W O R L D W I D E

wwPDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 21, 2023 – 10:14 PM EDT

PDB ID	:	2PO3
Title	:	Crystal Structure Analysis of DesI in the presence of its TDP-sugar product
Authors	:	Burgie, E.S.; Holden, H.M.
Deposited on	:	2007-04-25
Resolution	:	2.10 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*
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:

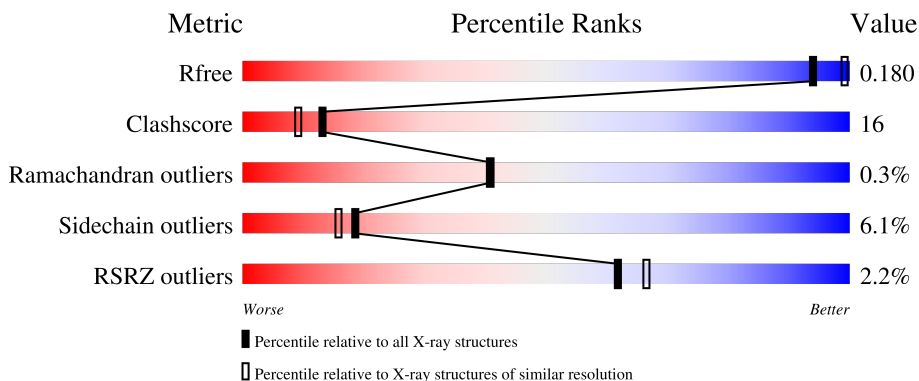|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & 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:
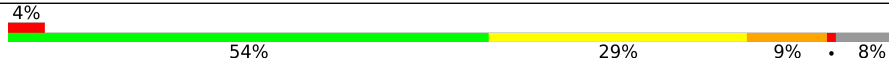*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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		0.180
Clashscore		16
Ramachandran outliers		0.3%
Sidechain outliers		6.1%
RSRZ outliers		2.2%

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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 >=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	424	62% 25% 5% 7%
1	B	424	4% 54% 29% 9% · 8%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 6635 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 4-dehydrase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	393	Total 2969	C 1847	N 561	O 551	S 10	0	1	0
1	B	392	Total 2973	C 1850	N 561	O 552	S 10	0	3	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1416	GLY	-	expression tag	UNP Q9ZGH0
A	1417	LEU	-	expression tag	UNP Q9ZGH0
A	1418	GLU	-	expression tag	UNP Q9ZGH0
A	1419	HIS	-	expression tag	UNP Q9ZGH0
A	1420	HIS	-	expression tag	UNP Q9ZGH0
A	1421	HIS	-	expression tag	UNP Q9ZGH0
A	1422	HIS	-	expression tag	UNP Q9ZGH0
A	1423	HIS	-	expression tag	UNP Q9ZGH0
A	1424	HIS	-	expression tag	UNP Q9ZGH0
B	2416	GLY	-	expression tag	UNP Q9ZGH0
B	2417	LEU	-	expression tag	UNP Q9ZGH0
B	2418	GLU	-	expression tag	UNP Q9ZGH0
B	2419	HIS	-	expression tag	UNP Q9ZGH0
B	2420	HIS	-	expression tag	UNP Q9ZGH0
B	2421	HIS	-	expression tag	UNP Q9ZGH0
B	2422	HIS	-	expression tag	UNP Q9ZGH0
B	2423	HIS	-	expression tag	UNP Q9ZGH0
B	2424	HIS	-	expression tag	UNP Q9ZGH0

