wwPDB X-ray Structure Validation Summary Report (i)

Oct 17, 2021 – 04:42 AM EDT

PDB ID	:	1LPQ
Title	:	Human DNA Topoisomerase I (70 Kda) In Non-Covalent Complex With A 22 Base Pair DNA Duplex Containing an 8-oxoG Lesion
Authors	:	Lesher, D.T.; Pommier, Y.; Stewart, L.; Redinbo, M.R.
Deposited on	:	2002-05-08
Resolution	:	3.14 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 following versions of software and data (see references (i)) were used in the production of this report:

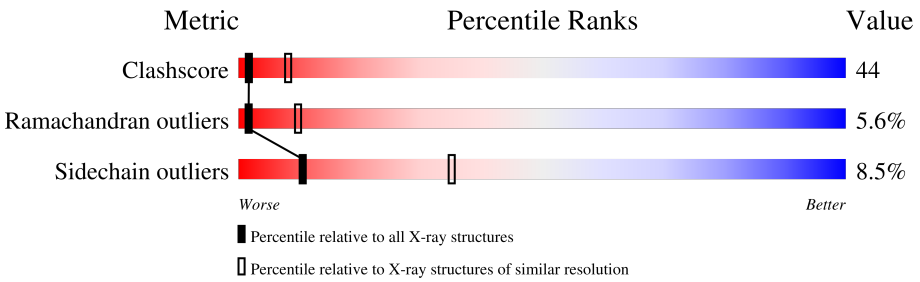
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 3.14 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))
Clashscore	141614	1735 (3.18-3.10)
Ramachandran outliers	138981	1677 (3.18-3.10)
Sidechain outliers	138945	1677 (3.18-3.10)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	B	22	77% 18% 5%
2	C	22	36% 59% 5%
3	A	564	40% 50% 8% ••

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5381 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 5'-D(*AP*AP*AP*AP*AP*GP*AP*CP*TP*TP*(8OG)P*GP*AP*AP*AP*AP*AP*TP*TP*TP*TP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	B	22	454	219	87	127	21	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	C	22	443	217	71	134	21	0	0	0

- Molecule 3 is a protein called DNA topoisomerase I.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	557	4457	2849	784	800	24	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	723	PHE	TYR	engineered mutation	UNP P11387

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	B	2	2	2	0	0
4	C	3	3	3	0	0
4	A	22	22	22	0	0

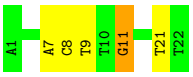
3 Residue-property plots (i)

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.

Note EDS was not executed.

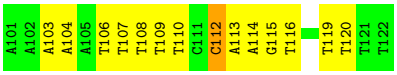
• Molecule 1: 5'-D(*AP*AP*AP*AP*AP*GP*AP*CP*TP*TP*(8OG)P*GP*AP*AP*AP*AP*AP*TP*TP*TP*TP*T)-3'

Chain B: 77% 18% 5%



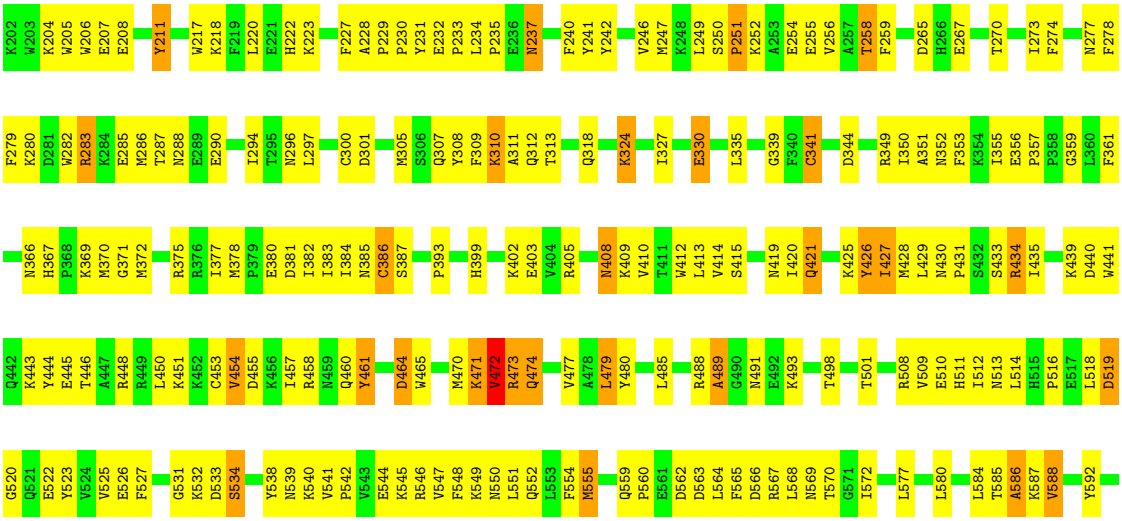
• Molecule 2: 5'-D(*AP*AP*AP*AP*AP*TP*TP*TP*TP*TP*CP*CP*AP*AP*GP*TP*CP*TP*TP*TP*TP*T)-3'

