



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 18, 2024 – 05:44 PM JST

PDB ID : 8XK9  
Title : ternary complex of DNA polymerase SFM4-3 recognizing C2 methoxy nucleotide  
Authors : Wen, C.; Liu, H.; Yang, L.; Gong, W.  
Deposited on : 2023-12-22  
Resolution : 2.40 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)  
Xtrriage (Phenix) : 1.13  
EDS : 3.0  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.003 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

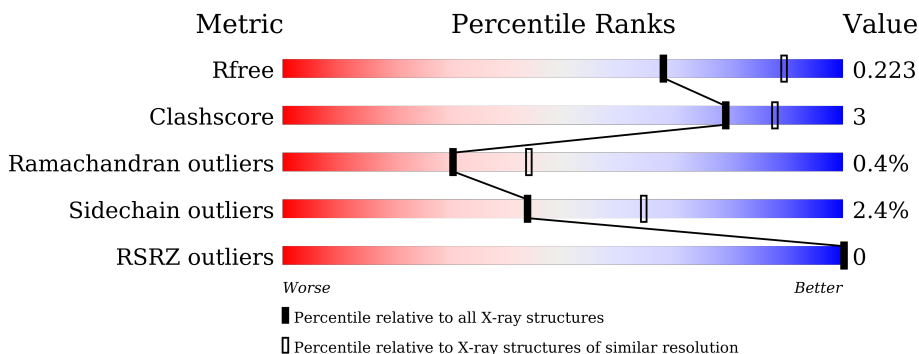
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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))
$R_{free}$	164625	4642 (2.40-2.40)
Clashscore	180529	5218 (2.40-2.40)
Ramachandran outliers	177936	5158 (2.40-2.40)
Sidechain outliers	177891	5159 (2.40-2.40)
RSRZ outliers	164620	4642 (2.40-2.40)

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	B	12	83% 17%
1	E	12	75% 25%
2	C	16	81% 19%
2	F	16	81% 19%
3	A	539	89% 9% .
3	D	539	89% 9% .

## 2 Entry composition i

There are 8 unique types of molecules in this entry. The entry contains 10336 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 DNA (5'-D(\*GP\*AP\*CP\*CP\*AP\*CP\*GP\*GP\*CP\*GP\*CP\*(DOC))-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	B	12	Total 240	C 114	N 48	O 67	P 11	0	0	0
1	E	12	Total 240	C 114	N 48	O 67	P 11	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	C	16	Total 329	C 156	N 64	O 94	P 15	0	0	0
2	F	16	Total 329	C 156	N 64	O 94	P 15	0	0	0

- Molecule 3 is a protein called DNA polymerase I, thermostable.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	537	Total 4289	C 2721	N 778	O 778	S 12	0	4	0
3	D	538	Total 4277	C 2714	N 775	O 776	S 12	0	2	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	294	MET	LEU	conflict	UNP P19821
A	518	ALA	VAL	conflict	UNP P19821
A	583	SER	ASN	conflict	UNP P19821
A	614	GLU	ILE	conflict	UNP P19821
A	615	GLY	GLU	conflict	UNP P19821
A	655	ASN	ASP	conflict	UNP P19821

*Continued on next page...*

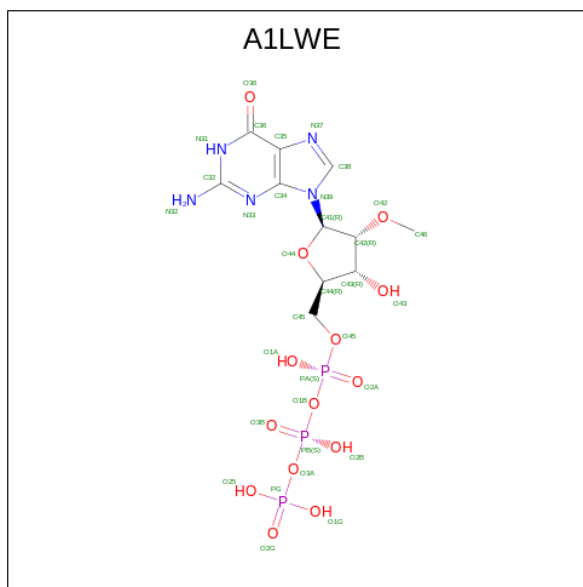
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	681	LYS	GLU	conflict	UNP P19821
A	742	GLN	GLU	conflict	UNP P19821
A	747	ARG	MET	conflict	UNP P19821
D	294	MET	LEU	conflict	UNP P19821
D	518	ALA	VAL	conflict	UNP P19821
D	583	SER	ASN	conflict	UNP P19821
D	614	GLU	ILE	conflict	UNP P19821
D	615	GLY	GLU	conflict	UNP P19821
D	655	ASN	ASP	conflict	UNP P19821
D	681	LYS	GLU	conflict	UNP P19821
D	742	GLN	GLU	conflict	UNP P19821
D	747	ARG	MET	conflict	UNP P19821

