



Full wwPDB X-ray Structure Validation Report ⓘ

Jul 14, 2025 – 08:09 PM JST

PDB ID : 8XXK / pdb_00008xxk
Title : Crystal structure of human 8-oxoguanine glycosylase K249H mutant bound to the reaction intermediate derived from the crystal soaked into the solution at pH 4.0 under 298 K for 3 weeks
Authors : Unno, M.; Koga, M.; Minowa, N.; Komuro, S.; Tanaka, Y.
Deposited on : 2024-01-18
Resolution : 1.70 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

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:

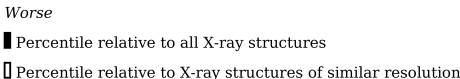
MolProbity	:	4-5-2 with Phenix2.0rc1
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.44

i

X-RAY DIFFRACTION

A.

Metric	Percentile Banks	Value
--------	------------------	-------



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R _{free}	164625	5161 (1.70-1.70)
Clashscore	180529	5671 (1.70-1.70)
Ramachandran outliers	177936	5594 (1.70-1.70)
Sidechain outliers	177891	5594 (1.70-1.70)
RSRZ outliers	164620	5159 (1.70-1.70)

Mol	Chain	Length	Quality of chain
1	C	8	<div><div></div>100%</div>
2	D	16	<div><div></div><div></div>6%</div> <div><div></div>94%</div> <div></div> 6%
3	A	336	<div><div></div>2%</div> <div><div></div>88%</div> <div></div> 7% <div></div> 6%
4	G	7	<div><div></div>100%</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	GOL	D	101	-	X	-	-

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 3590 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(*AP*GP*CP*GP*TP*CP*CP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	8	Total	C	N	O	P	0	0	0
			160	77	31	45	7			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	15	Total	C	N	O	P	0	0	0
			308	146	61	87	14			

- Molecule 3 is a protein called N-glycosylase/DNA lyase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	316	Total	C	N	O	S	0	9	0
			2534	1610	453	459	12			

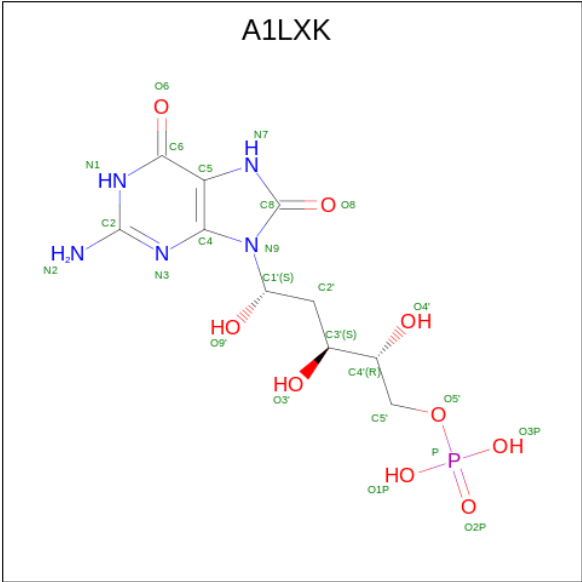
There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	10	GLY	-	expression tag	UNP O15527
A	11	PRO	-	expression tag	UNP O15527
A	249	HIS	LYS	engineered mutation	UNP O15527

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

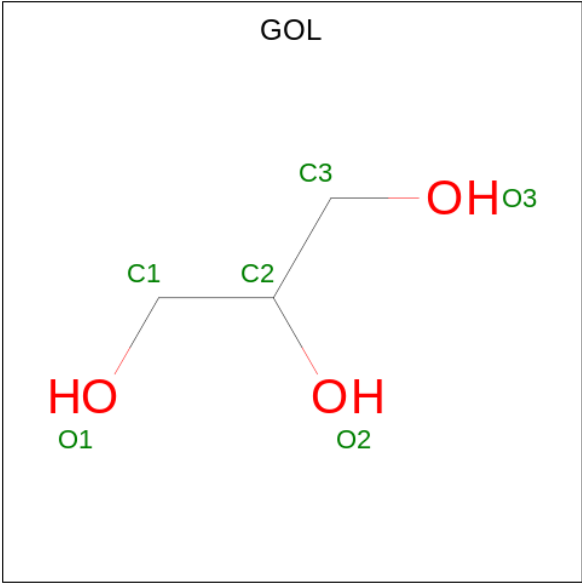
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	G	7	Total	C	N	O	P	0	0	0
			140	67	23	43	7			

