wwPDB X-ray Structure Validation Summary Report (i)

Oct 9, 2023 – 11:07 AM EDT

PDB ID	:	7JHB
Title	:	Self-assembly of a 3D DNA crystal lattice (4x6 junction version) containing the J24 immobile Holliday junction
Authors	:	Simmons, C.R.; MacCulloch, T.; Stephanopoulos, N.; Yan, H.
Deposited on	:	2020-07-20
Resolution	:	3.11 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

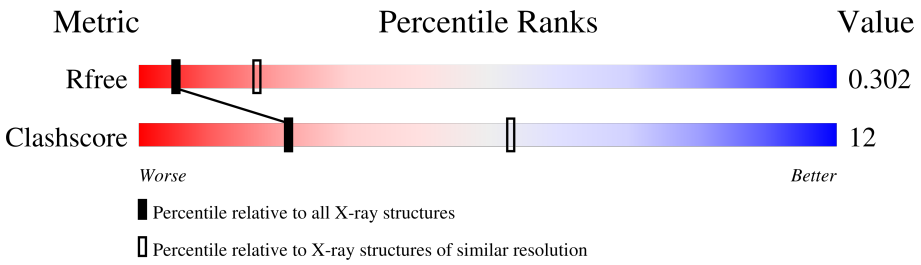
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 3.11 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	1292 (3.14-3.10)
Clashscore	141614	1389 (3.14-3.10)

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 of 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%

Mol	Chain	Length	Quality of chain
1	A	12	50% 50%
2	B	9	33% 67%
3	C	6	50% 50%
4	D	15	60% 40%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 857 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*AP*CP*CP*AP*GP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	A	12	247	117	54	65	11	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	9	180	86	31	54	9	0	0	0

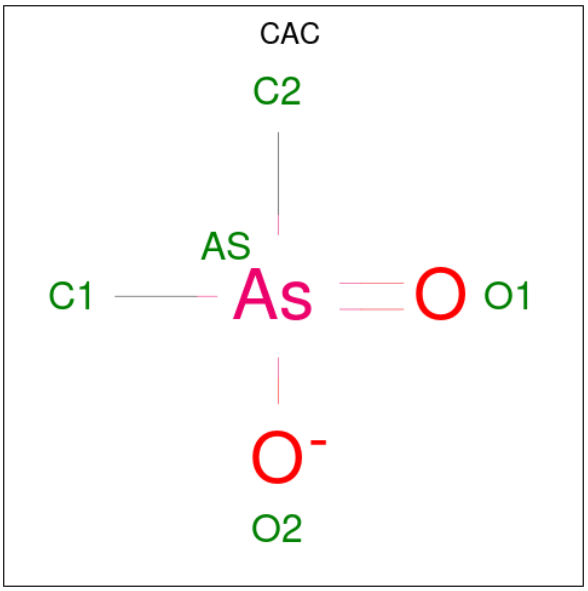
- Molecule 3 is a DNA chain called DNA (5'-D(P*TP*CP*TP*AP*CP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	C	6	121	58	20	37	6	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
4	D	15	307	147	54	92	14	0	0	0

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



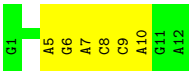
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	As		
5	A	1	1	1	0	0
			Total	As		
5	B	1	1	1	0	0

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. 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.

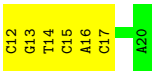
• Molecule 1: DNA (5'-D(*GP*AP*GP*CP*AP*GP*AP*CP*CP*AP*GP*A)-3')

Chain A:



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

Chain B:

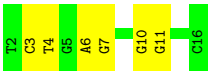


• Molecule 3: DNA (5'-D(P*TP*CP*TP*AP*CP*G)-3')

Chain C:

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

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α, β, γ	68.10Å 68.10Å 57.30Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	34.05 – 3.11 34.05 – 3.11	Depositor EDS
% Data completeness (in resolution range)	79.6 (34.05-3.11) 81.5 (34.05-3.11)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.74 (at 3.12Å)	Xtriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, R_{free}	0.271 , 0.298 0.272 , 0.302	Depositor DCC
R_{free} test set	217 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	92.0	Xtriage
Anisotropy	0.478	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.047 for -h,-k,l 0.087 for h,-h-k,-l 0.062 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	857	wwPDB-VP
Average B, all atoms (Å^2)	99.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.39% 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 (i)

5.1 Standard geometry (i)

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	A	0.57	0/279	0.72	0/429
2	B	0.50	0/200	0.88	0/305
3	C	0.57	0/134	0.95	0/204
4	D	0.59	0/343	1.02	0/529
All	All	0.56	0/956	0.90	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 (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	247	0	134	3	0
2	B	180	0	102	4	0
3	C	121	0	69	2	0
4	D	307	0	172	4	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
All	All	857	0	477	12	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 12.

The worst 5 of 12 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)
4:D:3:DC:H2'	4:D:4:DT:C2	2.31	0.65
2:B:16:DA:H2"	2:B:17:DC:H5"	1.86	0.58
2:B:12:DC:H2'	2:B:13:DG:H8	1.71	0.55
4:D:10:DG:H2'	4:D:11:DG:C8	2.41	0.55
3:C:2:DT:H2'	4:D:10:DG:H5"	1.89	0.54

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

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 (i)

6.1 Protein, DNA and RNA chains (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.