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

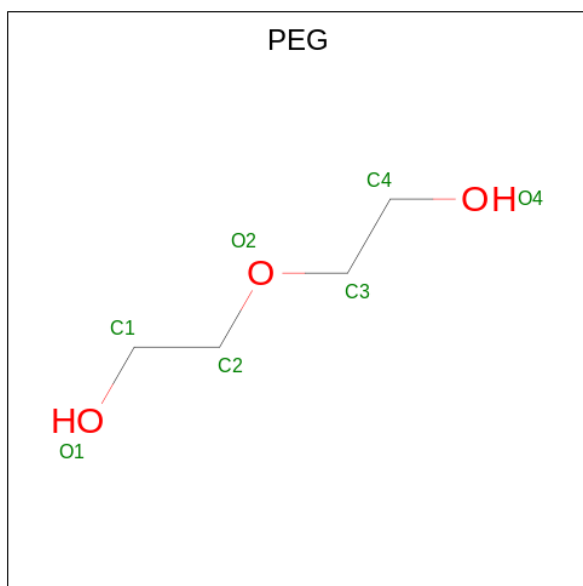
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Mg 1 1	0	0
4	A	2	Total Mg 2 2	0	0
4	E	1	Total Mg 1 1	0	0
4	D	2	Total Mg 2 2	0	0

- Molecule 5 is [(2 {R},3 {R},4 {R},5 {R})-5-(2-azanyl-6-oxidanylidene-1 {H}-purin-9-yl)-4-methoxy-3-oxidanyl-oxolan-2-yl]methoxy-oxidanyl-phosphoryl phosphono hydrogen phosphate (three-letter code: A1LWE) (formula: C<sub>11</sub>H<sub>18</sub>N<sub>5</sub>O<sub>14</sub>P<sub>3</sub>) (labeled as "Ligand of Interest" by depositor).



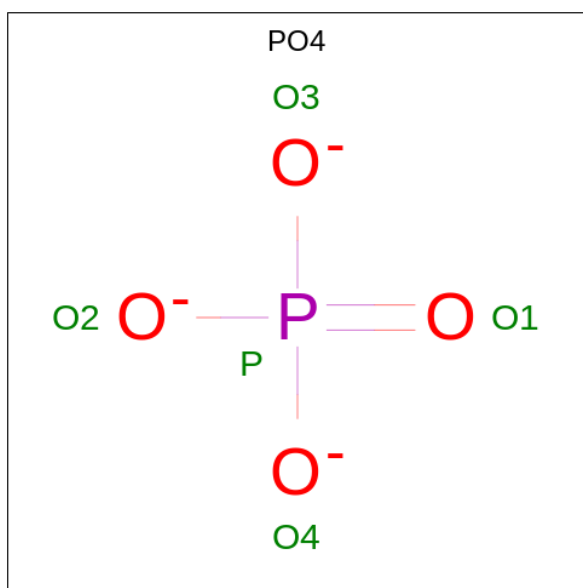
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
			Total	C	N	O			P
5	A	1	33	11	5	14	3	0	0
5	D	1	33	11	5	14	3	0	0

- Molecule 6 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	C O		
6	A	1	6	4 2	0	0
6	D	1	7	4 3	0	0

- Molecule 7 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



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


- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	B	22	Total	O	0	0
			22	22		
8	C	33	Total	O	0	0
			33	33		
8	A	221	Total	O	0	0
			221	221		
8	E	22	Total	O	0	0
			22	22		
8	F	30	Total	O	0	0
			30	30		
8	D	209	Total	O	0	0
			209	209		

### 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 (5'-D(\*GP\*AP\*CP\*CP\*AP\*CP\*GP\*GP\*CP\*GP\*CP\*(DOC))-3')

Chain B:  83% 17%




- Molecule 1: DNA (5'-D(\*GP\*AP\*CP\*CP\*AP\*CP\*GP\*GP\*CP\*GP\*CP\*(DOC))-3')

Chain E:  75% 25%




- Molecule 2: DNA (5'-D(\*AP\*AP\*CP\*GP\*GP\*CP\*GP\*CP\*CP\*GP\*TP\*GP\*GP\*TP\*C)-R(P\*(OMG))-3')

