



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 22, 2021 – 10:51 AM EST

PDB ID : 6VXU  
Title : Structure of Human Vaccinia-related Kinase 1 (VRK1) bound to ACH471  
Authors : dos Reis, C.V.; Dutra, L.A.; Gama, F.H.; Mascarello, A.; Azevedo, H.;  
Guimaraes, C.R.; Massirer, K.B.; Arruda, P.; Edwards, A.M.; Counago, R.M.;  
Structural Genomics Consortium (SGC)  
Deposited on : 2020-02-24  
Resolution : 2.00 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](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 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 : 2.17.1  
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) : 2.17.1

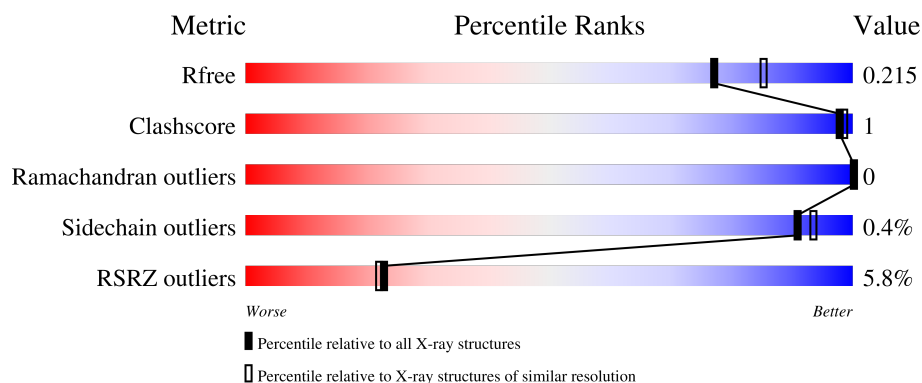
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	364	<div> <div>2%</div> <div> <div></div> <div>87%</div> <div>12%</div> </div> </div>
1	B	364	<div> <div>10%</div> <div> <div></div> <div>82%</div> <div>16%</div> </div> </div>
1	C	364	<div> <div></div> <div> <div></div> <div>85%</div> <div>15%</div> </div> </div>
1	D	364	<div> <div>8%</div> <div> <div></div> <div>86%</div> <div>13%</div> </div> </div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 10525 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 Serine/threonine-protein kinase VRK1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	320	Total	C	N	O	S	0	0	0
			2474	1587	416	459	12			
1	B	305	Total	C	N	O	S	0	0	0
			2388	1532	404	439	13			
1	C	311	Total	C	N	O	S	0	0	0
			2443	1563	410	458	12			
1	D	315	Total	C	N	O	S	0	0	0
			2438	1563	415	447	13			

There are 52 discrepancies between the modelled and reference sequences:

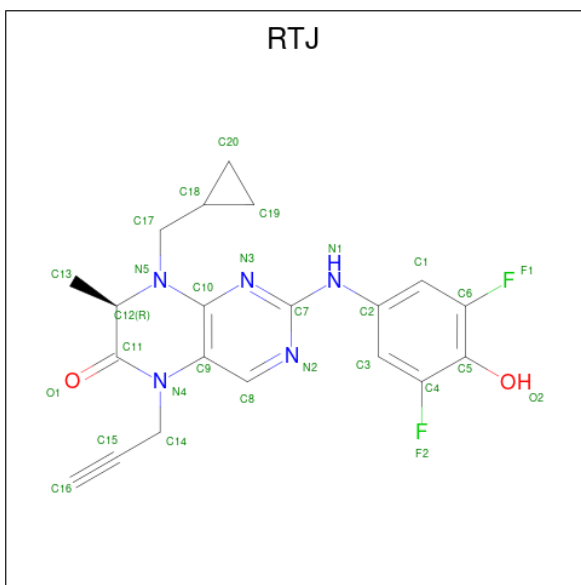
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	SER	-	expression tag	UNP Q99986
A	2	MET	-	expression tag	UNP Q99986
A	34	ALA	LYS	engineered mutation	UNP Q99986
A	35	ALA	LYS	engineered mutation	UNP Q99986
A	36	ALA	GLU	engineered mutation	UNP Q99986
A	212	ALA	GLU	engineered mutation	UNP Q99986
A	214	ALA	LYS	engineered mutation	UNP Q99986
A	215	ALA	GLU	engineered mutation	UNP Q99986
A	292	ALA	GLU	engineered mutation	UNP Q99986
A	293	ALA	LYS	engineered mutation	UNP Q99986
A	295	ALA	LYS	engineered mutation	UNP Q99986
A	359	ALA	LYS	engineered mutation	UNP Q99986
A	360	ALA	LYS	engineered mutation	UNP Q99986
B	1	SER	-	expression tag	UNP Q99986
B	2	MET	-	expression tag	UNP Q99986
B	34	ALA	LYS	engineered mutation	UNP Q99986
B	35	ALA	LYS	engineered mutation	UNP Q99986
B	36	ALA	GLU	engineered mutation	UNP Q99986
B	212	ALA	GLU	engineered mutation	UNP Q99986
B	214	ALA	LYS	engineered mutation	UNP Q99986
B	215	ALA	GLU	engineered mutation	UNP Q99986