- Molecule 2 is (2R,3R,4S,5S,6R)-3,4-DIHYDROXY-5-[({3-HYDROXY-2-METHYL-5-[(PHOSPHONOOXY)METHYL]PYRIDIN-4-YL}METHYL)IMINO]-6-METHYLTETRAHYDRO-2H-PYRAN-2-YL [(2R,3S,5R)-3-HYDROXY-5-(5-METHYL-2,4-DIOXO-3,4-DIHYDROPYRIMIDIN-1(2H)-YL)TETRAHYDROFURAN-2-YL]METHYL DIHYDROGEN DIPHOSPHATE (three-letter code: T4K) (formula: $C_{24}H_{35}N_4O_{19}P_3$).
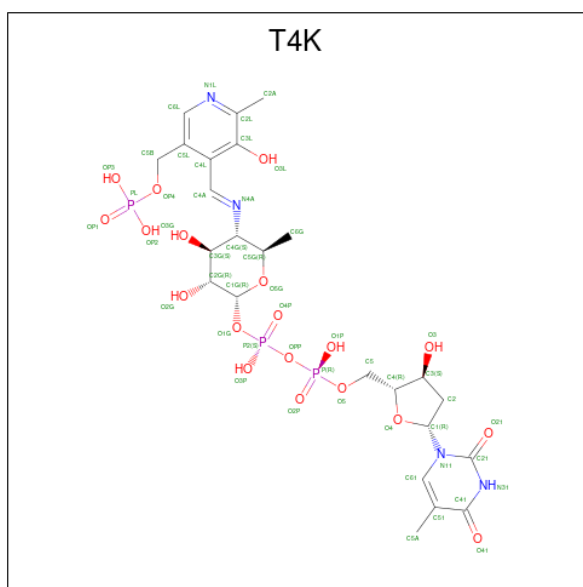
T4K

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			50	24	4	19	3		
2	B	1	Total	C	N	O	P	0	0
			50	24	4	19	3		

• Molecule 3 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	342	Total	O	0	0
			342	342		
3	B	251	Total	O	0	0
			251	251		

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: 4-dehydrase
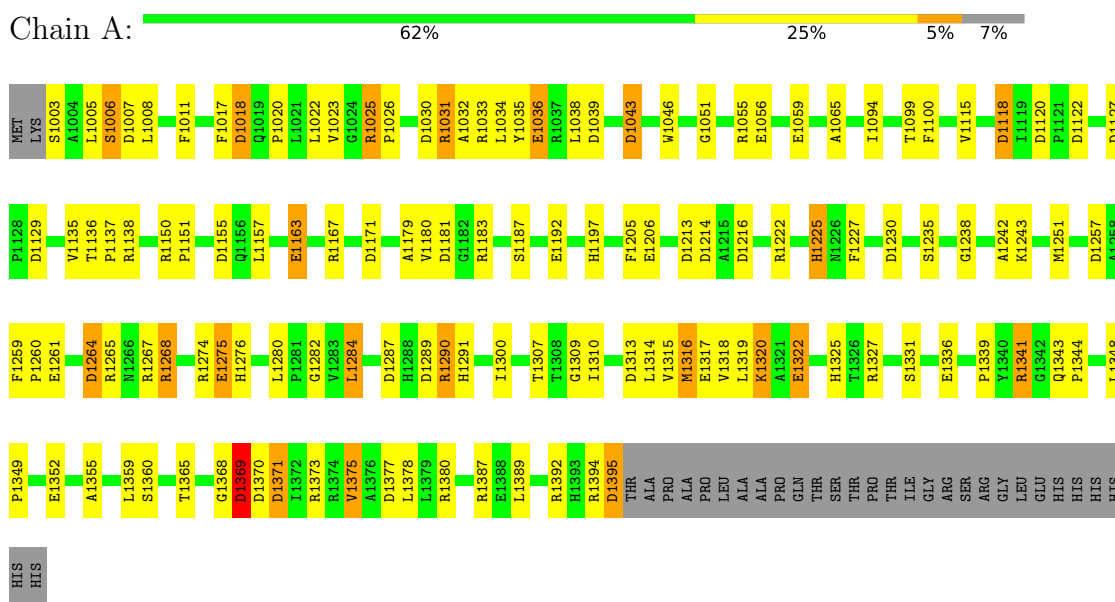
- Molecule 1: 4-dehydrase
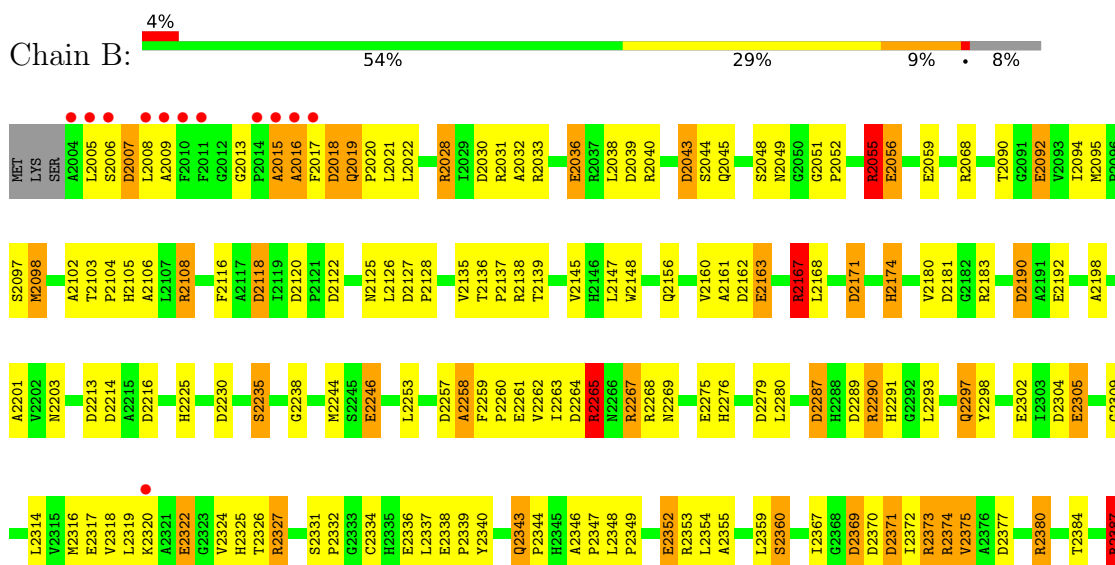
E2388 L2389 T2390 A2391 R2392 H2393 R2394 D2395 THR ALA PRO ALA PRO LEU ALA ALA PRO GLN THR SER THR PRO THR ILE GLY ARG SER ARG GLY LEU GLU HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	59.50Å 66.80Å 242.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.84 – 2.10 44.84 – 2.10	Depositor EDS
% Data completeness (in resolution range)	97.2 (44.84-2.10) 97.3 (44.84-2.10)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	0.63 (at 2.10Å)	Xtriage
Refinement program	TNT	Depositor
R, R_{free}	0.180 , 0.248 0.177 , 0.180	Depositor DCC
R_{free} test set	5599 reflections (10.03%)	wwPDB-VP
Wilson B-factor (Å^2)	25.2	Xtriage
Anisotropy	0.634	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 132.0	EDS
L-test for twinning[2]	$< |L| > = 0.47$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6635	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.03% 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 ⓘ

