



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 3, 2017 – 12:12 AM EDT

PDB ID : 1N1O  
Title : Crystal Structure of a B-form DNA Duplex Containing (L)-alpha-threofuranosyl (3'-2') Nucleosides: A Four-Carbon Sugar is Easily Accommodated into the Backbone of DNA  
Authors : Wilds, C.J.; Wawrzak, Z.; Krishnamurthy, R.; Eschenmoser, A.; Egli, M.  
Deposited on : unknown  
Resolution : 1.20 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](mailto: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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : rb-20030345

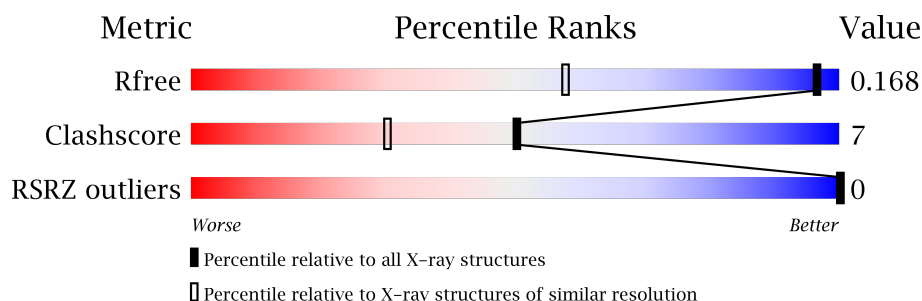
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.20 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))
$R_{free}$	100719	1131 (1.24-1.16)
Clashscore	112137	1201 (1.24-1.16)
RSRZ outliers	101464	1132 (1.24-1.16)

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  $\geq 3$ , 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	12	
1	B	12	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 682 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 DNA chain called 5'-D(\*CP\*GP\*CP\*GP\*AP\*AP\*(TFT)P\*TP\*CP\*GP\*CP\*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	12	Total	C	N	O	P	0	0	0
			242	115	46	70	11			
1	B	12	Total	C	N	O	P	0	0	0
			242	115	46	70	11			

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Mg	0	0
			1	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	106	Total	O	0	0
			106	106		
3	B	91	Total	O	0	0
			91	91		

### 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 the various outlier classes 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\*GP\*CP\*GP\*AP\*AP\*(TFT)P\*TP\*CP\*GP\*CP\*G)-3'

Chain A: 



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

Chain B: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	25.28Å 40.08Å 65.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.20 19.25 – 1.20	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-1.20) 96.3 (19.25-1.20)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.15 (at 1.20Å)	Xtriage
Refinement program	CNS, SHELXL-97	Depositor
R, $R_{free}$	0.160 , (Not available) 0.161 , 0.168	Depositor DCC
$R_{free}$ test set	1072 reflections (5.45%)	DCC
Wilson B-factor (Å <sup>2</sup> )	17.5	Xtriage
Anisotropy	0.437	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.25 , 85.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	682	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.38% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

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

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	2.35	14/249 (5.6%)	2.37	18/380 (4.7%)
1	B	2.48	15/249 (6.0%)	2.46	24/380 (6.3%)
All	All	2.42	29/498 (5.8%)	2.42	42/760 (5.5%)

