



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 28, 2023 – 02:28 AM EDT

PDB ID : 3LAN  
Title : Crystal structure of HIV-1 reverse transcriptase in complex with N1-butyl pyrimidinedione non-nucleoside inhibitor  
Authors : Lansdon, E.B.; Mitchell, M.L.  
Deposited on : 2010-01-06  
Resolution : 2.55 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 : 2.35  
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) : 2.35

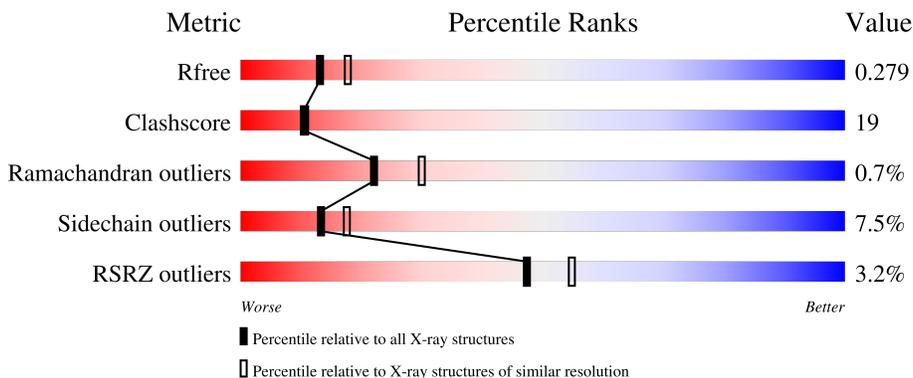
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.55 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1284 (2.56-2.52)
Clashscore	141614	1332 (2.56-2.52)
Ramachandran outliers	138981	1315 (2.56-2.52)
Sidechain outliers	138945	1315 (2.56-2.52)
RSRZ outliers	127900	1272 (2.56-2.52)

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	560	
1	B	560	

## 2 Entry composition [i](#)

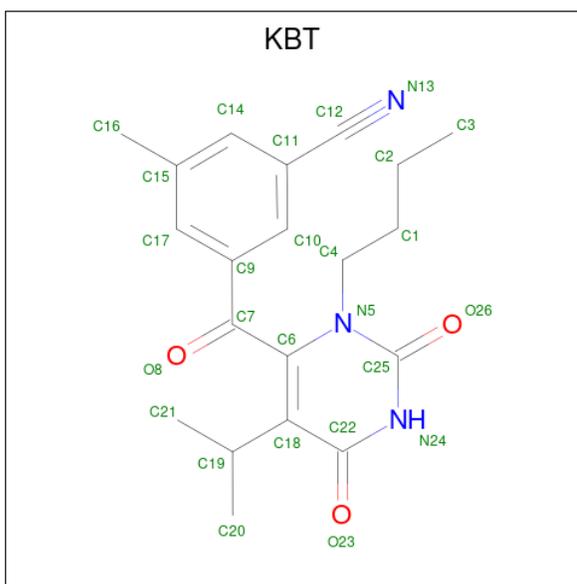
There are 4 unique types of molecules in this entry. The entry contains 8002 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 HIV Reverse transcriptase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	553	Total 4508	C 2914	N 752	O 834	S 8	0	0	0
1	B	403	Total 3334	C 2173	N 550	O 605	S 6	0	0	0

- Molecule 2 is 3-{{3-butyl-5-(1-methylethyl)-2,6-dioxo-1,2,3,6-tetrahydropyrimidin-4-yl}carbonyl}-5-methylbenzonitrile (three-letter code: KBT) (formula: C<sub>20</sub>H<sub>23</sub>N<sub>3</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	Total 26	C 20	N 3	O 3	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO<sub>4</sub>) (formula: O<sub>4</sub>S).

