



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 19, 2020 – 09:27 AM BST

PDB ID : 6UR9  
Title : DNA polymerase I Large Fragment from *Bacillus stearothermophilus* with DNA template, dideoxy primer, 3'-amino-ddGTP (nGTP), and Ca<sup>2+</sup>  
Authors : Zhang, W.; Lelyveld, V.S.; Szostak, J.W.  
Deposited on : 2019-10-22  
Resolution : 2.10 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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) : 1.13  
EDS : 2.14.6  
buster-report : 1.1.7 (2018)  
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) : 2.14.6

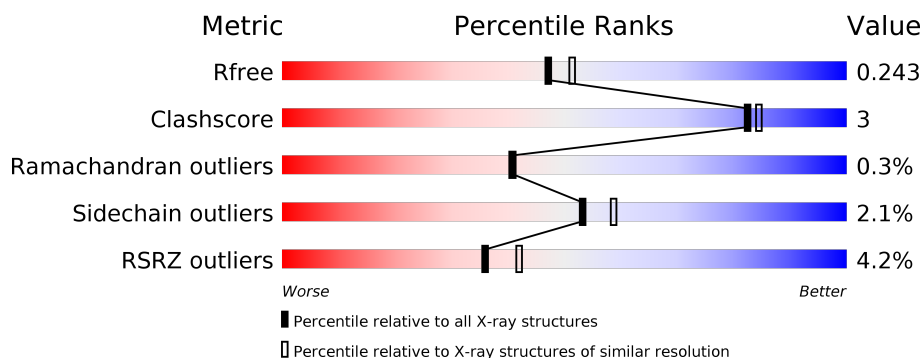
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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))
$R_{free}$	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 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  $\leq 5\%$ . 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
1	A	579	<div> <div>2%</div> <div> <div></div> <div>93%</div> <div>7%</div> </div> </div>
1	D	579	<div> <div>6%</div> <div> <div></div> <div>91%</div> <div>8%</div> </div> </div>
2	B	9	<div> <div></div> <div> <div>78%</div> <div>11%</div> <div>11%</div> </div> </div>
2	E	9	<div> <div></div> <div> <div>89%</div> <div>11%</div> </div> </div>
3	C	13	<div> <div></div> <div> <div>46%</div> <div>54%</div> </div> </div>
3	F	13	<div> <div></div> <div> <div>85%</div> <div>15%</div> </div> </div>

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
6	SO4	A	905	-	-	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 10551 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 DNA polymerase I.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	579	Total	C	N	O	S	0	0	0
			4646	2955	807	867	17			
1	D	579	Total	C	N	O	S	0	0	0
			4646	2955	807	867	17			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	710	TYR	PHE	engineered mutation	UNP D9N168
A	713	VAL	PRO	variant	UNP D9N168
D	710	TYR	PHE	engineered mutation	UNP D9N168
D	713	VAL	PRO	variant	UNP D9N168

- Molecule 2 is a DNA chain called DNA (5'-D(\*GP\*CP\*GP\*AP\*TP\*CP\*AP\*GP\*(DOC))-3').

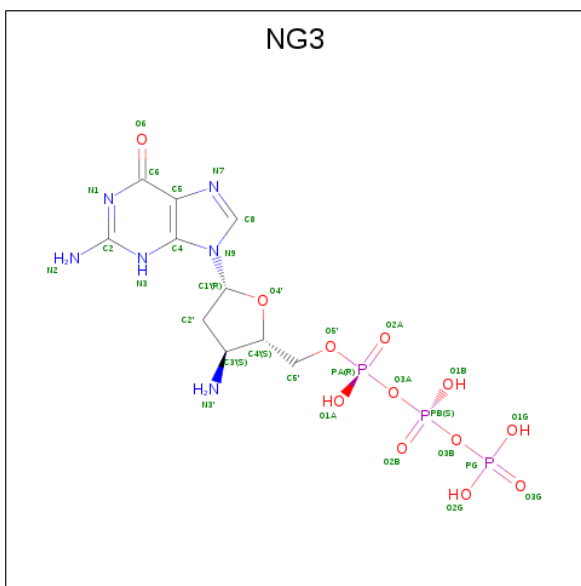
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	9	Total	C	N	O	P	0	0	0
			181	87	36	50	8			
2	E	9	Total	C	N	O	P	0	0	0
			181	87	36	50	8			