Chain C:  81% 19%




- Molecule 2: DNA (5'-D(\*AP\*AP\*CP\*GP\*GP\*CP\*GP\*CP\*CP\*GP\*TP\*GP\*GP\*TP\*C)-R(P\*(OMG))-3')

Chain F:  81% 19%



- Molecule 3: DNA polymerase I, thermostable

Chain A:  89% 9%





- Molecule 3: DNA polymerase I, thermostable

Chain D: 89% 9%





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	105.99Å 105.99Å 191.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.10 – 2.40 38.10 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.8 (38.10-2.40) 99.9 (38.10-2.40)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.69 (at 2.39Å)	Xtrriage
Refinement program	REFMAC 5.8.0352	Depositor
R, $R_{free}$	0.177 , 0.218 0.183 , 0.223	Depositor DCC
$R_{free}$ test set	4138 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	41.7	Xtrriage
Anisotropy	0.038	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 17.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.477 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	10336	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.44% 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: OMG, A1LWE, PO4, DOC, MG, PEG

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	0.78	0/249	1.31	1/382 (0.3%)
1	E	0.77	0/249	1.31	1/382 (0.3%)
2	C	0.78	0/342	1.33	2/527 (0.4%)
2	F	0.77	0/342	1.31	2/527 (0.4%)
3	A	0.38	0/4382	0.69	3/5934 (0.1%)
3	D	0.38	0/4367	0.68	1/5915 (0.0%)
All	All	0.45	0/9931	0.80	10/13667 (0.1%)

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
3	A	0	17
3	D	0	17
All	All	0	34

There are no bond length outliers.

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	457	ARG	NE-CZ-NH1	6.29	123.44	120.30
2	C	207	DC	P-O3'-C3'	-6.13	112.34	119.70
1	E	108	DG	C1'-O4'-C4'	-6.11	103.99	110.10
2	F	207	DC	P-O3'-C3'	-6.08	112.40	119.70
1	B	108	DG	C1'-O4'-C4'	-6.07	104.03	110.10
3	A	457	ARG	NE-CZ-NH2	-5.91	117.34	120.30
3	D	660	ARG	NE-CZ-NH1	-5.82	117.39	120.30
3	A	660	ARG	NE-CZ-NH2	-5.54	117.53	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	206	DG	P-O3'-C3'	-5.15	113.52	119.70
2	C	206	DG	P-O3'-C3'	-5.08	113.61	119.70

There are no chirality outliers.

All (34) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	A	328	ARG	Sidechain
3	A	392	ARG	Sidechain
3	A	393	ARG	Sidechain
3	A	411	ARG	Sidechain
3	A	419	ARG	Sidechain
3	A	431	ARG	Sidechain
3	A	457	ARG	Sidechain
3	A	487	ARG	Sidechain
3	A	492	ARG	Sidechain
3	A	523	ARG	Sidechain
3	A	556	ARG	Sidechain
3	A	573	ARG	Sidechain
3	A	587	ARG	Sidechain
3	A	636	ARG	Sidechain
3	A	651	ARG	Sidechain
3	A	677	ARG	Sidechain
3	A	716	ARG	Sidechain
3	D	328	ARG	Sidechain
3	D	393	ARG	Sidechain
3	D	411	ARG	Sidechain
3	D	431	ARG	Sidechain
3	D	457	ARG	Sidechain
3	D	487	ARG	Sidechain
3	D	492	ARG	Sidechain
3	D	523	ARG	Sidechain
3	D	556	ARG	Sidechain
3	D	573	ARG	Sidechain
3	D	596	ARG	Sidechain
3	D	651	ARG	Sidechain
3	D	660	ARG	Sidechain
3	D	677	ARG	Sidechain
3	D	716	ARG	Sidechain
3	D	717	ARG	Sidechain
3	D	771	ARG	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	240	0	134	1	0
1	E	240	0	134	1	0
2	C	329	0	179	0	0
2	F	329	0	179	0	0
3	A	4289	0	4331	25	0
3	D	4277	0	4315	24	0
4	A	2	0	0	0	0
4	B	1	0	0	0	0
4	D	2	0	0	0	0
4	E	1	0	0	0	0
5	A	33	0	0	0	0
5	D	33	0	0	1	0
6	A	6	0	7	1	0
6	D	7	0	10	0	0
7	A	5	0	0	0	0
7	D	5	0	0	0	0
8	A	221	0	0	1	0
8	B	22	0	0	0	0
8	C	33	0	0	0	0
8	D	209	0	0	1	0
8	E	22	0	0	0	0
8	F	30	0	0	0	0
All	All	10336	0	9289	50	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 (50) 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:694:GLU:O	3:D:698[B]:GLN:HG2	1.78	0.83
3:A:611:TYR:HB3	3:A:614:GLU:HB2	1.68	0.75
3:A:636:ARG:HD2	3:A:641:GLU:OE2	1.94	0.67
3:A:694:GLU:OE2	3:A:704[B]:ARG:NH1	2.28	0.62
3:D:297:ALA:HB1	3:D:298:PRO:CD	2.37	0.54