5.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: T4K

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	0.87	10/3041 (0.3%)	1.41	58/4140 (1.4%)
1	B	0.91	18/3053 (0.6%)	1.51	73/4156 (1.8%)
All	All	0.89	28/6094 (0.5%)	1.46	131/8296 (1.6%)

The worst 5 of 28 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	2056	GLU	CD-OE2	7.08	1.33	1.25
1	B	2163	GLU	CD-OE2	6.65	1.32	1.25
1	A	1261	GLU	CD-OE2	6.41	1.32	1.25
1	A	1059	GLU	CD-OE2	6.30	1.32	1.25
1	A	1275	GLU	CD-OE2	6.26	1.32	1.25

The worst 5 of 131 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	2394	ARG	NE-CZ-NH1	9.44	125.02	120.30
1	B	2214	ASP	CB-CG-OD2	-9.14	110.08	118.30
1	A	1043	ASP	CB-CG-OD2	-8.91	110.28	118.30
1	B	2214	ASP	CB-CG-OD1	8.84	126.26	118.30
1	B	2127	ASP	CB-CG-OD2	-8.64	110.52	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2   Too-close contacts ⓘ

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	2969	0	2896	71	0
1	B	2973	0	2901	126	0
2	A	50	0	31	6	0
2	B	50	0	31	1	0
3	A	342	0	0	11	0
3	B	251	0	0	13	0
All	All	6635	0	5859	194	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 16.

The worst 5 of 194 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:B:2201:ALA:HB2	1:B:2297[B]:GLN:HG3	1.41	1.02
1:B:2009:ALA:HA	1:B:2013:GLY:O	1.67	0.94
1:B:2017:PHE:HB2	1:B:2374:ARG:CZ	1.99	0.93
1:A:1022:LEU:HD23	1:A:1325:HIS:HB2	1.55	0.89
1:B:2167:ARG:HH11	1:B:2167:ARG:HG3	1.42	0.82

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	392/424 (92%)	379 (97%)	13 (3%)	0	100	100
1	B	393/424 (93%)	372 (95%)	19 (5%)	2 (0%)	29	26
All	All	785/848 (93%)	751 (96%)	32 (4%)	2 (0%)	41	41