- Molecule 3 is a DNA chain called DNA (5'-D(\*CP\*AP\*CP\*GP\*CP\*TP\*GP\*AP\*TP\*CP\*GP\*CP\*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	13	Total	C	N	O	P	0	0	1
			229	106	41	70	12			
3	F	13	Total	C	N	O	P	0	0	0
			244	115	44	73	12			

- Molecule 4 is 3'-amino-2',3'-dideoxyguanosine 5'-(tetrahydrogen triphosphate) (three-letter

code: NG3) (formula:  $\text{C}_{10}\text{H}_{17}\text{N}_6\text{O}_{12}\text{P}_3$ ) (labeled as "Ligand of Interest" by author).

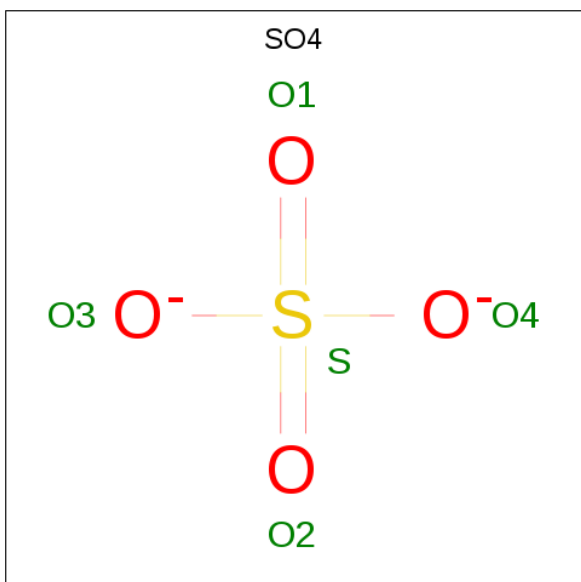


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			31	10	6	12	3		

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total Ca 1 1	0	0

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula:  $\text{O}_4\text{S}$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total O S 5 4 1	0	0
6	A	1	Total O S 5 4 1	0	0
6	A	1	Total O S 5 4 1	0	0
6	A	1	Total O S 5 4 1	0	0
6	A	1	Total O S 5 4 1	0	0
6	A	1	Total O S 5 4 1	0	0
6	A	1	Total O S 5 4 1	0	0
6	A	1	Total O S 5 4 1	0	0
6	D	1	Total O S 5 4 1	0	0

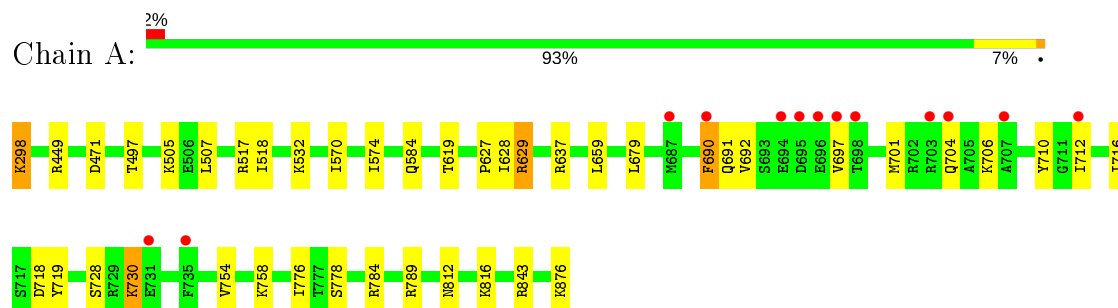
- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	193	Total O 193 193	0	0
7	B	13	Total O 13 13	0	0
7	C	17	Total O 17 17	0	0
7	D	100	Total O 100 100	0	0
7	E	14	Total O 14 14	0	0
7	F	10	Total O 10 10	0	0