All (29) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	9	DC	O4'-C1'	12.61	1.57	1.42
1	B	123	DC	O4'-C1'	11.61	1.56	1.42
1	B	124	DG	O4'-C1'	11.17	1.55	1.42
1	B	121	DC	O4'-C1'	10.93	1.55	1.42
1	A	11	DC	O4'-C1'	10.66	1.55	1.42
1	B	117	DA	O4'-C1'	10.38	1.54	1.42
1	B	123	DC	C2'-C1'	-10.25	1.42	1.52
1	B	124	DG	C2'-C1'	-9.37	1.42	1.52
1	A	3	DC	O4'-C1'	9.24	1.53	1.42
1	B	115	DC	O4'-C1'	8.87	1.52	1.42
1	A	9	DC	C2'-C1'	-8.75	1.43	1.52
1	B	120	DT	O4'-C1'	8.41	1.52	1.42
1	A	5	DA	O4'-C1'	8.36	1.52	1.42
1	B	121	DC	C2'-C1'	-7.87	1.44	1.52
1	A	12	DG	O4'-C1'	7.81	1.51	1.42
1	A	12	DG	C2'-C1'	-7.27	1.45	1.52
1	A	11	DC	C2'-C1'	-6.92	1.45	1.52
1	B	118	DA	O4'-C1'	6.76	1.50	1.42
1	B	113	DC	O4'-C1'	6.75	1.50	1.42
1	B	120	DT	C2'-C1'	-6.59	1.45	1.52
1	A	8	DT	O4'-C1'	6.59	1.50	1.42
1	A	1	DC	O4'-C1'	6.49	1.50	1.42
1	A	3	DC	C2'-C1'	-6.31	1.46	1.52
1	B	117	DA	C2'-C1'	-6.04	1.46	1.52

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	5	DA	C2'-C1'	-5.84	1.46	1.52
1	A	12	DG	C4'-O4'	-5.76	1.39	1.45
1	B	114	DG	O4'-C1'	5.46	1.48	1.42
1	A	8	DT	C2'-C1'	-5.30	1.47	1.52
1	B	122	DG	C4'-C3'	-5.25	1.47	1.52

All (42) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	9	DC	O4'-C1'-N1	-15.50	97.15	108.00
1	B	123	DC	O4'-C1'-N1	-13.04	98.87	108.00
1	A	11	DC	O4'-C1'-N1	-12.97	98.92	108.00
1	B	124	DG	O4'-C1'-N9	-11.76	99.77	108.00
1	A	12	DG	O4'-C1'-N9	-11.52	99.94	108.00
1	B	121	DC	O4'-C1'-N1	-11.47	99.97	108.00
1	B	120	DT	O4'-C1'-N1	-11.07	100.25	108.00
1	B	115	DC	O4'-C1'-N1	-10.39	100.73	108.00
1	B	117	DA	O4'-C1'-N9	-10.32	100.77	108.00
1	B	121	DC	C1'-O4'-C4'	-8.65	101.45	110.10
1	A	3	DC	O4'-C1'-N1	-7.93	102.45	108.00
1	B	116	DG	P-O3'-C3'	7.70	128.93	119.70
1	B	116	DG	O4'-C1'-N9	7.18	113.03	108.00
1	A	9	DC	N1-C1'-C2'	7.18	126.23	112.60
1	A	1	DC	O4'-C4'-C3'	6.68	110.01	106.00
1	A	11	DC	N1-C1'-C2'	6.62	125.18	112.60
1	B	113	DC	O4'-C1'-N1	-6.37	103.54	108.00
1	A	2	DG	OP1-P-OP2	6.10	128.75	119.60
1	A	5	DA	O4'-C1'-N9	-6.09	103.74	108.00
1	B	114	DG	O4'-C1'-N9	-6.01	103.79	108.00
1	B	121	DC	N1-C1'-C2'	5.86	123.74	112.60
1	A	10	DG	C1'-O4'-C4'	-5.82	104.28	110.10
1	B	123	DC	N1-C1'-C2'	5.78	123.59	112.60
1	B	124	DG	N9-C1'-C2'	5.70	123.44	112.60
1	A	2	DG	O4'-C1'-N9	-5.62	104.06	108.00
1	A	1	DC	N1-C1'-C2'	5.61	123.26	112.60
1	A	1	DC	O4'-C1'-N1	-5.59	104.08	108.00
1	A	4	DG	P-O3'-C3'	5.55	126.36	119.70
1	B	114	DG	OP1-P-OP2	5.46	127.79	119.60
1	A	2	DG	C1'-O4'-C4'	-5.45	104.65	110.10
1	B	114	DG	C1'-O4'-C4'	-5.44	104.66	110.10
1	B	122	DG	O4'-C1'-C2'	-5.43	101.56	105.90
1	B	117	DA	N9-C1'-C2'	5.23	122.54	112.60

