



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 26, 2024 – 04:54 AM EDT

PDB ID : 6SA1  
Title : Post catalytic complex of Prim-PolC from Mycobacterium smegmatis with gapped DNA and 3'-dUTP  
Authors : Brissett, N.C.; Doherty, A.J.  
Deposited on : 2019-07-15  
Resolution : 2.01 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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	:	2.37.1
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)	:	2.37.1

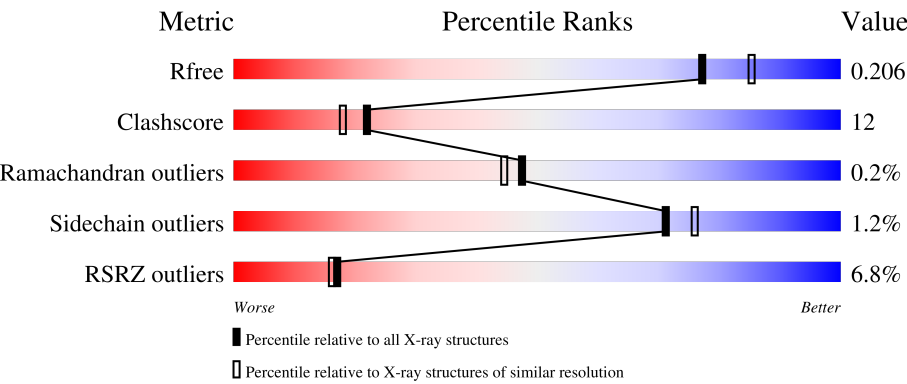
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.01 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.



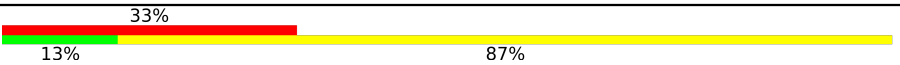

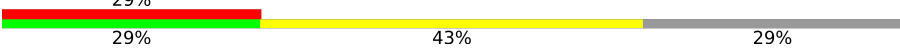
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R <sub>free</sub>	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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  $\geq 3$ , 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	350	<div><div>3%</div><div><div></div><div>83%</div><div>11%</div><div>6%</div></div></div>
1	D	350	<div><div>5%</div><div><div></div><div>83%</div><div>11%</div><div>5%</div></div></div>
2	B	7	<div><div></div><div><div>29%</div><div>71%</div></div></div>
2	E	7	<div><div>29%</div><div><div></div><div>57%</div><div>14%</div></div></div>
3	C	15	<div><div>27%</div><div><div></div><div>87%</div><div>13%</div></div></div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	15	
4	G	7	
4	H	7	

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	U3H	A	403	X	-	-	-

## 2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 6832 atoms, of which 77 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 LigD, polymerase domain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	328	Total	C	N	O	S	0	0	0
			2584	1637	454	486	7			
1	D	331	Total	C	N	O	S	0	0	0
			2614	1657	459	490	8			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	7	Total	C	N	O	P	0	0	0
			148	68	31	42	7			
2	E	7	Total	C	N	O	P	0	0	0
			148	68	31	42	7			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	13	Total	C	N	O	P	0	0	0
			260	124	50	74	12			
3	F	15	Total	C	N	O	P	0	0	0
			301	143	58	86	14			

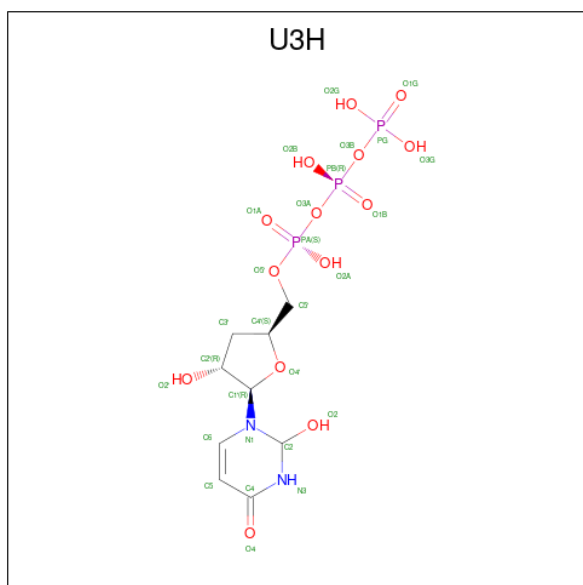
- Molecule 4 is a DNA chain called DNA (5'-D(P\*TP\*GP\*CP\*GP\*(DUI))-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	H	5	Total	C	N	O	P	0	0	0
			102	48	17	32	5			
4	G	7	Total	C	N	O	P	0	0	0
			140	67	25	42	6			

- Molecule 5 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

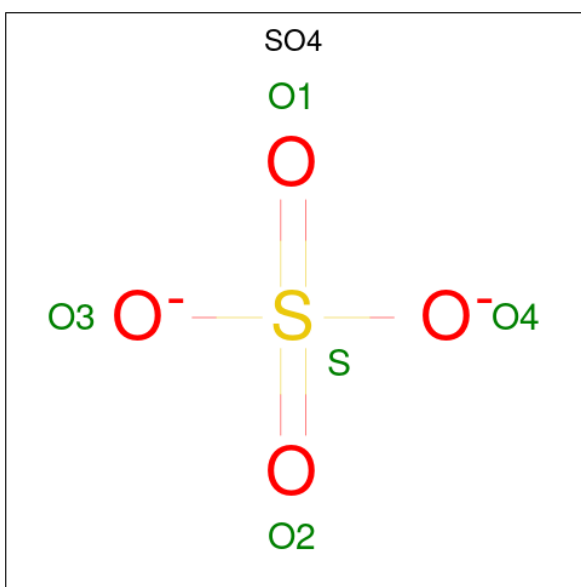
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	2	Total	Mn	0	0
			2	2		
5	D	2	Total	Mn	0	0
			2	2		