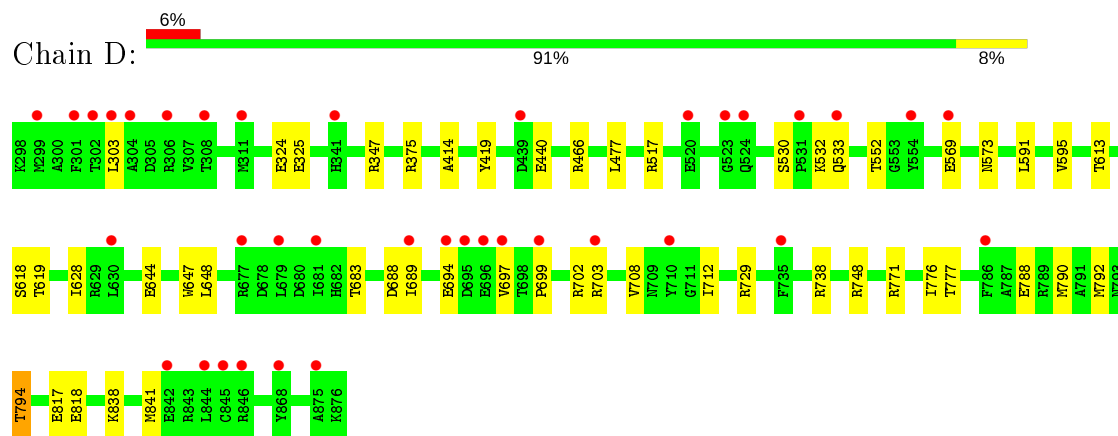
### 3 Residue-property plots [i](#)

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.

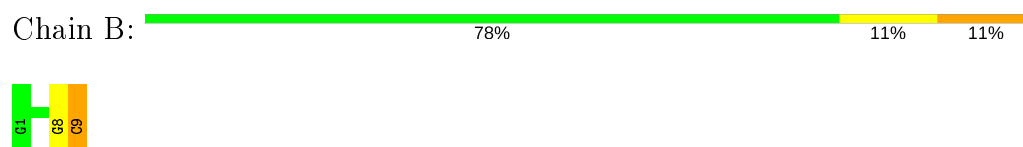
- Molecule 1: DNA polymerase I



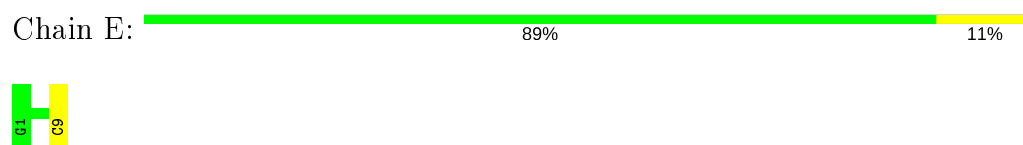
- Molecule 1: DNA polymerase I



- Molecule 2: DNA (5'-D(\*GP\*CP\*GP\*AP\*TP\*CP\*AP\*GP\*(DOC))-3')



- Molecule 2: DNA (5'-D(\*GP\*CP\*GP\*AP\*TP\*CP\*AP\*GP\*(DOC))-3')




- Molecule 3: DNA (5'-D(\*CP\*AP\*CP\*GP\*CP\*TP\*GP\*AP\*TP\*CP\*GP\*CP\*A)-3')

Chain C:  46% 54%



- Molecule 3: DNA (5'-D(\*CP\*AP\*CP\*GP\*CP\*TP\*GP\*AP\*TP\*CP\*GP\*CP\*A)-3')

Chain F:  85% 15%





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	93.63Å 108.07Å 149.75Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.81 – 2.10 46.81 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.7 (46.81-2.10) 99.6 (46.81-2.10)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.35 (at 2.10Å)	Xtriage
Refinement program	PHENIX 1.17_3644	Depositor
R, $R_{free}$	0.204 , 0.243 0.204 , 0.243	Depositor DCC
$R_{free}$ test set	4537 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.6	Xtriage
Anisotropy	0.035	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 49.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	10551	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.56% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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.

## 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: DOC, CA, SO4, NG3

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	A	0.25	0/4730	0.42	0/6393
1	D	0.24	0/4730	0.41	0/6393
2	B	0.54	0/183	0.88	0/281
2	E	0.52	0/183	0.88	0/281
3	C	0.55	0/255	0.86	0/393
3	F	0.57	0/272	0.85	0/418
All	All	0.28	0/10353	0.48	0/14159

There are no bond length outliers.

There are no bond angle outliers.