*Continued on next page...*

*Continued from previous page...*

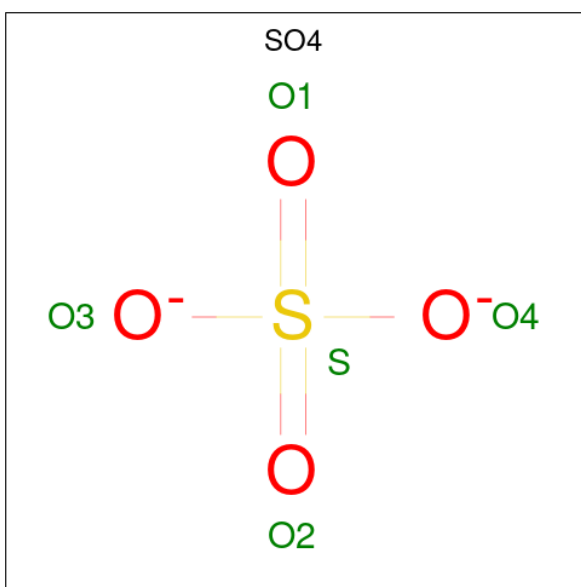
Chain	Residue	Modelled	Actual	Comment	Reference
B	292	ALA	GLU	engineered mutation	UNP Q99986
B	293	ALA	LYS	engineered mutation	UNP Q99986
B	295	ALA	LYS	engineered mutation	UNP Q99986
B	359	ALA	LYS	engineered mutation	UNP Q99986
B	360	ALA	LYS	engineered mutation	UNP Q99986
C	1	SER	-	expression tag	UNP Q99986
C	2	MET	-	expression tag	UNP Q99986
C	34	ALA	LYS	engineered mutation	UNP Q99986
C	35	ALA	LYS	engineered mutation	UNP Q99986
C	36	ALA	GLU	engineered mutation	UNP Q99986
C	212	ALA	GLU	engineered mutation	UNP Q99986
C	214	ALA	LYS	engineered mutation	UNP Q99986
C	215	ALA	GLU	engineered mutation	UNP Q99986
C	292	ALA	GLU	engineered mutation	UNP Q99986
C	293	ALA	LYS	engineered mutation	UNP Q99986
C	295	ALA	LYS	engineered mutation	UNP Q99986
C	359	ALA	LYS	engineered mutation	UNP Q99986
C	360	ALA	LYS	engineered mutation	UNP Q99986
D	1	SER	-	expression tag	UNP Q99986
D	2	MET	-	expression tag	UNP Q99986
D	34	ALA	LYS	engineered mutation	UNP Q99986
D	35	ALA	LYS	engineered mutation	UNP Q99986
D	36	ALA	GLU	engineered mutation	UNP Q99986
D	212	ALA	GLU	engineered mutation	UNP Q99986
D	214	ALA	LYS	engineered mutation	UNP Q99986
D	215	ALA	GLU	engineered mutation	UNP Q99986
D	292	ALA	GLU	engineered mutation	UNP Q99986
D	293	ALA	LYS	engineered mutation	UNP Q99986
D	295	ALA	LYS	engineered mutation	UNP Q99986
D	359	ALA	LYS	engineered mutation	UNP Q99986
D	360	ALA	LYS	engineered mutation	UNP Q99986

