



wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 01:08 pm BST

PDB ID : 6P5P
Title : Discovery of a Novel, Highly Potent, and Selective Thieno[3,2-d]pyrimidine-Based Cdc7 inhibitor with a Quinuclidine Moiety (TAK-931) as an Orally Active Investigational Anti-Tumor Agent
Authors : Hoffman, I.D.; Skene, R.J.
Deposited on : 2019-05-30
Resolution : 3.30 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 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 : 2.11
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

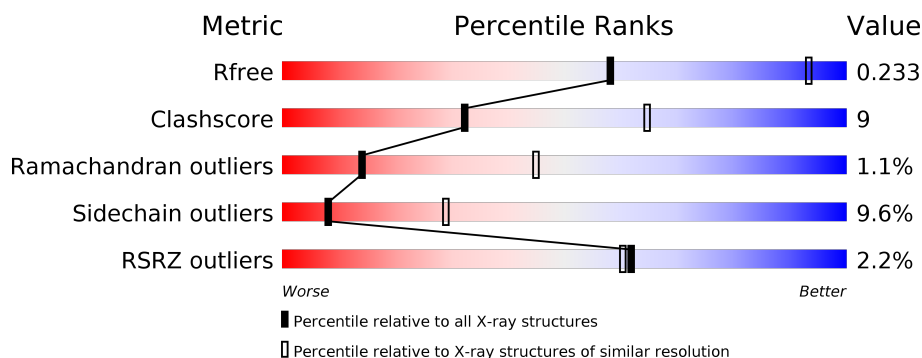
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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 $\leq 5\%$. 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	405	<div> <div>2%</div> <div> <div></div> <div>69%</div> <div>23%</div> <div>• 5%</div> </div> </div>
1	B	405	<div> <div>2%</div> <div> <div></div> <div>66%</div> <div>21%</div> <div>• 9%</div> </div> </div>
1	C	405	<div> <div>2%</div> <div> <div></div> <div>73%</div> <div>18%</div> <div>• 7%</div> </div> </div>
1	D	405	<div> <div>%</div> <div> <div></div> <div>72%</div> <div>20%</div> <div>• 6%</div> </div> </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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	O1V	B	2500	X	-	-	-

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 12378 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 Rho-associated protein kinase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	386	Total	C	N	O	S	0	0	0
			3131	2009	523	579	20			
1	B	369	Total	C	N	O	S	0	0	0
			3005	1938	504	544	19			
1	C	375	Total	C	N	O	S	0	0	0
			3044	1960	506	558	20			
1	D	382	Total	C	N	O	S	0	0	0
			3102	1994	519	570	19			

There are 24 discrepancies between the modelled and reference sequences:

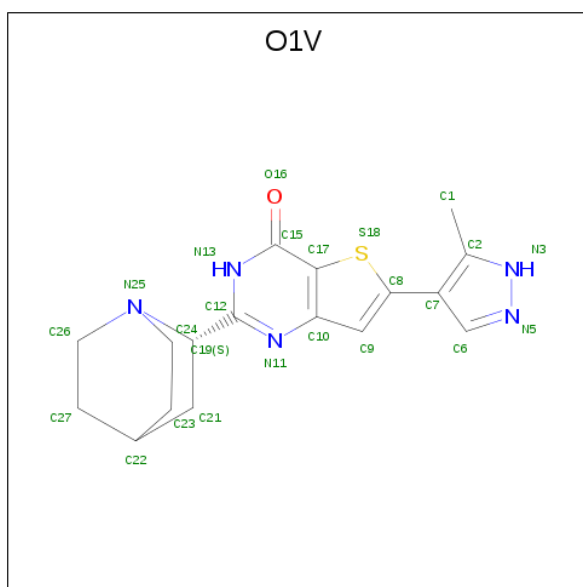
Chain	Residue	Modelled	Actual	Comment	Reference
A	13	GLY	-	expression tag	UNP O75116
A	14	ALA	-	expression tag	UNP O75116
A	15	MET	-	expression tag	UNP O75116
A	16	GLY	-	expression tag	UNP O75116
A	17	SER	-	expression tag	UNP O75116
A	270	TYR	PHE	engineered mutation	UNP O75116
B	13	GLY	-	expression tag	UNP O75116
B	14	ALA	-	expression tag	UNP O75116
B	15	MET	-	expression tag	UNP O75116
B	16	GLY	-	expression tag	UNP O75116
B	17	SER	-	expression tag	UNP O75116
B	270	TYR	PHE	engineered mutation	UNP O75116
C	13	GLY	-	expression tag	UNP O75116
C	14	ALA	-	expression tag	UNP O75116
C	15	MET	-	expression tag	UNP O75116
C	16	GLY	-	expression tag	UNP O75116
C	17	SER	-	expression tag	UNP O75116
C	270	TYR	PHE	engineered mutation	UNP O75116
D	13	GLY	-	expression tag	UNP O75116
D	14	ALA	-	expression tag	UNP O75116
D	15	MET	-	expression tag	UNP O75116

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	16	GLY	-	expression tag	UNP O75116
D	17	SER	-	expression tag	UNP O75116
D	270	TYR	PHE	engineered mutation	UNP O75116

- Molecule 2 is 2-[(2S)-1-azabicyclo[2.2.2]octan-2-yl]-6-(5-methyl-1H-pyrazol-4-yl)thieno[3,2-d]pyrimidin-4(3H)-one (three-letter code: O1V) (formula: C₁₇H₁₉N₅OS).

