



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 7, 2022 – 06:20 PM EDT

PDB ID : 1EG7
Title : THE CRYSTAL STRUCTURE OF FORMYLTETRAHYDROFOLATE SYNTHETASE FROM MOORELLA THERMOACETICA
Authors : Radfar, R.; Shin, R.; Sheldrick, G.M.; Minor, W.; Lovell, C.R.; Odom, J.D.; Dunlap, R.B.; Lebioda, L.
Deposited on : 2000-02-14
Resolution : 2.50 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtrriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.27

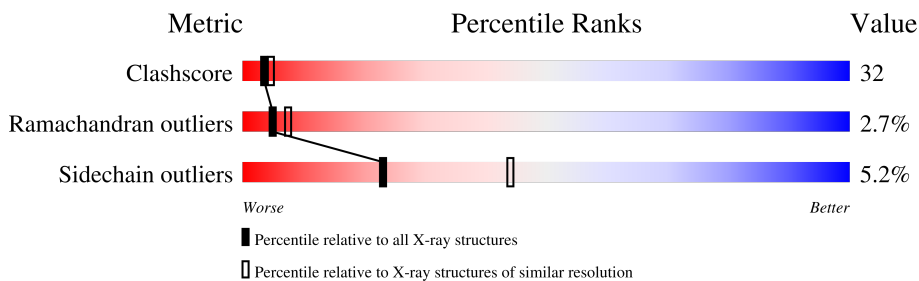
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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 $\leq 5\%$.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain	
1	A	557	61%	34% ..
1	B	557	44%	50% 5% .

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
2	SO4	A	3	-	-	X	-
2	SO4	A	5	-	-	X	-
2	SO4	A	7	-	-	X	-
2	SO4	A	8	-	-	X	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 8586 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 FORMYLTETRAHYDROFOLATE SYNTHETASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	549	4135	2618	716	780	21	0	0	0
1	B	548	4127	2614	715	777	21	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1174	GLU	ASP	conflict	UNP P21164
A	1225	LYS	ILE	conflict	UNP P21164
A	?	-	GLU	deletion	UNP P21164
A	?	-	VAL	deletion	UNP P21164
B	1174	GLU	ASP	conflict	UNP P21164
B	1225	LYS	ILE	conflict	UNP P21164
B	?	-	GLU	deletion	UNP P21164
B	?	-	VAL	deletion	UNP P21164

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	198	Total O 198 198	0	0

Continued on next page...

Continued from previous page...

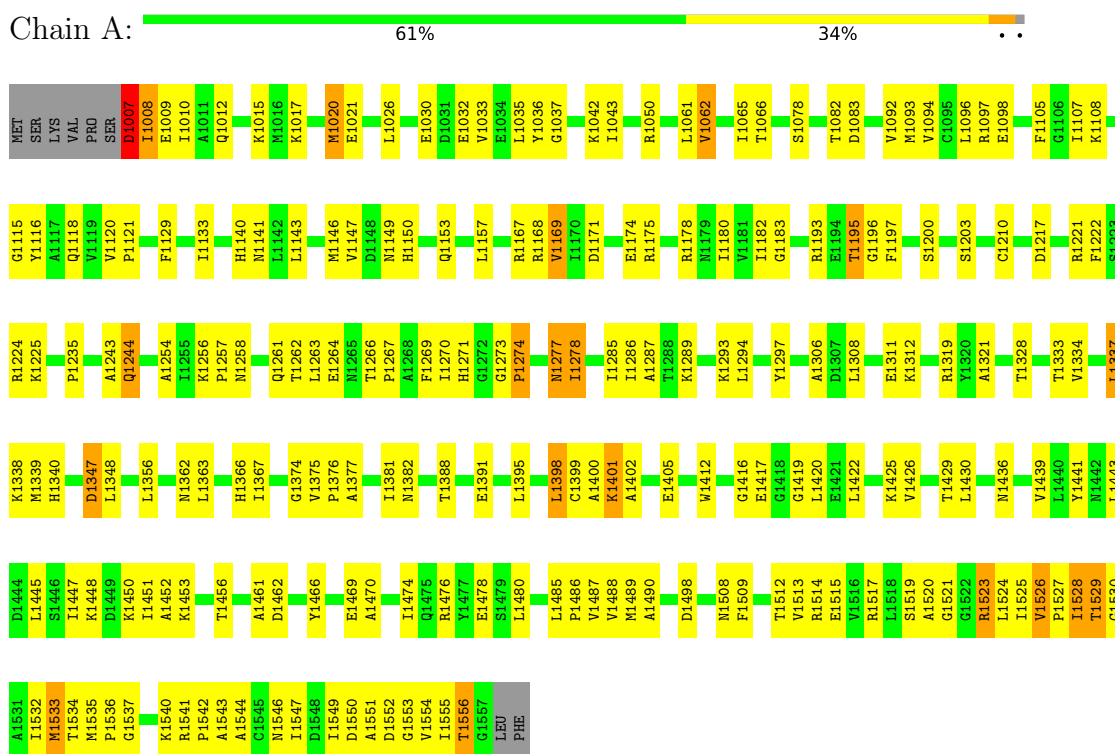
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	71	Total 71	O 71	0	0

