



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 21, 2017 – 12:48 AM EST

PDB ID : 4DFR  
Title : CRYSTAL STRUCTURES OF ESCHERICHIA COLI AND LACTO-BACILLUS CASEI DIHYDROFOLATE REDUCTASE REFINED AT 1.7 ANGSTROMS RESOLUTION. I. GENERAL FEATURES AND BINDING OF METHOTREXATE  
Authors : Filman, D.J.; Matthews, D.A.; Bolin, J.T.; Kraut, J.  
Deposited on : unknown  
Resolution : 1.70 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030345

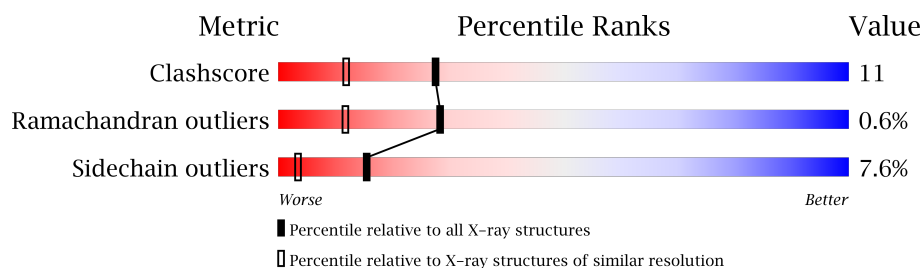
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.70 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))
Clashscore	112137	3876 (1.70-1.70)
Ramachandran outliers	110173	3815 (1.70-1.70)
Sidechain outliers	110143	3815 (1.70-1.70)

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 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	159	 13% 65% 18% .
1	B	159	 8% 70% 16% 6%

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 3035 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 DIHYDROFOLATE REDUCTASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	159	Total	C	N	O	S	0	1	0
			1252	797	213	235	7			
1	B	159	Total	C	N	O	S	0	3	0
			1286	817	222	240	7			

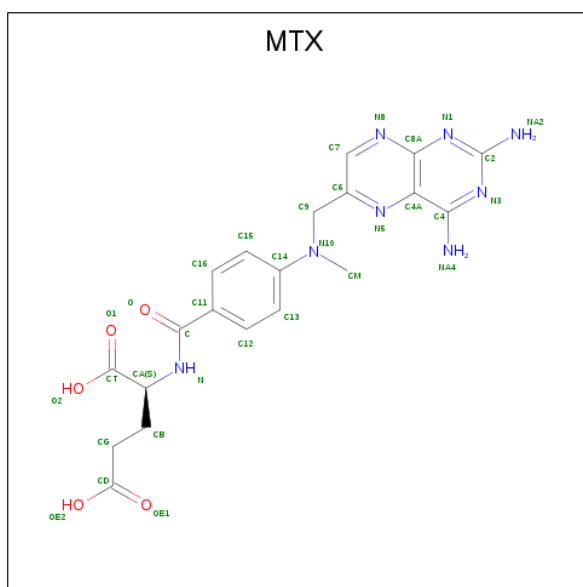
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	37	ASP	ASN	CONFLICT	UNP P0ABQ4
A	154	LYS	GLU	CONFLICT	UNP P0ABQ4
B	37	ASP	ASN	CONFLICT	UNP P0ABQ4
B	154	LYS	GLU	CONFLICT	UNP P0ABQ4

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

- Molecule 3 is METHOTREXATE (three-letter code: MTX) (formula: C<sub>20</sub>H<sub>22</sub>N<sub>8</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			33	20	8	5		
3	B	1	Total	C	N	O	0	0
			33	20	8	5		

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Ca	0	0
			1	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	239	Total	O	0	0
			239	239		
5	B	189	Total	O	0	0
			189	189		

### 3 Residue-property plots

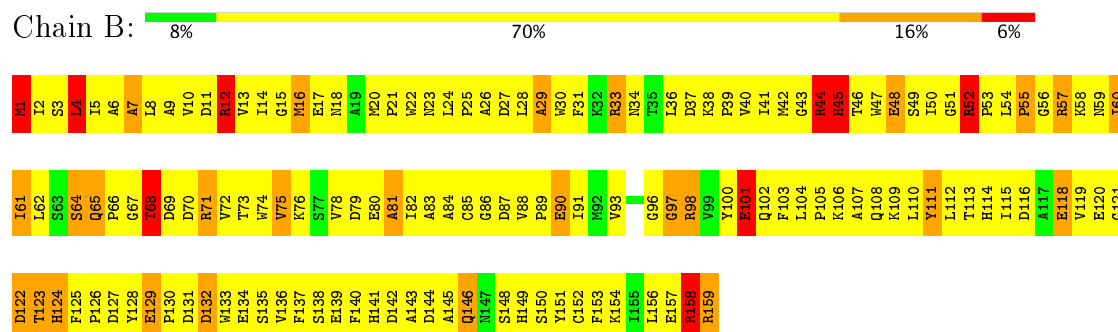
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are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