- Molecule 2 is (7R)-8-(cyclopropylmethyl)-2-[(3,5-difluoro-4-hydroxyphenyl)amino]-7-methyl-5-(prop-2-yn-1-yl)-7,8-dihydropteridin-6(5H)-one (three-letter code: RTJ) (formula: C<sub>20</sub>H<sub>19</sub>F<sub>2</sub>N<sub>5</sub>O<sub>2</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	F	N	O	0	0
			29	20	2	5	2		
2	B	1	Total	C	F	N	O	0	0
			29	20	2	5	2		
2	D	1	Total	C	F	N	O	0	0
			29	20	2	5	2		

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



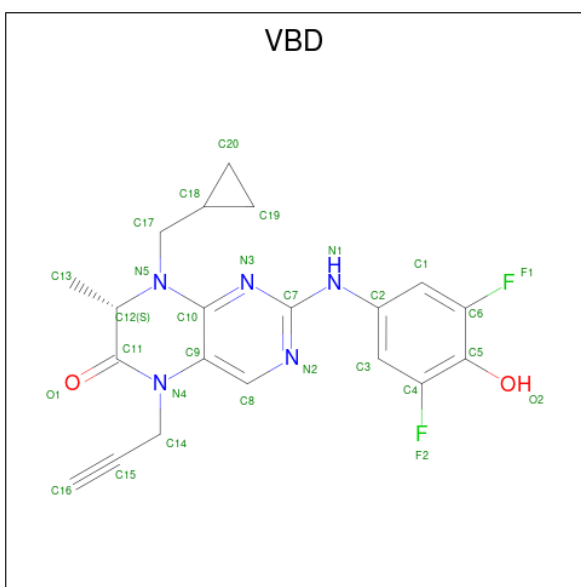
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	O S	0	0
			5	4 1		

*Continued on next page...*

*Continued from previous page...*

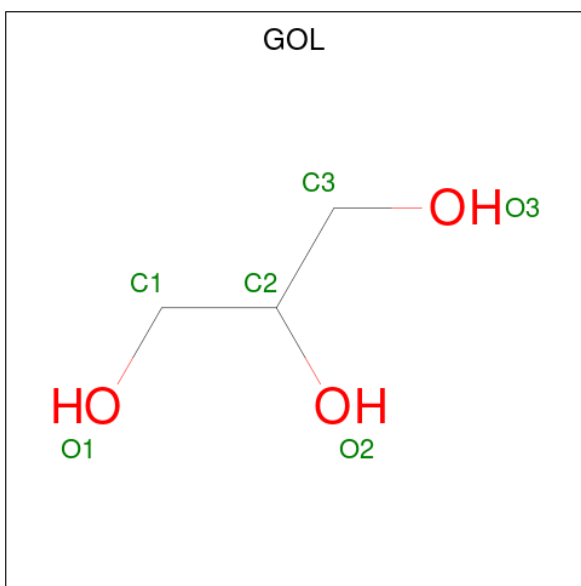
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total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

- Molecule 4 is (7S)-8-(cyclopropylmethyl)-2-[(3,5-difluoro-4-hydroxyphenyl)amino]-7-methyl-5-(prop-2-yn-1-yl)-7,8-dihydropteridin-6(5H)-one (three-letter code: VBD) (formula: C<sub>20</sub>H<sub>19</sub>F<sub>2</sub>N<sub>5</sub>O<sub>2</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	C	1	Total	C	F	N	O	0	0
			29	20	2	5	2		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



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

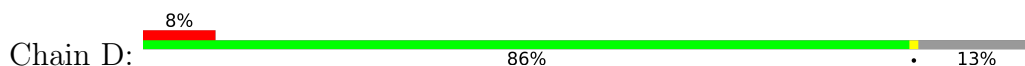
- Molecule 6 is water.

