



Full wwPDB X-ray Structure Validation Report

Mar 1, 2014 – 12:17 AM GMT

PDB ID : 1N4L
Title : A DNA analogue of the polypurine tract of HIV-1
Authors : Cote', M.L.; Pflomm, M.; Georgiadis, M.M.
Deposited on : 2002-10-31
Resolution : 2.00 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

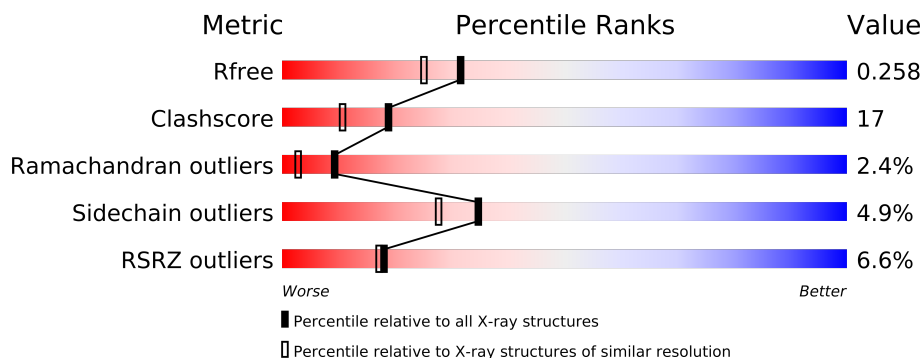
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.00 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	4888 (2.00-2.00)
Clashscore	79885	6188 (2.00-2.00)
Ramachandran outliers	78287	6102 (2.00-2.00)
Sidechain outliers	78261	6100 (2.00-2.00)
RSRZ outliers	66119	4890 (2.00-2.00)

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	B	16	
2	D	16	
3	A	255	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2834 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 DNA chain called 5'-D(*CP*TP*TP*TP*TP*TP*AP*AP*AP*AP*GP*AP*AP*AP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	16	Total	C	N	O	P	0	16	0
			328	159	63	91	15			

- Molecule 2 is a DNA chain called 5'-D(*CP*TP*TP*TP*TP*CP*TP*TP*TP*TP*AP*AP*AP*AP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	16	Total	C	N	O	P	0	16	0
			322	158	52	97	15			

- Molecule 3 is a protein called Reverse Transcriptase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	255	Total	C	N	O	S	0	0	0
			2039	1311	356	365	7			

- Molecule 4 is water.

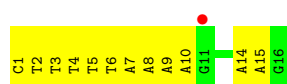
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	143	Total	O	0	0
			143	143		
4	D	2	Total	O	0	0
			2	2		

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 errors 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.

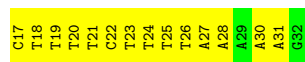
- Molecule 1: 5'-D(*CP*TP*TP*TP*TP*TP*AP*AP*AP*AP*GP*AP*AP*AP*AP*G)-3'

Chain B: 



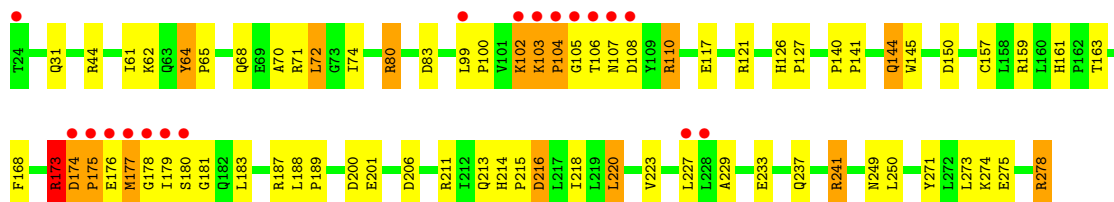
- Molecule 2: 5'-D(*CP*TP*TP*TP*TP*CP*TP*TP*TP*TP*AP*AP*AP*AP*AP*G)-3'

Chain D: 



- Molecule 3: Reverse Transcriptase

Chain A: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	54.41 Å 146.22 Å 46.78 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.00 19.92 – 2.00	Depositor EDS
% Data completeness (in resolution range)	93.2 (50.00-2.00) 93.8 (19.92-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.91 (at 2.01 Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.240 , 0.269 0.229 , 0.258	Depositor DCC
R_{free} test set	1249 reflections (5.40%)	DCC
Wilson B-factor (Å ²)	32.0	Xtriage
Anisotropy	0.178	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 51.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 24373 reflections	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	2834	wwPDB-VP
Average B, all atoms (Å ²)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.59% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

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	$\# Z > 5$
3	A	0.69	0/2095	1.29	23/2857 (0.8%)

There are no bond length outliers.