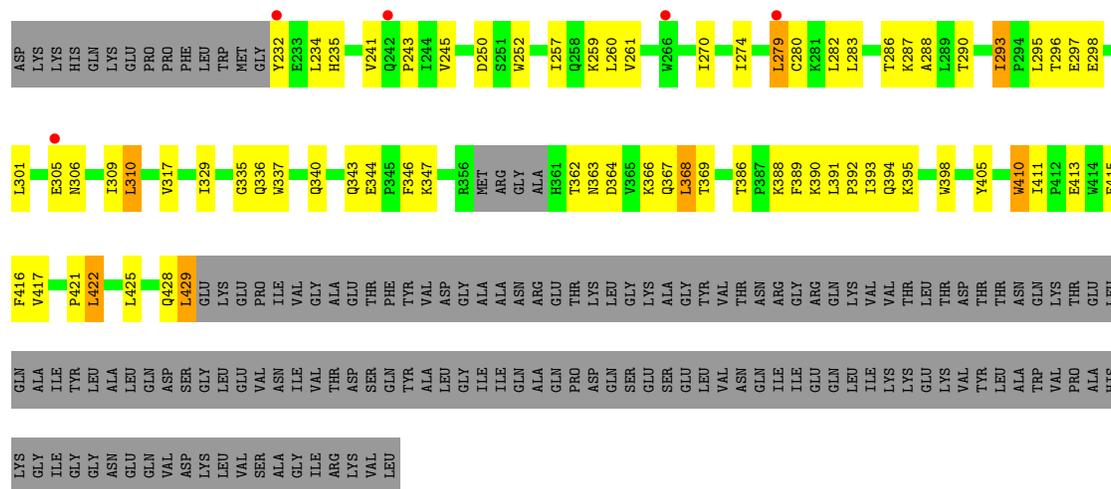


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	56	Total	O	0	0
			56	56		
4	B	48	Total	O	0	0
			48	48		





LYS  
GLY  
ILE  
GLY  
GLY  
ASN  
GLU  
GLN  
VAL  
ASP  
LYS  
LEU  
VAL  
SER  
ALA  
GLY  
ILE  
ARG  
LYS  
VAL  
LEU

GLN  
ALA  
ILE  
TYR  
LEU  
ALA  
LEU  
GLN  
ASP  
SER  
GLY  
LEU  
GLU  
VAL  
ASN  
ILE  
VAL  
THR  
ASP  
SER  
GLN  
TYR  
ALA  
LEU  
GLY  
ILE  
ILE  
GLN  
ALA  
GLN  
PRO  
ASP  
GLN  
SER  
GLU  
GLU  
VAL  
ASN  
GLN  
ILE  
GLU  
GLN  
LEU  
LYS  
VAL  
LYS  
LEU  
THR  
VAL  
TYR  
LEU  
ALA  
TRP  
VAL  
PRO  
ALA  
HIS

F416  
V417  
P421  
L422  
L425  
Q428  
L429  
GLU  
LYS  
GLU  
PRO  
VAL  
ASN  
ILE  
VAL  
GLY  
ALA  
GLU  
THR  
PHE  
TYR  
VAL  
ASP  
GLY  
ALA  
ALA  
ASN  
ASN  
ARG  
GLU  
THR  
LYS  
LEU  
LEU  
VAL  
ASN  
ARG  
GLN  
LEU  
TYR  
VAL  
THR  
THR  
ASN  
GLN  
LYS  
THR  
THR  
ASP  
THR  
ASP  
ARG  
ARG  
GLN  
LEU  
ILE  
VAL  
LYS  
LYS  
LEU  
THR  
THR  
ASP  
TYR  
THR  
LEU  
ALA  
TRP  
VAL  
PRO  
ALA  
HIS

L301  
E306  
N306  
I309  
L310  
V317  
I329  
G335  
Q336  
W337  
Q340  
Q343  
E344  
F345  
F346  
K347  
R356  
MET  
ARG  
GLY  
ASN  
ALA  
H361  
T362  
N363  
D364  
V365  
K366  
Q367  
L368  
T369  
T386  
F387  
K388  
F389  
K390  
L391  
F392  
I393  
Q394  
K395  
W396  
Y405  
W410  
I411  
F412  
E413  
W414  
E415

ASP  
LYS  
LYS  
HIS  
GLN  
LYS  
GLU  
PRO  
PRO  
PHE  
LEU  
TRP  
MET  
GLY  
E232  
E233  
L234  
H235  
V241  
Q242  
P243  
I244  
V245  
D250  
S251  
W252  
I257  
Q258  
K259  
L260  
V261  
W266  
I270  
I274  
L279  
C280  
K281  
L282  
L283  
T286  
K287  
A288  
L289  
T290  
I293  
P294  
L295  
T296  
E297  
E298