- Molecule 6 is 3'-DEOXY-URIDINE 5'-TRIPHOSPHATE (three-letter code: U3H) (formula:  $C_9H_{17}N_2O_{14}P_3$ ).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
6	A	1	Total	C	H	N	O	P	0	0
			39	9	11	2	14	3		

- Molecule 7 is SULFATE ION (three-letter code: SO4) (formula:  $O_4S$ ).

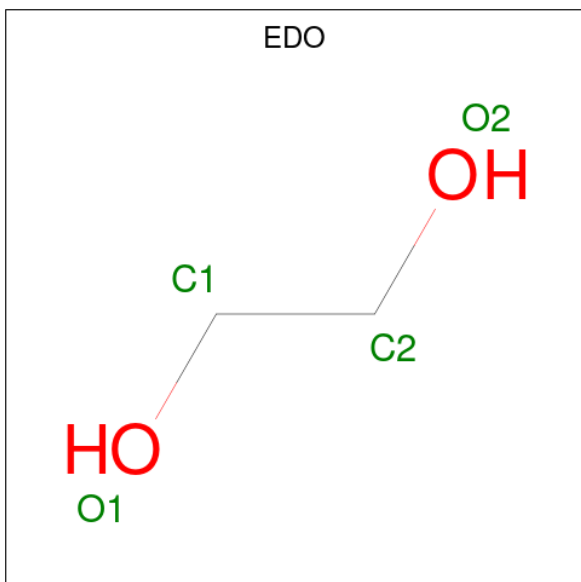


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	O	S	0	0
			5	4	1		
7	D	1	Total	O	S	0	0
			5	4	1		

- Molecule 8 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

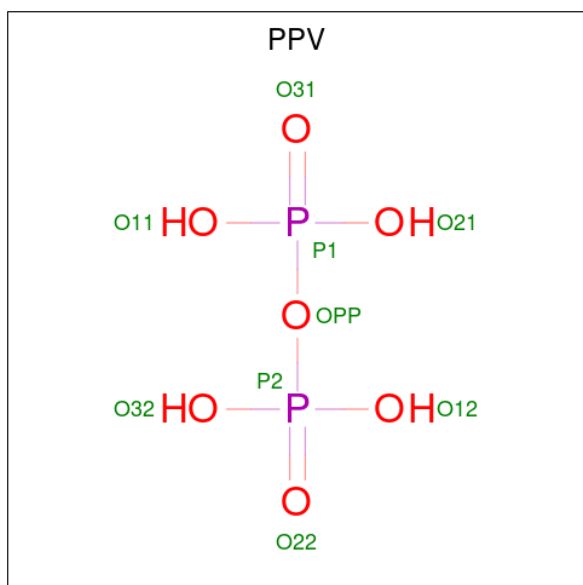
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	1	Total	Mg	0	0
			1	1		

- Molecule 9 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	1	Total C H O 10 2 6 2	0	0
9	A	1	Total C H O 10 2 6 2	0	0
9	A	1	Total C H O 10 2 6 2	0	0
9	A	1	Total C H O 10 2 6 2	0	0
9	A	1	Total C H O 10 2 6 2	0	0
9	D	1	Total C H O 10 2 6 2	0	0
9	D	1	Total C H O 10 2 6 2	0	0
9	D	1	Total C H O 10 2 6 2	0	0
9	D	1	Total C H O 10 2 6 2	0	0
9	D	1	Total C H O 10 2 6 2	0	0
9	D	1	Total C H O 10 2 6 2	0	0

- Molecule 10 is PYROPHOSPHATE (three-letter code: PPV) (formula:  $\text{H}_4\text{O}_7\text{P}_2$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	D	1	Total O P 9 7 2	0	0

- Molecule 11 is water.

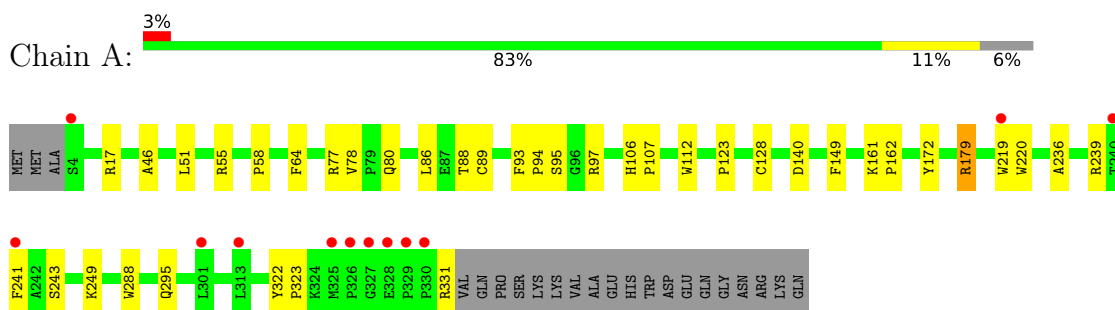
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	194	Total 194	O 194	0	0
11	B	1	Total 1	O 1	0	0
11	D	165	Total 165	O 165	0	0
11	E	2	Total 2	O 2	0	0