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	A	4646	0	4705	24	0
1	D	4646	0	4705	23	0
2	B	181	0	102	1	0
2	E	181	0	102	0	0
3	C	229	0	123	6	0
3	F	244	0	135	3	0
4	A	31	0	0	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	1	0	0	0	0
6	A	40	0	0	0	0
6	D	5	0	0	0	0
7	A	193	0	0	1	0
7	B	13	0	0	0	0
7	C	17	0	0	0	0
7	D	100	0	0	2	0
7	E	14	0	0	0	0
7	F	10	0	0	0	0
All	All	10551	0	9872	51	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 3.

All (51) 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:789:ARG:NH1	3:C:4:DG:OP1	2.20	0.75
1:A:843:ARG:NH2	7:A:1001:HOH:O	2.23	0.70
1:D:530:SER:HA	3:F:11:DG:H4'	1.76	0.66
1:A:730:LYS:HE3	1:A:730:LYS:H	1.60	0.64
1:A:697:VAL:HA	1:A:701:MET:HE1	1.79	0.63
1:D:517:ARG:NH1	7:D:1001:HOH:O	2.32	0.63
4:A:901:NG3:N2	3:C:3:DC:O2	2.32	0.63
1:D:790:MET:O	1:D:794:THR:HG23	2.06	0.55
1:A:518:ILE:HD13	1:A:574:ILE:HD13	1.89	0.54
1:D:613:THR:HG22	1:D:794:THR:HB	1.91	0.53
1:A:778:SER:O	1:A:784:ARG:NH1	2.38	0.53
1:D:324:GLU:O	1:D:375:ARG:NH2	2.42	0.52
2:B:8:DG:H2'	2:B:9:DOC:H6	1.92	0.52
1:A:812:ASN:O	1:A:816:LYS:HG2	2.10	0.51
1:D:648:LEU:HD12	1:D:841:MET:HG3	1.93	0.51
1:A:507:LEU:HD22	1:A:584:GLN:HB2	1.93	0.51
1:A:298:LYS:HB3	1:A:449:ARG:HH21	1.77	0.50
1:D:619:THR:HG22	3:F:7:DG:H3'	1.94	0.48
1:A:690:PHE:HZ	1:A:704:GLN:HB3	1.78	0.48
1:A:712:ILE:HA	1:A:716:ILE:HG22	1.95	0.48
1:D:414:ALA:HB1	1:D:419:TYR:HB3	1.96	0.47
1:D:794:THR:HG22	7:D:1040:HOH:O	2.16	0.46
3:C:2:DA:N3	3:C:2:DA:H2'	2.30	0.46
1:D:618:SER:O	3:F:7:DG:H4'	2.16	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:659:LEU:HG	1:A:710:TYR:CE1	2.51	0.46
1:A:627:PRO:O	1:A:637:ARG:HD2	2.17	0.45
1:A:690:PHE:O	1:A:692:VAL:N	2.49	0.45
1:D:817:GLU:HG3	1:D:818:GLU:N	2.31	0.45
1:A:532:LYS:NZ	3:C:13:DA:OP1	2.33	0.45
1:D:569:GLU:O	1:D:573:ASN:ND2	2.45	0.45
1:D:683:THR:HG23	1:D:697:VAL:HG21	1.99	0.44
1:D:689:ILE:O	1:D:738:ARG:NE	2.50	0.44
1:D:838:LYS:HA	1:D:841:MET:HG2	2.00	0.44
1:D:788:GLU:O	1:D:792:MET:HG3	2.18	0.44
1:D:644:GLU:HB2	1:D:647:TRP:CD1	2.54	0.43
1:D:771:ARG:HD3	1:D:794:THR:HG21	2.00	0.43
1:D:347:ARG:NE	1:D:440:GLU:OE2	2.51	0.42
1:A:570:ILE:O	1:A:574:ILE:HG12	2.19	0.42
1:A:471:ASP:N	1:A:471:ASP:OD1	2.49	0.42
1:D:699:PRO:O	1:D:703:ARG:HB2	2.19	0.42
1:D:776:ILE:HG23	1:D:777:THR:HG23	2.01	0.42
1:A:629:ARG:HD2	1:A:629:ARG:HA	1.75	0.42
1:A:728:SER:HB2	1:A:730:LYS:NZ	2.34	0.42
1:D:708:VAL:O	1:D:712:ILE:HG12	2.20	0.42
1:A:505:LYS:HA	1:A:505:LYS:HD2	1.84	0.41
3:C:10:DC:H2"	3:C:11:DG:C8	2.55	0.41
1:A:619:THR:HG22	3:C:7:DG:H3'	2.02	0.41
1:D:591:LEU:O	1:D:595:VAL:HG23	2.20	0.41
1:A:754:VAL:HG12	1:A:758:LYS:HE2	2.02	0.41
1:A:718:ASP:OD1	1:A:719:TYR:N	2.54	0.40
1:A:629:ARG:HH22	1:A:706:LYS:NZ	2.18	0.40