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	118.70Å 154.03Å 153.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.78 – 2.55 47.12 – 2.54	Depositor EDS
% Data completeness (in resolution range)	86.0 (29.78-2.55) 85.5 (47.12-2.54)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.34 (at 2.54Å)	Xtrriage
Refinement program	CNX 2005	Depositor
R, $R_{free}$	0.235 , 0.284 0.222 , 0.279	Depositor DCC
$R_{free}$ test set	1995 reflections (4.63%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	65.4	Xtrriage
Anisotropy	0.413	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 44.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	8002	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	69.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.43% 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, KBT

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	0.38	0/4625	0.53	0/6284
1	B	0.39	0/3426	0.53	0/4655
All	All	0.39	0/8051	0.53	0/10939

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
1	A	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	52	PRO	Peptide

### 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	4508	0	4558	194	0
1	B	3334	0	3365	112	0
2	A	26	0	23	1	0
3	A	20	0	0	0	0
3	B	10	0	0	1	0
4	A	56	0	0	5	0
4	B	48	0	0	2	0
All	All	8002	0	7946	294	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 19.

The worst 5 of 294 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:66:LYS:HE2	1:A:67:ASP:H	1.26	0.99
1:A:131:THR:CG2	1:A:143:ARG:HD2	1.92	0.98
1:A:199:ARG:HH11	1:A:219:LYS:HE2	1.29	0.94
1:B:194:GLU:HG3	1:B:197:GLN:HG3	1.49	0.94
1:A:131:THR:HG22	1:A:143:ARG:HD2	1.47	0.92

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	551/560 (98%)	518 (94%)	27 (5%)	6 (1%)	14 19
1	B	395/560 (70%)	371 (94%)	23 (6%)	1 (0%)	41 51
All	All	946/1120 (84%)	889 (94%)	50 (5%)	7 (1%)	22 30

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	230	MET
1	A	219	LYS
1	A	543	GLY
1	A	78	ARG
1	A	66	LYS

### 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	495/500 (99%)	458 (92%)	37 (8%)	13	17
1	B	367/500 (73%)	339 (92%)	28 (8%)	13	17
All	All	862/1000 (86%)	797 (92%)	65 (8%)	13	17

5 of 65 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	293	ILE
1	B	336	GLN
1	A	373	GLN
1	A	357	MET
1	B	343	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	336	GLN
1	B	367	GLN
1	B	407	GLN
1	B	348	ASN
1	A	524	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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

7 ligands 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
3	SO4	A	565	-	4,4,4	0.28	0	6,6,6	0.09	0
3	SO4	A	564	-	4,4,4	0.25	0	6,6,6	0.06	0
2	KBT	A	561	-	27,27,27	2.13	10 (37%)	33,38,38	1.41	7 (21%)
3	SO4	B	562	-	4,4,4	1.09	0	6,6,6	0.57	0
3	SO4	B	561	-	4,4,4	1.25	0	6,6,6	0.49	0
3	SO4	A	562	-	4,4,4	0.23	0	6,6,6	0.23	0
3	SO4	A	563	-	4,4,4	0.25	0	6,6,6	0.12	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	KBT	A	561	-	-	1/18/18/18	0/2/2/2

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	561	KBT	C22-C18	5.15	1.54	1.43
2	A	561	KBT	C17-C9	3.92	1.45	1.39

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	561	KBT	C6-C18	2.95	1.41	1.35
2	A	561	KBT	C25-N5	-2.83	1.34	1.39
2	A	561	KBT	C4-N5	2.77	1.53	1.47

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	561	KBT	C19-C18-C6	3.58	125.23	119.06
2	A	561	KBT	N24-C25-N5	2.71	120.15	115.29
2	A	561	KBT	C9-C7-C6	2.46	123.97	119.37
2	A	561	KBT	C18-C22-N24	2.43	119.03	116.13
2	A	561	KBT	C22-N24-C25	-2.37	124.28	127.35

There are no chirality outliers.

