



wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 08:06 PM GMT

PDB ID : 4DPD
Title : WILD TYPE PLASMODIUM FALCIPARUM DIHYDROFOLATE REDUCTASE-THYMIDYLATESYNTASE (PfDHFR-TS), DHF COMPLEX, NADP+, dUMP
Authors : Yuthavong, Y.; Vilaivan, T.; Kamchonwongpaisan, S.; Charman, S.A.; McLennan, D.N.; White, K.L.; Vivas, L.; Bongard, E.; Chitnumsub, P.; Tarnchompoo, B.; Thongphanchang, C.; Taweechai, S.; Vanichtanakul, J.; Arwon, U.; Fantauzzi, P.; Yuvaniyama, J.; Charman, W.N.; Matthews, D.
Deposited on : 2012-02-13
Resolution : 2.50 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

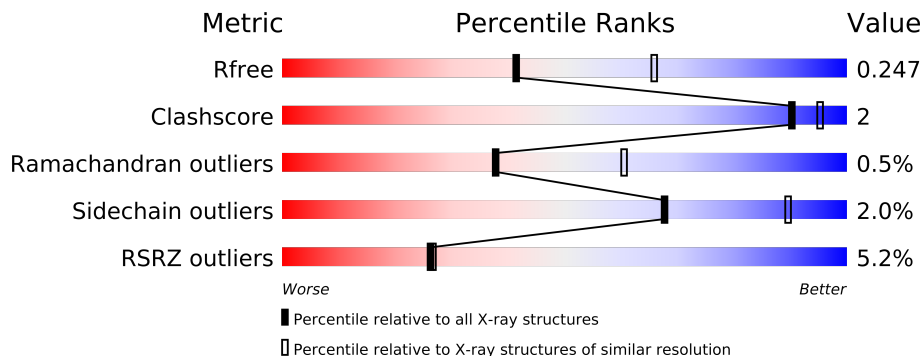
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

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))
R_{free}	66092	2784 (2.50-2.50)
Clashscore	79885	3562 (2.50-2.50)
Ramachandran outliers	78287	3480 (2.50-2.50)
Sidechain outliers	78261	3482 (2.50-2.50)
RSRZ outliers	66119	2785 (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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	608	
1	B	608	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	DHF	A	701	-	X

2 Entry composition i

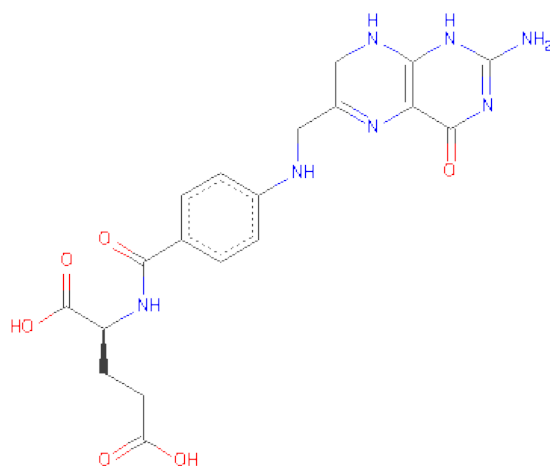
There are 5 unique types of molecules in this entry. The entry contains 9454 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Bifunctional dihydrofolate reductase-thymidylatesynthase.

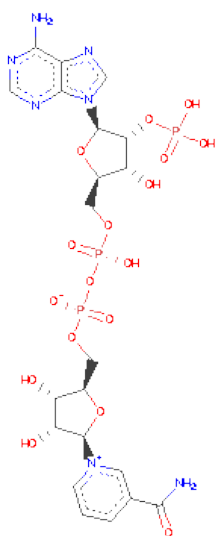
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	530	Total	C	N	O	S	0	0	0
			4413	2857	725	804	27			
1	B	543	Total	C	N	O	S	0	0	0
			4514	2914	745	829	26			

- Molecule 2 is DIHYDROFOLIC ACID (three-letter code: DHF) (formula: C₁₉H₂₁N₇O₆).



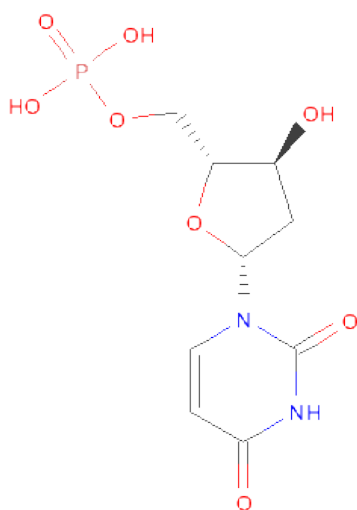
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			32	19	7	6		

- Molecule 3 is NADP NICOTINAMIDE-ADENINE-DINUCLEOTIDEPHOSPHATE (three-letter code: NAP) (formula: C₂₁H₂₈N₇O₁₇P₃).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			48	21	7	17	3		

