



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 14, 2020 – 03:24 pm BST

PDB ID : 2NMV  
Title : Damage detection by the UvrABC pathway: Crystal structure of UvrB bound to fluorescein-adducted DNA  
Authors : Waters, T.R.; Eryilmaz, J.; Geddes, S.; Barrett, T.E.  
Deposited on : 2006-10-23  
Resolution : 2.95 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) : 2.11

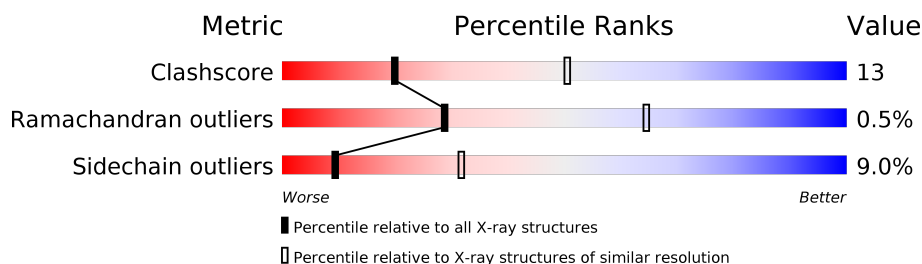
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.95 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	3462 (3.00-2.92)
Ramachandran outliers	138981	3340 (3.00-2.92)
Sidechain outliers	138945	3343 (3.00-2.92)

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  $\geq 3$ , 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	D	5	
2	A	661	
3	B	38	

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 5311 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(P\*TP\*TP\*TP\*TP\*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	D	5	Total	C	N	O	P	0	0	0
			84	40	8	31	5			

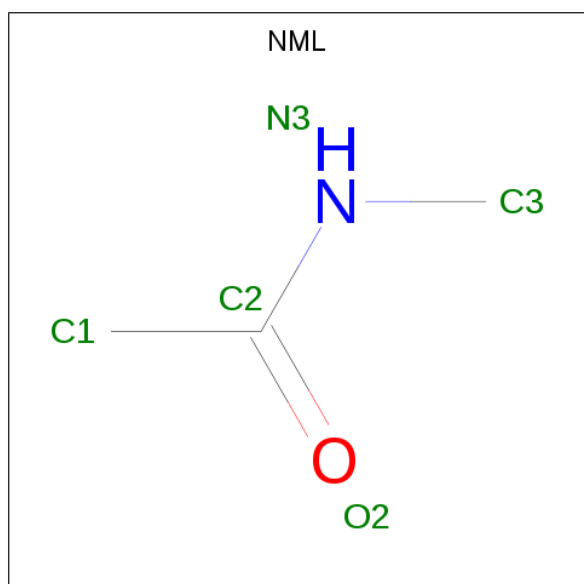
- Molecule 2 is a protein called UvrABC system protein B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	620	Total	C	N	O	S	0	0	0
			4878	3073	850	937	18			

- Molecule 3 is a protein called UvrABC system protein B.

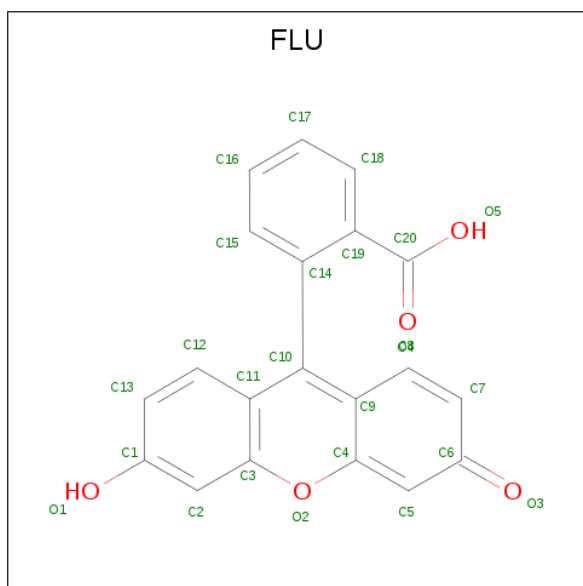
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	38	Total	C	N	O	S	0	0	0
			301	187	53	59	2			

- Molecule 4 is N-METHYLACETAMIDE (three-letter code: NML) (formula: C<sub>3</sub>H<sub>7</sub>NO).



