



Full wwPDB X-ray Structure Validation Report ⓘ

Oct 9, 2023 – 09:07 AM EDT

PDB ID : 7JH8
Title : Self-assembly of a 3D DNA crystal lattice (4x6 junction version) containing the J20 immobile Holliday junction
Authors : Simmons, C.R.; MacCulloch, T.; Stephanopoulos, N.; Yan, H.
Deposited on : 2020-07-20
Resolution : 3.09 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 ⓘ 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 ⓘ](#)) 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

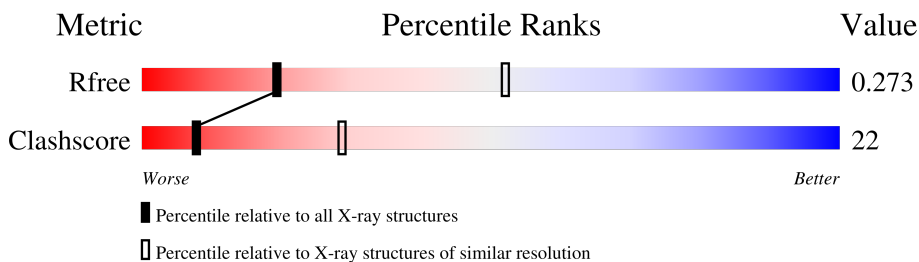
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.09 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	1447 (3.10-3.06)
Clashscore	141614	1546 (3.10-3.06)

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	12	
2	B	9	
3	C	6	
4	D	15	

2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 858 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*GP*TP*GP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	A	12	249	118	53	67	11	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	9	184	87	36	52	9	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
4	D	15	304	146	52	92	14	0	0	0

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. 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.

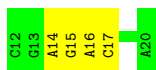
- Molecule 1: DNA (5'-D(*GP*AP*GP*CP*AP*GP*AP*CP*GP*TP*GP*A)-3')

Chain A: 



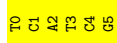
- Molecule 2: DNA (5'-D(P*CP*GP*AP*GP*AP*CP*TP*CP*A)-3')

Chain B: 



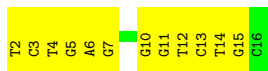
- Molecule 3: DNA (5'-D(P*TP*CP*AP*TP*CP*G)-3')

Chain C: 



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

Chain D: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α , β , γ	68.07Å 68.07Å 56.06Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	34.04 – 3.09 34.04 – 3.08	Depositor EDS
% Data completeness (in resolution range)	86.1 (34.04-3.09) 87.8 (34.04-3.08)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.06 (at 3.06Å)	Xtrriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, R_{free}	0.246 , 0.272 0.247 , 0.273	Depositor DCC
R_{free} test set	448 reflections (9.52%)	wwPDB-VP
Wilson B-factor (Å ²)	88.5	Xtrriage
Anisotropy	0.745	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.08 , -10.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.037 for -h,-k,l 0.063 for h,-h-k,-l 0.068 for -k,-h,-l	Xtrriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	858	wwPDB-VP
Average B, all atoms (Å ²)	100.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.94% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	0.58	0/281	0.90	0/433
2	B	0.59	0/206	0.80	0/315
3	C	0.67	0/134	1.06	0/204
4	D	0.59	0/339	1.00	0/522
All	All	0.60	0/960	0.94	0/1474

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	249	0	135	11	0
2	B	184	0	101	4	0
3	C	121	0	69	6	0
4	D	304	0	172	14	0
All	All	858	0	477	28	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 22.

All (28) 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:6:DA:H2''	4:D:7:DG:H5''	1.72	0.70
3:C:3:DT:H2''	3:C:4:DC:H5'	1.74	0.69
4:D:12:DT:H2''	4:D:13:DC:H5''	1.73	0.69
1:A:9:DG:H22	4:D:10:DG:N2	1.91	0.68
3:C:0:DT:H2''	3:C:1:DC:H5''	1.75	0.68
1:A:9:DG:H22	4:D:10:DG:H22	1.47	0.62
1:A:11:DG:H2'	1:A:12:DA:C8	2.34	0.62
1:A:6:DG:H1	4:D:13:DC:H42	1.47	0.61
3:C:3:DT:H2'	3:C:4:DC:C6	2.39	0.57
1:A:4:DC:H2''	1:A:5:DA:C8	2.39	0.57
1:A:9:DG:N2	4:D:10:DG:H22	2.03	0.57
4:D:2:DT:H2''	4:D:3:DC:O2	2.08	0.54
2:B:15:DG:H2''	2:B:16:DA:C8	2.43	0.53
4:D:4:DT:H2''	4:D:5:DG:C8	2.46	0.50
1:A:9:DG:H1	4:D:10:DG:H22	1.61	0.48
1:A:9:DG:H1	4:D:10:DG:H1	1.60	0.48
2:B:16:DA:H2''	2:B:17:DC:H5''	1.95	0.48
4:D:11:DG:H2'	4:D:12:DT:C6	2.49	0.46
2:B:15:DG:H2''	2:B:16:DA:H8	1.82	0.43
4:D:14:DT:H2'	4:D:15:DG:C8	2.54	0.42
1:A:9:DG:H1	4:D:10:DG:N2	2.16	0.42
3:C:1:DC:H2''	3:C:2:DA:O4'	2.19	0.41
1:A:12:DA:H8	1:A:12:DA:OP2	2.03	0.41
2:B:14:DA:H2''	2:B:15:DG:H8	1.85	0.41
1:A:11:DG:C6	1:A:12:DA:N6	2.89	0.41
3:C:4:DC:H2''	3:C:5:DG:C8	2.56	0.40
4:D:13:DC:H5''	4:D:13:DC:H6	1.87	0.40
3:C:0:DT:H2''	3:C:1:DC:C5'	2.49	0.40

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](#)

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.