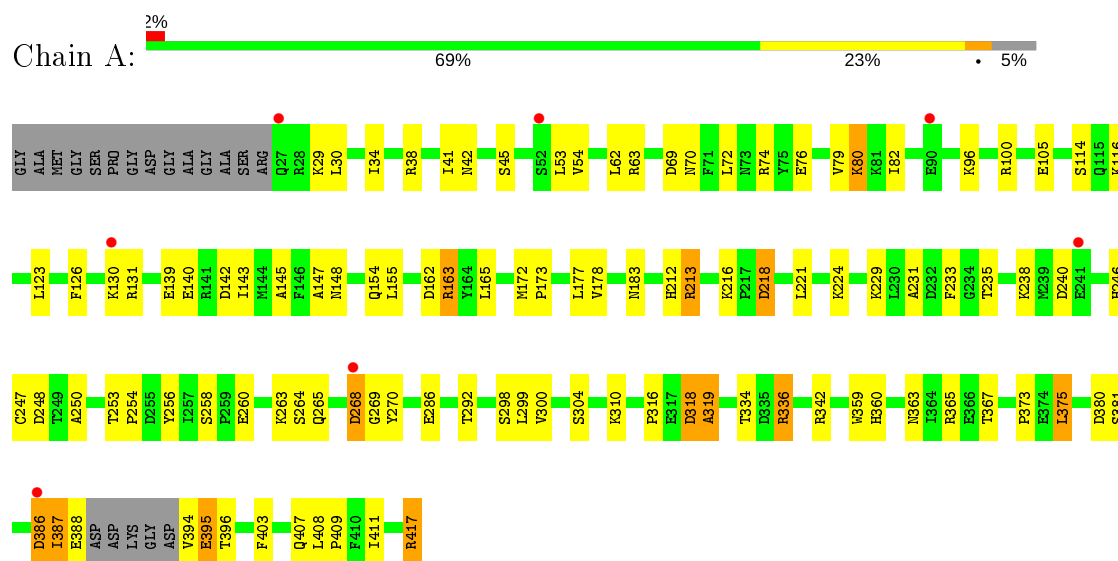


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			24	17	5	1	1		
2	B	1	Total	C	N	O	S	0	0
			24	17	5	1	1		
2	C	1	Total	C	N	O	S	0	0
			24	17	5	1	1		
2	D	1	Total	C	N	O	S	0	0
			24	17	5	1	1		

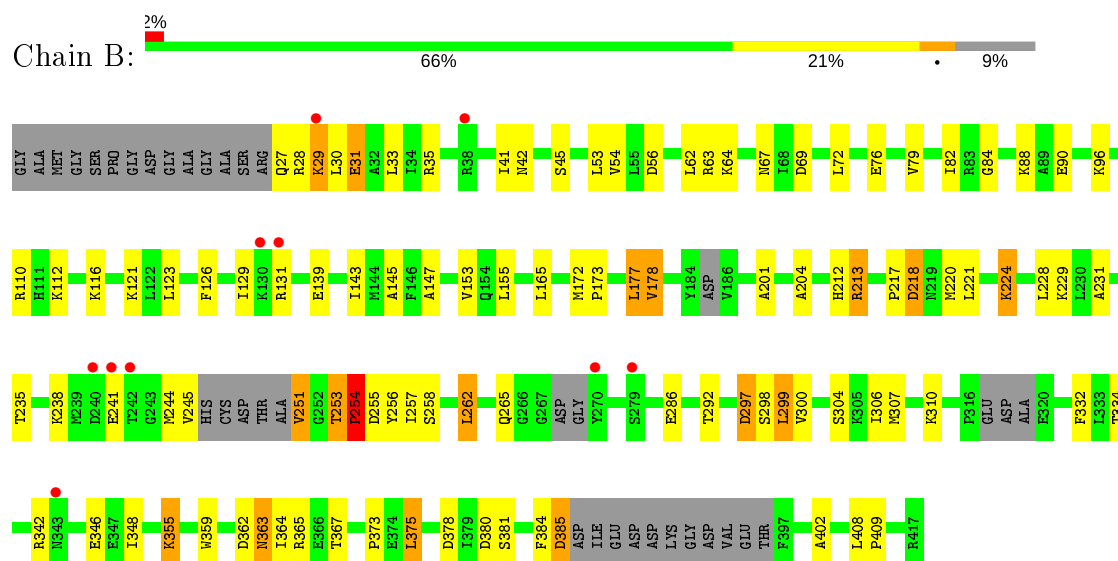
3 Residue-property plots

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.

