



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 03:52 PM GMT

PDB ID : 4DPD
Title : WILD TYPE PLASMODIUM FALCIPARUM DIHYDROFOLATE REDUCTASE-THYMIDYLATE SYNTHASE (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 Full wwPDB X-ray Structure Validation 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/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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

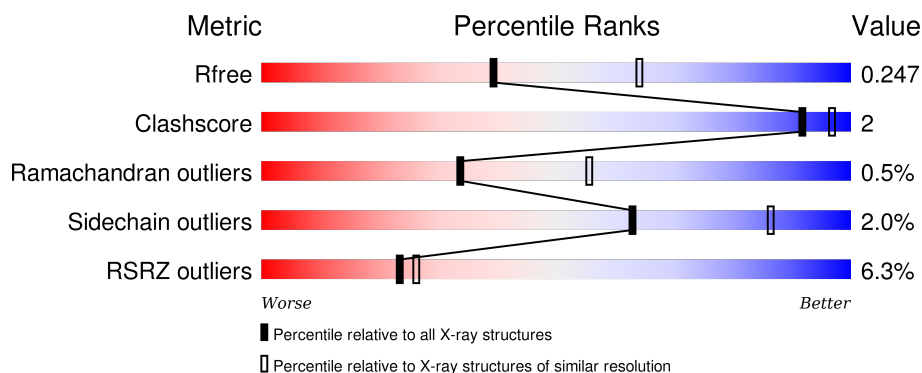
1 Overall quality at a glance

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))
R_{free}	91344	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (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. 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	608	
1	B	608	

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	DHF	A	701	-	-	-	X

2 Entry composition [i](#)

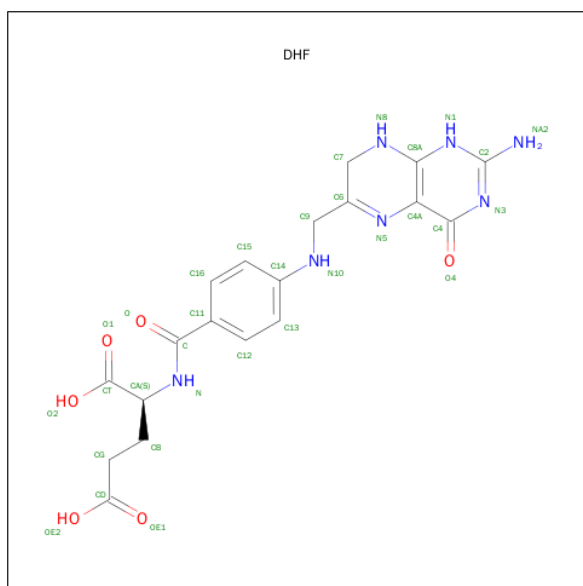
There are 5 unique types of molecules in this entry. The entry contains 9454 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Bifunctional dihydrofolate reductase-thymidylate synthase.

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).



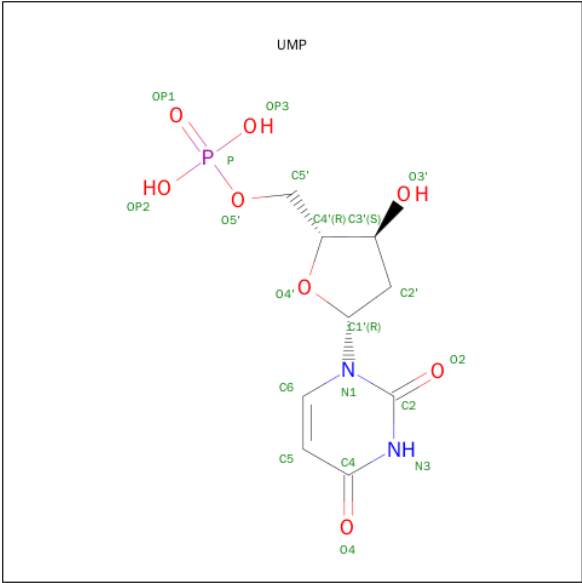
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			32	19	7	6		

- Molecule 3 is NADP NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (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: C₉H₁₃N₂O₈P).



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	58.06 Å 156.78 Å 165.03 Å 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 2.48 Å)	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 , 36.0	EDS
Estimated twinning fraction	0.002 for -h,l,k	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.36$	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.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: 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 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	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

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 2.

