Full wwPDB X-ray Structure Validation Report (i)

Jun 7, 2020 – 01:58 am BST

PDB ID	:	6GO5
Title	:	TdT chimera (Loop1 of pol mu) - Ternary complex with 1-nt gapped DNA substrate
Authors	:	Loc'h, J.; Gerodimos, C.A.; Rosario, S.; Lieber, M.R.; Delarue, M.
Deposited on	:	2018-06-01
Resolution	:	2.35 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

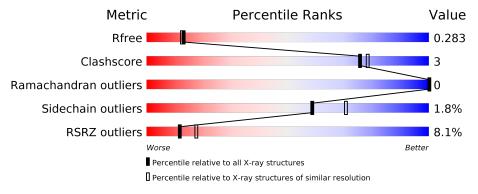|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

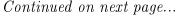The reported resolution of this entry is 2.35 Å.

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	Percentile Ranks	Value
Rfree		0.283
Clashscore		3
Ramachandran outliers		0
Sidechain outliers		1.8%
RSRZ outliers		8.1%

Worse   Better

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

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 >=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 <=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	401	8% / 79% / 6% / 14%
1	B	401	5% / 77% / 6% / 16%
2	H	12	8% / 75% / 25%
2	N	12	75% / 25%
3	F	6	100%
3	L	6	100%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	G	6	67% 33%
4	M	6	33% 67% 33%

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 6542 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 nucleotidylexotransferase,DNA-directed DNA/RNA polymerase mu,DNA nucleotidylexotransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	343	Total 2621	C 1667	N 453	O 485	S 16	0	0	0
1	B	335	Total 2598	C 1656	N 450	O 477	S 15	0	1	0

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	111	MET	-	initiating methionine	UNP P09838
A	112	GLY	-	expression tag	UNP P09838
A	113	SER	-	expression tag	UNP P09838
A	114	SER	-	expression tag	UNP P09838
A	115	HIS	-	expression tag	UNP P09838
A	116	HIS	-	expression tag	UNP P09838
A	117	HIS	-	expression tag	UNP P09838
A	118	HIS	-	expression tag	UNP P09838
A	119	HIS	-	expression tag	UNP P09838
A	120	HIS	-	expression tag	UNP P09838
A	121	SER	-	expression tag	UNP P09838
A	122	SER	-	expression tag	UNP P09838
A	123	GLY	-	expression tag	UNP P09838
A	124	LEU	-	expression tag	UNP P09838
A	125	VAL	-	expression tag	UNP P09838
A	126	PRO	-	expression tag	UNP P09838
A	127	ARG	-	expression tag	UNP P09838
A	128	GLY	-	expression tag	UNP P09838
A	129	SER	-	expression tag	UNP P09838
A	130	HIS	-	expression tag	UNP P09838
A	131	MET	-	expression tag	UNP P09838
A	401	VAL	ALA	conflict	UNP Q9JIW4
B	111	MET	-	initiating methionine	UNP P09838
B	112	GLY	-	expression tag	UNP P09838

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	113	SER	-	expression tag	UNP P09838
B	114	SER	-	expression tag	UNP P09838
B	115	HIS	-	expression tag	UNP P09838
B	116	HIS	-	expression tag	UNP P09838
B	117	HIS	-	expression tag	UNP P09838
B	118	HIS	-	expression tag	UNP P09838
B	119	HIS	-	expression tag	UNP P09838
B	120	HIS	-	expression tag	UNP P09838
B	121	SER	-	expression tag	UNP P09838
B	122	SER	-	expression tag	UNP P09838
B	123	GLY	-	expression tag	UNP P09838
B	124	LEU	-	expression tag	UNP P09838
B	125	VAL	-	expression tag	UNP P09838
B	126	PRO	-	expression tag	UNP P09838
B	127	ARG	-	expression tag	UNP P09838
B	128	GLY	-	expression tag	UNP P09838
B	129	SER	-	expression tag	UNP P09838
B	130	HIS	-	expression tag	UNP P09838
B	131	MET	-	expression tag	UNP P09838
B	401	VAL	ALA	conflict	UNP Q9JIW4

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	12	Total 243	C 116	N 49	O 67	P 11	0	0	0
2	N	12	Total 243	C 116	N 49	O 67	P 11	0	0	0

