



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 05:55 pm BST

PDB ID : 3UBI  
Title : The Absence of Tertiary Interactions in a Self-Assembled DNA Crystal Structure  
Authors : Nguyen, N.; Birktoft, J.J.; Sha, R.; Wang, T.; Zheng, J.; Constantinou, P.E.; Ginell, S.L.; Chen, Y.; Mao, C.; Seeman, N.C.  
Deposited on : 2011-10-24  
Resolution : 6.80 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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

<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  
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



## 2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 1265 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(\*GP\*AP\*GP\*CP\*AP\*GP\*CP\*CP\*CP\*GP\*TP\*AP\*CP\*TP\*CP\*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	A	16	324	154	62	93	15	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	D	17	351	165	72	97	17	0	0	0

- Molecule 3 is a DNA chain called DNA (5'-D(\*TP\*CP\*TP\*GP\*AP\*TP\*GP\*AP\*GP\*GP\*CP\*TP\*GP\*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	B	14	286	137	52	84	13	0	0	0


- Molecule 4 is a DNA chain called DNA (5'-D(P\*GP\*CP\*TP\*TP\*GP\*TP\*CP\*GP\*TP\*TP\*CP\*AP\*TP\*CP\*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
4	C	15	304	146	49	94	15	0	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.

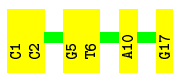
- Molecule 1: DNA (5'-D(\*GP\*AP\*GP\*CP\*AP\*GP\*CP\*CP\*CP\*GP\*TP\*AP\*CP\*TP\*CP\*G)-3')

Chain A:  81% 19%



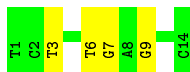
- Molecule 2: DNA (5'-D(P\*CP\*CP\*GP\*AP\*GP\*TP\*AP\*CP\*GP\*AP\*CP\*GP\*AP\*CP\*AP\*AP\*G)-3')

Chain D:  65% 35%



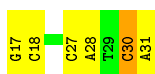
- Molecule 3: DNA (5'-D(\*TP\*CP\*TP\*GP\*AP\*TP\*GP\*AP\*GP\*GP\*CP\*TP\*GP\*C)-3')

Chain B:  71% 29%



- Molecule 4: DNA (5'-D(P\*GP\*CP\*TP\*TP\*GP\*TP\*CP\*GP\*TP\*TP\*CP\*AP\*TP\*CP\*A)-3')

Chain C:  60% 33% 7%



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	169.98 Å 169.98 Å 88.64 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.07 – 6.80 49.07 – 6.80	Depositor EDS
% Data completeness (in resolution range)	86.8 (49.07-6.80) 88.3 (49.07-6.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.48 (at 6.68 Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: dev_873)	Depositor
R, $R_{free}$	0.139 , 0.152 0.136 , 0.152	Depositor DCC
$R_{free}$ test set	145 reflections (10.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	239.4	Xtrriage
Anisotropy	1.394	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.09 , -110.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	0.053 for h,-h-k,-l	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	1265	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	236.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.49% 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

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	0.49	0/363	1.07	1/558 (0.2%)
2	D	0.47	0/395	0.95	1/607 (0.2%)
3	B	0.48	0/320	1.00	2/493 (0.4%)
4	C	0.49	0/338	1.19	1/519 (0.2%)
All	All	0.48	0/1416	1.05	5/2177 (0.2%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	3	DT	C5-C4-O4	-5.62	120.97	124.90
3	B	3	DT	N3-C4-O4	5.48	123.19	119.90
4	C	30	DC	O4'-C4'-C3'	-5.31	102.38	104.50
2	D	6	DT	C5-C4-O4	-5.07	121.35	124.90
1	A	13	DC	O4'-C1'-N1	5.04	111.53	108.00

There are no chirality outliers.

There are no planarity outliers.

### 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	324	0	180	1	1
2	D	351	0	189	4	0
3	B	286	0	160	2	0
4	C	304	0	172	4	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	1265	0	701	8	1

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 4.

The worst 5 of 8 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)
2:D:17:DG:N2	4:C:18:DC:N3	2.42	0.67
2:D:10:DA:H5'	3:B:9:DG:H5''	1.84	0.59
2:D:1:DC:OP1	2:D:2:DC:OP2	2.32	0.48
4:C:30:DC:H2'	4:C:31:DA:H8	1.80	0.46
4:C:27:DC:H2''	4:C:28:DA:C8	2.51	0.45

All (1) 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	Interatomic distance (Å)	Clash overlap (Å)
1:A:16:DG:O3'	4:C:17:DG:P[2_555]	2.11	0.09

## 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

### 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.