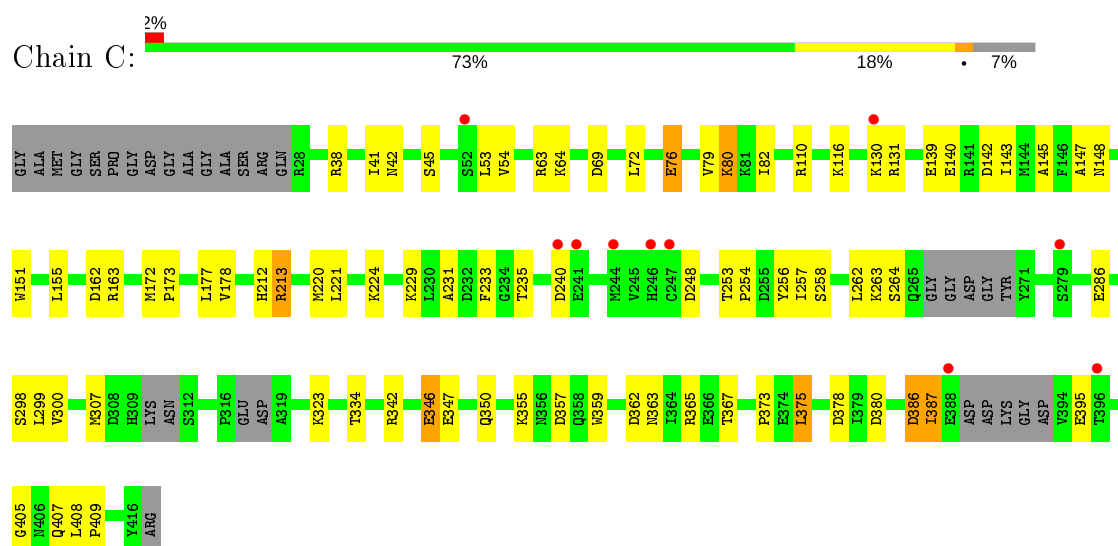
- Molecule 1: Rho-associated protein kinase 2



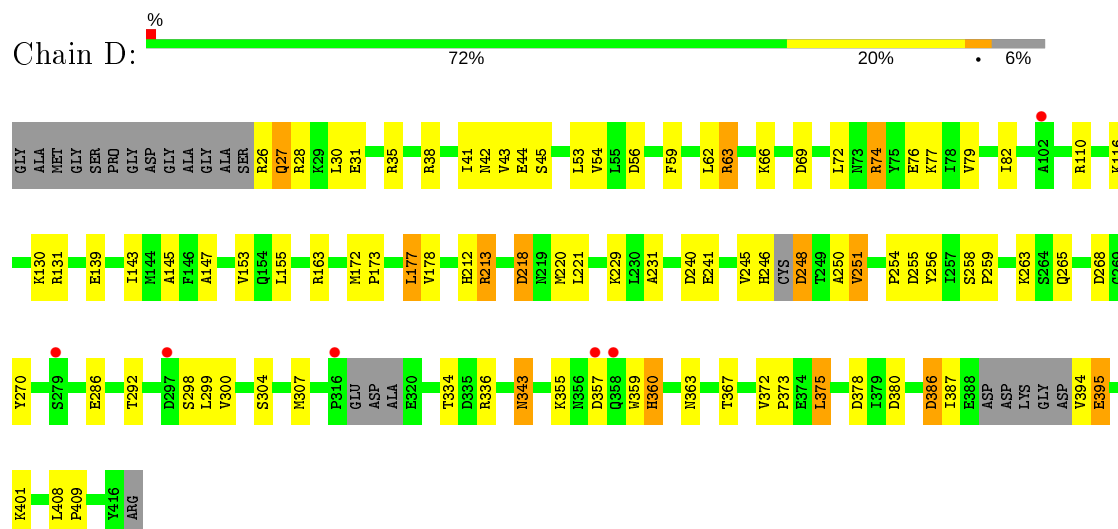
- Molecule 1: Rho-associated protein kinase 2



- Molecule 1: Rho-associated protein kinase 2



• Molecule 1: Rho-associated protein kinase 2



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	89.17Å 146.44Å 111.30Å 90.00° 97.13° 90.00°	Depositor
Resolution (Å)	25.00 – 3.30 24.85 – 3.28	Depositor EDS
% Data completeness (in resolution range)	98.6 (25.00-3.30) 98.9 (24.85-3.28)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.03 (at 3.30Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.206 , 0.236 0.206 , 0.233	Depositor DCC
R_{free} test set	2150 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å ²)	85.4	Xtriage
Anisotropy	0.107	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 50.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	12378	wwPDB-VP
Average B, all atoms (Å ²)	94.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.93% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: O1V

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	0.68	1/3209 (0.0%)	0.75	0/4337
1	B	0.66	0/3078	0.78	3/4151 (0.1%)
1	C	0.65	0/3118	0.74	0/4213
1	D	0.65	0/3178	0.75	0/4293
All	All	0.66	1/12583 (0.0%)	0.76	3/16994 (0.0%)