#### • Molecule 1: DIHYDROFOLATE REDUCTASE



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	93.22Å 93.22Å 73.56Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	(Not available) – 1.70	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-1.70)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	unknown	Depositor
R, $R_{free}$	0.155 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3035	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	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: MTX, CA, CL

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	3.24	144/1285 (11.2%)	4.42	311/1747 (17.8%)
1	B	3.44	163/1321 (12.3%)	4.72	340/1795 (18.9%)
All	All	3.34	307/2606 (11.8%)	4.58	651/3542 (18.4%)

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	7
1	B	0	15
All	All	0	22

The worst 5 of 307 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	138	SER	CB-OG	-20.55	1.15	1.42
1	B	57	ARG	CZ-NH1	16.57	1.54	1.33
1	B	100	TYR	CG-CD2	15.22	1.58	1.39
1	A	48	GLU	CD-OE2	14.75	1.41	1.25
1	B	118	GLU	CD-OE2	14.05	1.41	1.25

The worst 5 of 651 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	87	ASP	CB-CG-OD2	-39.34	82.89	118.30
1	B	37	ASP	CB-CG-OD2	-35.23	86.59	118.30
1	A	57	ARG	NE-CZ-NH1	33.75	137.17	120.30
1	A	98	ARG	NE-CZ-NH1	31.98	136.29	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	71	ARG	NE-CZ-NH2	-30.39	105.11	120.30

There are no chirality outliers.

5 of 22 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	33	ARG	Sidechain
1	A	44	ARG	Sidechain
1	A	52	ARG	Sidechain
1	A	65	GLN	Mainchain,Peptide
1	A	98	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	1252	0	1175	24	0
1	B	1286	0	1217	32	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	33	0	19	2	0
3	B	33	0	20	0	0
4	B	1	0	0	0	0
5	A	239	0	0	11	0
5	B	189	0	0	9	0
All	All	3035	0	2431	57	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 11.

The worst 5 of 57 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:57:ARG:CB	1:B:57:ARG:CA	1.74	1.61
1:A:58:LYS:CD	1:A:58:LYS:CE	1.75	1.57

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:109:LYS:NZ	1:A:109:LYS:CE	1.70	1.52
1:B:159:ARG:N	1:B:159:ARG:CA	1.70	1.49
1:B:122[B]:ASP:C	1:B:123:THR:N	1.82	1.32

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	158/159 (99%)	146 (92%)	10 (6%)	2 (1%)	14	2
1	B	160/159 (101%)	156 (98%)	4 (2%)	0	100	100
All	All	318/318 (100%)	302 (95%)	14 (4%)	2 (1%)	28	12

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	67	GLY
1	A	66	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	130/136 (96%)	117 (90%)	13 (10%)	9	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	B	137/136 (101%)	129 (94%)	8 (6%)	23 7
All	All	267/272 (98%)	246 (92%)	21 (8%)	15 3

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	104	LEU
1	A	120	GLU
1	B	68	THR
1	A	98	ARG
1	B	111	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	102	GLN
1	A	114	HIS
1	B	65	GLN
1	B	124	HIS

### 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 carbohydrates in this entry.

### 5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 3 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
3	MTX	A	161	-	28,35,35	2.91	12 (42%)	36,49,49	3.55	18 (50%)
3	MTX	B	162	-	28,35,35	2.67	12 (42%)	36,49,49	3.14	14 (38%)

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	MTX	A	161	-	-	0/19/25/25	0/3/3/3
3	MTX	B	162	-	-	0/19/25/25	0/3/3/3

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	162	MTX	C7-C6	-5.98	1.28	1.39
3	A	161	MTX	C16-C11	-3.62	1.33	1.39
3	B	162	MTX	CA-N	-3.19	1.42	1.46
3	B	162	MTX	CB-CA	-2.75	1.49	1.53
3	A	161	MTX	C4-NA4	-2.48	1.23	1.34

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	161	MTX	C2-N1-C8A	-8.95	104.71	115.16
3	A	161	MTX	N8-C8A-N1	-8.52	105.02	116.04
3	A	161	MTX	CB-CA-CT	-7.30	101.68	112.28
3	B	162	MTX	C15-C16-C11	-6.16	113.90	120.79
3	A	161	MTX	C11-C-N	-5.64	106.55	116.97

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	161	MTX	2	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	12
1	A	7

The worst 5 of 19 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	122[B]:ASP	C	123:THR	N	1.82
1	B	44:ARG	C	45[B]:HIS	N	1.65
1	A	30:TRP	C	31:PHE	N	1.20
1	B	34:ASN	C	35:THR	N	1.20
1	B	154:LYS	C	155:ILE	N	1.20

## 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.