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	187	ARG	NE-CZ-NH1	9.95	125.28	120.30
3	A	278	ARG	NE-CZ-NH1	9.35	124.97	120.30
3	A	44	ARG	NE-CZ-NH2	8.48	124.54	120.30
3	A	211	ARG	NE-CZ-NH2	-8.30	116.15	120.30
3	A	187	ARG	NE-CZ-NH2	-8.00	116.30	120.30
3	A	278	ARG	CD-NE-CZ	7.85	134.59	123.60
3	A	241	ARG	NE-CZ-NH1	7.57	124.08	120.30
3	A	150	ASP	CB-CG-OD1	7.29	124.86	118.30
3	A	80	ARG	NE-CZ-NH2	7.21	123.91	120.30
3	A	159	ARG	NE-CZ-NH2	-6.59	117.01	120.30
3	A	44	ARG	NE-CZ-NH1	-6.16	117.22	120.30
3	A	83	ASP	CB-CG-OD1	6.09	123.78	118.30
3	A	174	ASP	N-CA-C	-6.04	94.69	111.00
3	A	201	GLU	OE1-CD-OE2	-5.71	116.44	123.30
3	A	216	ASP	CB-CG-OD2	5.69	123.42	118.30
3	A	241	ARG	NH1-CZ-NH2	-5.62	113.22	119.40
3	A	110	ARG	NE-CZ-NH1	5.62	123.11	120.30
3	A	168	PHE	CB-CG-CD1	5.25	124.48	120.80
3	A	173	ARG	NE-CZ-NH1	5.24	122.92	120.30
3	A	200	ASP	CB-CG-OD1	5.18	122.96	118.30
3	A	157	CYS	CA-CB-SG	-5.12	104.79	114.00
3	A	121	ARG	NE-CZ-NH2	-5.11	117.74	120.30
3	A	211	ARG	NE-CZ-NH1	5.07	122.83	120.30

There are no chirality outliers.

There are no planarity outliers.

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	B	328	0	183	9	3
2	D	322	0	186	16	0
3	A	2039	0	2056	61	1
4	A	143	0	0	12	0
4	D	2	0	0	0	0
All	All	2834	0	2425	85	4

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 17.

All (85) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:5[B]:DT:H4'	1:B:6[B]:DT:OP1	1.38	1.07
3:A:177:MET:HE1	3:A:178:GLY:HA2	1.14	1.06
1:B:14[B]:DA:H2''	1:B:15[B]:DA:H5'	1.41	1.03
2:D:24[D]:DT:H4'	2:D:25[D]:DT:OP1	1.58	1.02
3:A:278:ARG:HD2	4:A:377:HOH:O	1.58	1.01
2:D:30[D]:DA:H2''	2:D:31[D]:DA:H5'	1.41	0.99
3:A:177:MET:HE1	3:A:178:GLY:CA	1.95	0.96
3:A:177:MET:CE	3:A:178:GLY:HA2	1.93	0.96
3:A:100:PRO:HG2	4:A:366:HOH:O	1.71	0.91
3:A:103:LYS:HB3	3:A:104:PRO:HD3	1.53	0.90
2:D:23[D]:DT:H1'	2:D:24[D]:DT:O5'	1.70	0.90
3:A:161:HIS:HD2	3:A:163:THR:H	1.20	0.89
3:A:278:ARG:CD	4:A:377:HOH:O	2.18	0.86
3:A:177:MET:CE	3:A:178:GLY:CA	2.56	0.79
2:D:23[D]:DT:C1'	2:D:24[D]:DT:O5'	2.31	0.79
3:A:103:LYS:CB	3:A:104:PRO:HD3	2.12	0.79
3:A:177:MET:CE	3:A:178:GLY:N	2.46	0.77
3:A:80:ARG:NH1	4:A:330:HOH:O	2.19	0.76
3:A:174:ASP:HB3	3:A:177:MET:O	1.89	0.72
3:A:177:MET:HE2	3:A:178:GLY:H	1.53	0.72
3:A:177:MET:HE2	3:A:178:GLY:N	2.05	0.70
3:A:161:HIS:CD2	3:A:163:THR:H	2.08	0.70

