



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 14, 2020 – 11:52 AM BST

PDB ID : 6UAL  
Title : A Self-Assembling DNA Crystal Scaffold with Cavities Containing 3 Helical Turns per Edge  
Authors : Simmons, C.R.; MacCulloch, T.; Stephanopoulos, N.; Yan, H.  
Deposited on : 2019-09-10  
Resolution : 4.51 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

<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 : 2.14.4.dev1  
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) : 2.14.4.dev1

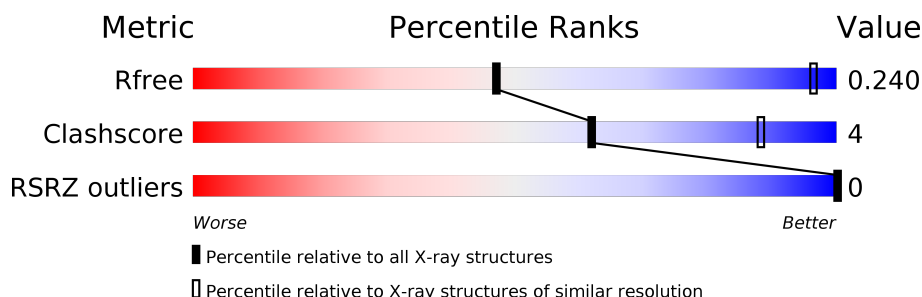
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.51 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	1056 (5.22-3.80)
Clashscore	141614	1124 (5.22-3.80)
RSRZ outliers	127900	1101 (5.30-3.70)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	B	32	<div> <div>88%</div> <div>13%</div> </div>
2	C	7	<div> <div>100%</div> </div>
3	D	5	<div> <div>40%</div> <div>60%</div> </div>
4	A	20	<div> <div>85%</div> <div>15%</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2033 atoms, of which 727 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 (32-MER).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	B	32	Total	C	H	N	O	P	0	0	0
			1008	310	365	107	195	31			

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

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	C	7	Total	C	H	N	O	P	0	0	0
			225	69	79	30	40	7			

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

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	D	5	Total	C	H	N	O	P	0	0	0
			161	49	57	20	30	5			


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

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
4	A	20	Total	C	H	N	O	P	0	0	0
			639	197	226	82	115	19			

### 3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: DNA (32-MER)

Chain B:  88% 13%



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

Chain C:  100%

There are no outlier residues recorded for this chain.

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

Chain D:  40% 60%



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

Chain A:  85% 15%



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	179.25Å 179.25Å 48.14Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.98 – 4.51 45.98 – 4.51	Depositor EDS
% Data completeness (in resolution range)	94.1 (45.98-4.51) 94.1 (45.98-4.51)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.39 (at 4.45Å)	Xtriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, $R_{free}$	0.197 , 0.243 0.197 , 0.240	Depositor DCC
$R_{free}$ test set	158 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	280.1	Xtriage
Anisotropy	0.703	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.14 , 98.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.097 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	2033	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	314.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.88% 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	B	0.54	0/717	0.96	0/1103
2	C	0.60	0/164	0.91	0/251
3	D	0.54	0/116	0.93	0/177
4	A	0.50	0/465	0.85	0/717
All	All	0.53	0/1462	0.92	0/2248

There are no bond length outliers.

There are no bond angle outliers.

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	B	643	365	365	2	0
2	C	146	79	79	0	0
3	D	104	57	57	4	1
4	A	413	226	226	1	1
All	All	1306	727	727	7	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.

All (7) 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)
3:D:26:DG:C2'	3:D:27:DT:H72	2.37	0.55
1:B:4:DT:H2''	1:B:5:DG:C8	2.47	0.49
3:D:27:DT:H2'	3:D:28:DC:C6	2.48	0.48
3:D:26:DG:C4	3:D:27:DT:H72	2.53	0.43
1:B:30:DT:H2'	1:B:31:DT:H71	1.99	0.43
3:D:26:DG:C2'	3:D:27:DT:C7	2.98	0.42
4:A:3:DG:H2''	4:A:4:DA:C8	2.56	0.41

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)
3:D:26:DG:OP1	4:A:20:DG:O3'[8_444]	2.05	0.15

## 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 monosaccharides 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](#)

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	B	32/32 (100%)	-1.39	0 100 100	231, 294, 366, 390	0
2	C	7/7 (100%)	-1.40	0 100 100	197, 211, 230, 231	0
3	D	5/5 (100%)	-1.38	0 100 100	258, 276, 303, 345	0
4	A	20/20 (100%)	-1.49	0 100 100	199, 316, 418, 421	0
All	All	64/64 (100%)	-1.42	0 100 100	197, 294, 372, 421	0

There are no RSRZ outliers to report.

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.