



wwPDB X-ray Structure Validation Summary Report ⓘ

May 26, 2020 – 04:15 am BST

PDB ID : 5I36
Title : Crystal structure of color device state A
Authors : Hao, Y.; Kristiansen, M.; Sha, R.; Birktoft, J.; Mao, C.; Seeman, N.C.
Deposited on : 2016-02-09
Resolution : 5.12 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 ⓘ symbol.

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	:	2.11
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)	:	2.11

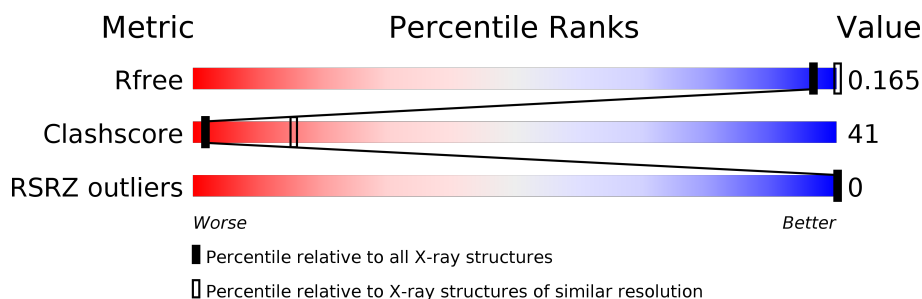
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 5.12 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	1144 (6.34-3.80)
Clashscore	141614	1006 (6.34-3.82)
RSRZ outliers	127900	1003 (6.50-3.74)

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 $\leq 5\%$. 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	A	21	<div> <div>14%</div> <div>86%</div> </div>
2	B	26	<div> <div>12%</div> <div>73%</div> <div>15%</div> </div>
3	C	14	<div> <div>7%</div> <div>86%</div> <div>7%</div> </div>
4	D	14	<div> <div>7%</div> <div>86%</div> <div>7%</div> </div>
5	E	14	<div> <div>7%</div> <div>86%</div> <div>7%</div> </div>
6	F	21	<div> <div>14%</div> <div>81%</div> <div>5%</div> </div>
7	G	21	<div> <div>19%</div> <div>76%</div> <div>5%</div> </div>

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 2664 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*TP*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
1	A	21	Total	C	N	O	P	0	0	0
			426	203	79	124	20			

- Molecule 2 is a DNA chain called DNA (26-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	26	Total	C	N	O	P	0	0	0
			520	249	93	153	25			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	14	Total	C	N	O	P	0	0	0
			290	139	53	85	13			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	14	Total	C	N	O	P	0	0	0
			285	137	52	83	13			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	14	Total	C	N	O	P	0	0	0
			287	137	52	85	13			

- Molecule 6 is a DNA chain called DNA (5'-D(*CP*AP*GP*AP*TP*AP*CP*CP*TP*GP*

AP*TP*CP*GP*GP*AP*CP*TP*AP*CP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	21	Total	C	N	O	P	0	0	0
			427	204	81	122	20			

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

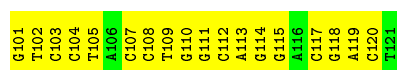
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	21	Total	C	N	O	P	0	0	0
			429	204	84	121	20			

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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.

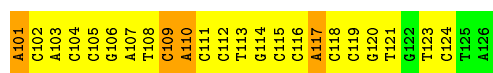
- Molecule 1: DNA (5'-D(*GP*TP*CP*CP*TP*AP*CP*CP*TP*GP*GP*CP*AP*GP*GP*AP*CP*GP*AP*CP*T)-3')

Chain A: 



- Molecule 2: DNA (26-MER)

Chain B: 



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

Chain C: 



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

Chain D: 



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

Chain E: 

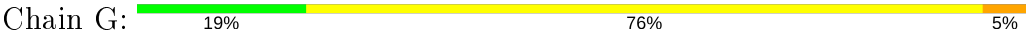


- Molecule 6: DNA (5'-D(*CP*AP*GP*AP*TP*AP*CP*CP*TP*GP*AP*TP*CP*GP*GP*AP*CP*TP*AP*CP*G)-3')

Chain F: 

C201	A202	G203	A204	T205	A206	C207	C208	T209	G210	A211	T212	C213	G214	G215	A216	C217	T218	A219	C220	G221
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

