wwPDB X-ray Structure Validation Summary Report (i)

Oct 15, 2023 – 01:47 AM EDT

PDB ID	:	7U3X
Title	:	[F233] Self-assembling tensegrity triangle with two turns, three turns and three turns of DNA per axis by extension with P1 symmetry
Authors	:	Woloszyn, K.; Vecchioni, S.; Seeman, N.C.; Sha, R.; Ohayon, Y.P.
Deposited on	:	2022-02-28
Resolution	:	5.68 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:

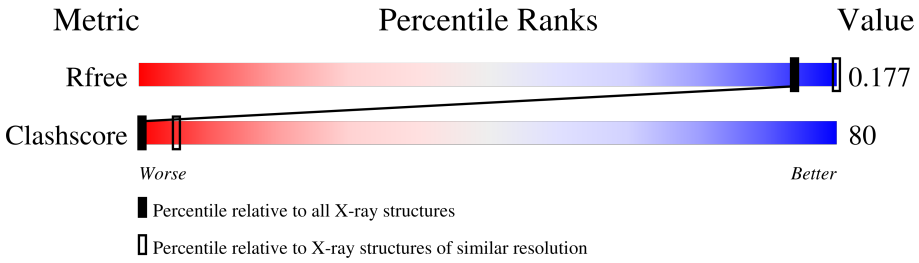
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 5.68 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	1001 (7.50-3.86)
Clashscore	141614	1018 (7.40-3.90)

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	31	. 23% 74%
2	E	14	14% 86%
3	D	24	. 50% 46%
4	B	31	29% 71%
5	F	24	21% 79%
6	C	21	14% 86%
7	M	21	5% 95%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 3385 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 (31-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	A	31	633	300	123	180	30	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	E	14	285	137	49	86	13	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	D	24	491	234	87	147	23	0	0	0

- Molecule 4 is a DNA chain called DNA (31-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
4	B	31	633	302	118	183	30	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
5	F	24	496	237	93	143	23	0	0	0

- Molecule 6 is a DNA chain called DNA (5'-D(*AP*AP*CP*CP*TP*AP*CP*CP*TP*GP*GP*CP*AP*GP*GP*AP*CP*GP*AP*CP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	C	21	Total	C	N	O	P	0	0	0
			426	203	82	121	20			

- Molecule 7 is a DNA chain called DNA (5'-D(P*TP*CP*AP*CP*CP*TP*GP*CP*CP*AP *CP*CP*GP*TP*AP*CP*AP*CP*CP*GP*A)-3').

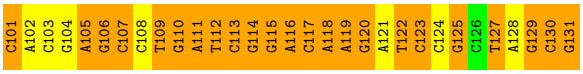
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	M	21	Total	C	N	O	P	0	0	0
			421	200	76	124	21			

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 (31-MER)

Chain A:



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

Chain E:

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

Chain D:

• Molecule 4: DNA (31-MER)

Chain B:

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

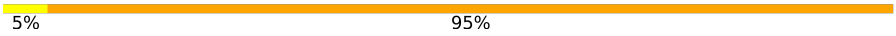
Chain F:

• Molecule 6: DNA (5'-D(*AP*AP*CP*CP*TP*AP*CP*CP*TP*GP*GP*CP*AP*GP*GP*AP *CP*GP*AP*CP*T)-3')

Chain C:

• Molecule 7: DNA (5'-D(P*TP*CP*AP*CP*CP*TP*GP*CP*CP*AP*CP*CP*GP*TP*AP*CP*AP*CP*CP*GP*A)-3')

Chain M:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	67.96Å 100.28Å 102.06Å 100.58° 100.81° 102.39°	Depositor
Resolution (Å)	41.05 – 5.68 97.49 – 5.68	Depositor EDS
% Data completeness (in resolution range)	68.8 (41.05-5.68) 62.4 (97.49-5.68)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.03 (at 5.76Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.124 , 0.179 0.121 , 0.177	Depositor DCC
R_{free} test set	273 reflections (5.33%)	wwPDB-VP
Wilson B-factor (Å^2)	334.3	Xtriage
Anisotropy	0.112	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.10 , 999.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.030 for -h,-l,-k	Xtriage
F_o,F_c correlation	0.86	EDS
Total number of atoms	3385	wwPDB-VP
Average B, all atoms (Å^2)	441.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.06% 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

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	2.84	53/711 (7.5%)	2.33	44/1095 (4.0%)
2	E	3.40	50/318 (15.7%)	2.59	33/490 (6.7%)
3	D	2.14	17/549 (3.1%)	1.81	13/847 (1.5%)
4	B	2.56	40/710 (5.6%)	2.14	41/1094 (3.7%)
5	F	2.56	30/557 (5.4%)	1.92	17/860 (2.0%)
6	C	3.08	48/478 (10.0%)	2.21	31/735 (4.2%)
7	M	3.64	54/470 (11.5%)	2.75	57/720 (7.9%)
All	All	2.86	292/3793 (7.7%)	2.24	236/5841 (4.0%)

The worst 5 of 292 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	M	113	DG	C3'-O3'	21.95	1.72	1.44
6	C	117	DC	C3'-O3'	18.64	1.68	1.44
2	E	107	DG	C3'-O3'	18.16	1.67	1.44
1	A	112	DT	C3'-O3'	16.75	1.65	1.44
7	M	101	DT	C1'-N1	16.08	1.70	1.49

The worst 5 of 236 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	112	DT	O5'-P-OP1	-17.77	89.37	110.70
4	B	112	DA	O5'-P-OP1	-17.34	89.89	110.70
1	A	119	DA	O5'-P-OP1	-15.89	91.40	105.70
4	B	127	DT	O4'-C1'-N1	15.52	118.86	108.00
7	M	107	DG	O4'-C1'-N9	14.30	118.01	108.00

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	A	633	0	347	90	0
2	E	285	0	160	63	0
3	D	491	0	272	70	0
4	B	633	0	350	60	0
5	F	496	0	273	57	0
6	C	426	0	236	52	0
7	M	421	0	235	80	0
All	All	3385	0	1873	402	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 80.

The worst 5 of 402 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)
7:M:101:DT:N1	7:M:101:DT:C1'	1.70	1.54
5:F:103:DT:N1	5:F:103:DT:C1'	1.69	1.51
7:M:102:DC:O3'	7:M:102:DC:C3'	1.64	1.45
5:F:111:DA:O5'	5:F:111:DA:C5'	1.63	1.45
1:A:112:DT:O3'	1:A:112:DT:C3'	1.65	1.43

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

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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.