- Molecule 4 is 2'-DEOXYURIDINE 5'-MONOPHOSPHATE (three-letter code: UMP) (formula: $C_9H_{13}N_2O_8P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			20	9	2	8	1		
4	B	1	Total	C	N	O	P	0	0
			20	9	2	8	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	222	Total 222	O 222	0	0
5	B	185	Total 185	O 185	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	58.06Å 156.78Å 165.03Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.82 – 2.50 49.82 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.7 (49.82-2.50) 99.6 (49.82-2.50)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.09 (at 2.48Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.212 , 0.266 0.202 , 0.247	Depositor DCC
R_{free} test set	2112 reflections (5.27%)	DCC
Wilson B-factor (Å ²)	36.8	Xtriage
Anisotropy	0.553	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 26.0	EDS
Estimated twinning fraction	0.002 for -h,l,k	Xtriage
L-test for twinning	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.36$	Xtriage
Outliers	0 of 53137 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	9454	wwPDB-VP
Average B, all atoms (Å ²)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.57% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: DHF, NAP, UMP

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	0.54	0/4517	0.65	0/6099
1	B	0.52	0/4619	0.63	0/6237
All	All	0.53	0/9136	0.64	0/12336

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	1
All	All	0	3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	165	GLY	Peptide
1	A	304	LYS	Peptide
1	B	165	GLY	Peptide

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit,

and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4413	0	4370	24	0
1	B	4514	0	4462	15	0
2	A	32	0	19	1	0
3	A	48	0	25	4	0
4	A	20	0	11	1	0
4	B	20	0	11	0	0
5	A	222	0	0	2	0
5	B	185	0	0	3	0
All	All	9454	0	8898	38	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 2.

The worst 5 of 38 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:50:CYS:SG	5:A:975:HOH:O	2.30	0.88
1:B:59:CYS:SG	5:B:984:HOH:O	2.34	0.85
1:A:59:CYS:SG	5:A:931:HOH:O	2.16	0.66
1:B:210:VAL:HG12	1:B:224:ILE:HG22	1.80	0.64
1:A:40:LEU:O	3:A:702:NAP:H2N	2.03	0.58

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	522/608 (86%)	495 (95%)	25 (5%)	2 (0%)	43	66
1	B	537/608 (88%)	506 (94%)	28 (5%)	3 (1%)	33	55
All	All	1059/1216 (87%)	1001 (94%)	53 (5%)	5 (0%)	38	60

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	304	LYS
1	A	139	ASP
1	B	26	GLY
1	A	82	ASN
1	B	429	ILE

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	496/570 (87%)	489 (99%)	7 (1%)	78	94
1	B	508/570 (89%)	495 (97%)	13 (3%)	59	84
All	All	1004/1140 (88%)	984 (98%)	20 (2%)	68	89

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

Mol	Chain	Res	Type
1	B	50	CYS
1	B	114	LYS
1	B	302	LYS
1	B	7	ASP
1	B	28	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA chains 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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

4 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	DHF	A	701	-	34,34,34	1.61	3 (8%)	45,47,47	1.47	5 (11%)
3	NAP	A	702	-	52,52,52	1.00	2 (3%)	80,80,80	1.44	10 (12%)
4	UMP	A	703	-	21,21,21	1.38	1 (4%)	26,31,31	1.45	6 (23%)
4	UMP	B	701	-	21,21,21	1.37	1 (4%)	26,31,31	1.52	5 (19%)

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
2	DHF	A	701	-	-	0/20/31/31	0/1/3/3
3	NAP	A	702	-	-	0/35/67/67	0/3/5/5
4	UMP	A	703	-	-	0/7/22/22	0/2/2/2
4	UMP	B	701	-	-	0/7/22/22	0/2/2/2

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	701	DHF	C6-N5	6.15	1.36	1.28
4	A	703	UMP	C2-N1	4.95	1.43	1.38
4	B	701	UMP	C2-N1	4.46	1.43	1.38
3	A	702	NAP	C5A-C4A	3.44	1.48	1.40
2	A	701	DHF	C4-C4A	3.22	1.48	1.42

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	701	DHF	C4-N3-C2	5.37	122.71	116.91
3	A	702	NAP	N3A-C2A-N1A	-5.18	124.38	128.71
3	A	702	NAP	N3A-C4A-N9A	4.93	134.33	125.43
4	B	701	UMP	N3-C2-N1	4.29	119.55	115.97
3	A	702	NAP	C3N-C7N-N7N	3.74	122.02	117.77

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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, 95th 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	530/608 (87%)	-0.11	14 (2%) 53 55	20, 31, 72, 90	0
1	B	543/608 (89%)	0.14	41 (7%) 14 13	20, 38, 90, 90	0
All	All	1073/1216 (88%)	0.02	55 (5%) 26 27	20, 34, 89, 90	0

The worst 5 of 55 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	85	THR	8.5
1	B	2	MET	7.5
1	A	302	LYS	7.3
1	A	1	MET	7.1
1	A	299	LYS	6.0

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 carbohydrates in this entry.

6.4 Ligands ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th 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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	DHF	A	701	32/32	0.58	14.91	25,34,36,37	32
4	UMP	B	701	20/20	0.15	0.34	28,39,50,51	0
3	NAP	A	702	48/48	0.09	-1.11	24,33,41,45	0
4	UMP	A	703	20/20	0.12	-1.28	29,36,41,42	0

6.5 Other polymers

There are no such residues in this entry.