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:14[B]:DA:C2'	1:B:15[B]:DA:H5'	2.21	0.69
2:D:30[D]:DA:C2'	2:D:31[D]:DA:H5'	2.21	0.68
3:A:61:ILE:HD11	3:A:117:GLU:HG3	1.77	0.67
1:B:5[B]:DT:C4'	1:B:6[B]:DT:OP1	2.28	0.65
3:A:71:ARG:HD2	3:A:175:PRO:HD3	1.81	0.63
3:A:103:LYS:CB	3:A:104:PRO:CD	2.77	0.63
3:A:100:PRO:HD2	4:A:366:HOH:O	1.99	0.62
3:A:103:LYS:HE3	3:A:110:ARG:NH1	2.14	0.62
3:A:220:LEU:HD22	3:A:227:LEU:HD23	1.81	0.62
3:A:179:ILE:HD13	3:A:183:LEU:HD21	1.84	0.60
3:A:103:LYS:HB3	3:A:104:PRO:CD	2.29	0.60
2:D:24[D]:DT:C4'	2:D:25[D]:DT:OP1	2.25	0.60
3:A:206:ASP:HB3	3:A:250:LEU:HD13	1.83	0.60
2:D:25[D]:DT:H1'	2:D:26[D]:DT:H5''	1.83	0.59
3:A:99:LEU:HD12	3:A:99:LEU:N	2.18	0.58
3:A:233:GLU:O	3:A:237:GLN:HG3	2.04	0.57
2:D:27[D]:DA:H2''	2:D:28[D]:DA:C8	2.41	0.56
3:A:100:PRO:CG	4:A:366:HOH:O	2.42	0.56
3:A:100:PRO:CD	4:A:366:HOH:O	2.53	0.56
2:D:23[D]:DT:C2'	2:D:24[D]:DT:O5'	2.55	0.55
2:D:17[D]:DC:H2'	2:D:18[D]:DT:H72	1.90	0.54
2:D:25[D]:DT:H1'	2:D:26[D]:DT:C5'	2.38	0.54
2:D:21[D]:DT:H1'	2:D:22[D]:DC:C6	2.44	0.53
3:A:173:ARG:CZ	3:A:180:SER:O	2.57	0.53
3:A:71:ARG:NH1	4:A:316:HOH:O	2.39	0.52
1:B:1[B]:DC:H2'	1:B:2[B]:DT:H72	1.90	0.52
3:A:216:ASP:HB3	4:A:381:HOH:O	2.08	0.52
3:A:71:ARG:NH1	3:A:175:PRO:HG3	2.27	0.50
3:A:102:LYS:HB2	3:A:105:GLY:O	2.12	0.50
3:A:274:LYS:HG2	3:A:275:GLU:HG3	1.93	0.50
3:A:233:GLU:HG2	3:A:237:GLN:HE21	1.77	0.49
3:A:173:ARG:HA	3:A:173:ARG:HE	1.78	0.48
3:A:70:ALA:HB1	3:A:100:PRO:CG	2.44	0.48
3:A:68:GLN:HG2	3:A:72:LEU:HD22	1.95	0.48
3:A:173:ARG:NH2	3:A:179:ILE:O	2.46	0.48
3:A:278:ARG:HG3	4:A:377:HOH:O	2.13	0.48
1:B:5[B]:DT:C6	1:B:6[B]:DT:H72	2.51	0.46
2:D:26[D]:DT:H2''	2:D:27[D]:DA:O5'	2.15	0.46
1:B:3[B]:DT:H1'	1:B:4[B]:DT:H5'	1.99	0.45
2:D:27[D]:DA:H2''	2:D:28[D]:DA:H8	1.79	0.45
3:A:173:ARG:NH1	3:A:180:SER:O	2.50	0.45
1:B:7[B]:DA:H2''	1:B:8[B]:DA:O5'	2.17	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:A:174:ASP:O	3:A:175:PRO:C	2.54	0.45
3:A:214:HIS:N	3:A:215:PRO:HD3	2.31	0.45
3:A:31:GLN:NE2	3:A:249:ASN:OD1	2.50	0.45
2:D:24[D]:DT:H2''	2:D:25[D]:DT:H71	1.98	0.44
3:A:106:THR:HG23	3:A:107:ASN:OD1	2.17	0.44
3:A:188:LEU:HA	3:A:189:PRO:HD3	1.80	0.44
3:A:213:GLN:HA	3:A:213:GLN:HE21	1.82	0.44
2:D:19[D]:DT:H1'	2:D:20[D]:DT:H5'	1.99	0.43
3:A:74:ILE:HD11	3:A:100:PRO:HB3	2.00	0.43
3:A:237:GLN:O	3:A:241:ARG:HG3	2.19	0.43
3:A:64:TYR:HE2	3:A:99:LEU:HG	1.84	0.42
3:A:177:MET:HE2	3:A:178:GLY:CA	2.41	0.42
3:A:145:TRP:CH2	3:A:233:GLU:HB2	2.55	0.42
3:A:144:GLN:HE21	3:A:144:GLN:HB3	1.61	0.41
1:B:1[B]:DC:N3	3:A:99:LEU:HD23	2.35	0.41
3:A:237:GLN:CD	4:A:398:HOH:O	2.58	0.41
3:A:216:ASP:N	3:A:216:ASP:OD2	2.54	0.41
3:A:218:ILE:HB	3:A:229:ALA:HB3	2.03	0.40
3:A:278:ARG:CG	4:A:377:HOH:O	2.59	0.40
3:A:140:PRO:HA	3:A:141:PRO:HD3	1.88	0.40
3:A:126:HIS:HA	3:A:127:PRO:HD3	1.98	0.40