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	A	577/579 (100%)	562 (97%)	13 (2%)	2 (0%)	41	41
1	D	577/579 (100%)	562 (97%)	13 (2%)	2 (0%)	41	41
All	All	1154/1158 (100%)	1124 (97%)	26 (2%)	4 (0%)	41	41

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	691	GLN
1	D	532	LYS
1	D	628	ILE
1	A	628	ILE

### 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	
1	A	495/495 (100%)	486 (98%)	9 (2%)	59	65
1	D	495/495 (100%)	483 (98%)	12 (2%)	49	53
All	All	990/990 (100%)	969 (98%)	21 (2%)	53	59

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	298	LYS
1	A	497	THR
1	A	517	ARG
1	A	629	ARG
1	A	679	LEU
1	A	690	PHE
1	A	730	LYS
1	A	776	ILE
1	A	876	LYS
1	D	303	LEU
1	D	325	GLU
1	D	466	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	477	LEU
1	D	533	GLN
1	D	552	THR
1	D	688	ASP
1	D	694	GLU
1	D	702	ARG
1	D	729	ARG
1	D	748	ARG
1	D	794	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	533	GLN
1	D	704	GLN

### 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$
2	DOC	B	9	3,2	14,19,20	3.41	5 (35%)	13,26,29	1.76	3 (23%)
2	DOC	E	9	3,2	14,19,20	3.47	5 (35%)	13,26,29	1.80	4 (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
2	DOC	B	9	3,2	-	0/4/18/19	0/2/2/2
2	DOC	E	9	3,2	-	0/4/18/19	0/2/2/2

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	9	DOC	O4'-C4'	8.69	1.62	1.44
2	B	9	DOC	O4'-C4'	8.57	1.61	1.44
2	E	9	DOC	O4'-C1'	-6.65	1.27	1.42
2	B	9	DOC	O4'-C1'	-6.52	1.27	1.42
2	E	9	DOC	C4-N4	3.99	1.46	1.35
2	B	9	DOC	C4-N4	3.97	1.46	1.35
2	B	9	DOC	C3'-C4'	-3.69	1.32	1.52
2	E	9	DOC	C3'-C4'	-3.65	1.33	1.52
2	E	9	DOC	C2'-C1'	3.00	1.58	1.51
2	B	9	DOC	C2'-C1'	2.69	1.58	1.51

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	9	DOC	C2-N3-C4	4.65	121.05	116.34
2	E	9	DOC	C2-N3-C4	4.64	121.04	116.34
2	E	9	DOC	N4-C4-N3	2.59	120.58	116.49
2	B	9	DOC	N4-C4-N3	2.45	120.36	116.49
2	E	9	DOC	C4'-O4'-C1'	-2.06	107.87	109.81
2	B	9	DOC	C4'-O4'-C1'	-2.04	107.88	109.81
2	E	9	DOC	C5-C4-N3	-2.04	119.36	121.72

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
2	B	9	DOC	1	0

## 5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 5.6 Ligand geometry

Of 11 ligands modelled in this entry, 1 is monoatomic - leaving 10 for Mogul analysis.

