



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 16, 2021 – 07:09 PM EDT

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](mailto: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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
buster-report : 1.1.7 (2018)  
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

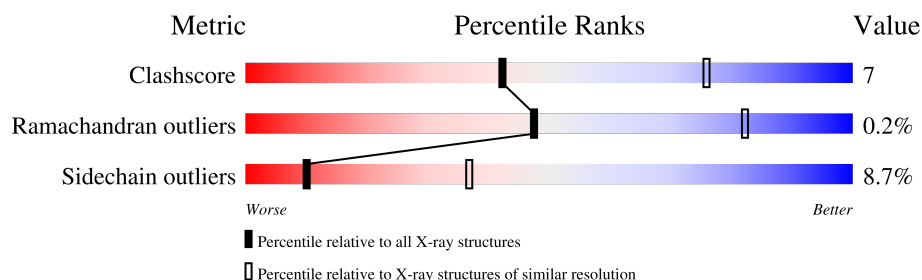
# 1 Overall quality at a glance

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	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	B	10	
2	C	12	
3	D	22	
4	A	592	

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			

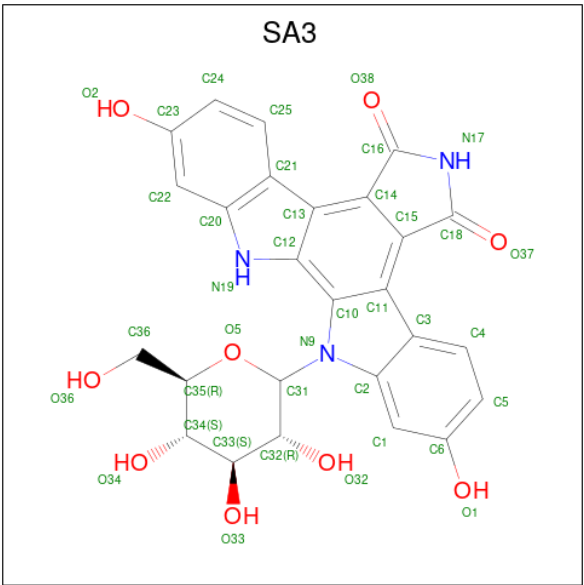
- Molecule 4 is a protein called DNA topoisomerase I.

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 mutation	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: C<sub>26</sub>H<sub>21</sub>N<sub>3</sub>O<sub>9</sub>).



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, 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\*T)-3'

Chain B:  100%

A1  
A2  
A3  
A4  
A5  
G6  
A7  
C8  
T9  
T10

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

Chain C:  33% 58% 8%


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  
C112  
A113  
A114  
G115  
T116  
C117  
T118  
T119  
T120  
T121  
T122

- Molecule 4: DNA topoisomerase I

Chain A:  80% 13% 5%

LYS  
LYS  
PRD  
LYS  
LYS  
ASN  
LYS  
ASP  
LYS  
ASP  
LYS  
LYS  
LYS  
VAL  
PRD  
GLU  
PRD  
ASP  
ASN  
ASN  
LYS  
LYS  
LYS  
LYS  
PRO  
LYS  
LYS  
GLU  
GLU  
GLU  
Q201  
Y211  
K216  
H222  
N237  
V238  
K239  
P240  
Y241  
K252  
A253  
E254  
E255  
V256  
A257  
T258  
R263  
K264  
E265  
S298  
D301  
Q307  
L335

C341  
E356  
N366  
H367  
M370  
L373  
G383  
L384  
R375  
D381  
I382  
D389  
P396  
P397  
G398  
H399  
V414  
E418  
M419  
I420  
Q421  
I424  
K425  
M428  
L429  
M430  
P431  
S432  
I435  
E438  
K439  
Q450  
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  
W732  
C733  
N745  
K746

T747  
G748  
R749  
A759  
F765

## 4 Data and refinement statistics

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, $\alpha$ , $\beta$ , $\gamma$	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 <sup>2</sup> )	78.0	wwPDB-VP

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: TGP, SA3, 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%)	47 82

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	759	ALA

### 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	
4	A	504/535 (94%)	460 (91%)	44 (9%)	10	37

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

Mol	Chain	Res	Type
4	A	606	THR
4	A	699	MET
4	A	609	ASP
4	A	623	ASN
4	A	703	VAL

Sometimes 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	578	GLN
4	A	593	ASN
4	A	599	GLN
4	A	419	ASN
4	A	430	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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	15,16,17	1.16	0	19,22,24	0.89	1 (5%)
2	TGP	C	11	3,2	18,21,25	1.16	2 (11%)	19,31,38	2.69	4 (21%)

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	-	0/10/11/13	0/1/1/1
2	TGP	C	11	3,2	-	0/2/18/22	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	11	TGP	C6-N1	3.72	1.39	1.33
2	C	11	TGP	C8-N7	-2.15	1.30	1.34

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	11	TGP	C5-C6-N1	-8.82	111.37	123.43
2	C	11	TGP	C6-N1-C2	5.81	125.17	115.93
2	C	11	TGP	C2-N3-C4	-3.15	111.77	115.36
2	C	11	TGP	N3-C2-N1	-2.58	123.78	127.22
4	A	723	PTR	O2P-P-O1P	2.01	118.56	110.68