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:538:LEU:HD23	3:D:590:LEU:HD22	1.89	0.54
3:A:297:ALA:HB1	3:A:298:PRO:CD	2.38	0.54
3:D:660:ARG:O	3:D:660:ARG:HD3	2.08	0.53
3:A:611:TYR:CB	3:A:614:GLU:HB2	2.39	0.51
3:A:771:ARG:HB3	3:A:802:LEU:HD21	1.92	0.51
3:D:297:ALA:HB1	3:D:298:PRO:HD2	1.94	0.49
3:A:297:ALA:HB1	3:A:298:PRO:HD2	1.95	0.49
3:A:549:LEU:N	3:A:550:PRO:HD2	2.28	0.48
3:A:547:ASP:HB2	3:A:548:PRO:HD3	1.96	0.48
3:D:525:ALA:HB3	3:D:526:HIS:CD2	2.49	0.48
3:A:299:TRP:CE3	3:A:335:ALA:HB2	2.49	0.47
3:D:794:GLU:H	3:D:794:GLU:HG3	1.44	0.47
3:D:354:LYS:NZ	3:D:564:PHE:O	2.48	0.47
3:D:299:TRP:CG	3:D:300:PRO:HA	2.50	0.46
3:D:549:LEU:N	3:D:550:PRO:HD2	2.29	0.46
3:D:547:ASP:HB2	3:D:548:PRO:HD3	1.96	0.46
3:D:299:TRP:CE3	3:D:335:ALA:HB2	2.50	0.46
3:A:354:LYS:NZ	3:A:564:PHE:O	2.48	0.46
3:A:299:TRP:CG	3:A:300:PRO:HA	2.51	0.46
3:A:831:LYS:HE3	8:A:1186:HOH:O	2.14	0.46
3:D:687:GLU:HA	3:D:690[A]:GLN:HE21	1.80	0.46
3:A:636:ARG:NH1	3:A:641:GLU:OE1	2.49	0.46
3:A:492:ARG:O	3:A:496:ASP:HB2	2.16	0.45
3:D:492:ARG:O	3:D:496:ASP:HB2	2.15	0.45
3:D:800:ALA:HB1	3:D:821:VAL:HG11	1.99	0.45
3:A:687:GLU:HA	3:A:690[A]:GLN:HE21	1.81	0.45
3:A:679:SER:HB2	3:A:684:ILE:O	2.17	0.45
3:D:679:SER:HB2	3:D:684:ILE:O	2.17	0.44
3:A:800:ALA:HB1	3:A:821:VAL:HG11	2.00	0.44
3:A:723:LEU:HD11	6:A:904:PEG:H42	2.01	0.42
3:A:334:ARG:HH21	3:A:401:GLU:CD	2.22	0.42
3:D:721:GLU:HA	3:D:726:ARG:O	2.19	0.42
3:D:563:ARG:O	3:D:576:SER:HA	2.20	0.42
3:A:563:ARG:O	3:A:576:SER:HA	2.19	0.42
3:D:334:ARG:HH21	3:D:401:GLU:CD	2.22	0.42
3:D:663:LYS:NZ	5:D:903:A1LWE:O1G	2.53	0.42
1:E:111:DC:H2'	1:E:112:DOC:H6	2.01	0.41
3:A:721:GLU:HA	3:A:726:ARG:O	2.19	0.41
1:B:111:DC:H5''	3:A:586:VAL:HG12	2.02	0.41
3:D:520:GLU:OE1	3:D:523:ARG:HD2	2.20	0.41
3:D:388:GLU:HG3	3:D:398:TRP:CD1	2.56	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:811:TYR:CG	3:D:812:PRO:HD2	2.56	0.41
3:D:737:VAL:HG23	8:D:1183:HOH:O	2.21	0.40
3:A:388:GLU:HG3	3:A:398:TRP:CD1	2.57	0.40
3:A:520:GLU:OE1	3:A:523:ARG:HD2	2.21	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	
3	A	539/539 (100%)	522 (97%)	15 (3%)	2 (0%)	30	44
3	D	538/539 (100%)	521 (97%)	15 (3%)	2 (0%)	30	44
All	All	1077/1078 (100%)	1043 (97%)	30 (3%)	4 (0%)	30	44