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	B	0	1
1	D	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	388	GLU	C-O	11.72	1.45	1.23

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	254	PRO	N-CA-C	6.30	128.48	112.10
1	B	254	PRO	N-CA-CB	-5.87	96.14	102.60
1	B	245	VAL	CA-C-O	5.58	131.82	120.10

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	241	GLU	Peptide
1	D	254	PRO	Peptide

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	3131	0	3056	64	0
1	B	3005	0	2955	65	0
1	C	3044	0	2981	37	0
1	D	3102	0	3034	59	0
2	A	24	0	0	0	0
2	B	24	0	0	2	0
2	C	24	0	0	0	0
2	D	24	0	0	1	0
All	All	12378	0	12026	214	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 9.

The worst 5 of 214 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:D:63:ARG:HD3	1:D:69:ASP:OD1	1.47	1.13
1:D:250:ALA:O	1:D:251:VAL:HG13	1.53	1.08
1:A:41:ILE:HD13	1:D:82:ILE:HD11	1.29	1.06
1:B:82:ILE:HD11	1:C:41:ILE:HD13	1.34	1.06
1:A:256:TYR:OH	1:A:286:GLU:OE1	1.82	0.98

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	382/405 (94%)	364 (95%)	14 (4%)	4 (1%)	15	46
1	B	357/405 (88%)	344 (96%)	10 (3%)	3 (1%)	19	51
1	C	365/405 (90%)	349 (96%)	13 (4%)	3 (1%)	19	51
1	D	374/405 (92%)	353 (94%)	15 (4%)	6 (2%)	9	36
All	All	1478/1620 (91%)	1410 (95%)	52 (4%)	16 (1%)	14	45

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	395	GLU
1	B	254	PRO
1	C	395	GLU
1	D	27	GLN
1	D	43	VAL

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	340/350 (97%)	311 (92%)	29 (8%)	10	35
1	B	326/350 (93%)	291 (89%)	35 (11%)	6	25
1	C	332/350 (95%)	302 (91%)	30 (9%)	9	32
1	D	337/350 (96%)	303 (90%)	34 (10%)	7	27
All	All	1335/1400 (95%)	1207 (90%)	128 (10%)	8	29

5 of 128 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	334	THR
1	C	142	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	304	SER
1	B	355	LYS
1	C	64	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	350	GLN
1	C	107	GLN
1	D	350	GLN
1	B	363	ASN
1	C	183	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

4 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	O1V	D	2500	-	25,28,28	1.53	4 (16%)	17,42,42	2.69	4 (23%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	O1V	B	2500	-	25,28,28	1.23	2 (8%)	17,42,42	2.58	6 (35%)
2	O1V	C	2500	-	25,28,28	1.32	3 (12%)	17,42,42	2.77	5 (29%)
2	O1V	A	2500	-	25,28,28	1.37	4 (16%)	17,42,42	2.76	4 (23%)

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	O1V	D	2500	-	-	0/0/26/26	0/6/5/5
2	O1V	B	2500	-	1/1/3/3	0/0/26/26	0/6/5/5
2	O1V	C	2500	-	-	0/0/26/26	0/6/5/5
2	O1V	A	2500	-	-	0/0/26/26	0/6/5/5

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	2500	O1V	C8-S18	4.64	1.77	1.72
2	A	2500	O1V	C8-S18	4.34	1.77	1.72
2	C	2500	O1V	C8-S18	4.06	1.76	1.72
2	B	2500	O1V	C8-S18	3.58	1.76	1.72
2	D	2500	O1V	C12-N11	-2.61	1.29	1.34

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	2500	O1V	C12-N11-C10	9.40	122.91	116.54
2	A	2500	O1V	C12-N11-C10	9.38	122.90	116.54
2	D	2500	O1V	C12-N11-C10	9.04	122.67	116.54
2	B	2500	O1V	C12-N11-C10	7.40	121.56	116.54
2	B	2500	O1V	C21-C19-C12	4.90	122.62	111.08

