Full wwPDB X-ray Structure Validation Report ⓘ

May 29, 2020 – 12:40 pm BST

PDB ID	:	1Y13
Title	:	Structural Analysis of Plasmodium falciparum 6-pyruvoyl tetrahydropterin synthase (PTPS)
Authors	:	Bosch, J.; Hol, W.G.J.; Structural Genomics of Pathogenic Protozoa Consortium (SGPP)
Deposited on	:	2004-11-17
Resolution	:	2.20 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
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:

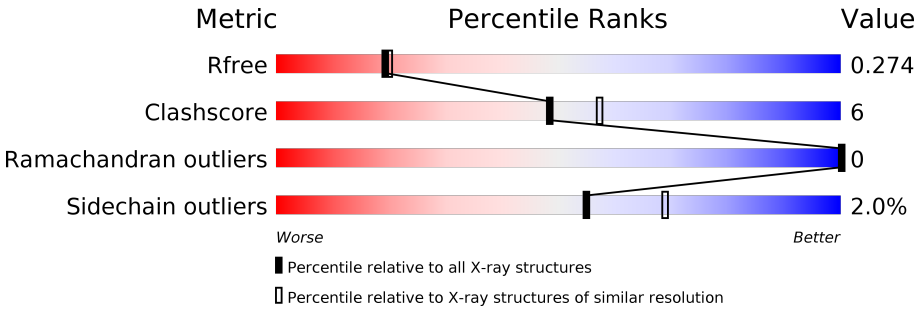
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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}	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)

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 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 <=5%

Mol	Chain	Length	Quality of chain
1	A	181	77% 13% • 10%
1	B	181	80% 10% 10%
1	C	181	76% 13% •• 10%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4150 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 6-pyruvoyl tetrahydropterin synthase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	163	Total	C	N	O	S	Se	0	0	0
			1319	848	222	244	4	1			
1	B	163	Total	C	N	O	S	Se	0	0	0
			1324	850	223	246	4	1			
1	C	163	Total	C	N	O	S	Se	0	0	0
			1329	851	224	249	4	1			

There are 27 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	MET	-	INITIATING METHIONINE	GB 23612358
A	-6	ALA	-	CLONING ARTIFACT	GB 23612358
A	-5	HIS	-	EXPRESSION TAG	GB 23612358
A	-4	HIS	-	EXPRESSION TAG	GB 23612358
A	-3	HIS	-	EXPRESSION TAG	GB 23612358
A	-2	HIS	-	EXPRESSION TAG	GB 23612358
A	-1	HIS	-	EXPRESSION TAG	GB 23612358
A	0	HIS	-	EXPRESSION TAG	GB 23612358
A	142	MSE	MET	MODIFIED RESIDUE	GB 23612358
B	-7	MET	-	INITIATING METHIONINE	GB 23612358
B	-6	ALA	-	CLONING ARTIFACT	GB 23612358
B	-5	HIS	-	EXPRESSION TAG	GB 23612358
B	-4	HIS	-	EXPRESSION TAG	GB 23612358
B	-3	HIS	-	EXPRESSION TAG	GB 23612358
B	-2	HIS	-	EXPRESSION TAG	GB 23612358
B	-1	HIS	-	EXPRESSION TAG	GB 23612358
B	0	HIS	-	EXPRESSION TAG	GB 23612358
B	142	MSE	MET	MODIFIED RESIDUE	GB 23612358
C	-7	MET	-	INITIATING METHIONINE	GB 23612358
C	-6	ALA	-	CLONING ARTIFACT	GB 23612358
C	-5	HIS	-	EXPRESSION TAG	GB 23612358
C	-4	HIS	-	EXPRESSION TAG	GB 23612358
C	-3	HIS	-	EXPRESSION TAG	GB 23612358

Continued on next page...

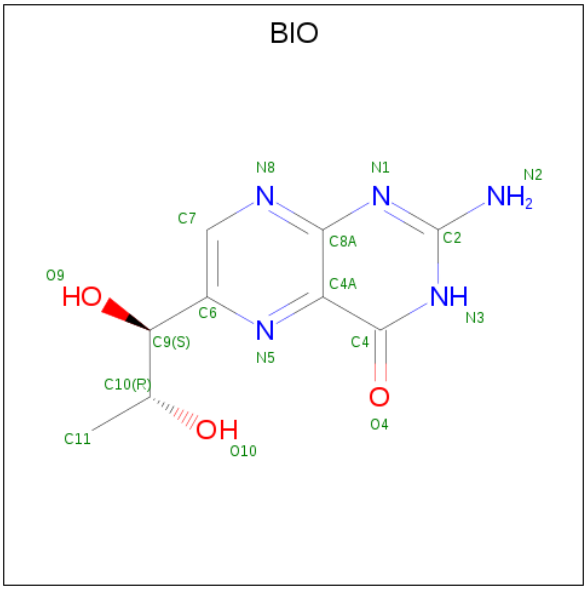
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-2	HIS	-	EXPRESSION TAG	GB 23612358
C	-1	HIS	-	EXPRESSION TAG	GB 23612358
C	0	HIS	-	EXPRESSION TAG	GB 23612358
C	142	MSE	MET	MODIFIED RESIDUE	GB 23612358