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	586	VAL
3	D	784	HIS
3	A	784	HIS
3	A	586	VAL

### 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	
3	A	439/439 (100%)	427 (97%)	12 (3%)	40	60
3	D	437/439 (100%)	427 (98%)	10 (2%)	45	66
All	All	876/878 (100%)	854 (98%)	22 (2%)	44	63

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	328	ARG
3	A	363	GLU
3	A	393	ARG
3	A	431	ARG
3	A	466	GLU
3	A	487	ARG
3	A	614	GLU
3	A	651	ARG
3	A	695	ARG
3	A	704[A]	ARG
3	A	704[B]	ARG
3	A	771	ARG
3	D	328	ARG
3	D	393	ARG
3	D	431	ARG
3	D	466	GLU
3	D	487	ARG
3	D	556	ARG
3	D	596	ARG
3	D	652	GLU
3	D	660	ARG
3	D	794	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	333	HIS
3	A	666	ASN
3	A	750	ASN
3	A	754	GLN
3	D	333	HIS
3	D	526	HIS
3	D	666	ASN
3	D	750	ASN

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type
3	D	754	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

4 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	OMG	F	217	2	18,26,27	0.97	1 (5%)	19,38,41	0.94	1 (5%)
2	OMG	C	217	2	18,26,27	0.97	1 (5%)	19,38,41	0.93	1 (5%)
1	DOC	B	112	1,2	16,19,20	0.44	0	20,26,29	0.53	0
1	DOC	E	112	1,2	16,19,20	0.44	0	20,26,29	0.51	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
2	OMG	F	217	2	-	1/5/27/28	0/3/3/3
2	OMG	C	217	2	-	1/5/27/28	0/3/3/3
1	DOC	B	112	1,2	-	0/7/18/19	0/2/2/2
1	DOC	E	112	1,2	-	0/7/18/19	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	217	OMG	C5-C6	-2.30	1.42	1.47
2	F	217	OMG	C5-C6	-2.27	1.42	1.47



All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	217	OMG	O6-C6-C5	2.38	129.02	124.37
2	C	217	OMG	O6-C6-C5	2.34	128.95	124.37

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	217	OMG	C3'-C2'-O2'-CM2
2	F	217	OMG	C3'-C2'-O2'-CM2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	E	112	DOC	1	0

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 6 are monoatomic - leaving 6 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
5	A1LWE	D	903	4	26,35,35	0.99	1 (3%)	32,55,55	1.28	4 (12%)
6	PEG	A	904	-	5,5,6	0.44	0	4,4,5	0.23	0
7	PO4	D	905	-	4,4,4	0.68	0	6,6,6	0.44	0
7	PO4	A	905	-	4,4,4	0.70	0	6,6,6	0.45	0
6	PEG	D	904	-	6,6,6	0.27	0	5,5,5	0.16	0
5	A1LWE	A	903	4	26,35,35	1.03	1 (3%)	32,55,55	1.38	7 (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
6	PEG	D	904	-	-	2/4/4/4	-
6	PEG	A	904	-	-	2/3/3/4	-
5	A1LWE	D	903	4	-	4/20/40/40	0/3/3/3
5	A1LWE	A	903	4	-	2/20/40/40	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	903	A1LWE	C35-C36	-3.05	1.41	1.47
5	D	903	A1LWE	C35-C36	-2.77	1.41	1.47

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	903	A1LWE	O42-C42-C41	3.66	116.35	109.09
5	D	903	A1LWE	O42-C42-C41	3.36	115.75	109.09
5	D	903	A1LWE	O43-C43-C42	3.23	120.34	111.17
5	A	903	A1LWE	O43-C43-C42	2.76	119.01	111.17
5	A	903	A1LWE	O1G-PG-O2G	2.69	121.22	110.68
5	D	903	A1LWE	O1G-PG-O2G	2.59	120.81	110.68
5	A	903	A1LWE	O1A-PA-O2A	2.58	124.97	112.24
5	A	903	A1LWE	O36-C36-C35	2.43	129.11	124.37
5	D	903	A1LWE	O36-C36-C35	2.37	129.00	124.37
5	A	903	A1LWE	PB-O1B-PA	-2.27	125.02	132.83
5	A	903	A1LWE	O2B-PB-O3B	2.10	122.60	112.24

There are no chirality outliers.