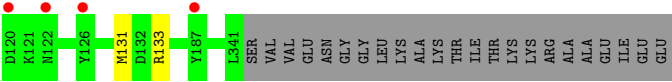
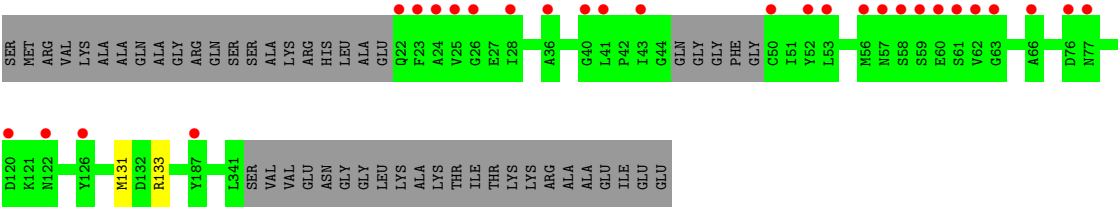
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	158	Total 158	O 158	0	0
6	B	138	Total 138	O 138	0	0
6	C	145	Total 145	O 145	0	0
6	D	149	Total 149	O 149	0	0





- Molecule 1: Serine/threonine-protein kinase VRK1





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	91.77Å 96.30Å 191.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.00 48.15 – 2.00	Depositor EDS
% Data completeness (in resolution range)	97.6 (30.00-2.00) 97.6 (48.15-2.00)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.71 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, $R_{free}$	0.184 , 0.203 0.197 , 0.215	Depositor DCC
$R_{free}$ test set	5598 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.6	Xtriage
Anisotropy	0.368	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 49.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.016 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	10525	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.39% 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, GOL, RTJ, VBD

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	0.62	0/2534	0.73	0/3443
1	B	0.59	0/2444	0.71	0/3316
1	C	0.62	0/2500	0.73	0/3394
1	D	0.64	0/2497	0.73	0/3395
All	All	0.62	0/9975	0.72	0/13548

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	A	2474	0	2357	1	0
1	B	2388	0	2280	5	0
1	C	2443	0	2350	1	0
1	D	2438	0	2310	1	0
2	A	29	0	0	1	0
2	B	29	0	0	2	0
2	D	29	0	0	1	0
3	A	10	0	0	0	0
3	B	15	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	15	0	0	0	0
3	D	30	0	0	0	0
4	C	29	0	0	1	0
5	C	6	0	8	0	0
6	A	158	0	0	0	0
6	B	138	0	0	0	0
6	C	145	0	0	0	0
6	D	149	0	0	0	0
All	All	10525	0	9305	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 1.

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)
2:B:401:RTJ:C18	2:B:401:RTJ:C13	2.91	0.48
2:B:401:RTJ:N3	2:B:401:RTJ:C1	2.77	0.48
4:C:401:VBD:N3	4:C:401:VBD:C3	2.78	0.47
1:B:263:ASP:OD1	1:B:263:ASP:N	2.46	0.46
2:D:401:RTJ:N3	2:D:401:RTJ:C1	2.79	0.46
1:B:41:LEU:HA	1:B:42:PRO:HD3	1.74	0.44
2:A:401:RTJ:N3	2:A:401:RTJ:C3	2.83	0.42
1:B:70:VAL:O	1:B:70:VAL:HG13	2.20	0.42
1:C:204:TYR:CE2	1:C:205:CYS:HB2	2.55	0.42
1:B:42:PRO:HG3	1:B:52:TYR:CE2	2.55	0.42
1:B:291:PRO:C	1:B:293:ALA:H	2.24	0.41
1:A:204:TYR:CE2	1:A:205:CYS:HB2	2.55	0.41
1:D:131:MET:HE3	1:D:131:MET:HB3	1.93	0.41

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	316/364 (87%)	309 (98%)	7 (2%)	0	100	100
1	B	295/364 (81%)	287 (97%)	8 (3%)	0	100	100
1	C	305/364 (84%)	300 (98%)	5 (2%)	0	100	100
1	D	311/364 (85%)	306 (98%)	5 (2%)	0	100	100
All	All	1227/1456 (84%)	1202 (98%)	25 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	247/304 (81%)	246 (100%)	1 (0%)	91	93
1	B	240/304 (79%)	239 (100%)	1 (0%)	91	93
1	C	252/304 (83%)	251 (100%)	1 (0%)	91	93
1	D	240/304 (79%)	239 (100%)	1 (0%)	91	93
All	All	979/1216 (80%)	975 (100%)	4 (0%)	91	93

