



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 10:10 PM GMT

PDB ID : 1SEU
Title : Human DNA Topoisomerase I (70 Kda) In Complex With The Indolocarbazole SA315F and Covalent Complex With A 22 Base Pair DNA Duplex
Authors : Staker, B.L.; Feese, M.D.; Cushman, M.; Pommier, Y.; Zembower, D.; Stewart, L.; Burgin, A.B.
Deposited on : 2004-02-18
Resolution : 3.00 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
<http://wwpdb.org/validation/2016/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
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

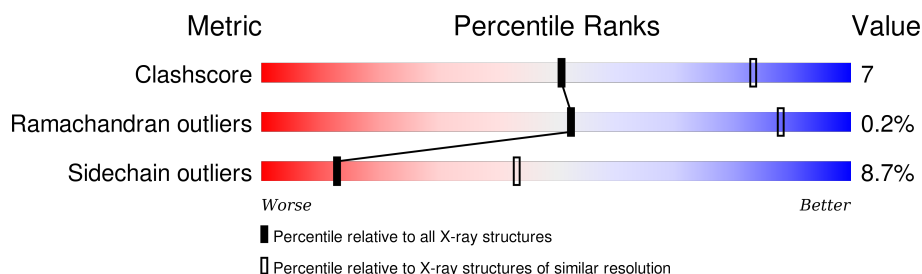
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	B	10	100%
2	C	12	33% 67%
3	D	22	32% 68%
4	A	592	80% 13% • 5%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	TGP	C	11	-	-	X	-
5	SA3	D	990	X	-	-	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 5615 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*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	10	Total	C	N	O	P	0	0	0
			203	99	42	53	9			

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

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	C	12	Total	C	N	O	P	S	0	0	0
			246	120	45	69	11	1			

- Molecule 3 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
3	D	22	Total	C	N	O	P	0	0	0
			443	217	71	134	21			

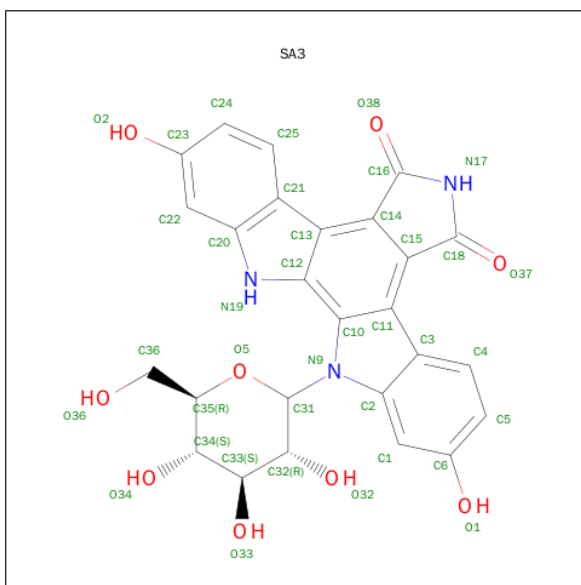
- Molecule 4 is a protein called DNA topoisomerase I.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
4	A	565	Total	C	N	O	P	S	0	0	0
			4685	2979	822	857	1	26			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	722	SER	ASN	ENGINEERED	UNP P11387
A	723	PTR	TYR	MODIFIED RESIDUE	UNP P11387

- Molecule 5 is 2,10-DIHYDROXY-12-(BETA-D-GLUCOPYRANOSYL)-6,7,12,13-TETRAHYDROINDOLO[2,3-A]PYRROLO[3,4-C]CARBAZOLE-5,7-DIONE (three-letter code: SA3) (formula: C₂₆H₂₁N₃O₉).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	D	1	Total	C	N	O	0	0
			38	26	3	9		

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 errors 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.

Note EDS was not executed.

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

Chain B:  100%

A1
A2
A3
A4
A5
A6
A7
A8
A9
A10

- Molecule 2: 5'-D(*(TGP)P*GP*AP*AP*AP*AP*AP*TP*TP*TP*TP*T)-3'

Chain C:  33% 67%


G11
G12
A13
A14
A15
A16
A17
T18
T21
T22

- Molecule 3: 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 D:  32% 68%

A101
A102
T108
T109
G112
A113
A114
G115
T116
C117
T118
T119
T120
T121
T122

- Molecule 4: DNA topoisomerase I

Chain A:  80% 13% 5%

LYS
LYS
PRO
LYS
ASN
LYS
ASP
LYS
ASP
LYS
VAL
P80
GLU
PRO
ASP
ASN
LYS
LYS
LYS
LYS
PRO
LYS
LYS
GLU
GLU
Q201
Y211
K216
H222
N237
Y238
K239
F240
Y241
K252
A253
E254
E255
V256
A257
T258
R283
K284
E285
S298
D301
Q307
L335
G341
E356
H366
H367
M370
L373
K374
R375
D381
T382
D389
F396
P397
G398
H399
V414
E418
N419
Y420
Q421
I424
K425
N428
L429
N430
P431
S432
I435
E438
K439
Q460
L479
L487
E492
K493
L518
N539
R540
V541
P542
F565
L568
N569
T570
G571
Q578
G583
L584
T585
Y592
Q599
L605
T606
D609
L617
R621
A622
N623
V626
D652
D660
D664
D671
M699
E702
V703
Q704
T714
A715
L716
G717
T718
S719
Y723
L724
V732
C733
N745
K746
T747
G748
R749
A759
F765