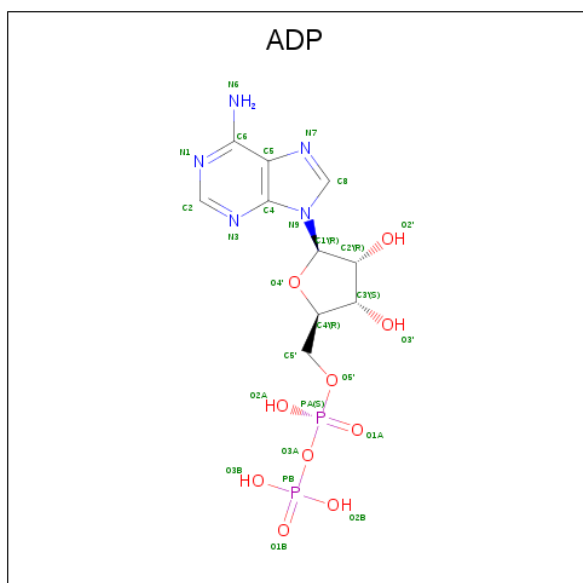
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	D	1	Total	C	N	O	0	0
			5	3	1	1		

- Molecule 5 is 2-(6-HYDROXY-3-OXO-3H-XANTHEN-9-YL)-BENZOIC ACID (three-letter code: FLU) (formula: C<sub>20</sub>H<sub>12</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	D	1	Total	C	O	0	0
			16	13	3		

- Molecule 6 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula: C<sub>10</sub>H<sub>15</sub>N<sub>5</sub>O<sub>10</sub>P<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	A	1	Total	C	N	O	P	0	0
			27	10	5	10	2		

### 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 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(P\*TP\*TP\*TP\*TP\*T)-3'

Chain D: 

T1  
T2  
T3  
T4  
T5

- Molecule 2: UvrABC system protein B

Chain A: 

MET  
LYS  
P3  
V8  
SER  
D106  
K10  
Y11  
Q12  
P13  
Q14  
Q17  
P18  
K19  
A20  
I21  
I28  
K33  
S141  
R34  
Q35  
L38  
G39  
K45  
T46  
P47  
T48  
S50  
N51  
L52  
L53  
K54  
E55  
V56  
N57  
K58  
P59  
T60  
K67  
Q72  
P80  
N84  
P89  
V90  
S91  
Y92  
Y96  
E99

A100  
Y101  
V102  
P103  
Q104  
T105  
D106  
T107  
F108  
I109  
D120  
K121  
S125  
R133  
R134  
D135  
A140  
S141  
V142  
I145  
Y146  
G147  
L148  
E152  
E156  
S160  
L161  
I162  
T163  
I167  
E168  
K58  
L172  
L176  
Y181  
F188  
Q189  
T192  
P199  
V200  
E201  
L202  
F203  
S206  
E207

D208  
E209  
H210  
C211  
V212  
R213  
T221  
E222  
R223  
D239  
S247  
V250  
L260  
Q261  
E266  
Q270  
L271  
M274  
H275  
K279  
E282  
A283  
Q284  
R285  
L286  
E287  
Q288  
R289  
L294  
C303  
I306  
R311  
T314  
L315  
R316  
T321  
P322  
Y323  
T324  
L325  
P330  
M334

T344  
P345  
Q346  
V347  
K350  
P351  
Q354  
D355  
K358  
Q359  
V360  
V362  
R367  
L368  
P369  
N374  
R375  
P376  
L377  
R378  
F379  
E380  
E381  
T389  
E409  
T415  
G416  
L417  
L418  
D419  
P420  
D423  
V424  
R425  
T427  
Q430  
I431  
T435  
G436  
E437  
R441  
V450  
L451  
T452

L453  
L454  
M457  
Y464  
L465  
T468  
G469  
I470  
V471  
V472  
N473  
S477  
E478  
I479  
K480  
T481  
L482  
E483  
L491  
L499  
V500  
L505  
L509  
D510  
I511  
V514  
E525  
R529  
S530  
E531  
R532  
S533  
T537  
A542  
R569  
R570  
Q573  
R576  
P585  
N589  
L589  
GLU

ILE  
ARG  
ASP  
VAL  
ILE  
ARG  
ALA  
THR  
VAL  
ALA  
ALA  
GLU  
ASP  
LYS  
ALA  
GLU  
TYR  
LYS  
THR  
LYS  
ALA  
ALA  
PRO  
LYS  
LYS  
LEU  
SER  
LYS  
MET  
THR  
K621  
M631  
K636  
K640  
R651  
P652  
L653  
L654  
LEU  
LEU  
LEU  
LYS  
ALA  
GLU  
GLY