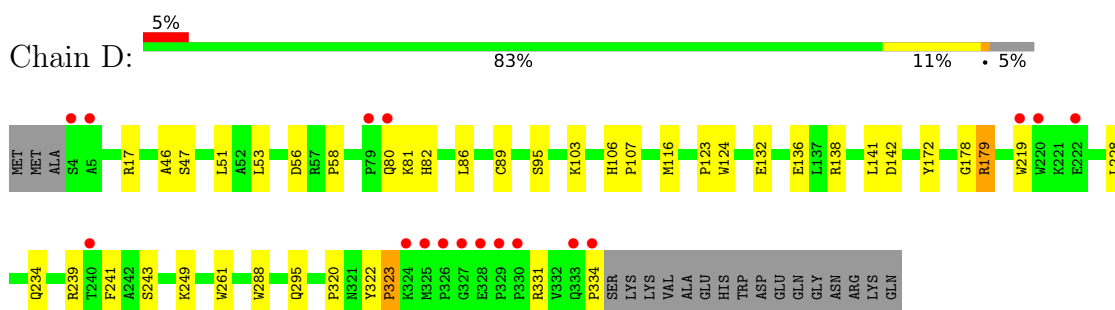
### 3 Residue-property plots [i](#)

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 LigD, polymerase domain



- Molecule 1: DNA polymerase LigD, polymerase domain



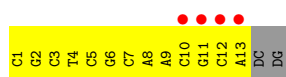
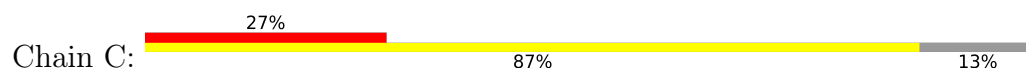
- Molecule 2: DNA (5'-D(P\*GP\*CP\*GP\*AP\*GP\*CP\*G)-3')



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



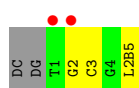
- Molecule 3: DNA (5'-D(\*CP\*GP\*CP\*TP\*CP\*GP\*CP\*AP\*AP\*CP\*GP\*CP\*A)-3')



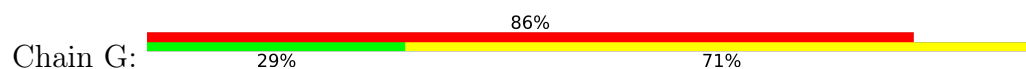
- Molecule 3: DNA (5'-D(\*CP\*GP\*CP\*TP\*CP\*GP\*CP\*AP\*AP\*CP\*GP\*CP\*A)-3')



- Molecule 4: DNA (5'-D(P\*TP\*GP\*CP\*GP\*(DUI))-3')



- Molecule 4: DNA (5'-D(P\*TP\*GP\*CP\*GP\*(DUI))-3')



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 4 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	198.25Å 198.25Å 85.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	88.66 – 2.01 99.12 – 2.01	Depositor EDS
% Data completeness (in resolution range)	99.9 (88.66-2.01) 100.0 (99.12-2.01)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.61 (at 2.02Å)	Xtriage
Refinement program	PHENIX 1.12-2829	Depositor
R, $R_{free}$	0.184 , 0.204 0.185 , 0.206	Depositor DCC
$R_{free}$ test set	5640 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	45.3	Xtriage
Anisotropy	0.337	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 59.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6832	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.78% 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: SO4, EDO, U3H, PPV, MN, MG, L2B

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	0.49	0/2654	0.64	0/3623
1	D	0.45	0/2685	0.63	0/3664
2	B	1.02	1/166 (0.6%)	0.85	0/253
2	E	1.00	1/166 (0.6%)	0.77	0/253
3	C	0.78	0/291	0.88	0/446
3	F	0.70	0/337	0.88	0/517
4	G	0.56	0/135	0.85	0/207
4	H	0.61	0/92	0.98	0/140
All	All	0.55	2/6526 (0.0%)	0.69	0/9103

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	1	DG	OP3-P	-9.67	1.49	1.61
2	B	1	DG	OP3-P	-9.40	1.49	1.61

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	2584	0	2530	32	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	2614	0	2572	33	0
2	B	148	0	78	4	0
2	E	148	0	78	6	0
3	C	260	0	146	33	0
3	F	301	0	168	28	0
4	G	140	0	68	14	0
4	H	102	0	45	2	0
5	A	2	0	0	0	0
5	D	2	0	0	0	0
6	A	28	11	11	1	0
7	A	5	0	0	0	0
7	D	5	0	0	0	0
8	A	1	0	0	0	0
9	A	20	30	30	3	0
9	D	24	36	36	3	0
10	D	9	0	0	2	0
11	A	194	0	0	3	0
11	B	1	0	0	0	0
11	D	165	0	0	9	0
11	E	2	0	0	0	0
All	All	6755	77	5762	142	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 12.