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 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	SA3	D	990	-	42,44,44	4.54	26 (61%)	41,70,70	3.94	22 (53%)

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	2/6/26/26	0/7/7/7

The worst 5 of 26 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	D	990	SA3	O1-C6	-12.36	1.06	1.43
5	D	990	SA3	O2-C23	-10.60	1.11	1.43
5	D	990	SA3	C20-N19	-8.40	1.35	1.48
5	D	990	SA3	C22-C20	-8.04	1.40	1.53
5	D	990	SA3	C4-C3	-7.38	1.39	1.53

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	990	SA3	C31-N9-C2	11.26	133.95	114.47
5	D	990	SA3	C1-C6-C5	8.09	120.21	110.55
5	D	990	SA3	C22-C23-C24	7.78	119.84	110.55
5	D	990	SA3	O1-C6-C1	6.59	122.97	109.85
5	D	990	SA3	C25-C24-C23	6.21	122.21	111.61

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	D	990	SA3	C31

All (2) torsion outliers are listed below:

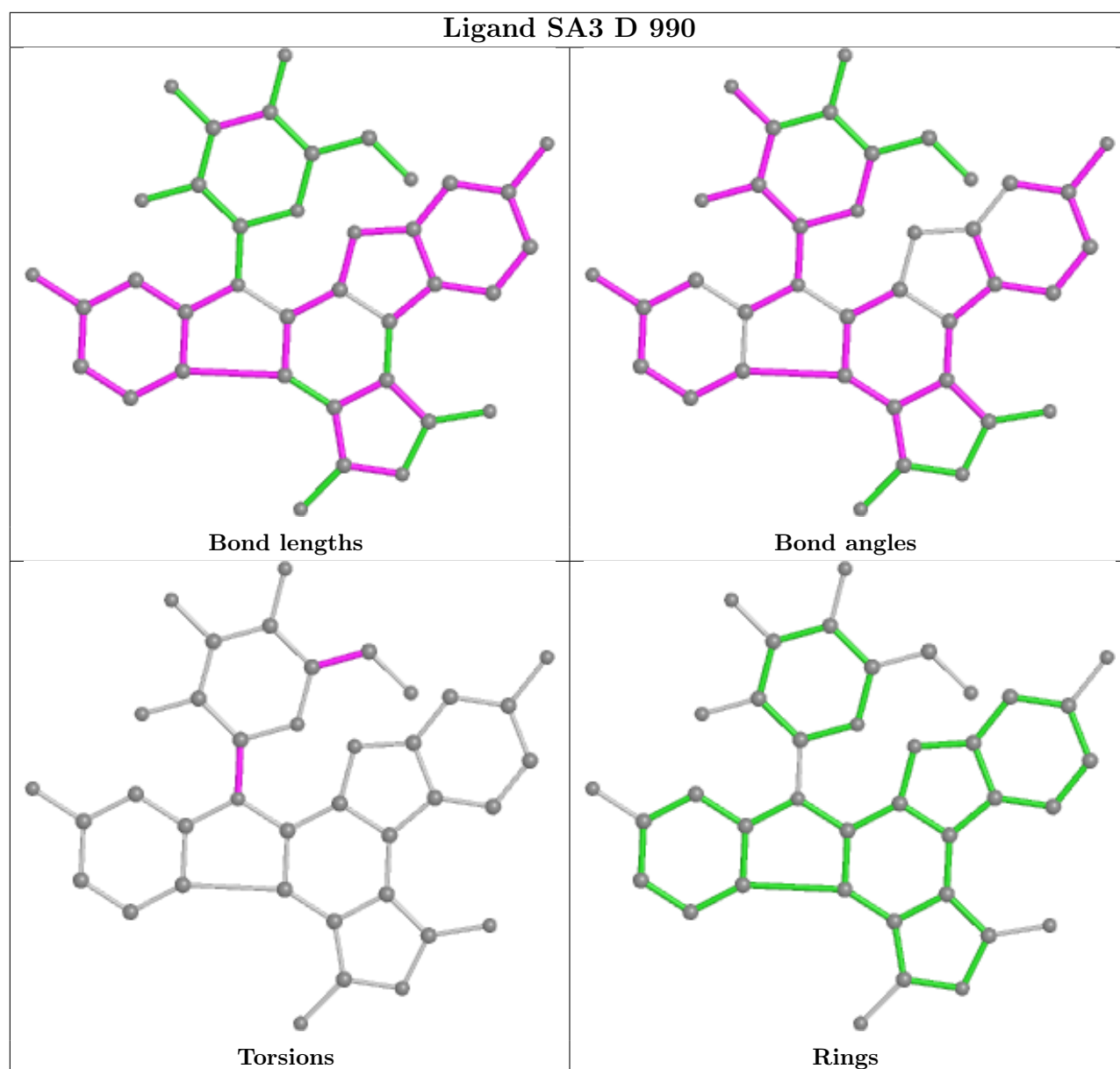
Mol	Chain	Res	Type	Atoms
5	D	990	SA3	O5-C31-N9-C2
5	D	990	SA3	O5-C35-C36-O36

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

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 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 is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.