- Molecule 5 is [(2 {R},3 {S},5 {S})-5-[2-azanyl-6,8-bis(oxidanylidene)-1,7-dihydropu rin-9-yl]-2,3,5-tris(oxidanyl)pentyl] dihydrogen phosphate (CCD ID: A1LXK) (formula: C₁₀H₁₆N₅O₉P) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	C	1	Total	C	N	O	P	0	0
			24	10	5	8	1		

- Molecule 6 is GLYCEROL (CCD ID: GOL) (formula: C₃H₈O₃).



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

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 7 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	2	Total	Mg	0	0
			2	2		

- Molecule 8 is SODIUM ION (CCD ID: NA) (formula: Na).

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

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	C	10	Total	O	0	0
			10	10		
9	D	52	Total	O	0	0
			52	52		
9	A	323	Total	O	0	3
			326	326		
9	G	9	Total	O	0	0
			9	9		

3 Residue-property plots

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(*AP*GP*CP*GP*TP*CP*CP*A)-3')

Chain C:  100%


There are no outlier residues recorded for this chain.

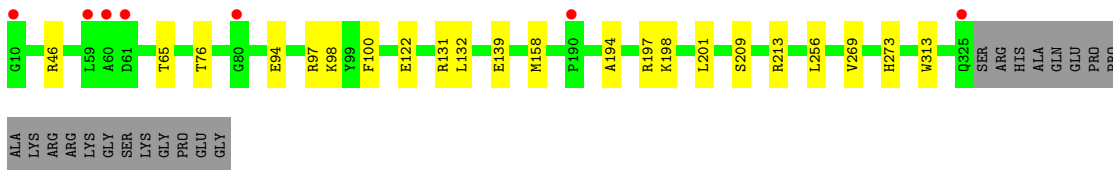
- Molecule 2: DNA (5'-D(*GP*GP*TP*AP*GP*AP*CP*CP*TP*GP*GP*AP*CP*GP*C)-3')

Chain D:  6% 94% 6%



- Molecule 3: N-glycosylase/DNA lyase

Chain A:  2% 88% 7% 6%



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

Chain G:  100%

There are no outlier residues recorded for this chain.

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	61.85Å 69.45Å 88.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.19 – 1.70 46.19 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.3 (46.19-1.70) 99.4 (46.19-1.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.50 (at 1.70Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.163 , 0.213 0.163 , 0.212	Depositor DCC
R_{free} test set	2139 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	14.7	Xtriage
Anisotropy	0.444	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 35.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	3590	wwPDB-VP
Average B, all atoms (Å ²)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.24% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: GOL, A1LXK, MG, NA

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	C	0.60	0/179	0.79	0/274
2	D	0.70	0/346	0.86	0/533
3	A	0.73	0/2633	0.81	0/3587
4	G	0.56	0/155	0.77	0/236
All	All	0.72	0/3313	0.81	0/4630

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	C	160	0	90	0	0
2	D	308	0	169	0	0
3	A	2534	0	2470	13	0
4	G	140	0	80	0	0
5	C	24	0	0	0	0
6	A	18	0	24	0	0
6	D	6	0	8	0	0
7	A	2	0	0	0	0
8	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	A	326	0	0	2	0
9	C	10	0	0	0	0
9	D	52	0	0	0	0
9	G	9	0	0	0	0
All	All	3590	0	2841	13	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 2.

