

wwPDB X-ray Structure Validation Summary Report (i)

Feb 3, 2024 – 05:59 PM EST

PDB ID	:	1MU5
Title	:	Structure of topoisomerase subunit
Authors	:	Corbett, K.D.; Berger, J.M.
Deposited on	:	2002-09-23
Resolution	:	2.00 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:

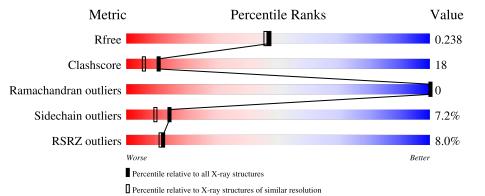
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: $X\text{-}RAY \, DIFFRACTION$

The reported resolution of this entry is 2.00 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	$\begin{array}{l} \textbf{Whole archive} \\ \textbf{(\#Entries)} \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
R_{free}	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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% 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				
			8%				
1	А	471	69%	25%	••		



1MU5

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 3923 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 Type II DNA topoisomerase VI Subunit B.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
1	А	460	Total 3684	C 2371	N 618	O 689	S 6	0	0	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
А	-1	GLY	-	expression tag	UNP 005207
А	0	ALA	-	expression tag	UNP 005207
А	?	-	MET	deletion	UNP 005207
А	303	TYR	ASP	SEE REMARK 999	UNP 005207
А	435	ASP	ASN	SEE REMARK 999	UNP 005207

• Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	А	2	Total Ca 2 2	0	0

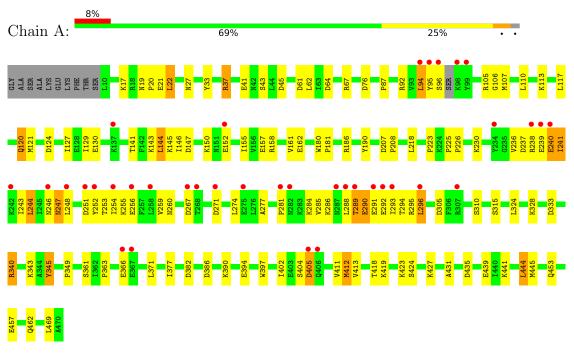
• Molecule 3 is water.

N	Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
	3	А	237	Total O 237 237	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 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: Type II DNA topoisomerase VI Subunit B



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants	94.09Å 110.97Å 54.54Å	Depositor
a, b, c, α , β , γ	90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 - 2.00	Depositor
Resolution (A)	19.89 - 2.00	EDS
% Data completeness	96.0 (20.00-2.00)	Depositor
(in resolution range)	96.0 (19.89-2.00)	EDS
R _{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) > 1$	$2.06 (at 2.01 \text{\AA})$	Xtriage
Refinement program	REFMAC 5.0	Depositor
B B.	0.214 , 0.238	Depositor
R, R_{free}	0.214 , 0.238	DCC
R_{free} test set	3130 reflections $(8.29%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	25.8	Xtriage
Anisotropy	0.608	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.40 , 55.0	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	0.94	EDS
Total number of atoms	3923	wwPDB-VP
Average B, all atoms $(Å^2)$	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 4.63% of the height of the origin peak. No significant pseudotranslation is detected.

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.



¹Intensities estimated from amplitudes.

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: CA

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		
	Unain	RMSZ	# Z > 5	RMSZ	# Z > 5	
1	А	0.51	0/3762	0.75	9/5091~(0.2%)	

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
1	А	382	ASP	CB-CG-OD2	6.54	124.18	118.30
1	А	45	ASP	CB-CG-OD2	6.21	123.89	118.30
1	А	305	ASP	CB-CG-OD2	5.82	123.54	118.30
1	А	267	ASP	CB-CG-OD2	5.66	123.39	118.30
1	А	386	ASP	CB-CG-OD2	5.54	123.28	118.30

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	А	3684	0	3746	137	0
2	А	2	0	0	0	0
3	А	237	0	0	14	0
All	All	3923	0	3746	137	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 18.

The worst 5 of 137 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:117:LEU:O	1:A:121:MET:HG3	1.48	1.13
1:A:254:ILE:HD11	1:A:288:LEU:HD21	1.33	1.10
1:A:94:LEU:H	1:A:94:LEU:CD1	1.65	1.10
1:A:277:ALA:O	1:A:295:ARG:HD2	1.54	1.07
1:A:94:LEU:H	1:A:94:LEU:HD12	0.90	1.05

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
1	А	456/471~(97%)	446 (98%)	10 (2%)	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 side chain 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	in Analysed Rotameric C		Outliers	Percentiles	
1	А	403/416~(97%)	374~(93%)	29~(7%)	14 9	

5 of 29 residues with a non-rotameric sidechain are listed below:



Mol	Chain	Res	Type
1	А	246	ASN
1	А	424	SER
1	А	290	GLU
1	А	405	ASP
1	А	289	THR

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

Mol	Chain	Res	Type
1	А	27	ASN
1	А	120	GLN
1	А	408	GLN
1	А	453	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)

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)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

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 (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, 95^{th} 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	А	460/471~(97%)	0.35	37 (8%) 12 11	4, 13, 45, 57	0

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	А	291	GLU	5.7
1	А	406	GLN	4.5
1	А	282	ASN	4.5
1	А	366	GLU	4.4
1	А	95	TYR	4.3

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 monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	CA	А	471	1/1	0.95	0.07	43,43,43,43	0
2	CA	А	1	1/1	0.97	0.09	48,48,48,48	0



6.5 Other polymers (i)

There are no such residues in this entry.