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	115	DC	C1'-O4'-C4'	-5.23	104.87	110.10
1	A	3	DC	C1'-O4'-C4'	-5.22	104.88	110.10
1	B	120	DT	N1-C1'-C2'	5.19	122.47	112.60
1	B	113	DC	N1-C1'-C2'	5.18	122.44	112.60
1	A	8	DT	P-O3'-C3'	5.16	125.89	119.70
1	A	2	DG	P-O3'-C3'	5.14	125.87	119.70
1	B	117	DA	P-O3'-C3'	5.09	125.81	119.70
1	B	114	DG	OP2-P-O3'	5.03	116.26	105.20
1	B	114	DG	O5'-P-OP2	-5.00	101.20	105.70

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	242	0	134	2	0
1	B	242	0	134	3	0
2	A	1	0	0	0	0
3	A	106	0	0	1	0
3	B	91	0	0	1	0
All	All	682	0	268	5	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 7.

All (5) 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:118:DA:OP1	3:B:858:HOH:O	2.10	0.69
1:B:123:DC:H2'	1:B:124:DG:C8	2.35	0.61
1:B:115:DC:H2'	1:B:116:DG:C8	2.41	0.56
1:A:7:TFT:OP1	3:A:903:HOH:O	2.20	0.51
1:A:3:DC:H2'	1:A:4:DG:C8	2.56	0.40



There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

There are no protein molecules in this entry.

### 5.3.2 Protein sidechains [i](#)

There are no protein molecules in this entry.

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues 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
1	TFT	A	7	1	12,20,21	1.45	4 (33%)	14,29,32	5.15	6 (42%)
1	TFT	B	119	1	12,20,21	1.50	3 (25%)	14,29,32	4.05	6 (42%)

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
1	TFT	A	7	1	-	0/1/21/22	0/2/2/2
1	TFT	B	119	1	-	0/1/21/22	0/2/2/2

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	119	TFT	O3T-C3T	-2.49	1.42	1.46
1	A	7	TFT	O3T-C3T	-2.40	1.42	1.46
1	A	7	TFT	C6-C5	-2.03	1.34	1.40
1	A	7	TFT	O4T-C1T	2.09	1.44	1.41
1	B	119	TFT	O4T-C1T	2.47	1.45	1.41
1	A	7	TFT	C4-N3	2.65	1.37	1.33
1	B	119	TFT	C4-N3	3.02	1.38	1.33

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	7	TFT	C5-C4-N3	-7.95	116.47	125.24
1	B	119	TFT	O4T-C1T-N1	-7.71	98.71	108.94
1	B	119	TFT	C5-C4-N3	-6.08	118.54	125.24
1	A	7	TFT	O4T-C1T-N1	-6.05	100.91	108.94
1	A	7	TFT	O2T-C2T-C1T	-3.38	101.03	111.61
1	B	119	TFT	O4T-C4T-C3T	-2.08	100.74	104.57
1	B	119	TFT	C5-C6-N1	2.59	124.95	122.15
1	B	119	TFT	C5M-C5-C6	2.66	123.98	118.67
1	A	7	TFT	C5M-C5-C6	3.84	126.33	118.67
1	A	7	TFT	C5-C6-N1	4.11	126.61	122.15
1	B	119	TFT	C4-N3-C2	10.27	124.14	115.16
1	A	7	TFT	C4-N3-C2	14.71	128.02	115.16

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	7	TFT	1	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

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.

No monomer is involved in short contacts.

## 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

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	11/12 (91%)	0.21	0 <span>100</span> <span>100</span>	17, 20, 23, 26	0
1	B	11/12 (91%)	0.35	0 <span>100</span> <span>100</span>	17, 21, 24, 27	0
All	All	22/24 (91%)	0.28	0 <span>100</span> <span>100</span>	17, 21, 24, 27	0

There are no RSRZ outliers to report.

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

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	TFT	A	7	19/20	0.97	0.07	-	15,16,21,23	0
1	TFT	B	119	19/20	0.95	0.08	-	17,21,41,41	0

### 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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	MG	A	901	1/1	1.00	0.05	-	17,17,17,17	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.