

# Full wwPDB X-ray Structure Validation Report (i)

#### May 27, 2020 - 01:51 am BST

PDB ID	:	330D
$\operatorname{Title}$	:	BASE-PAIRING SHIFT IN THE MAJOR GROOVE OF (CA)N TRACTS
		BY B-DNA CRYSTAL STRUCTURES
Authors	:	Timsit, Y.; Vilbois, E.; Moras, D.
Deposited on		
$\operatorname{Resolution}$	:	2.70  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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
$\mathrm{EDS}$	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
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 2.70 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	Percent	tile Ranks	Value
Clashscore 📕			57
W	orse	Bette	er
E.	Percentile relative to all X-ray stru	ictures	
01	Percentile relative to X-ray structu	ares of similar resolution	
Motria	Whole archive	Similar resol	ution

Metric	Whole archive	Similar resolution		
	$(\# \mathbf{Entries})$	$(\# { m Entries},  { m resolution}   { m range}({ m \AA}))$		
Clashscore	141614	3122 (2.70-2.70)		

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain		
1	А	12	8% 83% 8%		8%
2	В	12	33%	58%	8%



# 2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 538 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 DNA (5'-D(\*AP\*CP\*CP\*GP\*CP\*GP\*CP\*GP\*GP\*CP\*GP\*CP\*GP)-3').

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	А	12	Total 239	C 113	N 46	O 69	Р 11	0	0	0

• Molecule 2 is a DNA chain called DNA (5'-D(\*GP\*GP\*CP\*GP\*CP\*GP\*GP\*GP\*GP\*GP\*GP\*GP\*)-3').

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
2	В	12	Total 247	C 116	N 49	0 71	Р 11	0	0	0

• Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	А	29	Total         O           29         29	0	0
3	В	23	TotalO2323	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 colorcoded 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.

Note EDS was not executed.

• Molecule 1: DNA (5'-D(\*AP\*CP\*CP\*GP\*CP\*CP\*GP\*CP\*GP\*CP\*C)-3')

Chain A: 8%		83%	8%
A1 C C C C C C C C C C C C C C C C C C C	617 617		
• Molecule 2:	DNA (5'-D(*GF	»*GP*CP*GP*CP*CP*GP*G	P*CP*GP*GP*T)-3')
Chain B:	33%	58%	8%



# 4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	Н 3	Depositor
Cell constants	65.89Å $65.89$ Å $47.09$ Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $90.00^{\circ}$ $120.00^{\circ}$	Depositor
Resolution (Å)	7.00 - 2.70	Depositor
% Data completeness	69.7(7.00-2.70)	Depositor
(in resolution range)	03.1 (1.00 2.10)	Depositor
$R_{merge}$	0.05	Depositor
R <sub>sym</sub>	(Not available)	Depositor
Refinement program	X-PLOR, CORELS	Depositor
$R, R_{free}$	0.170 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	538	wwPDB-VP
Average B, all atoms $(Å^2)$	33.0	wwPDB-VP



# 5 Model quality (i)

# 5.1 Standard geometry (i)

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		
	Ullalli	RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	А	0.55	0/267	0.82	0/409	
2	В	0.58	0/277	0.86	0/427	
All	All	0.56	0/544	0.84	0/836	

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	А	0	1
2	В	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	А	2	DC	Sidechain
2	В	19	DG	Sidechain

### 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	А	239	0	134	24	0
2	В	247	0	135	22	0
3	А	29	0	0	1	0
3	В	23	0	0	0	0
All	All	538	0	269	41	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 57.

All (41) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic	Clash
Atom-1	Atom-2	distance (Å)	overlap (Å)
2:B:22:DG:H2"	2:B:23:DG:C5'	1.71	1.20
2:B:22:DG:H2"	2:B:23:DG:H5'	1.17	1.14
1:A:10:DG:H2"	1:A:11:DC:H5"	1.47	0.97
2:B:17:DC:H2"	2:B:18:DC:H5'	1.50	0.93
1:A:10:DG:H2"	1:A:11:DC:C5'	2.02	0.90
1:A:2:DC:H2"	1:A:3:DC:C5'	2.09	0.83
1:A:2:DC:H2"	1:A:3:DC:H5'	1.61	0.82
2:B:22:DG:H2"	2:B:23:DG:H5"	1.60	0.80
2:B:18:DC:H2"	2:B:19:DG:C8	2.16	0.80
1:A:3:DC:H2"	1:A:4:DG:H5'	1.71	0.72
2:B:22:DG:C2'	2:B:23:DG:C5'	2.62	0.69
1:A:6:DC:N4	2:B:18:DC:N4	2.43	0.66
1:A:2:DC:H1'	3:A:75:HOH:O	1.97	0.65
1:A:8:DG:H2"	1:A:9:DC:C6	2.32	0.64
2:B:22:DG:C2'	2:B:23:DG:H5"	2.30	0.61
2:B:18:DC:H2"	2:B:19:DG:N7	2.15	0.60
1:A:4:DG:H2"	1:A:5:DC:OP2	2.01	0.60
1:A:10:DG:H2"	1:A:11:DC:H5'	1.83	0.58
1:A:9:DC:H2'	1:A:10:DG:C8	2.39	0.58
2:B:23:DG:C8	2:B:24:DT:C7	2.89	0.56
1:A:11:DC:H2"	1:A:12:DC:H6	1.70	0.56
1:A:10:DG:C2'	1:A:11:DC:H5"	2.29	0.56
2:B:21:DC:N3	2:B:22:DG:C6	2.75	0.54
1:A:3:DC:H2"	1:A:4:DG:C5'	2.36	0.54
1:A:2:DC:H2'	1:A:3:DC:C6	2.44	0.53
1:A:12:DC:H42	2:B:13:DG:H1	1.56	0.53
1:A:1:DA:H2"	1:A:2:DC:O5'	2.09	0.52
1:A:1:DA:C2	1:A:2:DC:C2	2.99	0.51
2:B:17:DC:C2'	2:B:18:DC:H5'	2.32	0.50
2:B:18:DC:C2'	2:B:19:DG:C8	2.94	0.48

Continued on next page...



330D
------

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:21:DC:C2	2:B:22:DG:C6	3.01	0.48
1:A:11:DC:H2"	1:A:12:DC:C6	2.47	0.48
2:B:22:DG:C2'	2:B:23:DG:H5'	2.12	0.47
2:B:23:DG:C8	2:B:24:DT:H73	2.50	0.47
2:B:21:DC:O2	2:B:22:DG:C5	2.67	0.47
1:A:1:DA:N6	2:B:23:DG:O6	2.48	0.46
1:A:12:DC:N4	2:B:13:DG:H1	2.18	0.42
2:B:22:DG:H1'	2:B:23:DG:H5"	2.02	0.42
1:A:10:DG:H8	1:A:10:DG:OP2	2.04	0.40
1:A:2:DC:H2"	1:A:3:DC:O5'	2.21	0.40
1:A:3:DC:H42	2:B:21:DC:H42	1.68	0.40

Continued from previous page...

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)

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)

There are no ligands in this entry.

# 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 (i)

# 6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

# 6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

# 6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

## 6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.

