Full wwPDB X-ray Structure Validation Report ⓘ

Aug 27, 2023 – 07:45 AM EDT

PDB ID	:	3IGC
Title	:	Smallpox virus topoisomerase-DNA transition state
Authors	:	Perry, K.; Hwang, Y.; Bushman, F.D.; Van Duyne, G.D.
Deposited on	:	2009-07-27
Resolution	:	2.10 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:

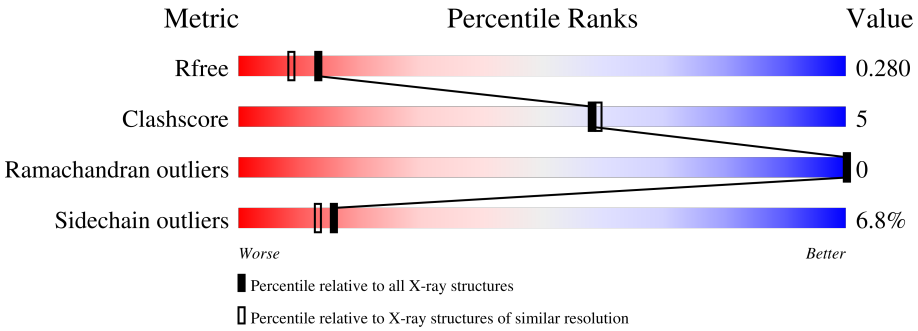
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 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 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)

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	314	84% 13% .
2	B	11	18% 82%
3	C	5	40% 60%
4	D	16	38% 50% 12%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 3663 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 protein called DNA topoisomerase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	313	Total	C	N	O	S	0	1	0
			2587	1676	439	463	9			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	100	SER	CYS	engineered mutation	UNP P32989
A	211	SER	CYS	engineered mutation	UNP P32989

- Molecule 2 is a DNA chain called 5'-D(*GP*TP*GP*TP*CP*GP*CP*CP*CP*TP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	11	Total	C	N	O	P	0	0	0
			219	106	35	68	10			

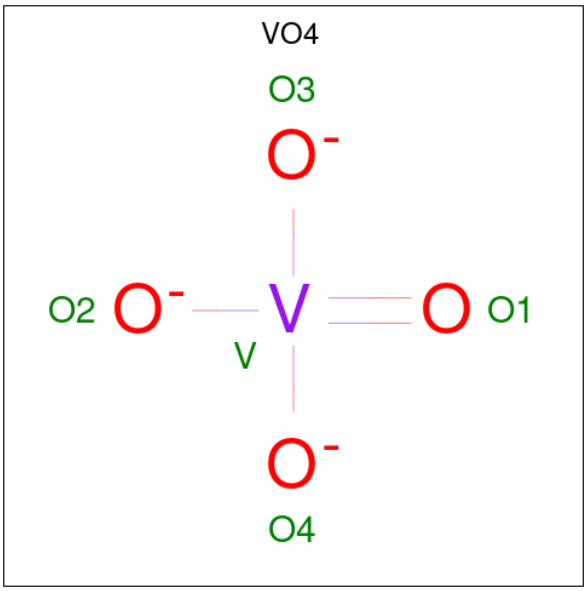
- Molecule 3 is a DNA chain called 5'-D(*AP*TP*TP*CP*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	5	Total	C	N	O	P	0	0	0
			96	48	15	29	4			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	16	Total	C	N	O	P	0	0	0
			332	157	71	89	15			

- Molecule 5 is VANADATE ION (three-letter code: VO4) (formula: O_4V).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total O V 3 2 1	0	0

- Molecule 6 is water.

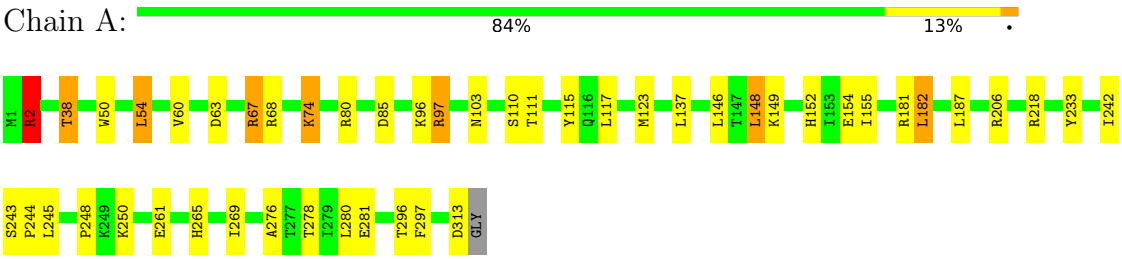
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	316	Total O 316 316	0	0
6	B	31	Total O 31 31	0	0
6	C	11	Total O 11 11	0	0
6	D	68	Total O 68 68	0	0