- Molecule 3 is a DNA chain called DNA (5'-D(*TP*GP*TP*TP*TP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	F	6	Total 121	C 60	N 18	O 38	P 5	0	0	0
3	L	6	Total 121	C 60	N 18	O 38	P 5	0	0	0

- Molecule 4 is a DNA chain called DNA (5'-D(*AP*CP*AP*GP*CP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	G	6	Total	C	N	O	P	0	0	1
			104	48	21	30	5			
4	M	6	Total	C	N	O	P	0	0	1
			104	48	21	30	5			

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	Na	0	0
			1	1		
5	A	1	Total	Na	0	0
			1	1		

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	1	Total	Mg	0	0
			1	1		
6	A	1	Total	Mg	0	0
			1	1		

- Molecule 7 is 2'-deoxy-5'-O-[(S)-hydroxy{[(S)-hydroxy(phosphonooxy)phosphoryl]methyl}phosphoryl]cytidine (three-letter code: XC5) (formula: $C_{10}H_{18}N_3O_{12}P_3$).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	A	1	Total	C	N	O	P	0	0
			28	10	3	12	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
7	B	1	28	10	3	12	3	0	0

- Molecule 8 is water.

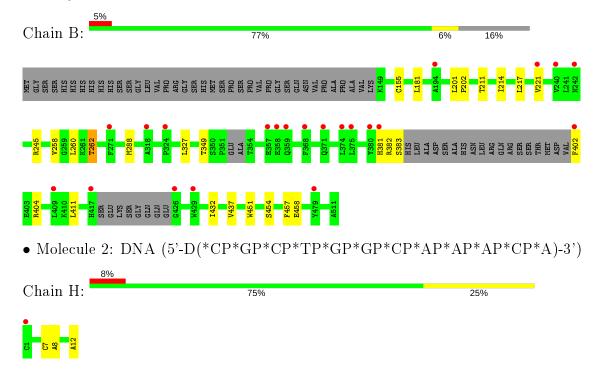| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	137	Total 137	O 137	0	0
8	B	131	Total 131	O 131	0	0
8	H	13	Total 13	O 13	0	0
8	F	3	Total 3	O 3	0	0
8	G	10	Total 10	O 10	0	0
8	L	7	Total 7	O 7	0	0
8	N	18	Total 18	O 18	0	0
8	M	8	Total 8	O 8	0	0

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 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 nucleotidylexotransferase,DNA-directed DNA/RNA polymerase mu,DNA nu cleotidylexotransferase

Chain A:

• Molecule 1: DNA nucleotidylexotransferase,DNA-directed DNA/RNA polymerase mu,DNA nu cleotidylexotransferase

Chain B:

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

Chain H:

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

Chain N:

75%   25%

• Molecule 3: DNA (5'-D(*TP*GP*TP*TP*TP*G)-3')

Chain F:

100%

There are no outlier residues recorded for this chain.

• Molecule 3: DNA (5'-D(*TP*GP*TP*TP*TP*G)-3')

Chain L:

100%

There are no outlier residues recorded for this chain.

• Molecule 4: DNA (5'-D(*AP*CP*AP*GP*CP*G)-3')

Chain G:

67%   33%

• Molecule 4: DNA (5'-D(*AP*CP*AP*GP*CP*G)-3')

33%

Chain M:

67%   33%

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	234.18Å 69.25Å 59.69Å 90.00° 95.20° 90.00°	Depositor
Resolution (Å)	44.86 – 2.35 43.73 – 2.35	Depositor EDS
% Data completeness (in resolution range)	98.3 (44.86-2.35) 98.3 (43.73-2.35)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.21 (at 2.34Å)	Xtriage
Refinement program	BUSTER 2.10.3	Depositor
R, R_{free}	0.218 , 0.265 0.231 , 0.283	Depositor DCC
R_{free} test set	1961 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	65.1	Xtriage
Anisotropy	0.120	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 64.2	EDS
L-test for twinning[2]	$< |L| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6542	wwPDB-VP
Average B, all atoms (Å^2)	72.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.13% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: NA, MG, XC5