All (38) 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: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
1:A:166:GLY:HA3	3:A:702:NAP:PA	2.44	0.57
1:A:127:LEU:HD23	1:A:143:ILE:HG13	1.88	0.54
1:A:19:LYS:HG2	1:A:36:THR:HG22	1.94	0.50
1:B:303:ASN:O	1:B:305:ASN:N	2.47	0.47
1:B:19:LYS:HG2	1:B:36:THR:HG22	1.96	0.46
1:A:181:LYS:NZ	1:B:289:ASP:OD2	2.46	0.45
1:A:315:TYR:HB2	1:A:564:LEU:O	2.16	0.45
1:A:319:LYS:HD3	1:B:286:GLU:HG3	1.98	0.45
1:A:456:LEU:O	1:A:459:ILE:HG13	2.17	0.45
1:B:192:GLU:O	5:B:850:HOH:O	2.21	0.44
1:B:109:TRP:CE2	1:B:117:LYS:HD2	2.53	0.44
1:B:415:ASN:ND2	5:B:829:HOH:O	2.51	0.43
1:B:114:LYS:HA	1:B:117:LYS:HB3	2.00	0.43
1:A:214:TYR:O	1:A:220:THR:HA	2.19	0.42
1:A:100:ASN:OD1	1:A:159:TYR:HB3	2.20	0.42
2:A:701:DHf:H72	3:A:702:NAP:C5N	2.50	0.42
1:B:165:GLY:HA3	1:B:170:TYR:CZ	2.55	0.42
1:A:153:LEU:HD22	1:A:158:TYR:CZ	2.55	0.42
1:A:166:GLY:HA3	3:A:702:NAP:O1A	2.20	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:41:GLY:HA2	1:A:47:PRO:HD3	2.02	0.41
1:B:569:TYR:HB3	1:B:570:PRO:HD2	2.01	0.41
1:A:491:HIS:CE1	4:A:703:UMP:O4	2.73	0.41
1:A:165:GLY:HA3	1:A:170:TYR:CZ	2.55	0.41
1:A:435:ARG:HB2	1:A:436:HIS:CD2	2.56	0.41
1:A:65:VAL:HG22	1:A:159:TYR:CB	2.51	0.41
1:A:387:ILE:O	1:A:435:ARG:NH1	2.50	0.41
1:A:124:ASN:HB2	1:A:140:VAL:HG12	2.02	0.41
1:B:165:GLY:CA	1:B:170:TYR:CZ	3.04	0.41
1:B:376:LEU:HD12	1:B:593:ILE:HG13	2.03	0.40
1:A:48:TRP:O	1:A:49:LYS:HB2	2.22	0.40
1:A:485:MET:SD	1:A:489:PRO:HD3	2.62	0.40
1:B:7:ASP:HA	1:B:180:LYS:HE2	2.03	0.40
1:A:289:ASP:HA	1:A:292:TYR:CD2	2.57	0.40

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%)	39	61
1	B	537/608 (88%)	506 (94%)	28 (5%)	3 (1%)	30	50
All	All	1059/1216 (87%)	1001 (94%)	53 (5%)	5 (0%)	34	55

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

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
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%)	74	91
1	B	508/570 (89%)	495 (97%)	13 (3%)	54	81
All	All	1004/1140 (88%)	984 (98%)	20 (2%)	63	86

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	MET
1	A	50	CYS
1	A	65	VAL
1	A	114	LYS
1	A	285	GLU
1	A	344	ASP
1	A	487	LEU
1	B	7	ASP
1	B	28	LYS
1	B	50	CYS
1	B	114	LYS
1	B	116	PHE
1	B	137	ASP
1	B	145	LYS
1	B	195	VAL
1	B	299	LYS
1	B	302	LYS
1	B	344	ASP
1	B	402	ARG
1	B	516	LEU

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 molecules 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	-	25,34,34	1.46	2 (8%)	24,47,47	1.80	4 (16%)
3	NAP	A	702	-	42,52,52	0.92	1 (2%)	54,80,80	1.49	6 (11%)
4	UMP	A	703	-	16,21,21	0.66	0	23,31,31	2.09	5 (21%)
4	UMP	B	701	-	16,21,21	0.89	0	23,31,31	2.11	4 (17%)

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	-	-	0/14/31/31	0/3/3/3
3	NAP	A	702	-	-	0/27/67/67	0/5/5/5
4	UMP	A	703	-	-	0/6/22/22	0/2/2/2
4	UMP	B	701	-	-	0/6/22/22	0/2/2/2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	701	DHF	C4A-C8A	2.41	1.48	1.41
3	A	702	NAP	C5A-C4A	3.44	1.48	1.40
2	A	701	DHF	C6-N5	5.94	1.36	1.28