All (142) 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:F:9:DA:H2''	3:F:10:DC:H5''	1.21	1.09
1:A:323:PRO:HG3	3:C:10:DC:H4'	1.11	1.08
3:F:10:DC:H2''	3:F:11:DG:H5'	1.35	1.06
3:C:9:DA:H2''	3:C:10:DC:H5''	1.36	1.02
3:C:6:DG:H2''	3:C:7:DC:H5''	1.39	1.01
3:F:5:DC:H2''	3:F:6:DG:C8	1.96	1.00
3:C:8:DA:H2''	3:C:9:DA:N7	1.81	0.96
1:A:323:PRO:CG	3:C:10:DC:H4'	1.98	0.94
1:D:136:GLU:OE1	11:D:502:HOH:O	1.88	0.91
1:D:295:GLN:OE1	11:D:501:HOH:O	1.86	0.91
1:D:323:PRO:HG3	3:F:10:DC:H4'	1.53	0.90
3:C:6:DG:C2'	3:C:7:DC:H5''	2.08	0.83
1:D:239:ARG:O	11:D:503:HOH:O	1.98	0.81

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
10:D:403:PPV:O11	11:D:504:HOH:O	1.99	0.80
1:D:249:LYS:HG3	1:D:249:LYS:O	1.84	0.78
4:G:2:DG:H2'	4:G:2:DG:OP2	1.83	0.78
1:D:261:TRP:HB2	9:D:405:EDO:H11	1.67	0.75
1:D:138:ARG:NH2	11:D:505:HOH:O	2.10	0.73
3:C:1:DC:H2'	3:C:2:DG:C8	2.23	0.73
3:C:5:DC:H2''	3:C:6:DG:C8	2.25	0.71
2:E:6:DC:H2'	2:E:7:DG:C8	2.26	0.71
3:F:11:DG:H2''	3:F:12:DC:H5'	1.73	0.69
3:C:10:DC:H2''	3:C:11:DG:C8	2.27	0.68
1:A:149:PHE:H	9:A:410:EDO:H12	1.59	0.67
2:B:6:DC:H2''	2:B:7:DG:H5'	1.77	0.67
1:D:138:ARG:NE	11:D:505:HOH:O	2.22	0.67
4:G:3:DT:H2''	4:G:4:DG:H5'	1.78	0.66
1:A:17:ARG:HH12	9:A:407:EDO:H12	1.61	0.65
4:G:3:DT:H2''	4:G:4:DG:C8	2.31	0.64
1:D:80:GLN:CD	1:D:80:GLN:H	2.01	0.64
3:F:13:DA:H2''	3:F:14:DC:C5	2.33	0.64
1:A:77:ARG:NH1	3:C:9:DA:OP1	2.30	0.64
3:F:3:DC:H2''	3:F:4:DT:O5'	1.98	0.63
3:F:9:DA:H2''	3:F:10:DC:C5'	2.13	0.63
3:F:13:DA:H2''	3:F:14:DC:C6	2.34	0.62
3:F:10:DC:H2''	3:F:11:DG:C5'	2.22	0.62
1:A:140:ASP:HB2	1:A:241:PHE:HZ	1.64	0.62
1:D:142:ASP:OD1	1:D:179:ARG:NH2	2.22	0.61
4:G:2:DG:H2''	4:G:3:DT:C7	2.30	0.61
3:F:12:DC:H2''	3:F:13:DA:H5'	1.82	0.61
4:G:2:DG:H2''	4:G:3:DT:H72	1.82	0.61
4:G:3:DT:H2''	4:G:4:DG:H8	1.65	0.61
1:D:219:TRP:CZ2	1:D:320:PRO:HG3	2.36	0.60
3:F:12:DC:H2''	3:F:13:DA:C5'	2.32	0.60
2:E:6:DC:H2''	2:E:7:DG:O5'	2.01	0.60
1:D:53:LEU:HD21	9:D:410:EDO:H22	1.85	0.59
3:C:3:DC:H2'	3:C:4:DT:H72	1.84	0.59
2:B:6:DC:H2''	2:B:7:DG:C5'	2.32	0.59
3:F:5:DC:C2'	3:F:6:DG:C8	2.82	0.57
3:C:9:DA:H2''	3:C:10:DC:C5'	2.25	0.57
1:A:64:PHE:CZ	1:A:249:LYS:HD3	2.39	0.57
3:C:11:DG:H1'	3:C:12:DC:H5''	1.85	0.57
2:E:6:DC:H4'	2:E:7:DG:OP1	2.04	0.56
3:C:12:DC:H6	3:C:12:DC:H5'	1.69	0.55