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	0.48	0/2675	0.69	1/3615 (0.0%)
1	B	0.47	0/2653	0.68	0/3578
2	H	1.03	0/273	0.90	0/419
2	N	1.12	0/273	0.99	0/419
3	F	1.06	0/134	0.98	0/206
3	L	1.17	0/134	1.08	0/206
4	G	1.07	0/116	0.87	0/178
4	M	0.90	0/116	0.80	0/178
All	All	0.61	0/6374	0.74	1/8799 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	401	VAL	C-N-CA	5.69	135.91	121.70

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	2621	0	2439	13	0
1	B	2598	0	2458	18	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	H	243	0	135	2	0
2	N	243	0	135	2	0
3	F	121	0	72	0	0
3	L	121	0	72	0	0
4	G	104	0	56	1	0
4	M	104	0	56	1	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
7	A	28	0	0	0	0
7	B	28	0	0	0	0
8	A	137	0	0	0	0
8	B	131	0	0	0	0
8	F	3	0	0	0	0
8	G	10	0	0	0	0
8	H	13	0	0	0	0
8	L	7	0	0	0	0
8	M	8	0	0	0	0
8	N	18	0	0	0	0
All	All	6542	0	5423	36	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 (36) 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:381:HIS:HB2	1:A:404:ARG:HB3	1.71	0.71
1:B:217:LEU:HA	1:B:221:VAL:CG2	2.28	0.64
2:N:7:DC:H2'	2:N:8:DA:C8	2.35	0.61
2:H:7:DC:H2'	2:H:8:DA:C8	2.35	0.61
1:B:217:LEU:HA	1:B:221:VAL:HG21	1.82	0.59
1:A:401:VAL:HB	1:A:440:PRO:HG3	1.85	0.59
1:B:404:ARG:HG3	1:B:437:VAL:HG22	1.85	0.59
1:B:402:PHE:CE2	1:B:404:ARG:HB2	2.40	0.56
1:A:383:SER:HB3	1:A:402:PHE:HB3	1.88	0.55
1:B:381:HIS:HB2	1:B:404:ARG:HB3	1.89	0.54
1:B:382:ARG:HG3	1:B:383:SER:H	1.73	0.52
1:B:155:CYS:HA	1:B:457:PHE:CE1	2.45	0.51
1:B:258:VAL:HG13	1:B:262:THR:HB	1.93	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:N:6:DG:H2"	2:N:7:DC:O5'	2.14	0.48
1:A:286:THR:HG22	2:H:12:DA:H5"	1.96	0.48
1:A:404:ARG:HG3	1:A:437:VAL:HG22	1.95	0.48
1:A:221:VAL:HA	1:A:224:ILE:HD12	1.96	0.47
1:A:361:LEU:HB3	1:A:438:MET:SD	2.54	0.47
1:B:454:SER:O	1:B:458:GLU:HG2	2.14	0.47
1:A:362:LEU:HD12	1:A:403:GLU:HB3	1.97	0.47
1:A:323:LEU:HD12	1:A:361:LEU:HD13	1.96	0.46
1:A:181:LEU:HA	1:A:260:LEU:HD13	1.98	0.45
1:B:211:THR:HA	1:B:214:ILE:HD12	1.99	0.45
1:B:402:PHE:HE2	1:B:404:ARG:HB2	1.79	0.45
1:B:181:LEU:HA	1:B:260:LEU:HD13	1.99	0.44
1:B:181:LEU:HD21	1:B:245:ARG:HG2	2.00	0.44
1:B:382:ARG:HG3	1:B:383:SER:N	2.32	0.44
1:B:411:LEU:HD22	1:B:432:ILE:HD12	2.00	0.43
4:M:5:DC:H2"	4:M:6:DG:C8	2.54	0.43
4:G:5:DC:H2"	4:G:6:DG:C8	2.55	0.42
1:B:382:ARG:CG	1:B:383:SER:H	2.32	0.42
1:B:404:ARG:HH22	1:B:451:TRP:HE1	1.68	0.41
1:A:256:PHE:CZ	1:A:406:PHE:HE2	2.39	0.41
1:A:377:TYR:HB3	1:A:408:ILE:HB	2.04	0.41
1:A:155:CYS:HA	1:A:457:PHE:CE2	2.57	0.40
1:B:201:LEU:HA	1:B:202:PRO:HD3	1.99	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	335/401 (84%)	323 (96%)	12 (4%)	0	100	100
1	B	328/401 (82%)	323 (98%)	5 (2%)	0	100	100