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	2394	ARG
1	B	2198	ALA

5.3.2 Protein sidechains ⓘ

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	297/321 (92%)	281 (95%)	16 (5%)	22	20
1	B	298/321 (93%)	277 (93%)	21 (7%)	15	12
All	All	595/642 (93%)	558 (94%)	37 (6%)	18	15

5 of 37 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	2265	ARG
1	B	2375	VAL
1	B	2297[A]	GLN
1	B	2369	ASP
1	A	1360	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	2019	GLN
1	B	2130	GLN
1	B	2276	HIS
1	B	2156	GLN
1	A	1343	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

2 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	$\#|Z| > 2$
2	T4K	B	2500	-	50,53,53	1.61	8 (16%)	70,81,81	1.85	16 (22%)
2	T4K	A	1500	-	50,53,53	1.55	7 (14%)	70,81,81	1.76	14 (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
2	T4K	B	2500	-	-	8/32/64/64	0/4/4/4
2	T4K	A	1500	-	-	3/32/64/64	0/4/4/4

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	2500	T4K	C4G-N4A	-5.27	1.39	1.46
2	A	1500	T4K	C4G-N4A	-4.28	1.41	1.46
2	A	1500	T4K	C4A-N4A	3.84	1.34	1.27
2	A	1500	T4K	PL-OP1	3.76	1.62	1.50
2	B	2500	T4K	C4A-N4A	3.56	1.33	1.27

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	2500	T4K	P-OPP-P2	-6.27	111.30	132.83
2	A	1500	T4K	C41-N31-C21	-5.09	120.77	127.35
2	B	2500	T4K	C41-N31-C21	-5.04	120.83	127.35
2	A	1500	T4K	C51-C41-N31	5.01	119.59	115.31
2	B	2500	T4K	C51-C41-N31	4.99	119.57	115.31

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

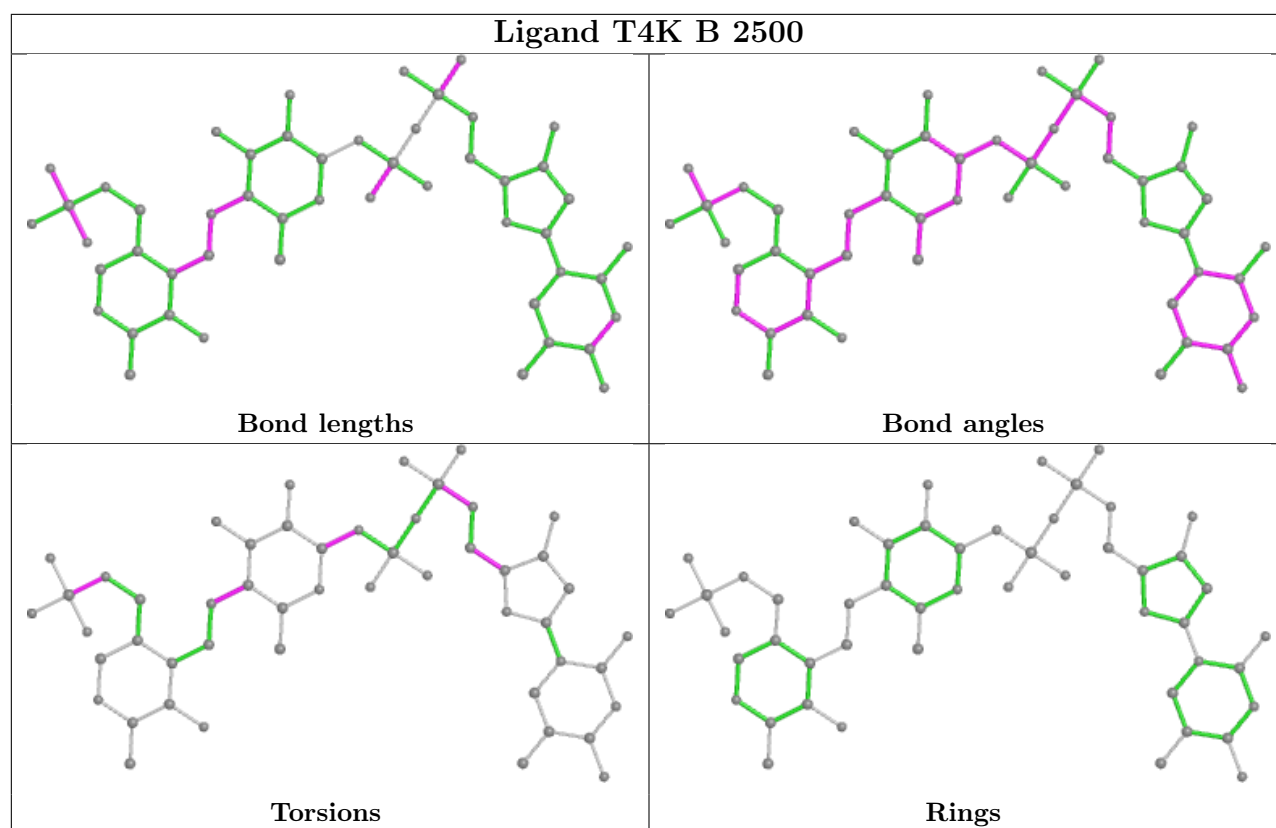| Mol	Chain	Res	Type	Atoms
2	A	1500	T4K	C3G-C4G-N4A-C4A
2	B	2500	T4K	C5B-OP4-PL-OP1
2	B	2500	T4K	C3G-C4G-N4A-C4A
2	B	2500	T4K	C5-O5-P-O1P
2	B	2500	T4K	C5-O5-P-O2P