All (10) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	904	PEG	O2-C3-C4-O4
6	D	904	PEG	O2-C3-C4-O4
6	D	904	PEG	O1-C1-C2-O2
6	A	904	PEG	C1-C2-O2-C3
5	D	903	A1LWE	PB-O3A-PG-O1G
5	A	903	A1LWE	PB-O1B-PA-O2A
5	D	903	A1LWE	PB-O1B-PA-O2A

*Continued on next page...*

*Continued from previous page...*

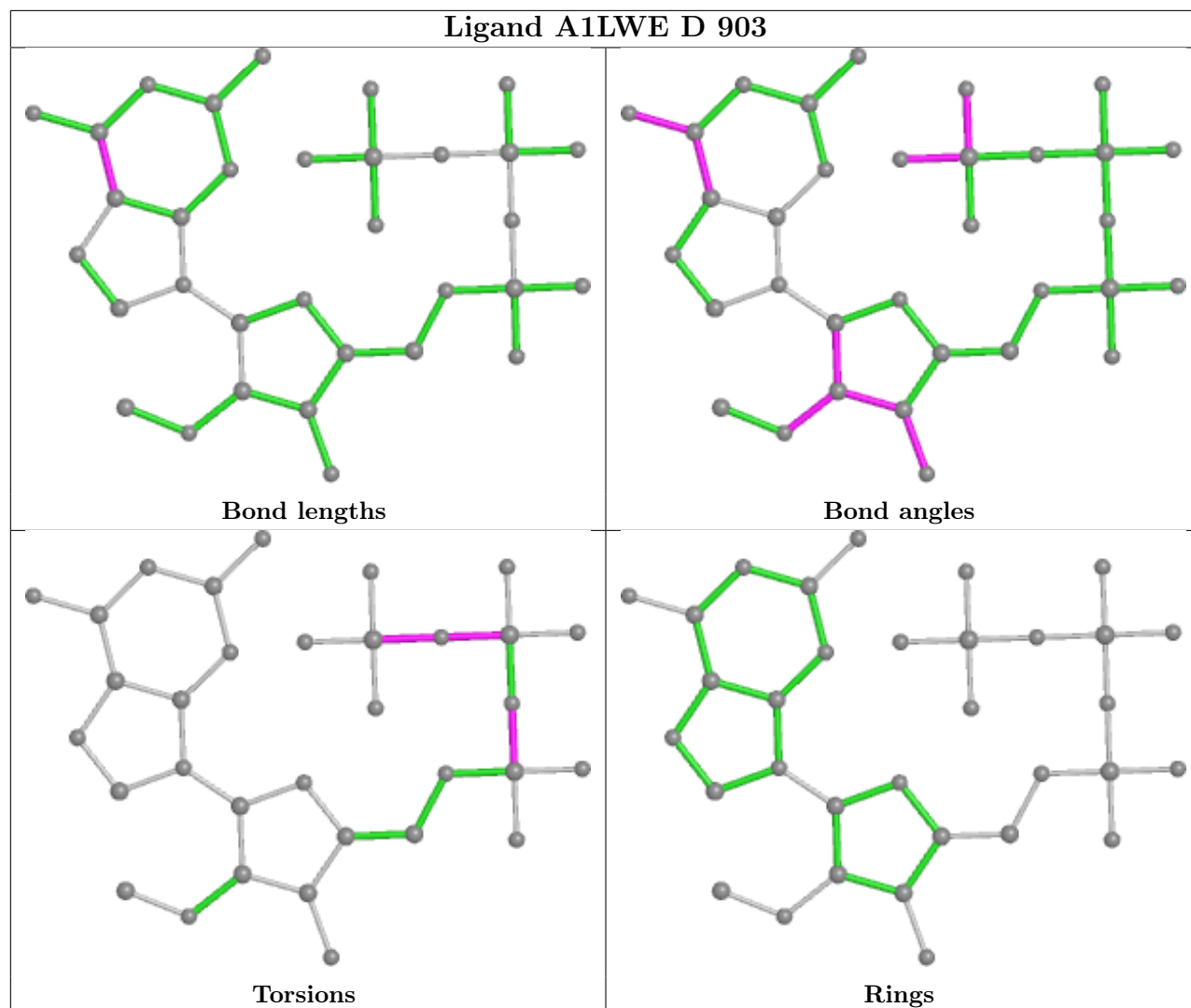
Mol	Chain	Res	Type	Atoms
5	D	903	A1LWE	PG-O3A-PB-O3B
5	A	903	A1LWE	PB-O3A-PG-O1G
5	D	903	A1LWE	PB-O1B-PA-O1A