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	B	2500	O1V	C19

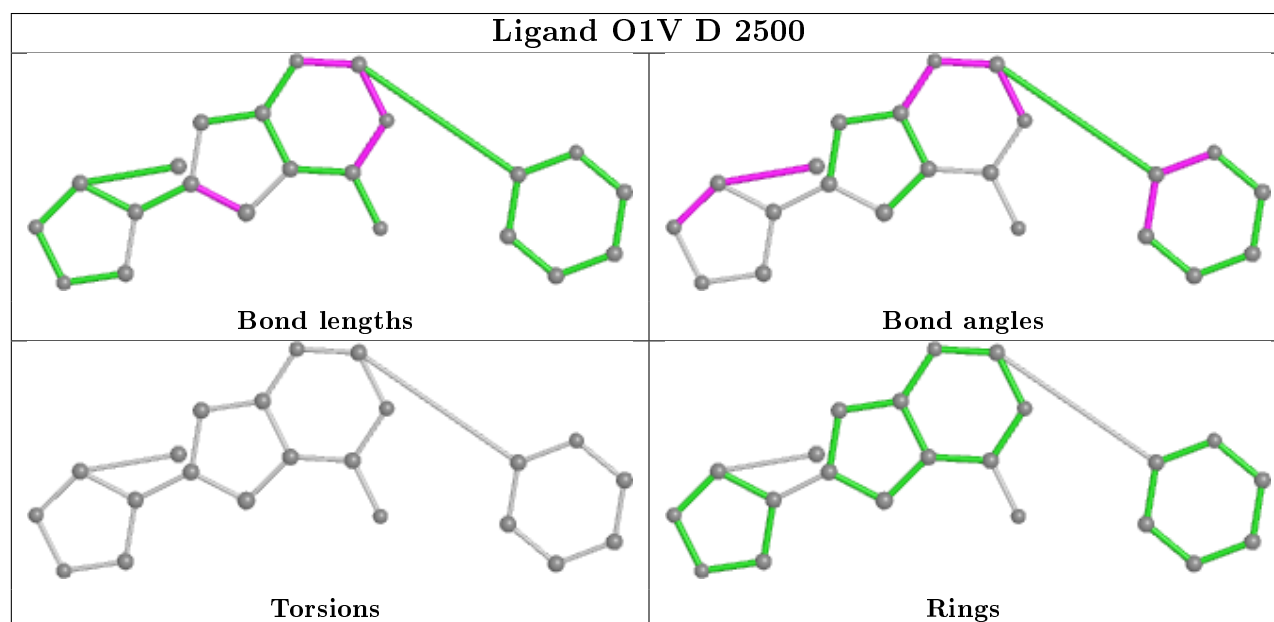
There are no torsion outliers.

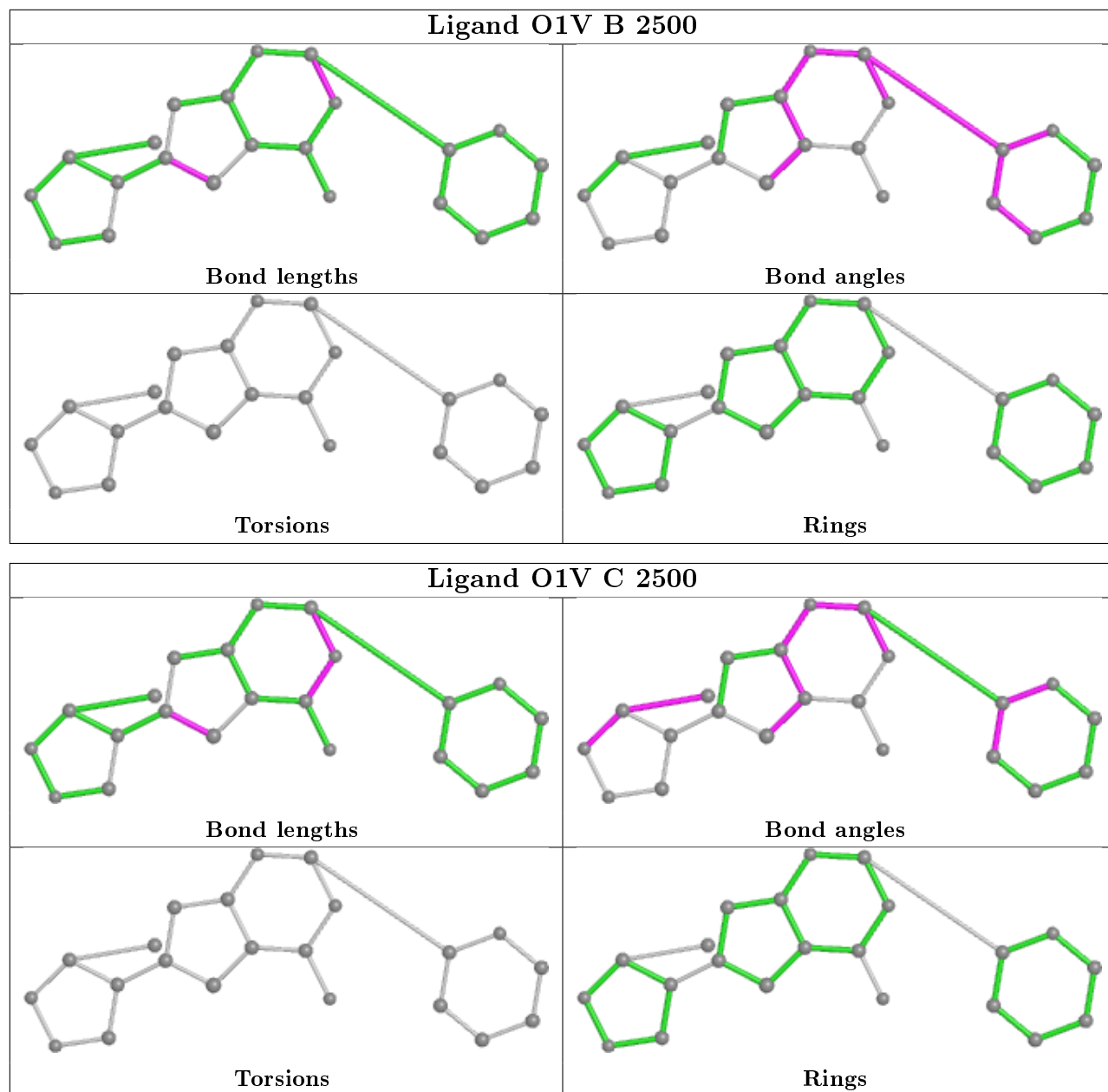
There are no ring outliers.