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
6	SO4	A	905	-	4,4,4	0.14	0	6,6,6	0.08	0
6	SO4	A	910	-	4,4,4	0.13	0	6,6,6	0.07	0
6	SO4	A	909	-	4,4,4	0.15	0	6,6,6	0.05	0
6	SO4	A	903	-	4,4,4	0.14	0	6,6,6	0.06	0
6	SO4	A	907	-	4,4,4	0.14	0	6,6,6	0.05	0
6	SO4	A	908	-	4,4,4	0.14	0	6,6,6	0.08	0
4	NG3	A	901	5	26,33,33	3.71	12 (46%)	27,52,52	4.05	13 (48%)
6	SO4	A	904	-	4,4,4	0.15	0	6,6,6	0.07	0
6	SO4	A	906	-	4,4,4	0.14	0	6,6,6	0.05	0
6	SO4	D	901	-	4,4,4	0.14	0	6,6,6	0.04	0

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	NG3	A	901	5	-	3/18/34/34	0/3/3/3

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	901	NG3	C6-C5	-10.31	1.23	1.41
4	A	901	NG3	C6-N1	-7.58	1.19	1.33
4	A	901	NG3	O6-C6	6.44	1.40	1.24
4	A	901	NG3	C4-N3	6.09	1.45	1.35
4	A	901	NG3	C2-N2	5.62	1.45	1.33
4	A	901	NG3	C8-N7	-5.04	1.25	1.34
4	A	901	NG3	C5-C4	3.25	1.49	1.40
4	A	901	NG3	C2-N3	2.44	1.46	1.34
4	A	901	NG3	C2-N1	-2.32	1.31	1.35

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	901	NG3	PG-O2G	-2.19	1.46	1.54
4	A	901	NG3	PG-O1G	-2.16	1.46	1.54
4	A	901	NG3	C1'-N9	-2.13	1.43	1.49

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	901	NG3	N3-C2-N1	-10.42	113.32	127.22
4	A	901	NG3	C4-C5-N7	-9.50	99.50	109.40
4	A	901	NG3	C6-C5-C4	-8.17	113.00	120.80
4	A	901	NG3	C6-N1-C2	8.09	128.78	115.93
4	A	901	NG3	C5-C6-N1	4.82	130.02	123.43
4	A	901	NG3	N2-C2-N1	4.66	124.50	117.25
4	A	901	NG3	C2'-C1'-N9	-3.59	105.98	114.27
4	A	901	NG3	O1G-PG-O3B	2.88	114.30	104.64
4	A	901	NG3	PB-O3B-PG	-2.87	122.97	132.83
4	A	901	NG3	O2G-PG-O3B	2.77	113.92	104.64
4	A	901	NG3	N2-C2-N3	2.69	122.18	117.79
4	A	901	NG3	PB-O3A-PA	-2.64	123.77	132.83
4	A	901	NG3	O1A-PA-O2A	-2.09	101.92	112.24

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	901	NG3	PG-O3B-PB-O2B
4	A	901	NG3	PB-O3A-PA-O1A
4	A	901	NG3	O4'-C4'-C5'-O5'

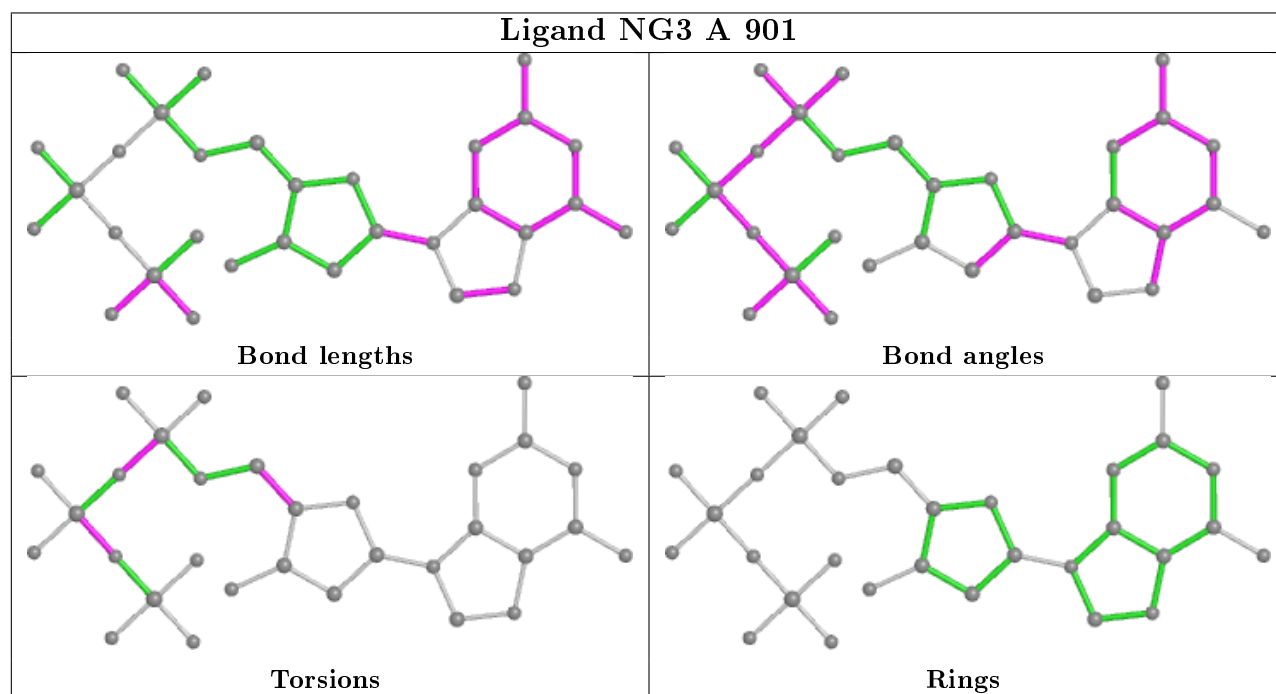
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	901	NG3	1	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 ⓘ