- Molecule 3: UvrABC system protein B

Chain B: 

K622  
L642  
E645  
R646  
L659

## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.41Å 95.60Å 97.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	68.36 – 2.95	Depositor
% Data completeness (in resolution range)	98.8 (68.36-2.95)	Depositor
$R_{merge}$	0.09	Depositor
$R_{sym}$	0.09	Depositor
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.224 , 0.281	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5311	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	50.0	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: NML, FLU, ADP

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	D	2.28	3/91 (3.3%)	3.45	13/139 (9.4%)
2	A	0.77	0/4957	0.84	2/6703 (0.0%)
3	B	0.73	0/302	0.72	0/402
All	All	0.82	3/5350 (0.1%)	0.96	15/7244 (0.2%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	3	DT	C5-C7	11.35	1.56	1.50
1	D	3	DT	N1-C6	-6.44	1.33	1.38
1	D	4	DT	C1'-N1	5.38	1.56	1.49

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	1	DT	O4'-C1'-N1	14.02	117.82	108.00
1	D	2	DT	OP2-P-O3'	12.29	132.24	105.20
1	D	3	DT	C2-N3-C4	-11.67	120.20	127.20
1	D	3	DT	P-O3'-C3'	-10.79	106.75	119.70
1	D	3	DT	O3'-P-O5'	-10.23	84.57	104.00

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	D	84	0	48	14	0
2	A	4878	0	4747	124	0
3	B	301	0	295	1	0
4	D	5	0	5	0	0
5	D	16	0	6	0	0
6	A	27	0	12	2	0
All	All	5311	0	5113	131	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 13.

The worst 5 of 131 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)
2:A:532:ARG:H	2:A:532:ARG:HD2	1.12	1.10
2:A:133:ARG:HD3	2:A:135:ASP:OD1	1.59	1.00
2:A:260:ILE:HG23	2:A:294:LEU:HD22	1.46	0.95
2:A:311:ARG:HD3	2:A:316:ARG:O	1.67	0.94
2:A:532:ARG:H	2:A:532:ARG:CD	1.88	0.86

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	
2	A	614/661 (93%)	581 (95%)	30 (5%)	3 (0%)	29	64
3	B	36/38 (95%)	36 (100%)	0	0	100	100
All	All	650/699 (93%)	617 (95%)	30 (5%)	3 (0%)	29	64

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	167	ILE
2	A	430	GLN
2	A	585	PRO

### 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	
2	A	513/586 (88%)	466 (91%)	47 (9%)	9	30
3	B	29/32 (91%)	27 (93%)	2 (7%)	15	44
All	All	542/618 (88%)	493 (91%)	49 (9%)	9	32

5 of 49 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	223	ARG
2	A	275	HIS
2	A	653	LEU
2	A	250	VAL
2	A	314	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
2	A	355	GLN
2	A	573	GLN
2	A	388	ASN
2	A	288	GLN
2	A	536	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

3 ligands 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$
5	FLU	D	600	-	15,18,28	1.51	2 (13%)	17,26,41	2.79	10 (58%)
4	NML	D	103	1	4,4,4	1.51	1 (25%)	4,4,4	1.68	1 (25%)
6	ADP	A	662	-	24,29,29	1.34	3 (12%)	29,45,45	1.46	5 (17%)

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	FLU	D	600	-	-	-	0/3/3/4
4	NML	D	103	1	-	0/2/2/2	-
6	ADP	A	662	-	-	6/12/32/32	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	D	600	FLU	C10-C9	3.69	1.47	1.38
6	A	662	ADP	O4'-C1'	3.49	1.45	1.41
6	A	662	ADP	C5-C4	3.35	1.49	1.40
5	D	600	FLU	C10-C11	2.93	1.48	1.42

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	103	NML	C1-C2	2.56	1.55	1.50

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	600	FLU	C2-C3-C11	-4.75	117.98	123.10
5	D	600	FLU	C9-C10-C11	-4.00	114.60	121.53
5	D	600	FLU	C12-C11-C3	3.83	121.04	116.61
5	D	600	FLU	O2-C3-C2	3.72	120.46	116.11
5	D	600	FLU	C8-C9-C10	-3.63	118.42	122.20