There are no ring outliers.

2 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	2500	T4K	1	0
2	A	1500	T4K	6	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 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 T4K B 2500

Bond lengths

Bond angles

Torsions

Rings

Ligand T4K A 1500

Bond lengths

Bond angles

Torsions

Rings
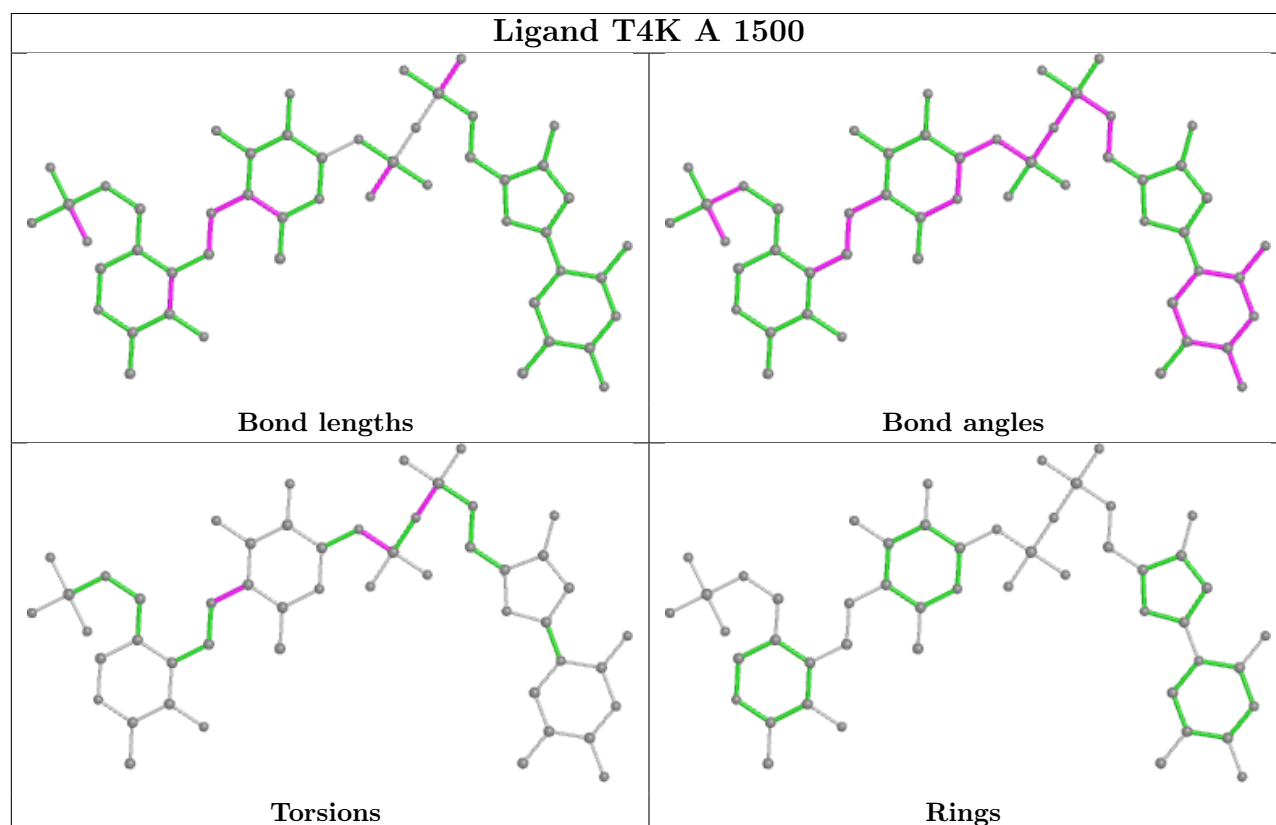
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	393/424 (92%)	-0.68	0	100	100	13, 27, 59, 93	0
1	B	392/424 (92%)	-0.30	17 (4%)	35	41	12, 32, 69, 99	0
All	All	785/848 (92%)	-0.49	17 (2%)	62	66	12, 30, 63, 99	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	2011	PHE	4.7
1	B	2005	LEU	4.6
1	B	2017	PHE	3.6
1	B	2008	LEU	3.5
1	B	2010	PHE	2.9