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	133	ARG
1	B	133	ARG
1	C	133	ARG
1	D	133	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 ⓘ

19 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	D	404	-	4,4,4	0.34	0	6,6,6	0.06	0
3	SO4	C	402	-	4,4,4	0.39	0	6,6,6	0.06	0
5	GOL	C	405	-	5,5,5	0.28	0	5,5,5	0.19	0
2	RTJ	D	401	-	31,32,32	0.65	0	36,47,47	1.06	5 (13%)
3	SO4	C	403	-	4,4,4	0.36	0	6,6,6	0.04	0
3	SO4	A	402	-	4,4,4	0.37	0	6,6,6	0.05	0
2	RTJ	A	401	-	31,32,32	0.35	0	36,47,47	1.06	4 (11%)
3	SO4	A	403	-	4,4,4	0.38	0	6,6,6	0.06	0
3	SO4	D	406	-	4,4,4	0.39	0	6,6,6	0.05	0
3	SO4	B	404	-	4,4,4	0.39	0	6,6,6	0.06	0
2	RTJ	B	401	-	31,32,32	0.36	0	36,47,47	1.01	4 (11%)
3	SO4	B	403	-	4,4,4	0.38	0	6,6,6	0.04	0
3	SO4	C	404	-	4,4,4	0.37	0	6,6,6	0.04	0
3	SO4	B	402	-	4,4,4	0.25	0	6,6,6	0.13	0
4	VBD	C	401	-	31,32,32	0.36	0	36,47,47	1.07	4 (11%)
3	SO4	D	403	-	4,4,4	0.18	0	6,6,6	0.10	0
3	SO4	D	407	-	4,4,4	0.38	0	6,6,6	0.05	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	SO4	D	402	-	4,4,4	0.37	0	6,6,6	0.05	0
3	SO4	D	405	-	4,4,4	0.39	0	6,6,6	0.05	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	RTJ	B	401	-	-	0/10/33/33	0/4/4/4
4	VBD	C	401	-	-	2/10/33/33	0/4/4/4
2	RTJ	A	401	-	-	0/10/33/33	0/4/4/4
5	GOL	C	405	-	-	2/4/4/4	-
2	RTJ	D	401	-	-	1/10/33/33	0/4/4/4

There are no bond length outliers.

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	401	VBD	F1-C6-C5	2.74	119.17	117.13
4	C	401	VBD	C1-C6-C5	-2.59	121.76	123.79
2	A	401	RTJ	C1-C6-C5	-2.57	121.78	123.79
2	D	401	RTJ	C3-C4-C5	-2.52	121.82	123.79
2	B	401	RTJ	C3-C4-C5	-2.46	121.87	123.79
2	D	401	RTJ	F2-C4-C5	2.42	118.93	117.13
2	A	401	RTJ	F1-C6-C5	2.38	118.90	117.13
2	B	401	RTJ	C1-C6-C5	-2.37	121.94	123.79
2	A	401	RTJ	C3-C4-C5	-2.33	121.97	123.79
4	C	401	VBD	C3-C4-C5	-2.31	121.98	123.79
2	B	401	RTJ	F2-C4-C5	2.20	118.77	117.13
4	C	401	VBD	F2-C4-C5	2.14	118.72	117.13
2	B	401	RTJ	F1-C6-C5	2.12	118.71	117.13
2	D	401	RTJ	C1-C6-C5	-2.10	122.15	123.79
2	D	401	RTJ	C9-C10-N3	-2.10	120.58	123.67
2	A	401	RTJ	F2-C4-C5	2.08	118.67	117.13
2	D	401	RTJ	F1-C6-C5	2.06	118.66	117.13

There are no chirality outliers.