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 topoisomerase 1

Chain A:



• Molecule 2: 5'-D(*GP*TP*GP*TP*CP*GP*CP*CP*CP*TP*T)-3'

Chain B:

• Molecule 3: 5'-D(*AP*TP*TP*CP*C)-3'

Chain C:

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

Chain D:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, α, β, γ	104.00Å 104.00Å 93.01Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.10 41.60 – 2.00	Depositor EDS
% Data completeness (in resolution range)	96.0 (50.00-2.10) 96.3 (41.60-2.00)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.73 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.5.0072	Depositor
R, R_{free}	0.177 , 0.244 0.230 , 0.280	Depositor DCC
R_{free} test set	1686 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	38.3	Xtriage
Anisotropy	0.067	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 36.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	3663	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.72% 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: VO4

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	0.76	0/2644	0.81	6/3562 (0.2%)
2	B	1.46	0/243	2.04	9/373 (2.4%)
3	C	1.23	0/106	1.75	2/161 (1.2%)
4	D	1.61	5/375 (1.3%)	2.00	15/578 (2.6%)
All	All	0.97	5/3368 (0.1%)	1.20	32/4674 (0.7%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	517	DC	N1-C6	11.62	1.44	1.37
4	D	517	DC	C2-O2	6.44	1.30	1.24
4	D	517	DC	N3-C4	5.80	1.38	1.33
4	D	517	DC	C4-C5	5.63	1.47	1.43
4	D	528	DC	C3'-O3'	-5.42	1.36	1.44

All (32) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	505	DC	O4'-C1'-N1	-11.59	99.89	108.00
1	A	67	ARG	NE-CZ-NH2	-10.06	115.27	120.30
4	D	531	DC	O4'-C1'-N1	-9.55	101.32	108.00
4	D	517	DC	O4'-C1'-N1	8.74	114.12	108.00
1	A	2	ARG	NE-CZ-NH2	-8.09	116.25	120.30
2	B	501	DG	O4'-C1'-N9	7.50	113.25	108.00
1	A	2	ARG	NE-CZ-NH1	7.43	124.02	120.30
4	D	520	DA	O4'-C1'-N9	-7.42	102.81	108.00
4	D	519	DG	O4'-C1'-N9	7.40	113.18	108.00
4	D	532	DA	O4'-C1'-N9	-7.15	102.99	108.00
4	D	519	DG	P-O3'-C3'	7.08	128.19	119.70
2	B	504	DT	O4'-C1'-N1	-6.91	103.17	108.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	D	522	DT	P-O3'-C3'	6.72	127.76	119.70
4	D	522	DT	O3'-P-O5'	-6.63	91.39	104.00
4	D	523	DA	O4'-C1'-N9	-6.53	103.43	108.00
2	B	502	DT	O4'-C4'-C3'	6.45	109.87	106.00
2	B	507	DC	C1'-O4'-C4'	-6.36	103.74	110.10
4	D	527	DG	O4'-C1'-N9	-6.29	103.60	108.00
4	D	522	DT	O4'-C1'-C2'	-6.08	101.04	105.90
2	B	511	DT	C3'-C2'-C1'	-6.05	95.24	102.50
1	A	80	ARG	NE-CZ-NH1	5.87	123.23	120.30
2	B	511	DT	N3-C4-O4	5.77	123.36	119.90
4	D	522	DT	C6-C5-C7	-5.67	119.50	122.90
3	C	512	DA	C5-C6-N1	5.65	120.53	117.70
4	D	523	DA	N1-C2-N3	-5.48	126.56	129.30
1	A	80	ARG	NE-CZ-NH2	-5.43	117.59	120.30
2	B	509	DC	P-O3'-C3'	5.37	126.14	119.70
1	A	148	LEU	CB-CG-CD1	5.35	120.09	111.00
2	B	506	DG	O4'-C1'-N9	-5.26	104.32	108.00
3	C	513	DT	C5-C4-O4	-5.21	121.25	124.90
4	D	521	DA	O5'-P-OP2	-5.19	101.03	105.70
4	D	527	DG	C2-N3-C4	5.04	114.42	111.90