All (13) 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:A:97:ARG:NH1	9:A:502:HOH:O	2.33	0.61
3:A:132[A]:LEU:HD21	3:A:256:LEU:HG	1.85	0.58
3:A:197:ARG:NH2	9:A:505:HOH:O	2.37	0.58
3:A:209:SER:OG	3:A:213:ARG:NH1	2.43	0.52
3:A:158:MET:HE3	3:A:201:LEU:HG	1.92	0.50
3:A:132[B]:LEU:HD12	3:A:313:TRP:CG	2.49	0.47
3:A:100:PHE:O	3:A:131:ARG:HD3	2.15	0.46
3:A:46:ARG:HG3	3:A:139:GLU:OE2	2.16	0.45
3:A:194:ALA:O	3:A:198:LYS:HG3	2.17	0.44
3:A:94:GLU:HG3	3:A:98:LYS:HE2	2.00	0.43
3:A:269:VAL:HG22	3:A:273:HIS:CE1	2.56	0.41
3:A:65:THR:HB	3:A:76:THR:HG22	2.02	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	323/336 (96%)	313 (97%)	10 (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	
3	A	271/280 (97%)	271 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	108	GLN
3	A	172	GLN
3	A	273	HIS
3	A	276	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](#)

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

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 3 are monoatomic - leaving 5 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	GOL	A	402	-	5,5,5	1.64	1 (20%)	5,5,5	1.04	0
5	A1LXK	C	101	1,4	22,25,26	3.17	6 (27%)	23,36,39	3.05	12 (52%)
6	GOL	A	401	-	5,5,5	0.87	0	5,5,5	0.74	0
6	GOL	D	101	-	5,5,5	1.72	2 (40%)	5,5,5	0.83	0
6	GOL	A	403	-	5,5,5	0.85	0	5,5,5	1.48	1 (20%)

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	GOL	A	402	-	-	4/4/4/4	-
5	A1LXK	C	101	1,4	-	2/11/17/18	0/2/2/2
6	GOL	A	401	-	-	2/4/4/4	-
6	GOL	D	101	-	-	4/4/4/4	-
6	GOL	A	403	-	-	2/4/4/4	-

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	101	A1LXK	O8-C8	9.05	1.39	1.23
5	C	101	A1LXK	O6-C6	8.17	1.39	1.23
5	C	101	A1LXK	C2-N2	5.55	1.47	1.34
5	C	101	A1LXK	C8-N9	-3.31	1.34	1.40
5	C	101	A1LXK	C6-N1	-3.16	1.33	1.38
6	A	402	GOL	C1-C2	2.96	1.63	1.51
6	D	101	GOL	C3-C2	2.79	1.63	1.51
5	C	101	A1LXK	C5-C4	2.28	1.40	1.37
6	D	101	GOL	O3-C3	2.25	1.51	1.42

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	101	A1LXK	C5-C4-N3	-8.58	118.94	127.80
5	C	101	A1LXK	N9-C4-N3	5.91	132.57	125.81
5	C	101	A1LXK	C2-N3-C4	5.11	121.41	112.30
5	C	101	A1LXK	N7-C8-N9	4.65	112.10	106.58
6	A	403	GOL	C3-C2-C1	-2.80	100.81	111.70
5	C	101	A1LXK	C2-N1-C6	-2.76	120.07	125.10
5	C	101	A1LXK	O8-C8-N9	-2.63	122.32	125.99
5	C	101	A1LXK	C5-C6-N1	2.57	119.96	112.31
5	C	101	A1LXK	O3'-C3'-C2'	2.35	113.82	109.11
5	C	101	A1LXK	O6-C6-C5	-2.15	122.32	127.24
5	C	101	A1LXK	N2-C2-N1	2.11	121.20	116.71
5	C	101	A1LXK	O4'-C4'-C5'	-2.09	105.21	109.92
5	C	101	A1LXK	C5-N7-C8	-2.06	106.51	109.47

There are no chirality outliers.