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	579/579 (100%)	0.19	13 (2%) 62 66	21, 34, 61, 102	0
1	D	579/579 (100%)	0.35	37 (6%) 19 24	28, 44, 70, 108	0
2	B	8/9 (88%)	-0.39	0 100 100	26, 33, 45, 49	0
2	E	8/9 (88%)	-0.39	0 100 100	31, 35, 46, 48	0
3	C	13/13 (100%)	-0.21	0 100 100	25, 31, 67, 76	0
3	F	13/13 (100%)	-0.24	0 100 100	28, 38, 79, 98	0
All	All	1200/1202 (99%)	0.25	50 (4%) 36 42	21, 40, 67, 108	0

All (50) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	303	LEU	6.1
1	D	694	GLU	4.8
1	D	695	ASP	4.8
1	D	710	TYR	4.7
1	A	694	GLU	3.8
1	D	679	LEU	3.8
1	D	697	VAL	3.7
1	D	523	GLY	3.7
1	A	735	PHE	3.7
1	D	703	ARG	3.6
1	A	712	ILE	3.5
1	A	695	ASP	3.5
1	A	697	VAL	3.2
1	D	735	PHE	3.0
1	D	699	PRO	3.0
1	D	302	THR	3.0
1	D	311	MET	2.9
1	D	868	TYR	2.9
1	D	301	PHE	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	696	GLU	2.7
1	D	531	PRO	2.7
1	A	687	MET	2.6
1	D	681	ILE	2.6
1	D	846	ARG	2.6
1	D	341	HIS	2.5
1	A	707	ALA	2.5
1	D	533	GLN	2.4
1	D	299	MET	2.3
1	A	690	PHE	2.3
1	A	696	GLU	2.3
1	D	844	LEU	2.3
1	A	704	GLN	2.3
1	D	308	THR	2.2
1	D	677	ARG	2.2
1	D	554	TYR	2.2
1	D	304	ALA	2.1
1	D	439	ASP	2.1
1	D	306	ARG	2.1
1	D	845	CYS	2.1
1	A	698	THR	2.1
1	A	731	GLU	2.1
1	D	520	GLU	2.1
1	D	875	ALA	2.1
1	D	524	GLN	2.1
1	D	842	GLU	2.1
1	D	786	PHE	2.1
1	D	569	GLU	2.0
1	A	703	ARG	2.0
1	D	630	LEU	2.0
1	D	689	ILE	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [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<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	DOC	B	9	18/19	0.96	0.16	21,29,32,32	0
2	DOC	E	9	18/19	0.98	0.12	27,33,37,38	0