4 Data and refinement statistics

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	57.38 Å 115.97 Å 73.55 Å 90.00° 93.71° 90.00°	Depositor
Resolution (Å)	46.63 – 3.00	Depositor
% Data completeness (in resolution range)	80.4 (46.63-3.00)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
Refinement program	CNX 2002	Depositor
R, R_{free}	0.237 , 0.295	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5615	wwPDB-VP
Average B, all atoms (Å ²)	78.0	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: SA3, TGP, PTR

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	0.58	0/229	0.81	0/351
2	C	0.37	0/254	0.72	0/390
3	D	0.50	0/494	0.75	0/760
4	A	0.65	0/4768	0.77	3/6398 (0.0%)
All	All	0.63	0/5745	0.76	3/7899 (0.0%)

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
4	A	0	1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
4	A	617	LEU	CA-CB-CG	7.20	131.87	115.30
4	A	335	LEU	CA-CB-CG	5.76	128.55	115.30
4	A	487	LEU	CA-CB-CG	5.17	127.19	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	A	592	TYR	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	203	0	112	11	0
2	C	246	0	138	15	0
3	D	443	0	255	18	0
4	A	4685	0	4720	35	0
5	D	38	0	18	7	0
All	All	5615	0	5243	75	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 7.

The worst 5 of 75 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:D:119:DT:H2''	3:D:120:DT:H5''	1.44	0.98
3:D:115:DG:H2'	3:D:116:DT:H71	1.46	0.95
1:B:9:DT:C6	1:B:10:DT:H72	2.11	0.86
3:D:112:DC:H2''	3:D:113:DA:H5''	1.60	0.81
1:B:1:DA:H2'	1:B:2:DA:C8	2.15	0.81

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
4	A	562/592 (95%)	525 (93%)	36 (6%)	1 (0%)	52 88

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	759	ALA

5.3.2 Protein sidechains ⓘ

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	
4	A	504/535 (94%)	460 (91%)	44 (9%)	13	43

5 of 44 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	A	438	GLU
4	A	518	LEU
4	A	716	LEU
4	A	439	LYS
4	A	479	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
4	A	430	ASN
4	A	539	ASN
4	A	578	GLN
4	A	419	ASN
4	A	569	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues are 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$
4	PTR	A	723	1,4	14,16,17	1.23	0	18,22,24	1.00	1 (5%)
2	TGP	C	11	3,2	17,21,25	1.09	1 (5%)	20,31,38	2.65	4 (20%)

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
4	PTR	A	723	1,4	-	0/9/11/13	0/1/1/1
2	TGP	C	11	3,2	-	0/2/18/22	0/3/3/3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	11	TGP	C6-N1	3.45	1.39	1.33

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	11	TGP	C5-C6-N1	-8.94	111.37	123.59
2	C	11	TGP	N3-C2-N1	-2.41	123.78	127.44
2	C	11	TGP	C4'-C5'-S5'	-2.10	110.59	113.89
4	A	723	PTR	O2P-P-O1P	2.48	118.56	110.58
2	C	11	TGP	C6-N1-C2	6.65	125.17	115.94

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	723	PTR	2	0
2	C	11	TGP	7	0

5.5 Carbohydrates [i](#)

There are no carbohydrates 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	SA3	D	990	-	36,44,44	3.33	12 (33%)	40,70,70	2.30	12 (30%)

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
5	SA3	D	990	-	1/1/5/5	0/2/26/26	0/7/7/7

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	D	990	SA3	O1-C6	-12.78	1.06	1.37
5	D	990	SA3	O2-C23	-10.59	1.11	1.37
5	D	990	SA3	C16-C14	2.00	1.47	1.42
5	D	990	SA3	C34-C33	2.05	1.57	1.52
5	D	990	SA3	C21-C20	2.13	1.48	1.42

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

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
5	D	990	SA3	O33-C33-C32	-4.60	99.97	110.34
5	D	990	SA3	C11-C10-C12	-4.24	117.33	121.91
5	D	990	SA3	C1-C2-C3	-4.08	114.73	120.92
5	D	990	SA3	C33-C32-C31	-4.01	102.69	109.21
5	D	990	SA3	C31-N9-C10	-3.84	118.73	125.54

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	D	990	SA3	C31

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	D	990	SA3	7	0

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 ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.