All (1) torsion outliers are listed below:

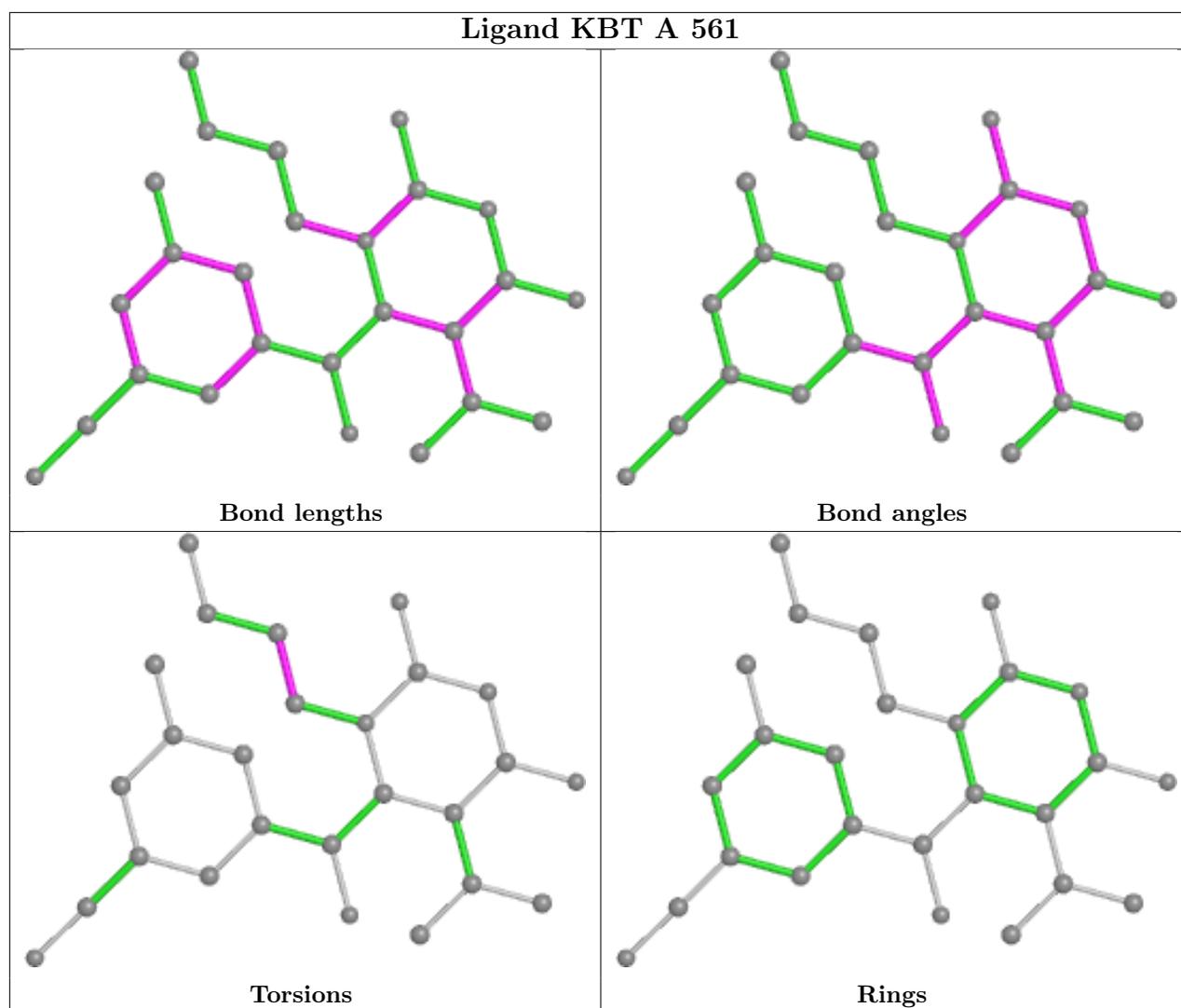
Mol	Chain	Res	Type	Atoms
2	A	561	KBT	C2-C1-C4-N5

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	561	KBT	1	0
3	B	562	SO4	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	A	553/560 (98%)	0.28	19 (3%) 45 52	46, 68, 97, 117	0
1	B	403/560 (71%)	0.29	12 (2%) 50 57	41, 63, 96, 112	0
All	All	956/1120 (85%)	0.29	31 (3%) 47 55	41, 66, 97, 117	0

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	214	LEU	4.4
1	A	193	LEU	4.2
1	A	195	ILE	3.4
1	B	14	PRO	3.4
1	B	232	TYR	3.1