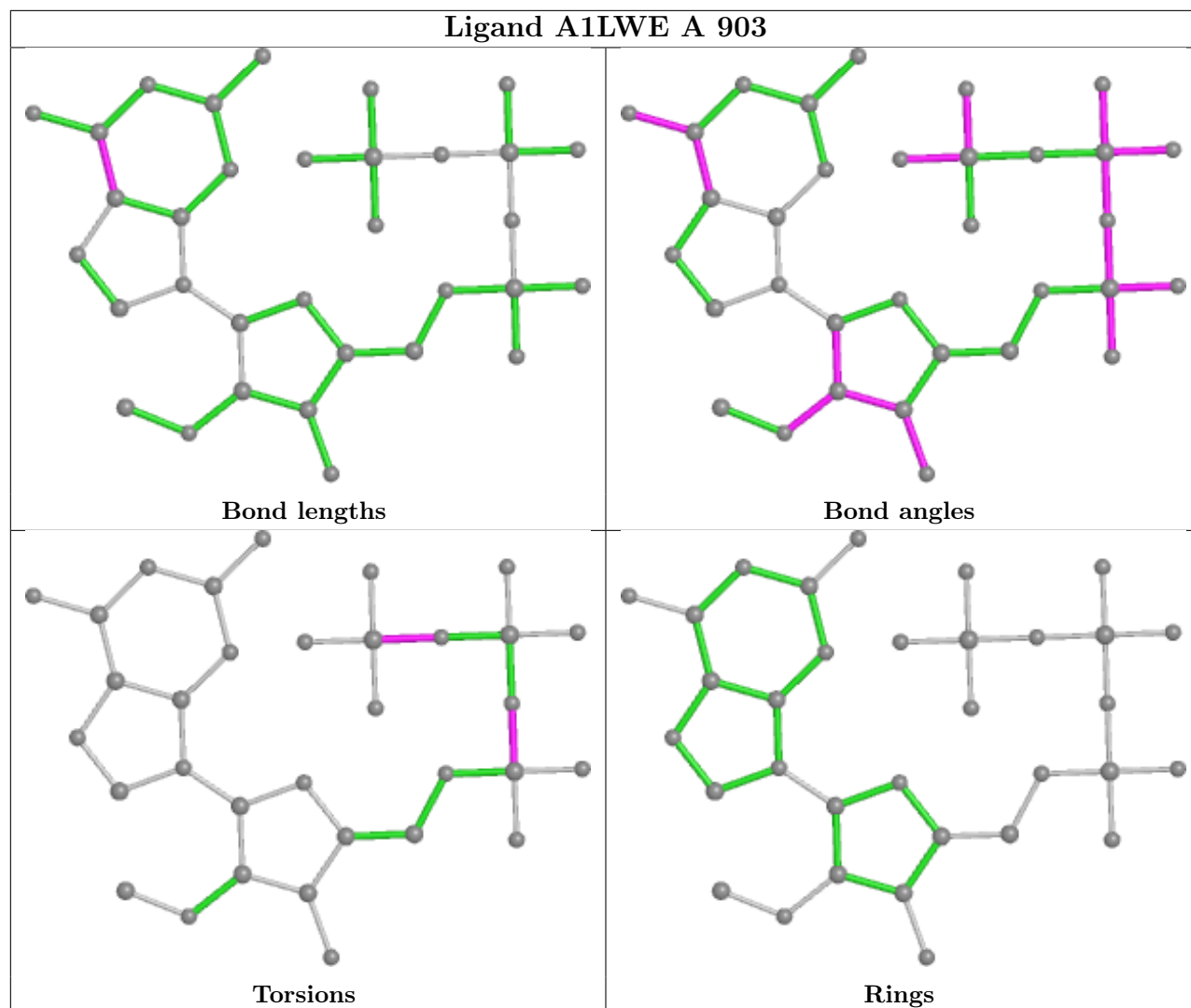
There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	D	903	A1LWE	1	0
6	A	904	PEG	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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	B	11/12 (91%)	-1.97	0 100 100	30, 47, 69, 72	0
1	E	11/12 (91%)	-2.00	0 100 100	30, 47, 68, 71	0
2	C	15/16 (93%)	-1.67	0 100 100	30, 40, 96, 138	0
2	F	15/16 (93%)	-1.70	0 100 100	30, 40, 94, 136	0
3	A	537/539 (99%)	-1.59	0 100 100	23, 44, 78, 107	4 (0%)
3	D	538/539 (99%)	-1.58	0 100 100	23, 44, 79, 117	2 (0%)
All	All	1127/1134 (99%)	-1.59	0 100 100	23, 44, 79, 138	6 (0%)

There are no RSRZ outliers to report.