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ
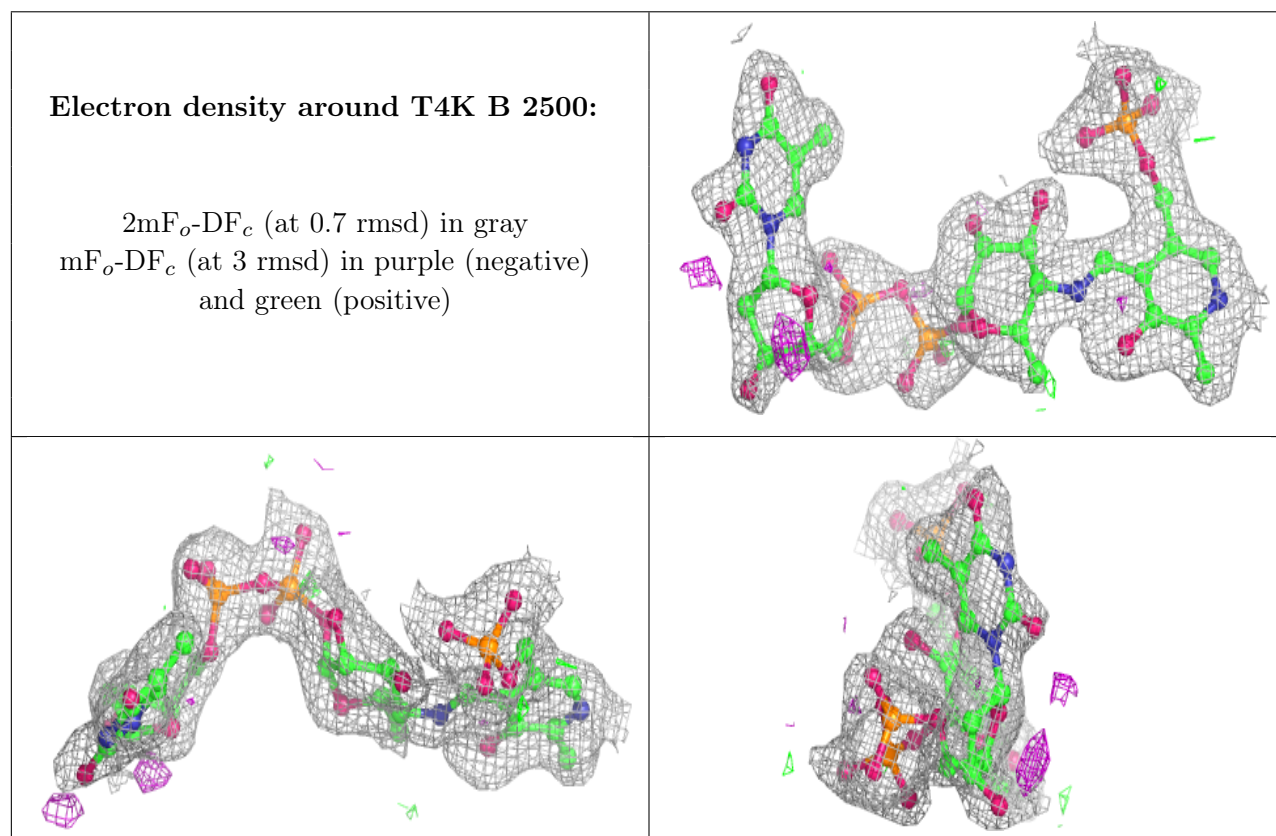
There are no monosaccharides in this entry.
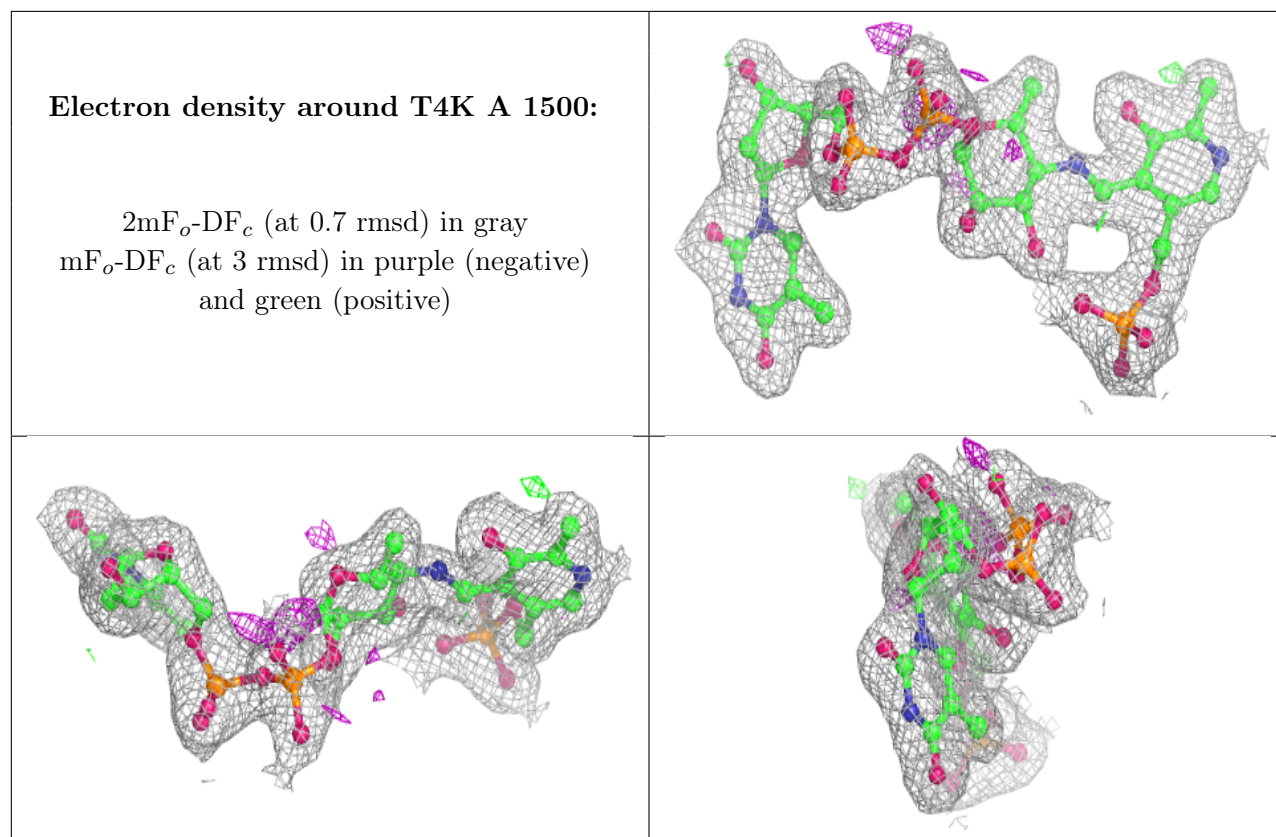
6.4 Ligands ⓘ

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, 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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
2	T4K	B	2500	50/50	0.95	0.12	14,54,100,100	0
2	T4K	A	1500	50/50	0.97	0.09	13,28,67,100	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 T4K B 2500:

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

Electron density around T4K A 1500:

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

6.5 Other polymers ⓘ

There are no such residues in this entry.