### 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](#)

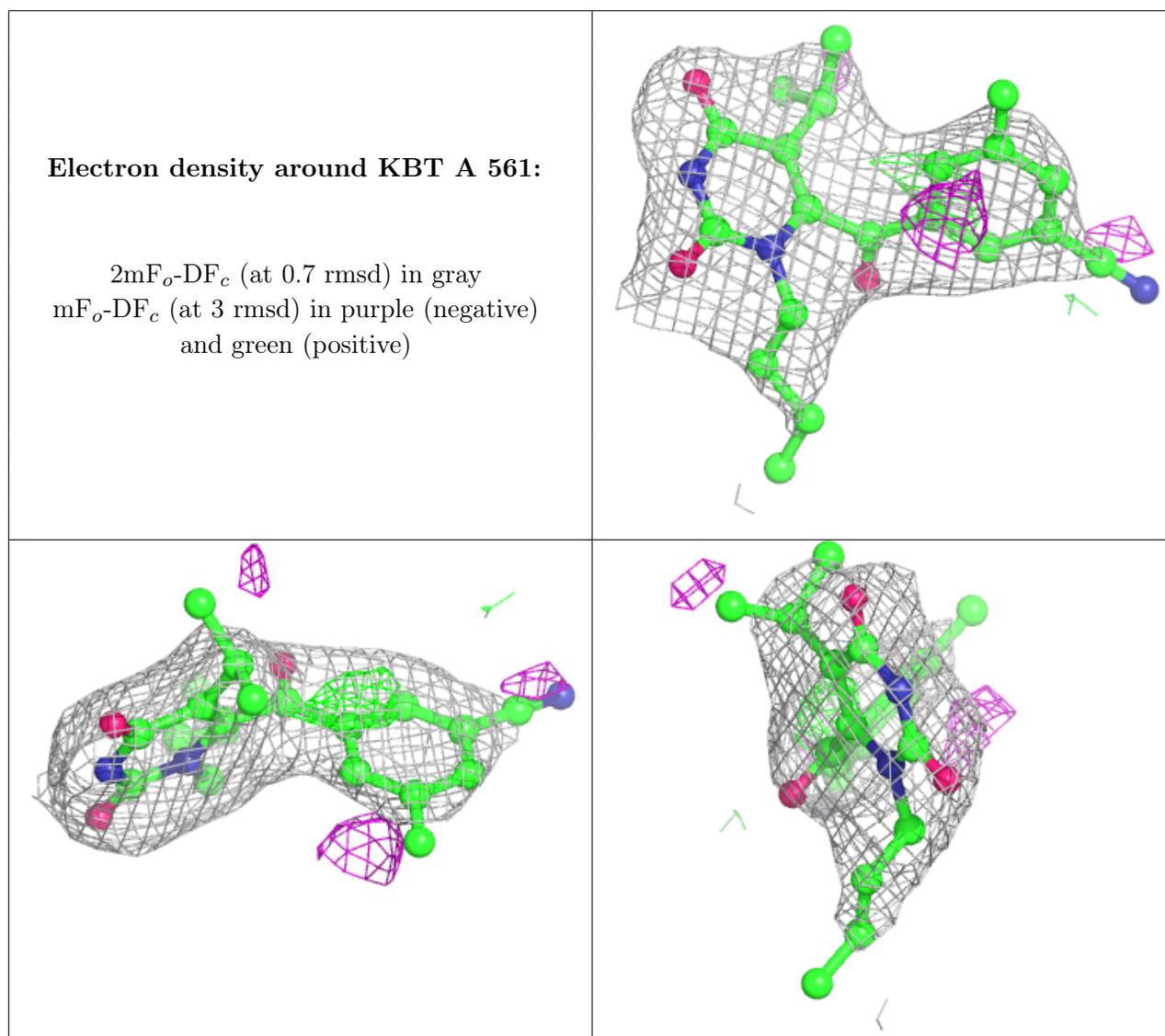
There are no monosaccharides in this entry.

### 6.4 Ligands [i](#)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	SO4	A	564	5/5	0.86	0.27	163,163,164,164	0
3	SO4	A	565	5/5	0.87	0.19	132,132,133,133	0
3	SO4	B	561	5/5	0.91	0.31	97,99,99,100	0
2	KBT	A	561	26/26	0.94	0.30	69,71,75,78	0
3	SO4	A	563	5/5	0.95	0.12	81,82,83,84	0
3	SO4	A	562	5/5	0.96	0.12	75,78,79,80	0
3	SO4	B	562	5/5	0.96	0.10	84,85,86,87	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.