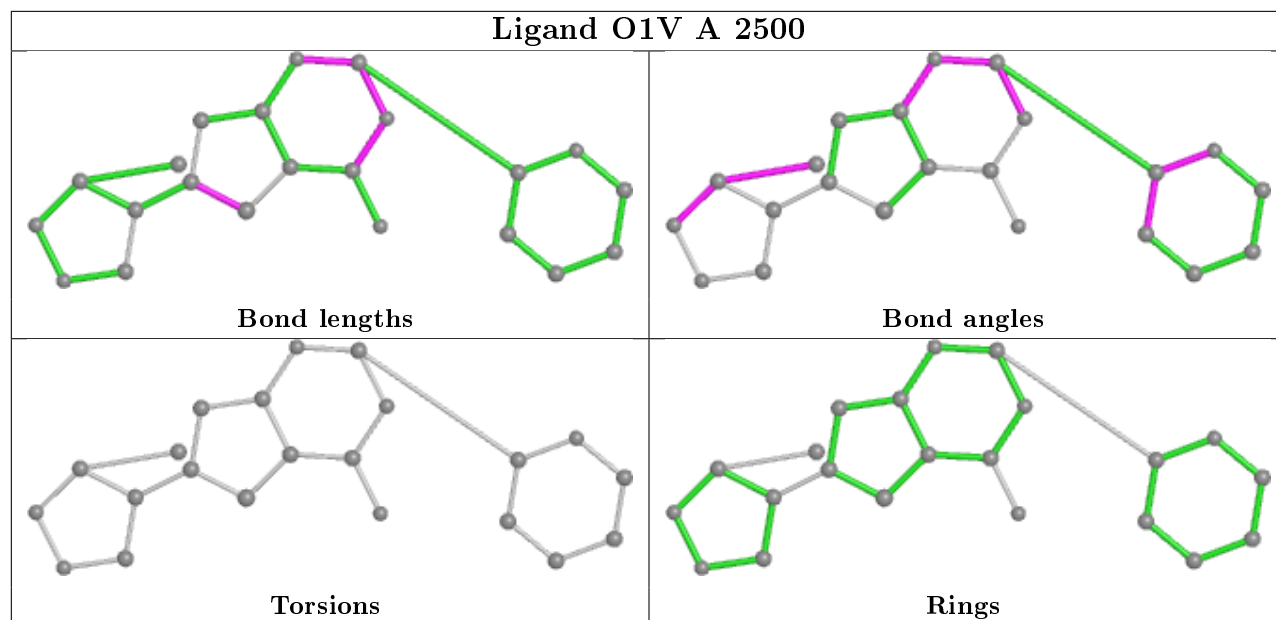
2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	2500	O1V	1	0
2	B	2500	O1V	2	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, 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	A	386/405 (95%)	-0.13	7 (1%) 68 67	62, 87, 118, 146	0
1	B	369/405 (91%)	-0.07	10 (2%) 54 52	63, 89, 127, 165	0
1	C	375/405 (92%)	-0.03	10 (2%) 54 52	65, 93, 126, 151	0
1	D	382/405 (94%)	-0.06	6 (1%) 72 70	63, 95, 134, 160	0
All	All	1512/1620 (93%)	-0.07	33 (2%) 62 60	62, 91, 128, 165	0

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	270	TYR	5.2
1	D	357	ASP	3.9
1	C	246	HIS	3.8
1	B	130	LYS	3.3
1	B	242	THR	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](#)