All (14) torsion outliers are listed below:

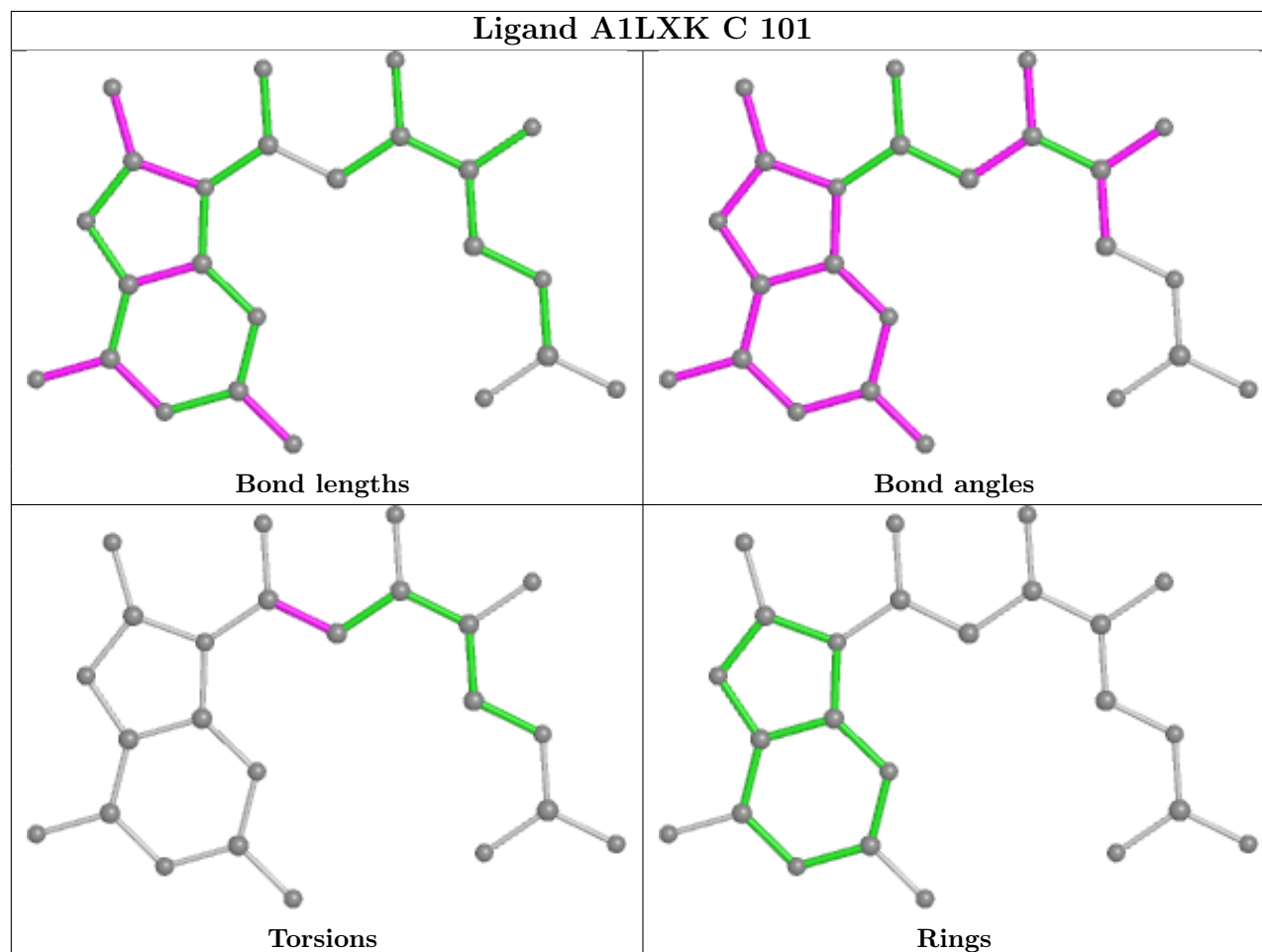
Mol	Chain	Res	Type	Atoms
5	C	101	A1LXK	N9-C1'-C2'-C3'
6	D	101	GOL	O1-C1-C2-C3
6	A	402	GOL	O1-C1-C2-C3
6	A	402	GOL	C1-C2-C3-O3
6	D	101	GOL	O2-C2-C3-O3
6	A	402	GOL	O2-C2-C3-O3
6	A	403	GOL	O2-C2-C3-O3
6	D	101	GOL	C1-C2-C3-O3
6	A	401	GOL	O1-C1-C2-C3
6	A	403	GOL	C1-C2-C3-O3
6	D	101	GOL	O1-C1-C2-O2
6	A	401	GOL	O1-C1-C2-O2
6	A	402	GOL	O1-C1-C2-O2
5	C	101	A1LXK	O9'-C1'-C2'-C3'