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	663/802 (83%)	646 (97%)	17 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	258/350 (74%)	253 (98%)	5 (2%)	57	68
1	B	257/350 (73%)	253 (98%)	4 (2%)	62	75
All	All	515/700 (74%)	506 (98%)	9 (2%)	59	72

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	288	MET
1	A	323	LEU
1	A	327	LEU
1	A	349	THR
1	A	400	ASP
1	B	262	THR
1	B	288	MET
1	B	327	LEU
1	B	349	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	184	ASN

5.3.3   RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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	#$|Z| > 2$
7	XC5	A	603	6	24,29,29	1.30	4 (16%)	30,45,45	1.12	2 (6%)
7	XC5	B	603	6	24,29,29	1.38	3 (12%)	30,45,45	1.16	3 (10%)

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
7	XC5	A	603	6	-	2/16/34/34	0/2/2/2
7	XC5	B	603	6	-	3/16/34/34	0/2/2/2

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	B	603	XC5	PA-O1A	-4.36	1.46	1.56
7	A	603	XC5	PA-O1A	-3.40	1.48	1.56
7	B	603	XC5	C6-N1	2.64	1.39	1.35
7	A	603	XC5	PA-O2A	2.45	1.57	1.51
7	A	603	XC5	C6-N1	2.29	1.38	1.35
7	A	603	XC5	PB-O2B	-2.09	1.51	1.56

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	B	603	XC5	PB-O2B	-2.04	1.51	1.56

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	B	603	XC5	C2-N3-C4	4.50	120.90	116.34
7	A	603	XC5	C2-N3-C4	4.16	120.55	116.34
7	A	603	XC5	O1G-PG-O3B	2.94	114.49	104.64
7	B	603	XC5	O1G-PG-O3B	2.54	113.16	104.64
7	B	603	XC5	O1A-PA-C3A	2.07	115.05	106.58

There are no chirality outliers.

All (5) torsion outliers are listed below:

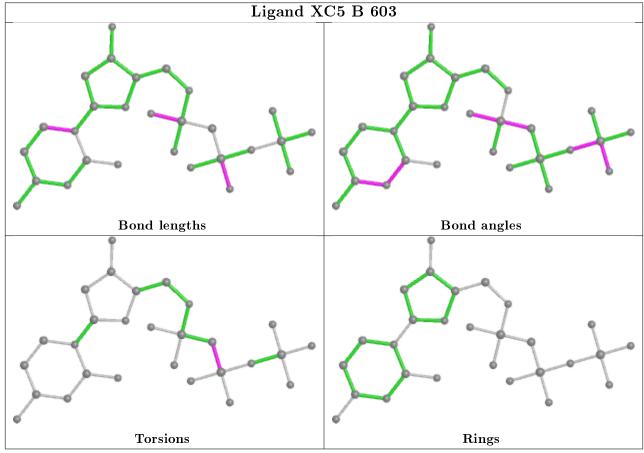| Mol	Chain	Res	Type	Atoms
7	A	603	XC5	PB-O3B-PG-O1G
7	B	603	XC5	PA-C3A-PB-O1B
7	B	603	XC5	PA-C3A-PB-O2B
7	B	603	XC5	PA-C3A-PB-O3B
7	A	603	XC5	PB-O3B-PG-O2G

There are no ring outliers.

No monomer is involved in short contacts.

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 then 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.

Ligand XC5 A 603

Bond lengths

Bond angles

Torsions

Rings

Ligand XC5 B 603

Bond lengths

Bond angles

Torsions

Rings

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, 95th 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(Å²)	Q<0.9
1	A	343/401 (85%)	0.85	34 (9%)	7	11	43, 71, 107, 130	0
1	B	335/401 (83%)	0.77	22 (6%)	18	26	48, 67, 94, 124	0
2	H	12/12 (100%)	0.57	1 (8%)	11	16	62, 85, 118, 129	0
2	N	12/12 (100%)	0.52	0	100	100	59, 77, 102, 115	0
3	F	6/6 (100%)	0.26	0	100	100	65, 69, 80, 91	0
3	L	6/6 (100%)	0.33	0	100	100	61, 65, 77, 94	0
4	G	6/6 (100%)	0.09	0	100	100	73, 78, 85, 87	0
4	M	6/6 (100%)	0.81	2 (33%)	0	0	71, 77, 84, 85	0
All	All	726/850 (85%)	0.79	59 (8%)	12	17	43, 69, 102, 130	0