3 Residue-property plots

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. 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. 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: FORMYLTETRAHYDROFOLATE SYNTHETASE



- Molecule 1: FORMYLTETRAHYDROFOLATE SYNTHETASE



K1540	F1222
R1541	S1223
P1542	R1224
A1543	G1225
A1544	V1226
C1545	V1227
M1546	V1236
I1547	
D1548	Q1244
I1549	L1250
D1550	
A1551	D1253
D1552	A1254
	I1255
I1555	K1256
T1556	P1257
G1557	M1258
LEU	L1259
PHE	V1260
	Q1261
	T1262
	L1263
	F1264
	M1265
	T1266
	P1267
	A1268
	F1269
	I1270
	H1271
	G1272
	G1273
	P1274
	F1275
	A1276
	N1277
	I1278
	A1279
	H1280
	G1281
	C1282
	M1283
	S1284
	I1285
	I1286
	A1287
	T1288
	K1289
	T1290
	A1291
	L1292
	K1293
	L1294
	A1295
	D1296
	Y1297
	V1298
	V1299
	T1300
	E1301
	A1302
	G1303
	F1304
	E1311
	K1312
	F1313
	V1314
	D1315
	V1316
	R1319
	Y1320
	A1321
	I1322
	P1325
	D1326
	V1329
	I1330
	V1331
	A1332
	T1333
	V1334
	R1335
	A1336
	P1337
	K1338
	V1343
	T1350
	E1351
	M1352
	L1353
	L1356
	R1357
	F1360
	A1361
	M1362
	L1363
	E1364
	K1365
	H1366
	I1367
	E1368
	M1369
	I1370
	G1371
	K1372
	F1373
	G1374
	V1375
	P1376
	A1377
	V1378
	V1379
	A1380
	I1381
	N1382
	A1383
	F1384
	P1385
	T1386
	D1387
	T1388
	E1389
	M1393
	L1394
	L1398
	C1399
	A1400
	K1401
	E1405
	L1408
	S1409
	K1414
	G1415
	G1416
	E1417
	G1418
	G1419
	L1420
	E1421
	L1422
	K1425
	V1426
	L1430
	E1431
	S1432
	R1433
	P1434
	S1435
	M1436
	F1437
	L1440
	L1443
	D1444
	L1445
	S1446
	I1447
	K1448
	D1449
	K1450
	I1451
	I1454
	A1455
	I1458
	Y1459
	G1460
	A1461
	D1462
	Y1466
	K1472
	R1476
	Y1477
	E1478
	S1479
	L1480
	G1481
	Y1482
	G1483
	M1484
	L1485
	P1486
	V1487
	V1488
	M1489
	A1490
	D1498
	P1506
	F1509
	T1510
	I1511
	V1512
	V1513
	R1514
	E1515
	V1516
	R1517
	L1518
	S1519
	A1520
	R1523
	L1524
	I1525
	V1526
	P1527
	I1528
	T1529
	G1530
	A1531
	I1532
	M1533
	T1534
	M1535
	P1536
	G1537