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 3 is BIOPTERIN (three-letter code: BIO) (formula: $C_9H_{11}N_5O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O 15 8 5 2	0	0
3	B	1	Total C N O 17 9 5 3	0	0
3	C	1	Total C N O 17 9 5 3	0	0

- Molecule 4 is water.

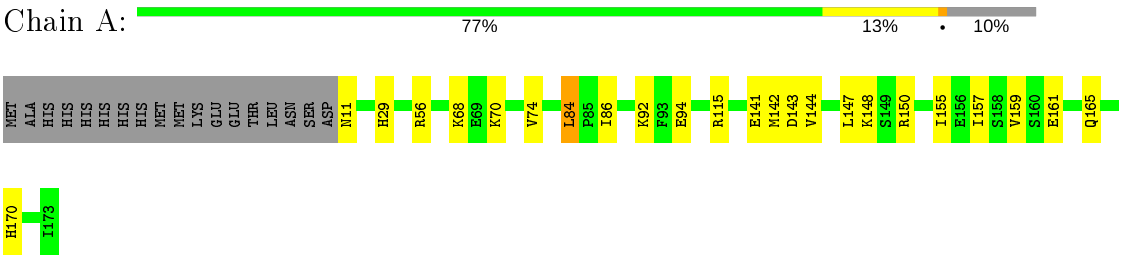
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	57	Total O 57 57	0	0
4	B	50	Total O 50 50	0	0
4	C	19	Total O 19 19	0	0

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. 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.

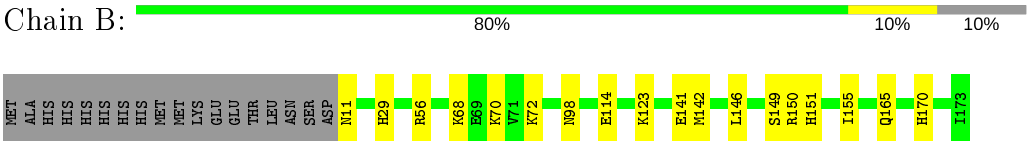
- Molecule 1: 6-pyruvoyl tetrahydropterin synthase

Chain A:



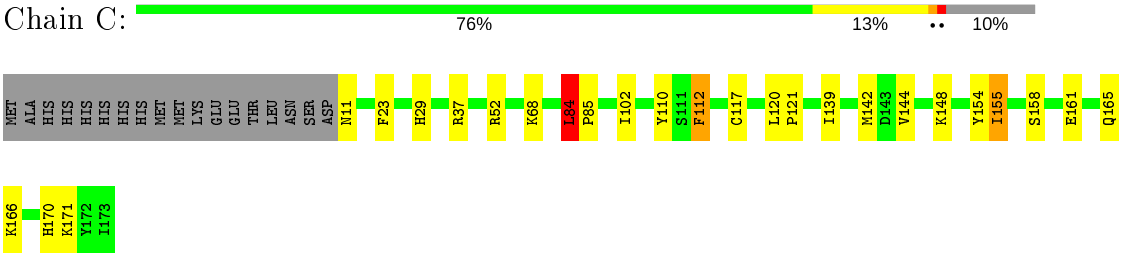
- Molecule 1: 6-pyruvoyl tetrahydropterin synthase

Chain B:



- Molecule 1: 6-pyruvoyl tetrahydropterin synthase

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	103.32Å 103.32Å 131.11Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.20 19.55 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.8 (20.00-2.20) 99.9 (19.55-2.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.16	Depositor
$< I/\sigma(I) >$ [1]	2.06 (at 2.21Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.195 , 0.235 0.235 , 0.274	Depositor DCC
R_{free} test set	1847 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	52.8	Xtriage
Anisotropy	0.085	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 57.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	4150	wwPDB-VP
Average B, all atoms (Å^2)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.49% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: BIO, ZN

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	0.60	0/1347	0.71	1/1817 (0.1%)
1	B	0.60	0/1353	0.69	0/1823
1	C	0.47	0/1358	0.60	1/1831 (0.1%)
All	All	0.56	0/4058	0.67	2/5471 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	84	LEU	CA-CB-CG	7.38	132.28	115.30
1	C	84	LEU	CA-CB-CG	6.14	129.43	115.30

There are no chirality outliers.

There are no planarity outliers.