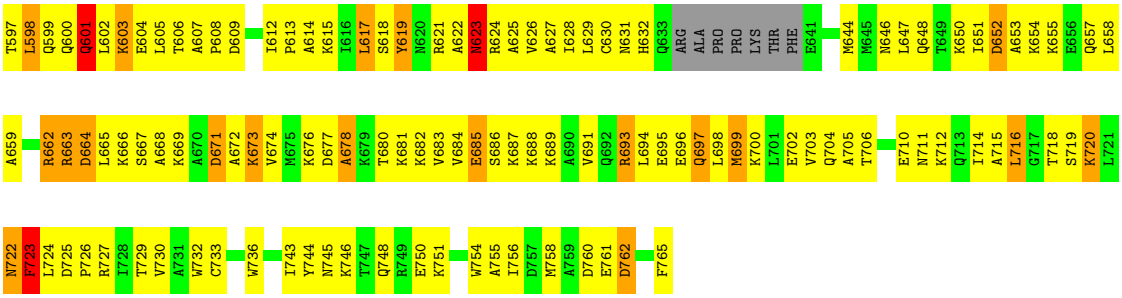
Chain C: 36% 59% 5%



• Molecule 3: DNA topoisomerase I

Chain A: 40% 50% 8% ••




T597 L598 Q599 Q600 Q601 L602 K603 E604 L605 T606 A607 P608 D609
I612 P613 A614 K615 I616 L617 S618 Y619 N620 R621 A622 N623 R624 A625 V626 A627 I628 L629 C630 N631 H632 Q633 ARG ALA PRO PRO LYS THR PHE E641
M644 M645 N646 L647 Q648 T649 K650 I651 D652 A653 K654 K655 E656 Q657 L658
A659 R662 R663 D664 L665 K666 S667 A668 K669 A670 D671 A672 K673 V674 M675 K676 D677 A678 K679 T680 K681 K682 V683 V684 E685 S686 K687 K688 K689 A690 V691 Q692 R693 L694 E695 E696 Q697 L698 M699 K700 L701 E702 V703 Q704 A705 T706
E710 N711 K712 Q713 I714 A715 L716 G717 T718 S719 K720 L721
N722 F723 L724 D725 P726 R727 I728 T729 V730 A731 W732 C733
W736
I743 Y744 N745 K746 T747 Q748 R749 E750 K751
W754 A755 I756 D757 M758 A759 D760 E761 D762
F765

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	57.20Å 122.50Å 72.00Å 90.00° 97.40° 90.00°	Depositor
Resolution (Å)	19.87 – 3.14	Depositor
% Data completeness (in resolution range)	95.5 (19.87-3.14)	Depositor
R_{merge}	0.33	Depositor
R_{sym}	0.20	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.256 , 0.298	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5381	wwPDB-VP
Average B, all atoms (Å^2)	62.0	wwPDB-VP

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: 8OG

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	B	0.55	0/484	0.78	0/743
2	C	0.64	0/494	0.86	1/760 (0.1%)
3	A	0.45	0/4555	0.69	5/6142 (0.1%)
All	All	0.48	0/5533	0.71	6/7645 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	C	0	1

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	664	ASP	N-CA-C	-7.79	89.98	111.00
3	A	472	VAL	N-CA-C	-6.99	92.12	111.00
3	A	426	TYR	N-CA-C	5.43	125.66	111.00
3	A	663	ARG	N-CA-C	-5.40	96.42	111.00
3	A	723	PHE	N-CA-C	5.26	125.19	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	112	DC	Sidechain

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	B	454	0	251	6	0
2	C	443	0	255	16	0
3	A	4457	0	4375	433	0
4	A	22	0	0	5	0
4	B	2	0	0	0	0
4	C	3	0	0	1	0
All	All	5381	0	4881	451	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 44.

The worst 5 of 451 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)
3:A:473:ARG:CZ	3:A:550:ASN:HB3	1.64	1.26
3:A:473:ARG:NH2	3:A:550:ASN:HB3	1.57	1.17
3:A:473:ARG:CZ	3:A:550:ASN:CB	2.37	1.01
3:A:273:ILE:H	3:A:273:ILE:HD12	1.28	0.96
3:A:267:GLU:O	3:A:270:THR:HG22	1.68	0.91

There are no symmetry-related clashes.

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
3	A	553/564 (98%)	412 (74%)	110 (20%)	31 (6%)	2 10

5 of 31 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	386	CYS
3	A	473	ARG
3	A	519	ASP
3	A	673	LYS
3	A	723	PHE

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
3	A	457/508 (90%)	418 (92%)	39 (8%)	10 35

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	652	ASP
3	A	722	ASN
3	A	662	ARG
3	A	693	ARG
3	A	761	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
3	A	632	HIS
3	A	748	GLN
3	A	722	ASN
3	A	408	ASN
3	A	601	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue 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	$\#\|Z\| > 2$
1	8OG	B	11	1,2	18,25,26	1.29	2 (11%)	21,37,40	2.52	5 (23%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	8OG	B	11	1,2	-	0/3/21/22	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	11	8OG	C6-N1	3.80	1.39	1.33
1	B	11	8OG	C8-N7	-3.24	1.30	1.34

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	11	8OG	C5-C6-N1	-8.63	111.63	123.43
1	B	11	8OG	C6-N1-C2	5.70	124.98	115.93
1	B	11	8OG	C2-N3-C4	-3.17	111.74	115.36
1	B	11	8OG	N3-C2-N1	-2.42	124.00	127.22
1	B	11	8OG	C6-C5-C4	-2.06	118.83	120.80

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	11	8OG	1	0

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 (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.