4 Data and refinement statistics

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

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α , β , γ	160.88Å 160.88Å 256.12Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.99 – 2.50	Depositor
% Data completeness (in resolution range)	90.6 (19.99-2.50)	Depositor
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 0.9	Depositor
R, R_{free}	0.253 , 0.301	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	8586	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

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: 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	0.43	3/4203 (0.1%)	0.70	7/5691 (0.1%)
1	B	0.39	0/4195	0.64	1/5680 (0.0%)
All	All	0.41	3/8398 (0.0%)	0.67	8/11371 (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	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1008	ILE	N-CA	9.27	1.64	1.46
1	A	1008	ILE	CA-C	5.74	1.67	1.52
1	A	1007	ASP	CA-C	5.03	1.66	1.52

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1007	ASP	N-CA-C	7.38	130.92	111.00
1	B	1015	LYS	N-CA-C	7.38	130.92	111.00
1	A	1007	ASP	CA-C-O	-6.80	105.83	120.10
1	A	1008	ILE	CB-CA-C	-6.64	98.32	111.60
1	A	1007	ASP	CB-CA-C	6.52	123.44	110.40

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1007	ASP	Peptide,Mainchain

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	4135	0	4223	218	1
1	B	4127	0	4219	317	0
2	A	35	0	0	12	0
2	B	20	0	0	2	0
3	A	198	0	0	14	0
3	B	71	0	0	4	0
All	All	8586	0	8442	535	1

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 32.

The worst 5 of 535 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:1009:GLU:OE2	1:B:1111:ALA:HB2	1.33	1.27
1:A:1175:ARG:HD3	2:A:5:SO4:O3	1.55	1.07
1:A:1244:GLN:H	1:A:1244:GLN:NE2	1.55	1.03
1:B:1277:ASN:ND2	1:B:1278:ILE:H	1.59	1.01
1:B:1244:GLN:H	1:B:1244:GLN:NE2	1.59	0.99

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1462:ASP:O	1:A:1462:ASP:O[12_555]	2.02	0.18

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	547/557 (98%)	485 (89%)	54 (10%)	8 (2%)	10	18
1	B	546/557 (98%)	444 (81%)	81 (15%)	21 (4%)	3	4
All	All	1093/1114 (98%)	929 (85%)	135 (12%)	29 (3%)	5	7

5 of 29 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1015	LYS
1	A	1401	LYS
1	B	1015	LYS
1	A	1399	CYS
1	B	1065	ILE

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	432/440 (98%)	409 (95%)	23 (5%)	22	43
1	B	431/440 (98%)	409 (95%)	22 (5%)	24	45
All	All	863/880 (98%)	818 (95%)	45 (5%)	23	44

5 of 45 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1195	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	1311	GLU
1	B	1244	GLN
1	B	1275	PHE
1	B	1382	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1140	HIS
1	B	1158	ASN
1	B	1382	ASN
1	B	1150	HIS
1	B	1189	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](#)

11 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
2	SO4	A	1	-	4,4,4	0.82	0	6,6,6	0.50	0
2	SO4	A	3	-	4,4,4	0.82	0	6,6,6	0.50	0
2	SO4	B	2	-	4,4,4	0.82	0	6,6,6	0.50	0
2	SO4	A	5	-	4,4,4	0.82	0	6,6,6	0.50	0
2	SO4	A	6	-	4,4,4	0.82	0	6,6,6	0.50	0
2	SO4	A	4	-	4,4,4	0.82	0	6,6,6	0.50	0
2	SO4	B	11	-	4,4,4	0.82	0	6,6,6	0.50	0
2	SO4	B	9	-	4,4,4	0.82	0	6,6,6	0.50	0
2	SO4	B	10	-	4,4,4	0.82	0	6,6,6	0.50	0
2	SO4	A	7	-	4,4,4	0.82	0	6,6,6	0.50	0
2	SO4	A	8	-	4,4,4	0.82	0	6,6,6	0.50	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	3	SO4	6	0
2	B	2	SO4	1	0
2	A	5	SO4	2	0
2	B	9	SO4	1	0
2	A	7	SO4	2	0
2	A	8	SO4	2	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.