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	1319	0	1296	20	0
1	B	1324	0	1307	15	0
1	C	1329	0	1314	22	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	15	0	8	1	0
3	B	17	0	10	0	0
3	C	17	0	11	0	0
4	A	57	0	0	2	0
4	B	50	0	0	0	0
4	C	19	0	0	1	0
All	All	4150	0	3946	48	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 6.

All (48) 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:142:MSE:HE1	1:B:155:ILE:CD1	1.99	0.92
1:A:68:LYS:HE3	1:B:165:GLN:OE1	1.74	0.87
3:A:175:BIO:N3	3:A:175:BIO:C4A	2.43	0.80
1:B:142:MSE:HE1	1:B:155:ILE:HD12	1.64	0.80
1:A:68:LYS:CE	1:B:165:GLN:OE1	2.30	0.79
1:A:165:GLN:OE1	1:C:68:LYS:HE3	1.83	0.78
1:A:161:GLU:OE2	1:A:165:GLN:NE2	2.23	0.69
1:A:11:ASN:HB3	1:B:170:HIS:NE2	2.09	0.67
1:B:68:LYS:HE2	1:C:165:GLN:OE1	1.94	0.67
1:A:56:ARG:NH1	1:A:150:ARG:HA	2.10	0.65
1:C:102:ILE:HD12	1:C:112:PHE:HE1	1.65	0.61
1:A:84:LEU:HD12	1:A:86:ILE:HG23	1.86	0.57
1:A:56:ARG:HG3	1:A:56:ARG:HH11	1.70	0.57
1:C:161:GLU:OE2	1:C:165:GLN:NE2	2.36	0.56
1:A:84:LEU:HD12	1:A:86:ILE:CG2	2.36	0.55
1:A:68:LYS:HE2	1:B:165:GLN:OE1	2.05	0.55
1:A:144:VAL:O	1:A:148:LYS:HG2	2.07	0.54
1:A:165:GLN:OE1	1:C:68:LYS:CE	2.56	0.52
1:A:157:ILE:HD12	1:A:159:VAL:HG23	1.92	0.52
1:B:142:MSE:O	1:B:146:LEU:HD23	2.10	0.51
1:A:142:MSE:O	1:A:143:ASP:HB3	2.11	0.51
1:A:56:ARG:NH1	4:A:184:HOH:O	2.43	0.50
1:C:142:MSE:HE1	1:C:155:ILE:HD12	1.93	0.49
1:B:70:LYS:HD2	1:B:141:GLU:HB3	1.94	0.49
1:A:147:LEU:HD13	1:A:155:ILE:HD11	1.94	0.48
1:C:23:PHE:HZ	1:C:68:LYS:HG2	1.80	0.46
1:B:98:ASN:O	1:B:114:GLU:HG2	2.17	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:144:VAL:O	1:C:148:LYS:HG2	2.17	0.45
1:A:170:HIS:NE2	1:C:11:ASN:HB3	2.31	0.45
1:C:139:ILE:CD1	1:C:171:LYS:HD2	2.47	0.45
1:A:70:LYS:HD2	1:A:141:GLU:HB3	2.00	0.44
1:A:74:VAL:HG13	4:A:230:HOH:O	2.17	0.44
1:B:68:LYS:CE	1:C:165:GLN:OE1	2.64	0.44
1:C:37:ARG:NH1	1:C:110:TYR:OH	2.51	0.44
1:C:23:PHE:N	1:C:23:PHE:CD1	2.83	0.44
1:C:84:LEU:HA	1:C:85:PRO:HD3	1.85	0.44
1:A:92:LYS:HE3	1:A:94:GLU:HG2	2.00	0.44
1:B:11:ASN:HB3	1:C:170:HIS:NE2	2.33	0.43
1:B:56:ARG:HD3	1:B:150:ARG:O	2.18	0.43
1:C:23:PHE:CZ	1:C:68:LYS:HG2	2.53	0.43
1:C:84:LEU:HD23	1:C:117:CYS:HB3	2.00	0.43
1:C:158:SER:HB3	1:C:166:LYS:HD2	2.01	0.42
1:B:123:LYS:HB3	1:B:123:LYS:HE2	1.83	0.42
1:C:166:LYS:NZ	4:C:184:HOH:O	2.54	0.41
1:C:102:ILE:HD12	1:C:112:PHE:CE1	2.52	0.41
1:C:120:LEU:HA	1:C:121:PRO:HD3	1.94	0.40
1:C:52:ARG:HB3	1:C:154:TYR:HB3	2.03	0.40
1:B:150:ARG:O	1:B:151:HIS:HB2	2.20	0.40