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:239:ARG:HB2	11:D:510:HOH:O	2.05	0.55
3:F:8:DA:H2''	3:F:9:DA:N7	2.22	0.55
1:A:46:ALA:O	1:A:51:LEU:HB2	2.08	0.54
2:B:6:DC:H2'	2:B:7:DG:C8	2.42	0.54
1:A:80:GLN:H	1:A:80:GLN:CD	2.11	0.53
1:D:132:GLU:OE2	11:D:506:HOH:O	2.19	0.53
3:F:1:DC:H2''	3:F:2:DG:O5'	2.08	0.53
1:D:56:ASP:HA	1:D:103:LYS:HE2	1.90	0.53
4:G:4:DG:H1'	4:G:5:DC:H5''	1.90	0.53
1:D:219:TRP:HZ2	1:D:320:PRO:HG3	1.75	0.52
3:C:8:DA:O3'	3:C:9:DA:C8	2.63	0.52
3:C:11:DG:H2''	3:C:12:DC:H5'	1.91	0.52
3:F:12:DC:H1'	3:F:13:DA:H5''	1.91	0.52
1:D:106:HIS:CG	1:D:107:PRO:HD2	2.46	0.51
1:D:323:PRO:HG3	3:F:10:DC:C4'	2.35	0.50
1:A:95:SER:HB3	3:C:10:DC:P	2.51	0.50
4:H:2:DG:H2''	4:H:3:DC:O5'	2.11	0.50
3:F:1:DC:H6	3:F:1:DC:H5'	1.77	0.50
3:F:13:DA:H5'	3:F:13:DA:C8	2.47	0.50
3:C:6:DG:C3'	3:C:7:DC:H5''	2.42	0.49
3:C:13:DA:H8	3:C:13:DA:O5'	1.96	0.49
1:A:179:ARG:HH11	9:A:410:EDO:H11	1.78	0.49
1:A:236:ALA:HB3	1:A:239:ARG:HG3	1.95	0.49
3:C:3:DC:H2''	3:C:4:DT:O5'	2.12	0.48
3:F:1:DC:H2'	3:F:2:DG:C8	2.47	0.48
3:F:13:DA:C2'	3:F:14:DC:C5	2.96	0.48
4:G:2:DG:C2'	4:G:3:DT:H72	2.44	0.48
1:D:123:PRO:HD3	1:D:243:SER:HB3	1.96	0.48
1:D:172:TYR:HB3	1:D:288:TRP:CE2	2.49	0.48
3:F:13:DA:H5'	3:F:13:DA:H8	1.79	0.47
3:C:9:DA:N1	4:H:5:L2B:N3	2.62	0.47
1:D:17:ARG:HH12	9:D:408:EDO:H12	1.80	0.47
4:G:4:DG:H5'	4:G:4:DG:C8	2.49	0.47
1:A:78:VAL:HG11	1:A:88:THR:HG21	1.95	0.47
1:A:123:PRO:HD3	1:A:243:SER:HB3	1.97	0.47
1:A:140:ASP:HB2	1:A:241:PHE:CZ	2.47	0.46
3:C:1:DC:H2''	3:C:2:DG:C5'	2.45	0.46
1:A:295:GLN:HG3	11:A:508:HOH:O	2.14	0.46
1:A:323:PRO:HG2	3:C:10:DC:O5'	2.16	0.46
3:C:5:DC:C2'	3:C:6:DG:C8	2.96	0.46
4:G:5:DC:H6	4:G:5:DC:H5'	1.81	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:178:GLY:HA3	10:D:403:PPV:O12	2.15	0.46
1:A:106:HIS:CG	1:A:107:PRO:HD2	2.51	0.45
3:C:12:DC:H2''	3:C:13:DA:C8	2.51	0.45
4:G:2:DG:H2''	4:G:3:DT:C5	2.52	0.45
1:A:86:LEU:HD11	1:A:112:TRP:CD2	2.52	0.45
1:D:219:TRP:CZ3	1:D:334:PRO:HB3	2.52	0.45
3:F:11:DG:H1'	3:F:12:DC:H5''	1.98	0.45
3:C:1:DC:H6	3:C:1:DC:H5'	1.82	0.44
2:E:6:DC:C2'	2:E:7:DG:C8	3.00	0.44
3:F:11:DG:H1'	3:F:12:DC:C5'	2.47	0.44
3:F:11:DG:C2'	3:F:12:DC:H5'	2.46	0.44
4:G:1:DC:H3'	4:G:2:DG:C8	2.53	0.44
3:C:11:DG:H1'	3:C:12:DC:C5'	2.47	0.44
1:D:46:ALA:O	1:D:51:LEU:HB2	2.18	0.44
1:A:172:TYR:HB3	1:A:288:TRP:CE2	2.52	0.44
1:D:95:SER:HB3	3:F:10:DC:P	2.58	0.44
1:A:55:ARG:NE	11:A:501:HOH:O	2.44	0.43
2:B:2:DC:H2''	2:B:3:DG:C8	2.53	0.43
1:D:47:SER:HB2	11:D:513:HOH:O	2.19	0.43
1:D:234:GLN:NE2	1:D:241:PHE:HD2	2.16	0.43
1:A:58:PRO:HD3	1:A:128:CYS:SG	2.58	0.43
3:C:10:DC:C2'	3:C:11:DG:C8	2.98	0.43
3:C:12:DC:H5'	3:C:12:DC:C6	2.51	0.43
1:D:141:LEU:HB3	1:D:228:LEU:HD21	2.01	0.43
1:D:116:MET:HA	2:E:1:DG:H4'	2.01	0.42
3:F:10:DC:C2'	3:F:11:DG:H5'	2.26	0.42
4:G:2:DG:H2''	4:G:3:DT:C6	2.53	0.42
1:A:86:LEU:HD11	1:A:112:TRP:CE2	2.53	0.42
1:D:322:TYR:HA	1:D:323:PRO:HD3	1.93	0.42
1:A:161:LYS:HB3	1:A:162:PRO:HD3	2.02	0.42
4:G:1:DC:O2	4:G:1:DC:C2'	2.66	0.42
3:C:12:DC:H2''	3:C:13:DA:N7	2.35	0.42
3:C:1:DC:C2'	3:C:2:DG:O5'	2.68	0.41
3:F:1:DC:C2'	3:F:2:DG:O5'	2.68	0.41
1:D:58:PRO:HG2	1:D:124:TRP:HB2	2.02	0.41
1:A:93:PHE:HB3	1:A:94:PRO:HD2	2.02	0.41
1:D:82:HIS:HB2	1:D:86:LEU:HD12	2.03	0.41
1:A:55:ARG:NH2	11:A:501:HOH:O	2.24	0.41
1:A:239:ARG:HA	6:A:403:U3H:O2	2.20	0.41
1:D:179:ARG:H	1:D:179:ARG:HG3	1.64	0.41
1:A:322:TYR:HA	1:A:323:PRO:HD3	1.95	0.41