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

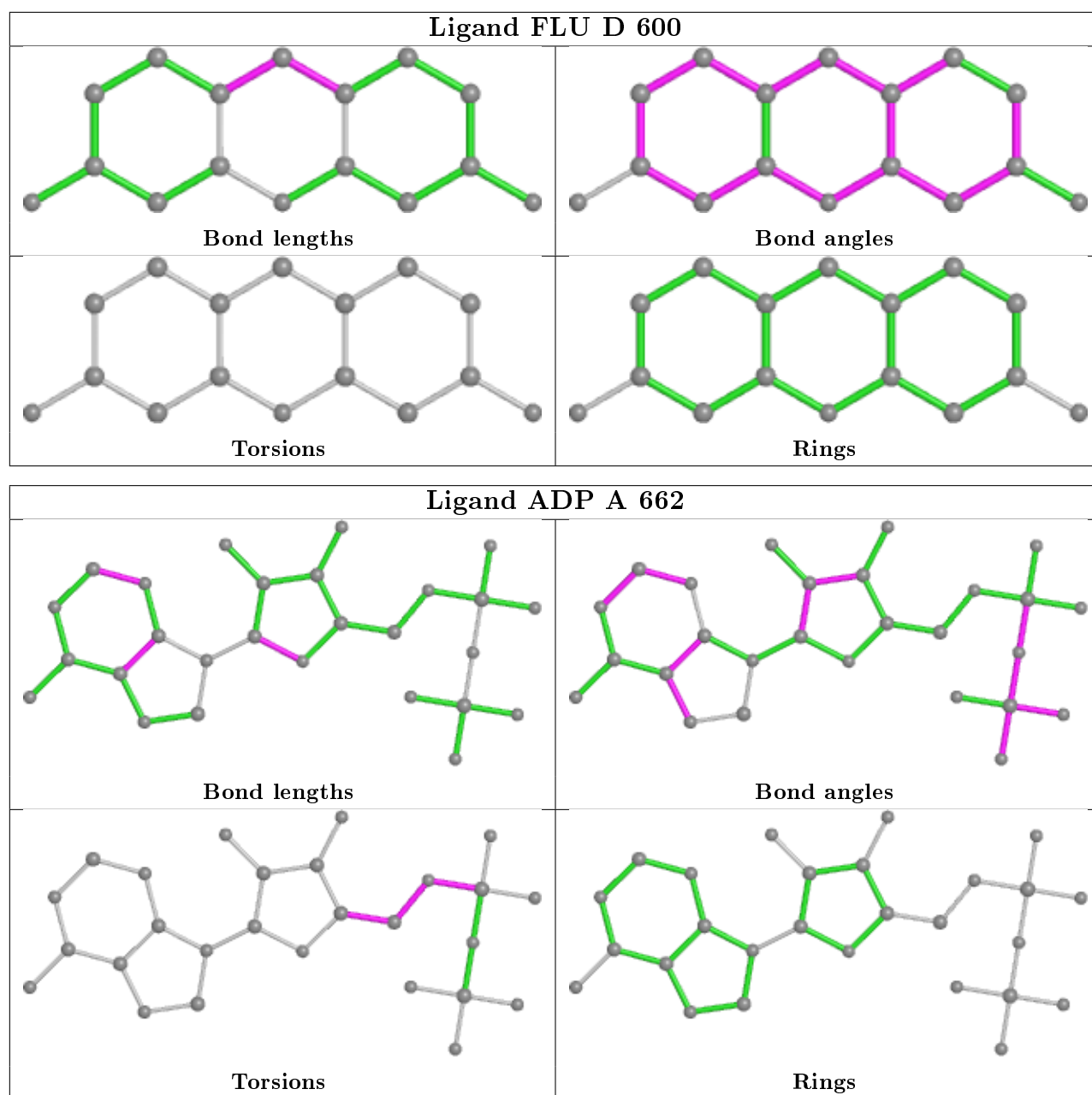
Mol	Chain	Res	Type	Atoms
6	A	662	ADP	C5'-O5'-PA-O2A
6	A	662	ADP	O4'-C4'-C5'-O5'
6	A	662	ADP	C3'-C4'-C5'-O5'
6	A	662	ADP	C5'-O5'-PA-O3A
6	A	662	ADP	C5'-O5'-PA-O1A

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	662	ADP	2	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.