## 6.3 Carbohydrates ⓘ

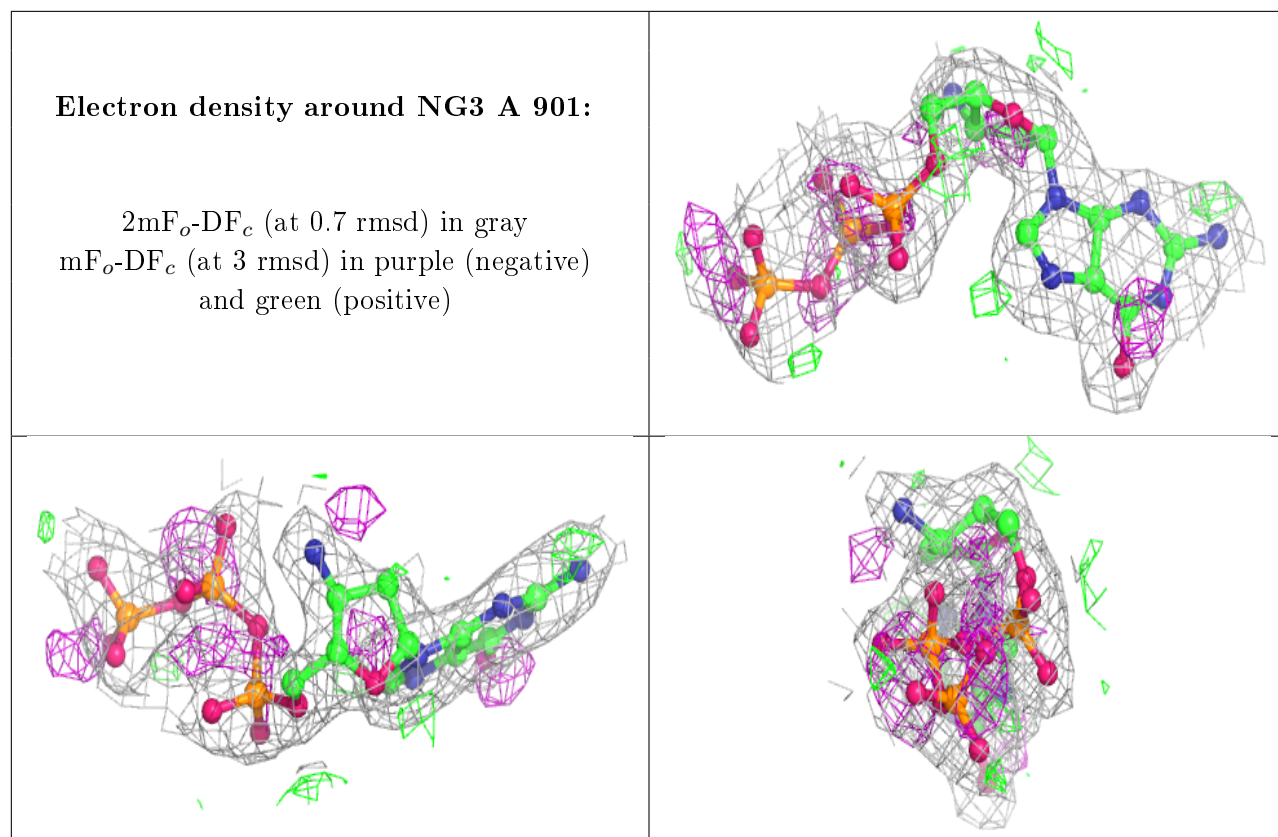
There are no monosaccharides in this entry.

## 6.4 Ligands ⓘ

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
6	SO4	A	905	5/5	0.64	0.41	28,43,45,48	5
6	SO4	A	908	5/5	0.82	0.19	56,56,66,87	0
6	SO4	A	909	5/5	0.87	0.26	65,69,78,80	0
6	SO4	D	901	5/5	0.89	0.38	74,74,80,93	0
6	SO4	A	904	5/5	0.90	0.32	53,53,54,85	0
6	SO4	A	906	5/5	0.90	0.19	60,61,85,87	0
4	NG3	A	901	31/31	0.90	0.16	30,40,62,63	0
6	SO4	A	910	5/5	0.94	0.25	52,55,78,79	0
5	CA	A	902	1/1	0.96	0.05	45,45,45,45	0
6	SO4	A	903	5/5	0.98	0.11	42,47,56,68	0
6	SO4	A	907	5/5	0.98	0.14	48,57,60,65	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



## 6.5 Other polymers [i](#)

There are no such residues in this entry.