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:1:DC:H2''	3:C:2:DG:H5'	2.02	0.41
2:E:2:DC:H2''	2:E:3:DG:C8	2.55	0.41
1:A:95:SER:HG	1:A:97:ARG:HD3	1.86	0.40
1:A:95:SER:N	3:C:10:DC:OP1	2.36	0.40
1:A:219:TRP:HB3	1:A:220:TRP:CE3	2.57	0.40
1:A:219:TRP:HE3	1:A:220:TRP:CE2	2.39	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	326/350 (93%)	323 (99%)	3 (1%)	0	100	100
1	D	329/350 (94%)	325 (99%)	3 (1%)	1 (0%)	41	37
All	All	655/700 (94%)	648 (99%)	6 (1%)	1 (0%)	47	44

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	323	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	
1	A	278/300 (93%)	275 (99%)	3 (1%)	73	78
1	D	283/300 (94%)	279 (99%)	4 (1%)	67	72
All	All	561/600 (94%)	554 (99%)	7 (1%)	71	76

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	89	CYS
1	A	179	ARG
1	A	331	ARG
1	D	81	LYS
1	D	89	CYS
1	D	179	ARG
1	D	331	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 20 ligands modelled in this entry, 5 are monoatomic - leaving 15 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	$\# Z  > 2$
9	EDO	D	407	-	3,3,3	0.43	0	2,2,2	0.30	0
9	EDO	A	407	-	3,3,3	0.40	0	2,2,2	0.41	0
7	SO4	A	404	-	4,4,4	0.16	0	6,6,6	0.08	0
6	U3H	A	403	5	23,29,29	2.32	5 (21%)	27,45,45	1.87	7 (25%)
9	EDO	D	405	-	3,3,3	0.68	0	2,2,2	0.33	0
9	EDO	D	406	-	3,3,3	0.34	0	2,2,2	0.63	0
9	EDO	A	408	-	3,3,3	0.50	0	2,2,2	0.36	0
7	SO4	D	404	-	4,4,4	0.14	0	6,6,6	0.18	0
9	EDO	A	406	-	3,3,3	0.48	0	2,2,2	0.35	0
9	EDO	D	410	-	3,3,3	0.51	0	2,2,2	0.36	0
10	PPV	D	403	5	6,8,8	0.74	0	13,13,13	1.22	2 (15%)
9	EDO	D	409	-	3,3,3	0.46	0	2,2,2	0.51	0
9	EDO	A	410	-	3,3,3	0.59	0	2,2,2	0.20	0
9	EDO	A	409	-	3,3,3	0.44	0	2,2,2	0.35	0
9	EDO	D	408	-	3,3,3	0.47	0	2,2,2	0.53	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
9	EDO	D	407	-	-	0/1/1/1	-
9	EDO	A	407	-	-	0/1/1/1	-
6	U3H	A	403	5	1/1/12/12	8/22/47/47	0/2/2/2
9	EDO	D	405	-	-	0/1/1/1	-
9	EDO	D	406	-	-	0/1/1/1	-
9	EDO	A	408	-	-	1/1/1/1	-
9	EDO	A	406	-	-	1/1/1/1	-
9	EDO	D	410	-	-	0/1/1/1	-
10	PPV	D	403	5	-	0/6/6/6	-
9	EDO	D	409	-	-	1/1/1/1	-
9	EDO	A	410	-	-	1/1/1/1	-
9	EDO	A	409	-	-	0/1/1/1	-
9	EDO	D	408	-	-	1/1/1/1	-

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	403	U3H	C2-N3	-7.84	1.34	1.46
6	A	403	U3H	O2-C2	-4.30	1.26	1.40
6	A	403	U3H	PG-O1G	3.29	1.61	1.50
6	A	403	U3H	PB-O1B	2.86	1.61	1.50
6	A	403	U3H	PG-O3G	2.19	1.63	1.54

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	403	U3H	C5-C6-N1	-5.18	118.74	124.65
6	A	403	U3H	PB-O3B-PG	-3.78	119.84	132.83
6	A	403	U3H	PB-O3A-PA	-3.57	120.57	132.83
6	A	403	U3H	C5-C4-N3	3.43	120.39	116.05
10	D	403	PPV	P2-OPP-P1	-3.18	121.92	132.83
6	A	403	U3H	O4-C4-C5	-2.94	120.00	125.16
6	A	403	U3H	O4-C4-N3	-2.35	119.61	122.21
10	D	403	PPV	O21-P1-OPP	2.13	111.79	104.64
6	A	403	U3H	C2'-C1'-N1	2.09	118.11	114.00

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
6	A	403	U3H	C2

All (13) torsion outliers are listed below:

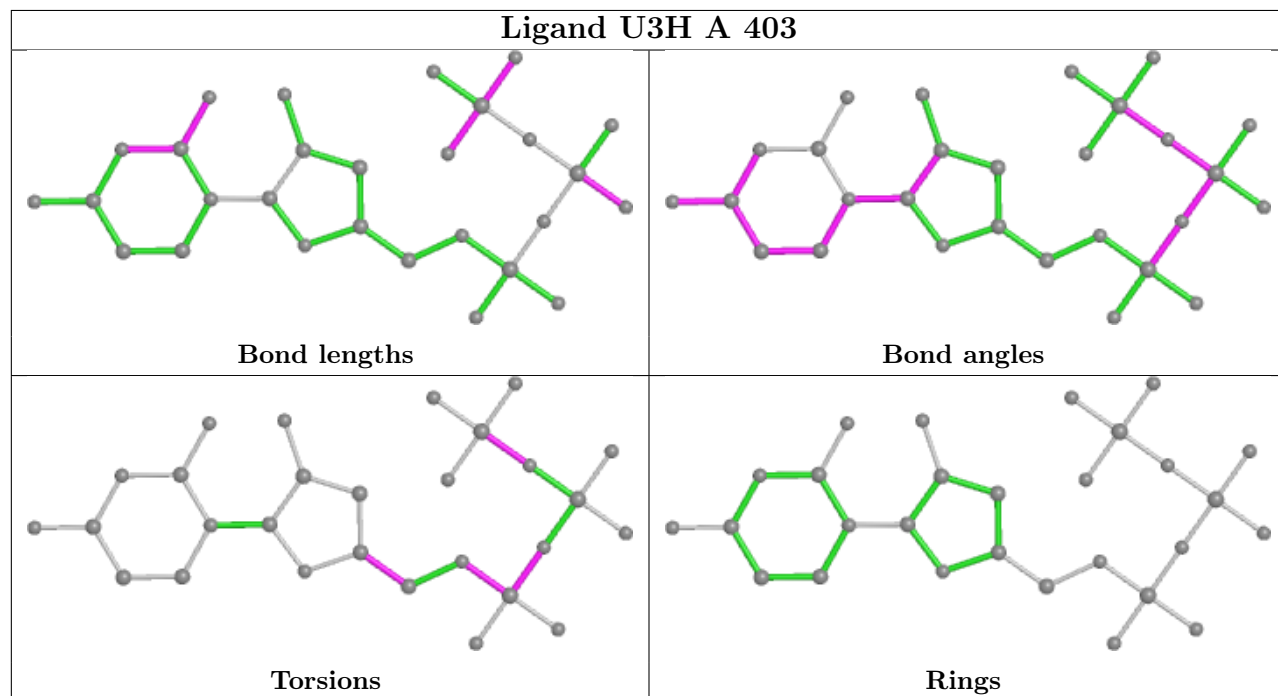
Mol	Chain	Res	Type	Atoms
6	A	403	U3H	C3'-C4'-C5'-O5'
6	A	403	U3H	C5'-O5'-PA-O2A
6	A	403	U3H	C5'-O5'-PA-O3A
6	A	403	U3H	PB-O3B-PG-O3G
6	A	403	U3H	O4'-C4'-C5'-O5'
9	D	408	EDO	O1-C1-C2-O2
6	A	403	U3H	PB-O3A-PA-O5'
6	A	403	U3H	PB-O3B-PG-O2G
6	A	403	U3H	C5'-O5'-PA-O1A
9	A	406	EDO	O1-C1-C2-O2
9	A	410	EDO	O1-C1-C2-O2
9	D	409	EDO	O1-C1-C2-O2
9	A	408	EDO	O1-C1-C2-O2

There are no ring outliers.

7 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	A	407	EDO	1	0
6	A	403	U3H	1	0
9	D	405	EDO	1	0
9	D	410	EDO	1	0
10	D	403	PPV	2	0
9	A	410	EDO	2	0
9	D	408	EDO	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 ⓘ

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	328/350 (93%)	0.66	12 (3%) 41 41	39, 50, 87, 198	0
1	D	331/350 (94%)	0.47	17 (5%) 28 27	40, 55, 98, 177	0
2	B	7/7 (100%)	0.45	0 100 100	63, 89, 123, 130	0
2	E	7/7 (100%)	1.02	2 (28%) 0 0	74, 97, 131, 132	0
3	C	13/15 (86%)	1.56	4 (30%) 0 0	91, 124, 199, 218	0
3	F	15/15 (100%)	2.22	5 (33%) 0 0	117, 134, 219, 252	0
4	G	6/7 (85%)	4.91	6 (100%) 0 0	179, 199, 235, 243	0
4	H	4/7 (57%)	4.38	2 (50%) 0 0	199, 222, 230, 230	0
All	All	711/758 (93%)	0.68	48 (6%) 17 16	39, 54, 133, 252	0

All (48) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	326	PRO	23.8
1	A	327	GLY	11.5
3	F	15	DG	9.2
1	D	325	MET	8.5
4	H	1	DT	8.0
1	A	325	MET	7.6
1	D	326	PRO	7.6
4	G	4	DG	7.4
1	D	327	GLY	7.3
1	A	328	GLU	6.9
4	H	2	DG	6.8
1	D	328	GLU	6.3
4	G	5	DC	5.7
4	G	1	DC	5.5
3	C	13	DA	4.9
1	A	330	PRO	4.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	4	SER	4.6
1	A	329	PRO	4.3
4	G	3	DT	4.2
4	G	2	DG	4.0
3	F	14	DC	4.0
1	D	330	PRO	4.0
1	D	80	GLN	3.8
1	D	324	LYS	3.7
3	C	12	DC	3.6
1	D	329	PRO	3.5
1	D	222	GLU	3.5
1	A	240	THR	3.1
1	D	334	PRO	3.1
1	A	241	PHE	3.1
1	D	333	GLN	3.0
1	D	220	TRP	3.0
1	D	5	ALA	3.0
3	C	11	DG	2.9
3	F	13	DA	2.9
3	F	12	DC	2.9
1	A	219	TRP	2.8
1	D	79	PRO	2.8
1	A	4	SER	2.8
4	G	6	DG	2.6
1	D	240	THR	2.3
1	D	219	TRP	2.2
1	A	313	LEU	2.2
3	C	10	DC	2.2
2	E	7	DG	2.1
2	E	6	DC	2.1
1	A	301	LEU	2.0
3	F	11	DG	2.0