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	2587	0	2653	28	0
2	B	219	0	126	1	0
3	C	96	0	58	1	0
4	D	332	0	179	1	0
5	A	3	0	0	0	0
6	A	316	0	0	7	1
6	B	31	0	0	0	0
6	C	11	0	0	1	0
6	D	68	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	3663	0	3016	30	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 5.

All (30) 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)
1:A:74:LYS:CD	1:A:74:LYS:H	1.57	1.18
1:A:74:LYS:HD2	1:A:74:LYS:N	1.49	1.17
1:A:74:LYS:H	1:A:74:LYS:HD2	0.82	0.97
1:A:111:THR:HG23	6:A:403:HOH:O	1.66	0.96
1:A:149:LYS:H	1:A:152:HIS:HD2	1.20	0.88
3:C:516:DC:OP2	6:C:332:HOH:O	1.91	0.87
1:A:248:PRO:HA	6:A:421:HOH:O	1.82	0.80
1:A:74:LYS:HG2	6:A:471:HOH:O	1.82	0.79
1:A:206:ARG:NH2	2:B:503:DG:N7	2.34	0.75
1:A:2:ARG:HD2	6:A:442:HOH:O	1.85	0.74
1:A:123:MET:HG2	1:A:182:LEU:HD11	1.84	0.59
1:A:38:THR:HG21	6:A:607:HOH:O	2.05	0.55
1:A:149:LYS:H	1:A:152:HIS:CD2	2.12	0.52
1:A:110:SER:HA	1:A:115:TYR:CD1	2.44	0.52
1:A:97:ARG:NH1	6:A:545:HOH:O	2.40	0.49
1:A:155:ILE:HD13	1:A:187:LEU:HD12	1.95	0.48
1:A:233:TYR:HA	1:A:297:PHE:CE1	2.48	0.48
1:A:60:VAL:HG22	1:A:68:ARG:NH2	2.31	0.46
1:A:38:THR:HG23	6:A:366:HOH:O	2.16	0.45
1:A:74:LYS:CD	1:A:74:LYS:N	2.32	0.45
1:A:103:ASN:ND2	1:A:181:ARG:HH21	2.16	0.44
1:A:243:SER:HB3	1:A:244:PRO:HD3	2.00	0.44
1:A:296:THR:HG22	1:A:297:PHE:N	2.33	0.43
1:A:54:LEU:HD12	1:A:74:LYS:HD3	2.00	0.43
1:A:63:ASP:OD2	1:A:67:ARG:HD3	2.18	0.42
4:D:531:DC:H2"	4:D:532:DA:C8	2.54	0.42
1:A:50:TRP:CH2	1:A:74:LYS:HE2	2.55	0.42
1:A:269:ILE:HD12	1:A:269:ILE:HA	1.96	0.41
1:A:85:ASP:OD1	1:A:276:ALA:HB1	2.21	0.41
1:A:278:THR:HA	1:A:281:GLU:HG2	2.02	0.40

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)
6:A:326:HOH:O	6:A:582:HOH:O[8_555]	0.77	1.43

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	312/314 (99%)	304 (97%)	8 (3%)	0	100 100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	294/293 (100%)	274 (93%)	20 (7%)	16 13

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	ARG
1	A	38	THR
1	A	54	LEU
1	A	74	LYS
1	A	96	LYS
1	A	97	ARG
1	A	117	LEU
1	A	137	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	146	LEU
1	A	148	LEU
1	A	154	GLU
1	A	182	LEU
1	A	218	ARG
1	A	242	ILE
1	A	245	LEU
1	A	250	LYS
1	A	261	GLU
1	A	265	HIS
1	A	280	LEU
1	A	313	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	79	ASN
1	A	93	ASN
1	A	103	ASN
1	A	152	HIS

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)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and

the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
5	VO4	A	600	3,2	0,2,4	-	-	-		

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

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.