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 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, 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(Å ²)	Q<0.9
1	C	8/8 (100%)	0.15	0 100 100	16, 34, 36, 39	0
2	D	15/16 (93%)	-0.02	1 (6%) 25 26	15, 20, 41, 51	0
3	A	316/336 (94%)	-0.35	7 (2%) 62 65	6, 13, 29, 49	9 (2%)
4	G	7/7 (100%)	0.15	0 100 100	13, 22, 60, 63	0
All	All	346/367 (94%)	-0.31	8 (2%) 61 64	6, 14, 36, 63	9 (2%)

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	60	ALA	4.2
3	A	190	PRO	3.6
3	A	80	GLY	3.4
2	D	2	DG	2.8
3	A	325	GLN	2.7
3	A	10	GLY	2.6
3	A	61	ASP	2.5
3	A	59	LEU	2.2

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 oligosaccharides in this entry.

6.4 Ligands

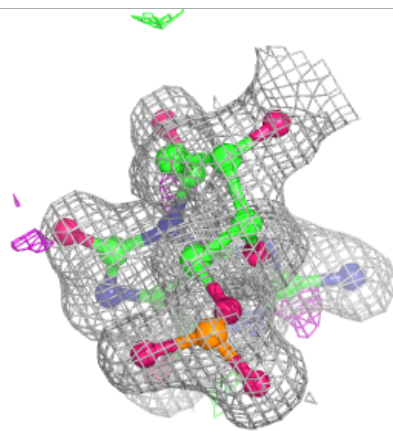
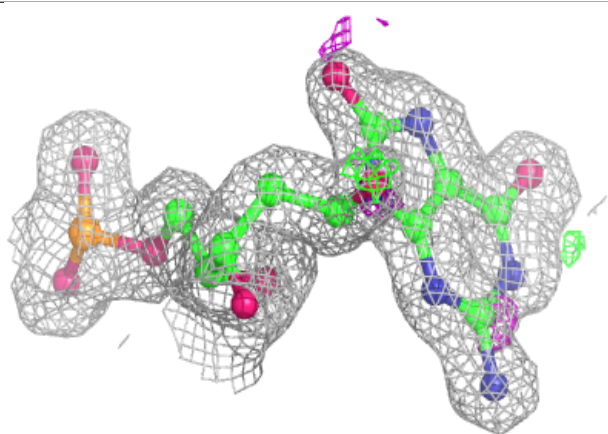
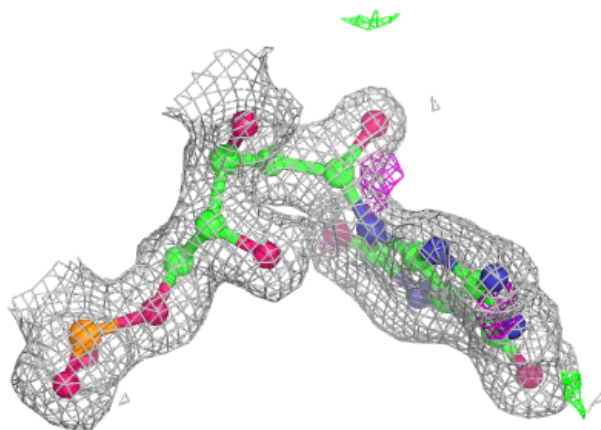
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
6	GOL	D	101	6/6	0.84	0.13	23,29,33,37	0
6	GOL	A	401	6/6	0.85	0.12	27,28,35,39	0
6	GOL	A	402	6/6	0.92	0.10	14,19,22,27	0
6	GOL	A	403	6/6	0.92	0.11	20,30,36,37	0
7	MG	A	404	1/1	0.92	0.17	37,37,37,37	0
5	A1LXK	C	101	24/25	0.96	0.07	12,17,23,26	0
8	NA	A	406	1/1	0.96	0.09	33,33,33,33	0
7	MG	A	405	1/1	0.99	0.11	18,18,18,18	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 A1LXK C 101:

$2mF_o-DF_c$ (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.