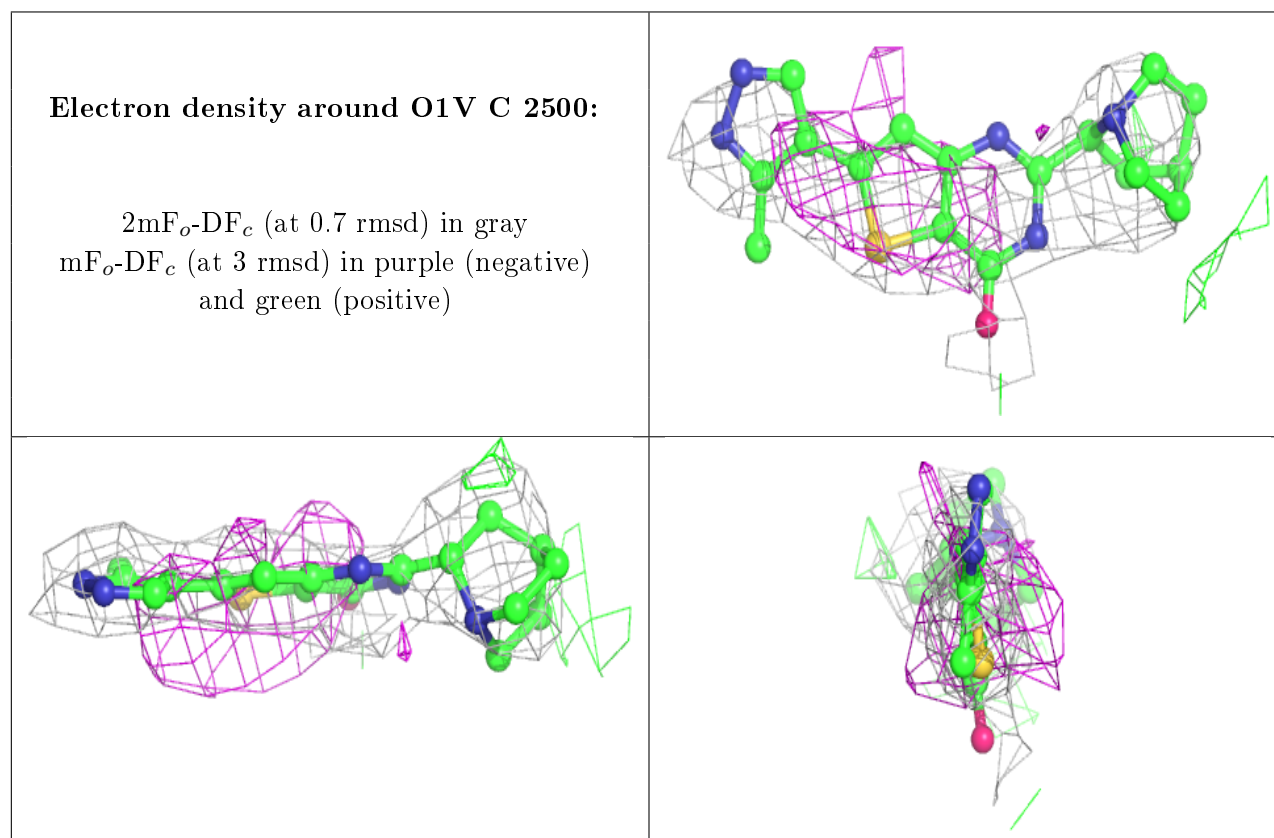
There are no carbohydrates in this entry.

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.

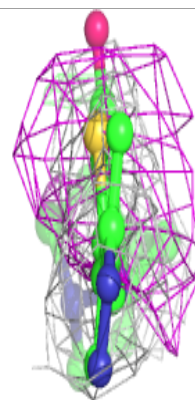
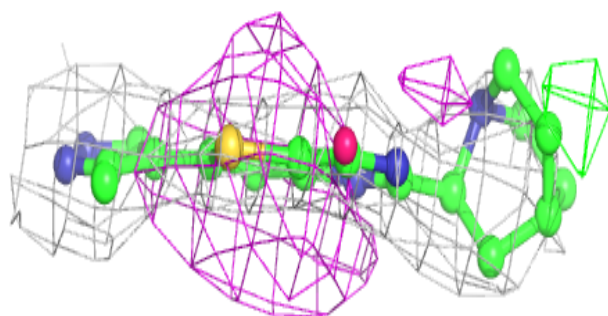
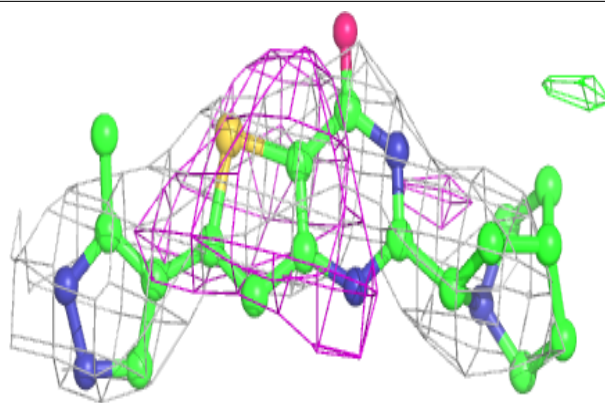
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	O1V	C	2500	24/24	0.86	0.37	82,138,161,165	0
2	O1V	A	2500	24/24	0.88	0.37	86,133,161,165	0
2	O1V	B	2500	24/24	0.92	0.21	74,94,123,126	0
2	O1V	D	2500	24/24	0.93	0.22	75,97,117,122	0

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.