### 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(Å <sup>2</sup> )	Q<0.9
2	OMG	C	217	24/25	0.94	0.09	93,125,144,150	0
2	OMG	F	217	24/25	0.94	0.10	90,119,142,147	0
1	DOC	B	112	18/19	1.00	0.03	23,28,30,33	0
1	DOC	E	112	18/19	1.00	0.02	23,26,28,31	0

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands

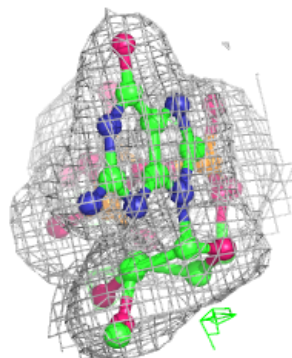
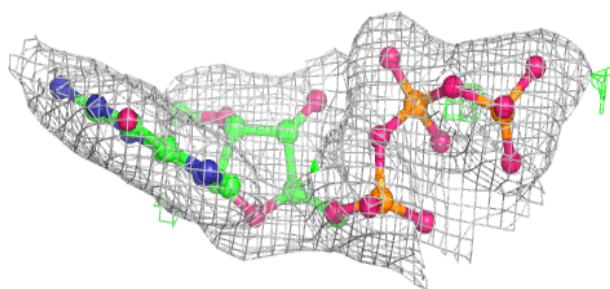
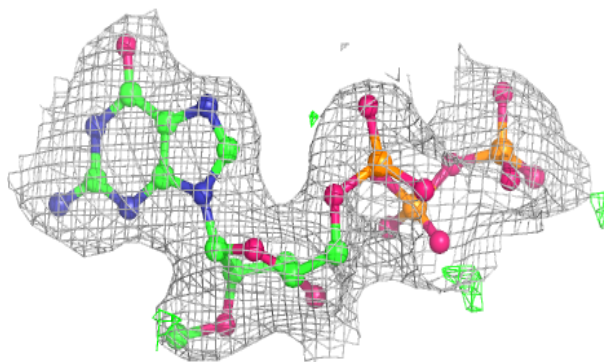
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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
4	MG	B	201	1/1	0.97	0.07	60,60,60,60	0
6	PEG	A	904	6/7	0.97	0.08	65,72,73,77	0
4	MG	D	902	1/1	0.98	0.07	53,53,53,53	0
4	MG	E	201	1/1	0.98	0.07	63,63,63,63	0
6	PEG	D	904	7/7	0.98	0.10	80,91,100,101	0
7	PO4	A	905	5/5	0.98	0.06	89,95,106,109	0
7	PO4	D	905	5/5	0.98	0.05	86,87,96,100	0
4	MG	A	902	1/1	0.99	0.06	46,46,46,46	0
4	MG	D	901	1/1	1.00	0.02	34,34,34,34	0
4	MG	A	901	1/1	1.00	0.02	32,32,32,32	0
5	A1LWE	A	903	33/33	1.00	0.02	25,32,41,48	0
5	A1LWE	D	903	33/33	1.00	0.02	24,32,41,46	0

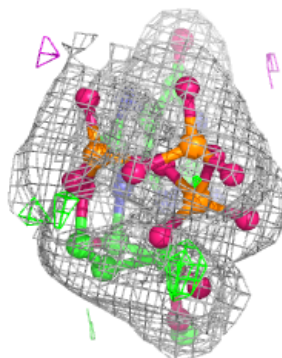
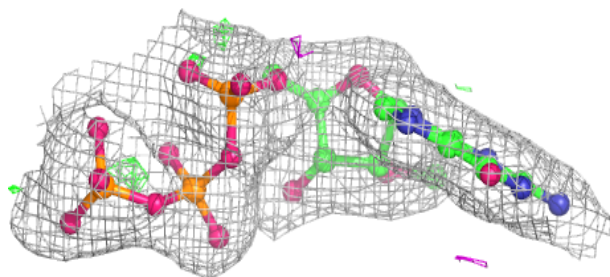
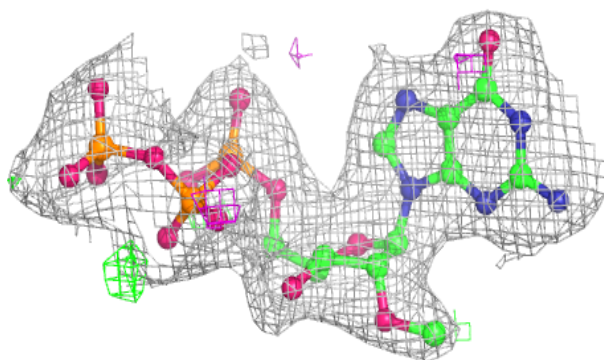
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around A1LWE A 903:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

**Electron density around A1LWE D 903:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





## 6.5 Other polymers [i](#)

There are no such residues in this entry.