wwPDB X-ray Structure Validation Summary Report (i)

Nov 30, 2020 – 12:45 PM EST

PDB ID	:	6V6R
Title	:	Crystal Structure of a Bromine Derivatized Self-Assembling DNA Crystal Scaffold with Rhombohedral Symmetry.
Authors	:	Simmons, C.R.; MacCulloch, T.; Stephanopoulos, N.; Yan, H.
Deposited on	:	2019-12-05
Resolution	:	2.70 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*
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:

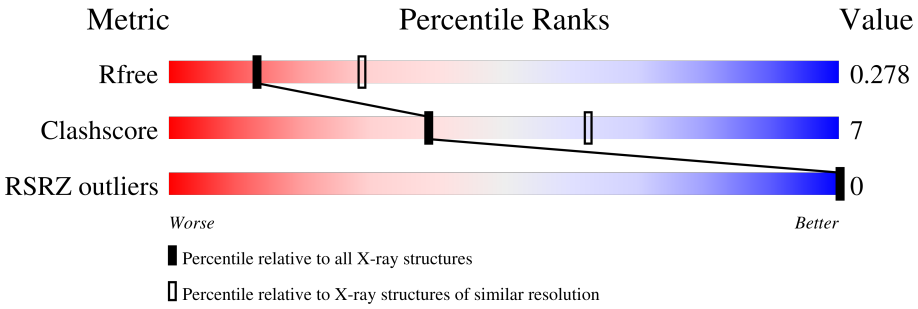
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.14.6
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.14.6

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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
RSRZ outliers	127900	2737 (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% 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	21	76% 24%
2	A	7	100%
3	C	9	56% 44%
4	D	5	60% 40%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 1309 atoms, of which 452 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(*CP*AP*CP*TP*GP*AP*CP*TP*CP*AP*TP*GP*CP*TP*CP*AP*TP*CP*TP*GP*A)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	P			
1	B	21	648	203	227	73	125	20	0	0	0

- 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
			Total	C	H	N	O	P			
2	A	7	224	69	78	30	40	7	0	0	0

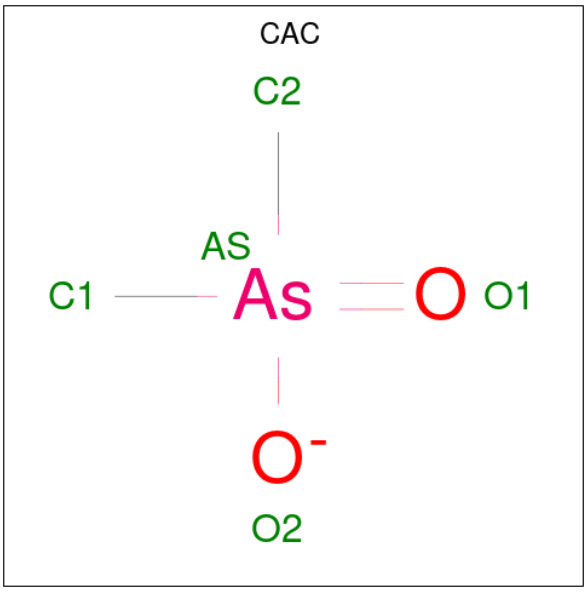
- Molecule 3 is a DNA chain called DNA (5'-D(*TP*GP*TP*CP*AP*GP*AP*TP*G)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	P			
3	C	9	274	89	90	34	53	8	0	0	0

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

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

- Molecule 5 is CACODYLATE ION (three-letter code: CAC) (formula: $C_2H_6AsO_2$).



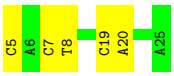
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total As 1 1	0	0
5	D	1	Total As 1 1	0	0

3 Residue-property plots (i)

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.

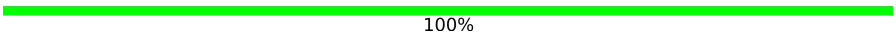
• Molecule 1: DNA (5'-D(*CP*AP*CP*TP*GP*AP*CP*TP*CP*AP*TP*GP*CP*TP*CP*AP*TP*CP*TP*GP*A)-3')

Chain B:



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

Chain A:



There are no outlier residues recorded for this chain.

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

Chain C:

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

Chain D:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α, β, γ	115.81Å 115.81Å 44.75Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	40.87 – 2.70 40.86 – 2.70	Depositor EDS
% Data completeness (in resolution range)	98.0 (40.87-2.70) 94.8 (40.86-2.70)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.80 (at 2.69Å)	Xtriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, R_{free}	0.242 , 0.279 0.242 , 0.278	Depositor DCC
R_{free} test set	305 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	77.4	Xtriage
Anisotropy	0.522	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.24 , 71.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.059 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	1309	wwPDB-VP
Average B, all atoms (Å^2)	109.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.30% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CAC

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	0.60	0/470	0.99	0/722
2	A	0.52	0/164	0.85	0/251
3	C	0.51	0/206	0.95	0/317
4	D	0.49	0/116	0.85	0/177
All	All	0.55	0/956	0.94	0/1467

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

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	421	227	239	6	0
2	A	146	78	79	0	0
3	C	184	90	104	3	0
4	D	104	57	57	1	0
5	A	1	0	0	0	0
5	D	1	0	0	0	0
All	All	857	452	479	10	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.

The worst 5 of 10 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:C:56:DT:H2”	3:C:57:DG:H5’	1.87	0.57
3:C:56:DT:H1’	3:C:57:DG:H5”	1.91	0.52
1:B:19:DC:H4’	1:B:20:DA:OP1	2.12	0.48
1:B:7:DC:H4’	1:B:8:DT:OP1	2.14	0.48
1:B:19:DC:H2”	1:B:20:DA:O5’	2.17	0.45

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

There are no protein molecules in this entry.

5.3.2 Protein sidechains

There are no protein molecules in this entry.

5.3.3 RNA

There are no RNA molecules 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 monosaccharides in this entry.

5.6 Ligand geometry

Of 2 ligands modelled in this entry, 2 are modelled with single atom - 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^{th} 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	21/21 (100%)	-0.40	0	100	100	96, 104, 112, 114	0
2	A	7/7 (100%)	-0.63	0	100	100	85, 95, 106, 108	0
3	C	9/9 (100%)	-0.11	0	100	100	90, 98, 106, 108	0
4	D	5/5 (100%)	0.06	0	100	100	91, 100, 119, 119	0
All	All	42/42 (100%)	-0.32	0	100	100	85, 103, 113, 119	0

There are no RSRZ outliers to report.

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 monosaccharides 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. The B-factors column lists the minimum, median, 95^{th} 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	B-factors(Å^2)	Q<0.9
5	CAC	D	101	1/5	0.40	0.14	176,176,176,176	0
5	CAC	A	101	1/5	0.85	0.31	168,168,168,168	0

6.5 Other polymers (i)

There are no such residues in this entry.