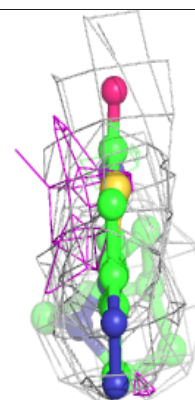
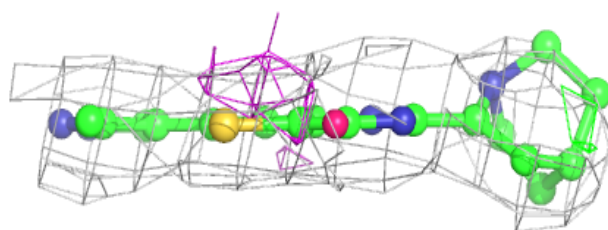
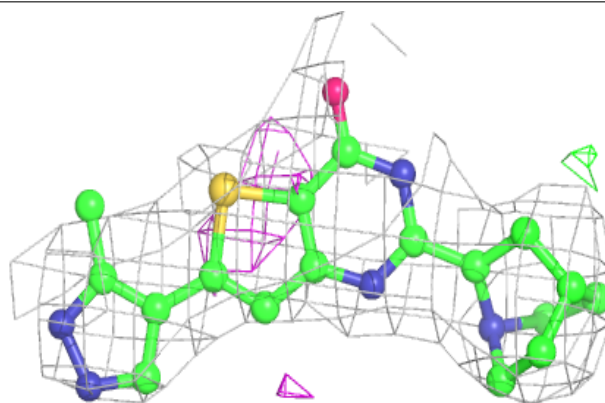


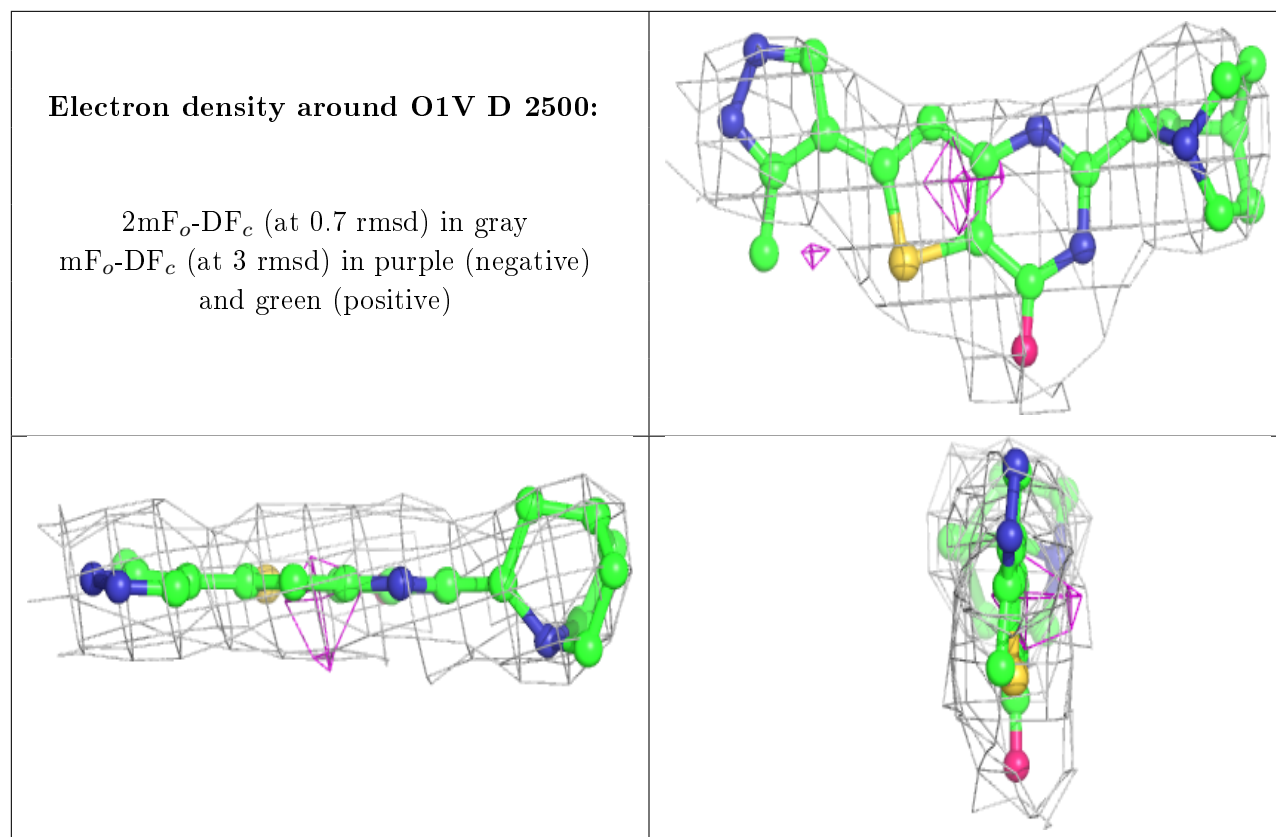
Electron density around O1V A 2500:

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

**Electron density around O1V 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)





6.5 Other polymers [i](#)

There are no such residues in this entry.