## 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.

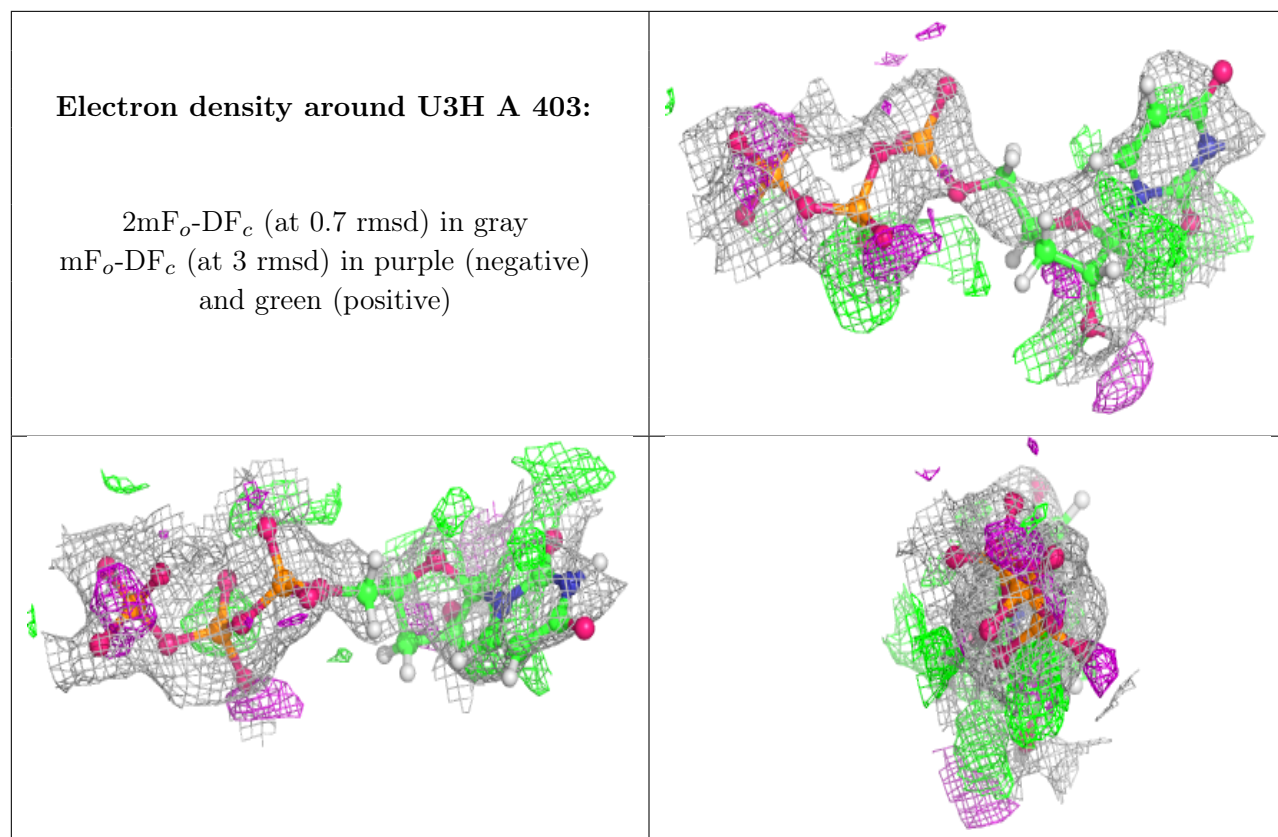


## 6.4 Ligands ⓘ

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
5	MN	D	401	1/1	0.55	0.24	146,146,146,146	0
5	MN	D	402	1/1	0.59	0.09	119,119,119,119	0
9	EDO	A	409	4/4	0.72	0.14	93,111,120,126	0
9	EDO	A	408	4/4	0.75	0.15	71,85,92,92	0
9	EDO	D	405	4/4	0.81	0.23	47,68,78,82	0
8	MG	A	405	1/1	0.83	0.43	84,84,84,84	0
5	MN	A	402	1/1	0.84	0.11	93,93,93,93	0
9	EDO	D	410	4/4	0.84	0.24	63,75,83,83	0
9	EDO	D	407	4/4	0.85	0.12	67,81,85,91	0
6	U3H	A	403	28/28	0.88	0.22	57,97,117,124	39
10	PPV	D	403	9/9	0.88	0.23	98,100,117,122	0
9	EDO	A	406	4/4	0.89	0.20	62,78,94,94	0
9	EDO	D	409	4/4	0.89	0.17	70,93,104,112	0
9	EDO	D	408	4/4	0.90	0.17	77,93,96,98	0
9	EDO	A	410	4/4	0.90	0.26	55,67,78,94	0
9	EDO	D	406	4/4	0.91	0.18	64,77,81,82	0
5	MN	A	401	1/1	0.91	0.12	99,99,99,99	0
9	EDO	A	407	4/4	0.91	0.09	76,91,102,106	0
7	SO4	D	404	5/5	0.95	0.17	64,66,73,73	5
7	SO4	A	404	5/5	0.95	0.15	122,123,127,127	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.