All (4) 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	Distance(Å)	Clash(Å)
1:B:8[B]:DA:N1	1:B:9[B]:DA:N1[2_765]	1.74	0.46
1:B:7[B]:DA:N1	1:B:10[B]:DA:N1[2_765]	1.84	0.36
3:A:271:TYR:OH	3:A:271:TYR:OH[2_665]	2.07	0.13
1:B:7[B]:DA:C2	1:B:10[B]:DA:N1[2_765]	2.10	0.10

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
3	A	253/255 (99%)	241 (95%)	6 (2%)	6 (2%)	9 3

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	103	LYS
3	A	175	PRO
3	A	104	PRO
3	A	223	VAL
3	A	108	ASP
3	A	181	GLY

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
3	A	224/224 (100%)	213 (95%)	11 (5%)	35 28

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	62	LYS
3	A	64	TYR
3	A	65	PRO
3	A	72	LEU
3	A	102	LYS
3	A	144	GLN
3	A	173	ARG
3	A	176	GLU
3	A	177	MET
3	A	220	LEU
3	A	273	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	31	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	A	77	HIS
3	A	79	GLN
3	A	134	ASN
3	A	144	GLN
3	A	161	HIS
3	A	213	GLN
3	A	237	GLN
3	A	245	GLN
3	A	249	ASN
3	A	277	GLN

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 ⓘ

There are no ligands in this entry.

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	B	16/16 (100%)	0.92	1 (6%) 19 19	46, 69, 72, 72	16 (100%)
2	D	16/16 (100%)	0.81	0 100 100	46, 71, 72, 73	16 (100%)
3	A	255/255 (100%)	0.59	18 (7%) 16 15	25, 40, 73, 88	0
All	All	287/287 (100%)	0.62	19 (6%) 18 17	25, 42, 72, 88	32 (11%)

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	106	THR	13.3
3	A	178	GLY	11.4
3	A	105	GLY	9.4
3	A	177	MET	9.3
3	A	176	GLU	8.7
3	A	107	ASN	8.0
3	A	108	ASP	7.1
3	A	179	ILE	7.1
3	A	175	PRO	6.4
3	A	104	PRO	5.8
3	A	180	SER	5.3
3	A	99	LEU	3.7
3	A	174	ASP	3.4
3	A	102	LYS	3.1
3	A	103	LYS	2.8
3	A	227	LEU	2.7
3	A	228	LEU	2.5
1	B	11[B]	DG	2.3
3	A	24	THR	2.3

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 ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.