All (59) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	268	ARG	5.4
1	A	240	VAL	5.3
1	B	324	PRO	5.3
1	A	228	ILE	5.0
1	A	273	THR	4.7
1	A	359	GLN	4.6
1	B	357	GLU	3.8
1	B	380	TYR	3.8
1	B	374	LEU	3.8
1	A	221	VAL	3.7
1	B	271	PHE	3.5
1	A	271	PHE	3.3
1	A	296	TYR	3.3
1	B	368	PHE	3.2
1	B	426	GLY	3.2
2	H	1	DC	3.2

Continued on next page...

Full wwPDB X-ray Structure Validation Report

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	399	MET	3.2
1	B	417	HIS	3.1
1	A	225	ILE	3.1
1	B	381	HIS	3.1
4	M	6	DG	3.0
1	B	402	PHE	3.0
1	A	210	ASP	2.9
1	A	411	LEU	2.8
1	A	352	GLU	2.8
1	A	322	PHE	2.8
1	B	194	ALA	2.8
1	B	242	ASN	2.8
1	A	260	LEU	2.8
1	A	384	HIS	2.7
1	A	211	THR	2.6
1	A	361	LEU	2.6
1	A	409	LEU	2.6
1	A	241	LEU	2.5
1	A	214	ILE	2.4
1	A	203	PHE	2.4
1	A	249	PHE	2.4
1	B	358	GLU	2.4
1	A	223	SER	2.4
1	A	402	PHE	2.4
1	B	371	GLN	2.4
1	B	221	VAL	2.3
1	B	375	LEU	2.3
1	A	367	ASP	2.3
1	A	401	VAL	2.3
4	M	1	DA	2.2
1	B	429	TRP	2.2
1	B	318	ALA	2.2
1	A	428	GLY	2.1
1	A	368	PHE	2.1
1	B	409	LEU	2.1
1	B	359	GLN	2.1
1	A	430	LYS	2.1
1	B	479	TYR	2.1
1	A	266	TRP	2.0
1	A	417	HIS	2.0
1	B	240	VAL	2.0
1	A	406	PHE	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	426	GLY	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 carbohydrates in this entry.

6.4 Ligands (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, 95th 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.

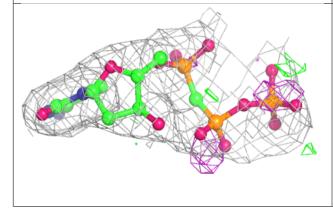| Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\mathring{A}^2$)	Q<0.9
5	NA	A	601	1/1	0.47	0.18	72,72,72,72	0
5	NA	B	602	1/1	0.84	0.09	55,55,55,55	0
7	XC5	A	603	28/28	0.95	0.18	45,52,72,75	0
6	MG	B	601	1/1	0.95	0.09	51,51,51,51	0
7	XC5	B	603	28/28	0.96	0.18	43,50,59,63	0
6	MG	A	602	1/1	0.97	0.19	52,52,52,52	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.
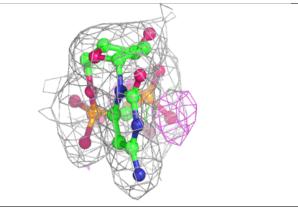
Electron density around XC5 A 603:

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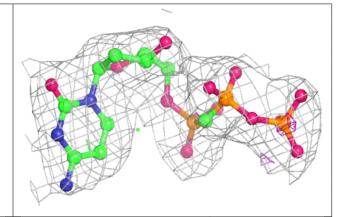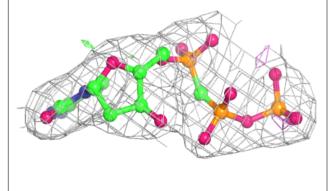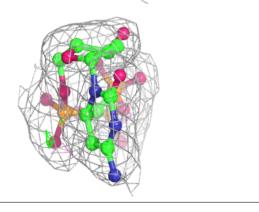mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

Electron density around XC5 B 603:

$2mF_o$-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.