0
			118	118		

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

These plots are drawn for all protein, RNA and DNA 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.

Note EDS was not executed.

#### • Molecule 1: THYMIDYLATE SYNTHASE



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 63	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	126.33Å 126.33Å 67.12Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	10.00 – 2.15	Depositor
% Data completeness (in resolution range)	96.7 (10.00-2.15)	Depositor
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.196 , 0.267	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4682	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, CXM, TP2, SO4

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	0.74	0/2211	1.80	52/3002 (1.7%)
1	B	0.71	1/2209 (0.0%)	1.74	42/2999 (1.4%)
All	All	0.73	1/4420 (0.0%)	1.77	94/6001 (1.6%)

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	2
1	B	0	2
All	All	0	4

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	149	PHE	CB-CG	-5.20	1.42	1.51

The worst 5 of 94 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	126	ARG	NE-CZ-NH2	-20.50	110.05	120.30
1	B	127	ARG	NE-CZ-NH1	16.18	128.39	120.30
1	B	126	ARG	NE-CZ-NH2	-15.89	112.35	120.30
1	A	127	ARG	NE-CZ-NH2	-14.52	113.04	120.30
1	B	99	ARG	NE-CZ-NH1	-13.72	113.44	120.30

There are no chirality outliers.

All (4) planarity outliers are listed below:



Mol	Chain	Res	Type	Group
1	A	118	GLN	Mainchain
1	A	210	SER	Mainchain
1	B	40	ASP	Mainchain
1	B	61	TRP	Mainchain

## 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	2156	0	2076	39	0
1	B	2155	0	2077	46	0
2	A	10	0	0	0	0
2	B	10	0	0	0	0
3	A	21	0	19	3	0
3	B	21	0	19	1	0
4	A	6	0	8	4	0
4	B	6	0	8	5	0
5	A	179	0	0	4	0
5	B	118	0	0	5	0
All	All	4682	0	4207	76	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 76 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:B:3[A]:GLN:OE1	5:B:627:HOH:O	1.61	1.14
1:A:166:ARG:HH12	4:A:501:GOL:H11	1.49	0.77
1:B:233:LYS:HD2	1:B:248:GLU:HG2	1.69	0.74
1:B:103:THR:HB	1:B:104:PRO:HD2	1.70	0.73
5:A:678:HOH:O	1:B:18:LYS:HE3	1.90	0.71

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	263/264 (100%)	255 (97%)	6 (2%)	2 (1%)	21	13
1	B	263/264 (100%)	250 (95%)	12 (5%)	1 (0%)	36	31
All	All	526/528 (100%)	505 (96%)	18 (3%)	3 (1%)	27	19

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	94	TYR
1	B	94	TYR
1	A	228	PRO

### 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	233/232 (100%)	224 (96%)	9 (4%)	35	33
1	B	233/232 (100%)	222 (95%)	11 (5%)	29	25
All	All	466/464 (100%)	446 (96%)	20 (4%)	32	29

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	18	LYS
1	B	43	PRO
1	B	149	PHE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	229	LYS
1	A	237	GLU

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

Mol	Chain	Res	Type
1	A	190	GLN
1	A	219	GLN
1	B	73	HIS
1	B	190	GLN
1	B	219	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 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	CXM	A	1	1	7,10,11	1.83	3 (42%)	5,11,13	2.36	2 (40%)
1	CXM	B	1	1	7,10,11	2.32	2 (28%)	5,11,13	1.81	2 (40%)

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	CXM	A	1	1	-	0/6/10/12	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	CXM	B	1	1	-	0/6/10/12	0/0/0/0

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1	CXM	CA-N	-4.82	1.39	1.46
1	A	1	CXM	CA-N	-2.99	1.42	1.46
1	A	1	CXM	CA-C	2.30	1.53	1.50
1	A	1	CXM	O-C	2.35	1.29	1.19
1	B	1	CXM	CA-C	2.78	1.53	1.50

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1	CXM	CB-CA-C	-4.16	103.92	111.85
1	A	1	CXM	O-C-CA	-3.18	117.68	125.09
1	B	1	CXM	O-C-CA	-3.13	117.80	125.09
1	B	1	CXM	CB-CA-C	-2.20	107.66	111.85

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

8 ligands 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
3	TP2	A	401	1	21,22,22	4.60	3 (14%)	29,31,31	2.74	10 (34%)
4	GOL	A	501	-	5,5,5	0.85	0	5,5,5	1.38	1 (20%)
2	SO4	A	511	-	4,4,4	0.57	0	6,6,6	0.36	0
2	SO4	A	512	-	4,4,4	0.83	0	6,6,6	1.15	1 (16%)
3	TP2	B	402	1	21,22,22	4.35	3 (14%)	29,31,31	1.95	8 (27%)
4	GOL	B	502	-	5,5,5	0.65	0	5,5,5	2.65	3 (60%)
2	SO4	B	513	-	4,4,4	0.63	0	6,6,6	0.29	0
2	SO4	B	514	-	4,4,4	0.61	0	6,6,6	0.46	0

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
3	TP2	A	401	1	-	0/20/30/30	0/2/2/2
4	GOL	A	501	-	-	0/4/4/4	0/0/0/0
2	SO4	A	511	-	-	0/0/0/0	0/0/0/0
2	SO4	A	512	-	-	0/0/0/0	0/0/0/0
3	TP2	B	402	1	-	0/20/30/30	0/2/2/2
4	GOL	B	502	-	-	0/4/4/4	0/0/0/0
2	SO4	B	513	-	-	0/0/0/0	0/0/0/0
2	SO4	B	514	-	-	0/0/0/0	0/0/0/0

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	401	TP2	S8-N7	9.74	1.78	1.63
3	B	402	TP2	S8-N7	9.99	1.78	1.63
3	B	402	TP2	O17-S8	11.66	1.57	1.43
3	A	401	TP2	O17-S8	11.90	1.58	1.43
3	B	402	TP2	O16-S8	12.39	1.58	1.43

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	401	TP2	O17-S8-O16	-7.37	107.39	119.50
3	B	402	TP2	O17-S8-O16	-5.66	110.20	119.50
3	A	401	TP2	O16-S8-N7	-3.68	100.09	106.95
4	B	502	GOL	O2-C2-C1	-3.61	92.49	109.00
4	B	502	GOL	O1-C1-C2	-3.59	92.69	110.11

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	401	TP2	3	0
4	A	501	GOL	4	0
3	B	402	TP2	1	0
4	B	502	GOL	5	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

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.