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	702	NAP	N3A-C2A-N1A	-5.90	124.38	128.89
2	A	701	DHF	C4A-C4-N3	-4.41	117.56	123.59
3	A	702	NAP	C4A-C5A-N7A	-2.91	106.80	109.48
3	A	702	NAP	PN-O3-PA	-2.84	124.74	132.73
3	A	702	NAP	O7N-C7N-N7N	-2.42	119.18	122.59
4	A	703	UMP	O5'-P-OP1	-2.21	101.52	107.14
4	B	701	UMP	O4'-C4'-C3'	-2.04	100.54	105.67
4	A	703	UMP	OP3-P-OP1	2.08	117.26	110.58
3	A	702	NAP	O2A-PA-O1A	2.17	124.31	112.53
4	A	703	UMP	OP3-P-OP2	2.35	116.31	107.38
4	A	703	UMP	C2'-C1'-N1	2.39	119.98	114.16
4	B	701	UMP	OP2-P-OP1	2.48	118.55	110.58
4	B	701	UMP	C2'-C1'-N1	2.56	120.39	114.16
2	A	701	DHF	C4-C4A-C8A	2.60	116.26	114.52
2	A	701	DHF	C2-N1-C8A	3.12	121.54	114.54
3	A	702	NAP	C3N-C7N-N7N	3.84	122.02	117.82
2	A	701	DHF	C4-N3-C2	4.88	122.71	115.94
4	B	701	UMP	C4-N3-C2	7.92	121.98	114.14
4	A	703	UMP	C4-N3-C2	8.11	122.17	114.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	701	DHF	1	0
3	A	702	NAP	4	0
4	A	703	UMP	1	0

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%)	-0.13	15 (2%) 56 61	20, 31, 72, 90	0
1	B	543/608 (89%)	0.23	53 (9%) 10 10	20, 38, 90, 90	0
All	All	1073/1216 (88%)	0.05	68 (6%) 23 26	20, 34, 89, 90	0

All (68) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1	MET	8.3
1	B	2	MET	7.8
1	A	2	MET	7.4
1	B	85	THR	7.1
1	B	116	PHE	6.4
1	B	231	ASN	5.9
1	A	302	LYS	5.4
1	A	299	LYS	5.3
1	B	24	ASN	5.2
1	A	301	GLU	5.1
1	A	300	GLU	4.9
1	B	75	TYR	4.9
1	B	4	GLN	4.8
1	B	27	LYS	4.8
1	B	151	VAL	4.7
1	B	299	LYS	4.7
1	B	138	GLU	4.7
1	B	3	GLU	4.6
1	B	118	PRO	4.4
1	B	23	LYS	4.0
1	B	86	VAL	3.9
1	B	300	GLU	3.8
1	B	136	PHE	3.8
1	B	301	GLU	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	605	ASP	3.7
1	B	9	PHE	3.7
1	B	230	ASN	3.6
1	B	26	GLY	3.6
1	B	306	SER	3.6
1	B	8	VAL	3.6
1	B	119	LEU	3.5
1	B	127	LEU	3.5
1	B	7	ASP	3.2
1	B	117	LYS	3.2
1	A	3	GLU	3.0
1	B	114	LYS	2.9
1	B	178	LEU	2.8
1	B	25	GLU	2.8
1	B	5	VAL	2.8
1	B	307	ILE	2.7
1	B	284	ASP	2.7
1	A	312	PHE	2.6
1	A	306	SER	2.6
1	B	103	VAL	2.6
1	B	82	ASN	2.5
1	B	286	GLU	2.5
1	A	303	ASN	2.5
1	B	109	TRP	2.5
1	A	4	GLN	2.4
1	B	285	GLU	2.4
1	A	284	ASP	2.4
1	A	605	ASP	2.4
1	B	153	LEU	2.4
1	B	303	ASN	2.3
1	B	134	GLU	2.3
1	A	285	GLU	2.3
1	B	110	GLU	2.3
1	B	310	ASN	2.3
1	B	149	LEU	2.2
1	B	130	THR	2.2
1	B	111	SER	2.2
1	B	283	ASP	2.2
1	B	102	VAL	2.1
1	B	146	VAL	2.1
1	B	84	GLU	2.1
1	B	141	TYR	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	302	LYS	2.1
1	A	202	GLU	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [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. 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	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
2	DHF	A	701	32/32	0.59	0.58	21.04	25,34,36,37	32
4	UMP	B	701	20/20	0.97	0.17	1.08	28,39,50,51	0
4	UMP	A	703	20/20	0.97	0.13	-0.78	29,36,41,42	0
3	NAP	A	702	48/48	0.98	0.10	-1.13	24,33,41,45	0

6.5 Other polymers [i](#)

There are no such residues in this entry.