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	161/181 (89%)	155 (96%)	6 (4%)	0	100	100
1	B	161/181 (89%)	155 (96%)	6 (4%)	0	100	100
1	C	161/181 (89%)	154 (96%)	7 (4%)	0	100	100
All	All	483/543 (89%)	464 (96%)	19 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	149/170 (88%)	147 (99%)	2 (1%)	69	81
1	B	152/170 (89%)	149 (98%)	3 (2%)	55	69
1	C	153/170 (90%)	149 (97%)	4 (3%)	46	58
All	All	454/510 (89%)	445 (98%)	9 (2%)	55	69

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	29	HIS
1	A	115	ARG
1	B	29	HIS
1	B	72	LYS
1	B	149	SER
1	C	29	HIS
1	C	84	LEU
1	C	112	PHE
1	C	155	ILE

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 (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 6 ligands modelled in this entry, 3 are monoatomic - leaving 3 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	BIO	B	175	2	17,18,18	2.25	4 (23%)	17,26,26	3.88	13 (76%)
3	BIO	A	175	2	13,15,18	1.68	3 (23%)	11,20,26	4.12	5 (45%)
3	BIO	C	175	2	17,18,18	2.63	6 (35%)	17,26,26	3.68	10 (58%)

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	BIO	B	175	2	-	2/8/8/8	0/2/2/2
3	BIO	A	175	2	-	3/10/8/8	0/1/1/2
3	BIO	C	175	2	-	1/8/8/8	0/2/2/2

All (13) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	175	BIO	C6-N5	6.66	1.38	1.32
3	B	175	BIO	C6-N5	6.31	1.38	1.32
3	C	175	BIO	C4-C4A	5.32	1.50	1.41
3	C	175	BIO	C4A-C8A	4.15	1.48	1.40
3	B	175	BIO	C4-C4A	3.80	1.47	1.41
3	A	175	BIO	C10-C9	3.41	1.57	1.53
3	A	175	BIO	C4A-C8A	3.27	1.48	1.41
3	A	175	BIO	O9-C9	2.84	1.48	1.42

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	175	BIO	C4A-C8A	2.68	1.45	1.40
3	C	175	BIO	C10-C9	2.65	1.56	1.53
3	C	175	BIO	C4A-N5	2.61	1.37	1.33
3	B	175	BIO	C8A-N8	-2.47	1.33	1.37
3	C	175	BIO	C8A-N8	-2.05	1.34	1.37

All (28) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	175	BIO	C4A-N5-C6	8.07	123.84	117.95
3	A	175	BIO	C4A-C8A-N8	-8.01	117.58	121.27
3	B	175	BIO	C4-C4A-C8A	-7.88	114.73	119.95
3	C	175	BIO	C4-C4A-C8A	-7.17	115.20	119.95
3	C	175	BIO	C4-C4A-N5	6.78	126.35	118.60
3	B	175	BIO	C4-C4A-N5	6.08	125.55	118.60
3	B	175	BIO	C2-N1-C8A	5.12	121.21	115.36
3	B	175	BIO	N8-C8A-N1	5.07	121.61	115.82
3	C	175	BIO	C2-N1-C8A	5.06	121.14	115.36
3	B	175	BIO	C7-N8-C8A	5.04	121.75	116.69
3	A	175	BIO	C7-N8-C8A	4.78	122.67	117.82
3	A	175	BIO	C8A-C4A-N5	-4.77	116.77	122.41
3	B	175	BIO	N1-C2-N3	-4.54	121.17	127.22
3	C	175	BIO	C4-N3-C2	4.52	123.10	115.93
3	C	175	BIO	N8-C8A-N1	4.09	120.49	115.82
3	B	175	BIO	C4-N3-C2	3.98	122.26	115.93
3	C	175	BIO	C7-N8-C8A	3.97	120.68	116.69
3	C	175	BIO	N1-C2-N3	-3.87	122.06	127.22
3	C	175	BIO	C4A-C4-N3	-3.81	118.22	123.43
3	C	175	BIO	C8A-C4A-N5	-3.43	118.44	122.33
3	B	175	BIO	O10-C10-C11	3.00	118.62	109.74
3	B	175	BIO	C4A-C4-N3	-2.87	119.51	123.43
3	B	175	BIO	N2-C2-N3	2.83	121.66	117.25
3	B	175	BIO	O9-C9-C10	-2.43	102.69	110.06
3	B	175	BIO	C6-N5-C4A	2.35	120.18	118.07
3	B	175	BIO	C8A-C4A-N5	-2.30	119.72	122.33
3	C	175	BIO	C6-N5-C4A	2.13	119.98	118.07
3	A	175	BIO	C11-C10-C9	2.12	116.67	111.55

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	175	BIO	O10-C10-C9-C6
3	A	175	BIO	C11-C10-C9-C6
3	B	175	BIO	C7-C6-C9-O9
3	A	175	BIO	C11-C10-C9-O9
3	B	175	BIO	C11-C10-C9-O9
3	C	175	BIO	C11-C10-C9-O9

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	175	BIO	1	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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.