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



G301	A302	G303	C304	G305	A306	C307	C308	T309	G310	T311	A312	C313	G314	G315	A316	C317	A318	T319	C320	A321
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	67.53Å 69.13Å 70.11Å 98.19° 97.34° 99.74°	Depositor
Resolution (Å)	33.58 – 5.12 33.58 – 5.12	Depositor EDS
% Data completeness (in resolution range)	96.2 (33.58-5.12) 87.6 (33.58-5.12)	Depositor EDS
R_{merge}	0.90	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.54 (at 5.07Å)	Xtriage
Refinement program	PHENIX (1.10_2152)	Depositor
R, R_{free}	0.134 , 0.165 0.134 , 0.165	Depositor DCC
R_{free} test set	464 reflections (9.83%)	wwPDB-VP
Wilson B-factor (Å ²)	251.7	Xtriage
Anisotropy	0.215	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.07 , 99.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.095 for k,l,h 0.095 for l,h,k 0.037 for -k,-h,-l 0.048 for -l,-k,-h 0.056 for -h,-l,-k	Xtriage
F_o, F_c correlation	0.99	EDS
Total number of atoms	2664	wwPDB-VP
Average B, all atoms (Å ²)	352.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.42% 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	1.26	1/477 (0.2%)	1.21	1/734 (0.1%)
2	B	1.09	1/581 (0.2%)	1.15	7/892 (0.8%)
3	C	1.17	1/325 (0.3%)	1.18	0/502
4	D	0.99	0/319	1.23	1/491 (0.2%)
5	E	1.11	1/321 (0.3%)	1.11	0/495
6	F	1.17	1/479 (0.2%)	1.12	1/737 (0.1%)
7	G	1.16	1/482 (0.2%)	1.07	2/742 (0.3%)
All	All	1.14	6/2984 (0.2%)	1.15	12/4593 (0.3%)

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	112	DC	C1'-N1	7.22	1.58	1.49
3	C	103	DT	C1'-N1	6.03	1.57	1.49
7	G	319	DT	C3'-O3'	5.93	1.51	1.44
5	E	114	DC	C1'-N1	5.20	1.56	1.49
6	F	212	DT	C1'-N1	5.17	1.55	1.49

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	117	DA	O4'-C1'-N9	6.07	112.25	108.00
2	B	110	DA	O4'-C1'-N9	6.01	112.20	108.00
2	B	110	DA	C1'-O4'-C4'	-5.88	104.22	110.10
1	A	112	DC	O4'-C1'-N1	5.62	111.93	108.00
2	B	109	DC	O4'-C1'-N1	5.61	111.92	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	426	0	237	32	0
2	B	520	0	293	34	0
3	C	290	0	161	24	1
4	D	285	0	160	25	0
5	E	287	0	160	18	0
6	F	427	0	237	29	0
7	G	429	0	236	20	1
All	All	2664	0	1484	153	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 41.

The worst 5 of 153 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:109:DG:O6	6:F:208:DC:N4	2.06	0.88
3:C:112:DA:H2''	3:C:113:DG:H5'	1.64	0.80
1:A:104:DC:N4	3:C:113:DG:O6	2.17	0.78
1:A:113:DA:H1'	1:A:114:DG:C5	2.19	0.77
5:E:110:DG:N2	7:G:308:DC:N3	2.32	0.77

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:C:102:DC:N4	7:G:301:DG:O6[1_556]	2.15	0.05

5.3 Torsion angles

5.3.1 Protein backbone

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 carbohydrates 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, 95th 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	A	21/21 (100%)	-1.49	0 100 100	258, 313, 382, 389	0
2	B	26/26 (100%)	-1.11	0 100 100	257, 301, 809, 868	0
3	C	14/14 (100%)	-1.50	0 100 100	280, 370, 436, 443	0
4	D	14/14 (100%)	-1.53	0 100 100	294, 367, 403, 420	0
5	E	14/14 (100%)	-1.57	0 100 100	297, 336, 406, 419	0
6	F	21/21 (100%)	-1.50	0 100 100	277, 329, 362, 374	0
7	G	21/21 (100%)	-1.49	0 100 100	318, 339, 395, 404	0
All	All	131/131 (100%)	-1.43	0 100 100	257, 335, 430, 868	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 carbohydrates 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.