All (5) torsion outliers are listed below:



Mol	Chain	Res	Type	Atoms
5	C	405	GOL	C1-C2-C3-O3
4	C	401	VBD	C18-C17-N5-C10
4	C	401	VBD	C18-C17-N5-C12
5	C	405	GOL	O2-C2-C3-O3
2	D	401	RTJ	C18-C17-N5-C12

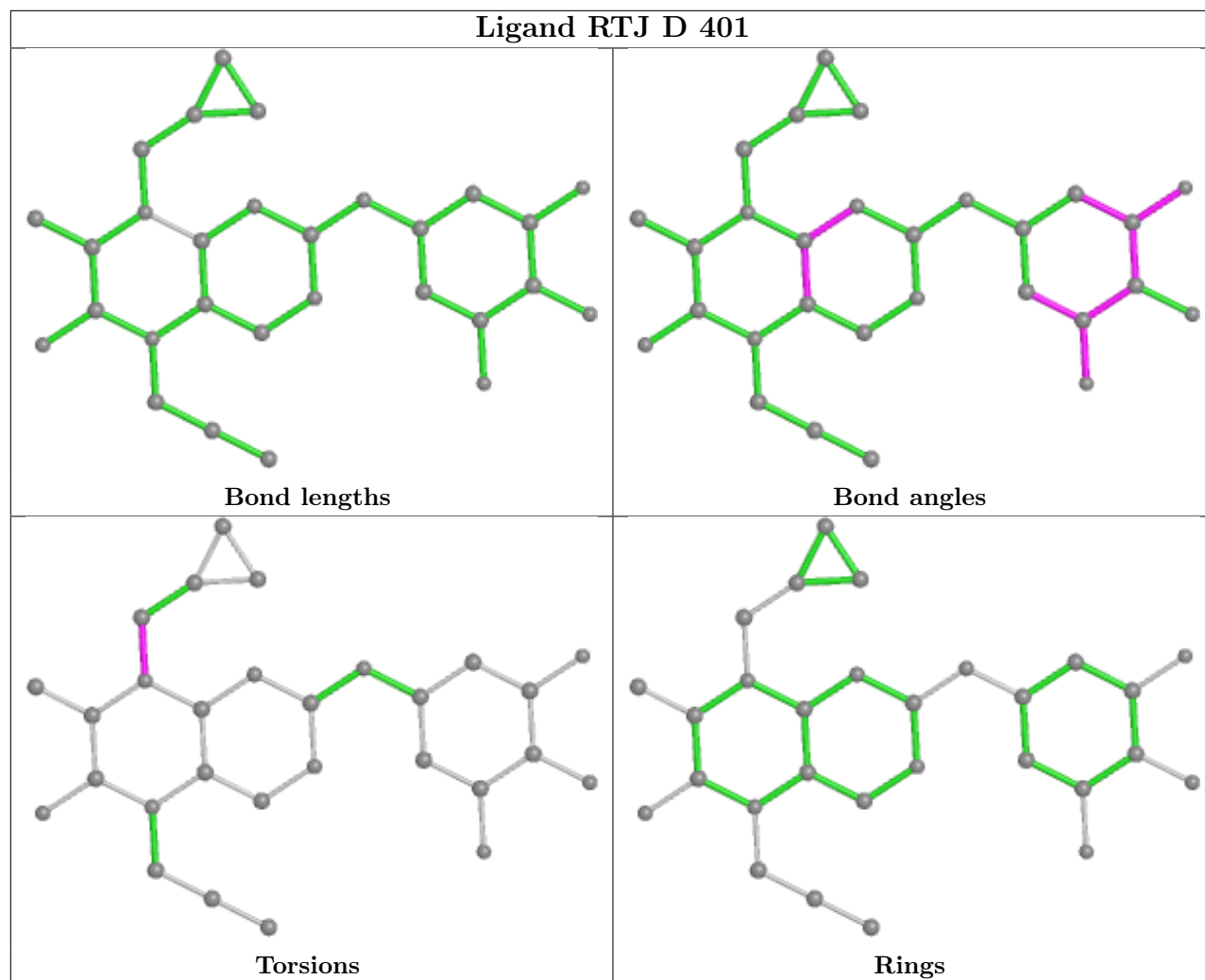
There are no ring outliers.

4 monomers are involved in 5 short contacts:

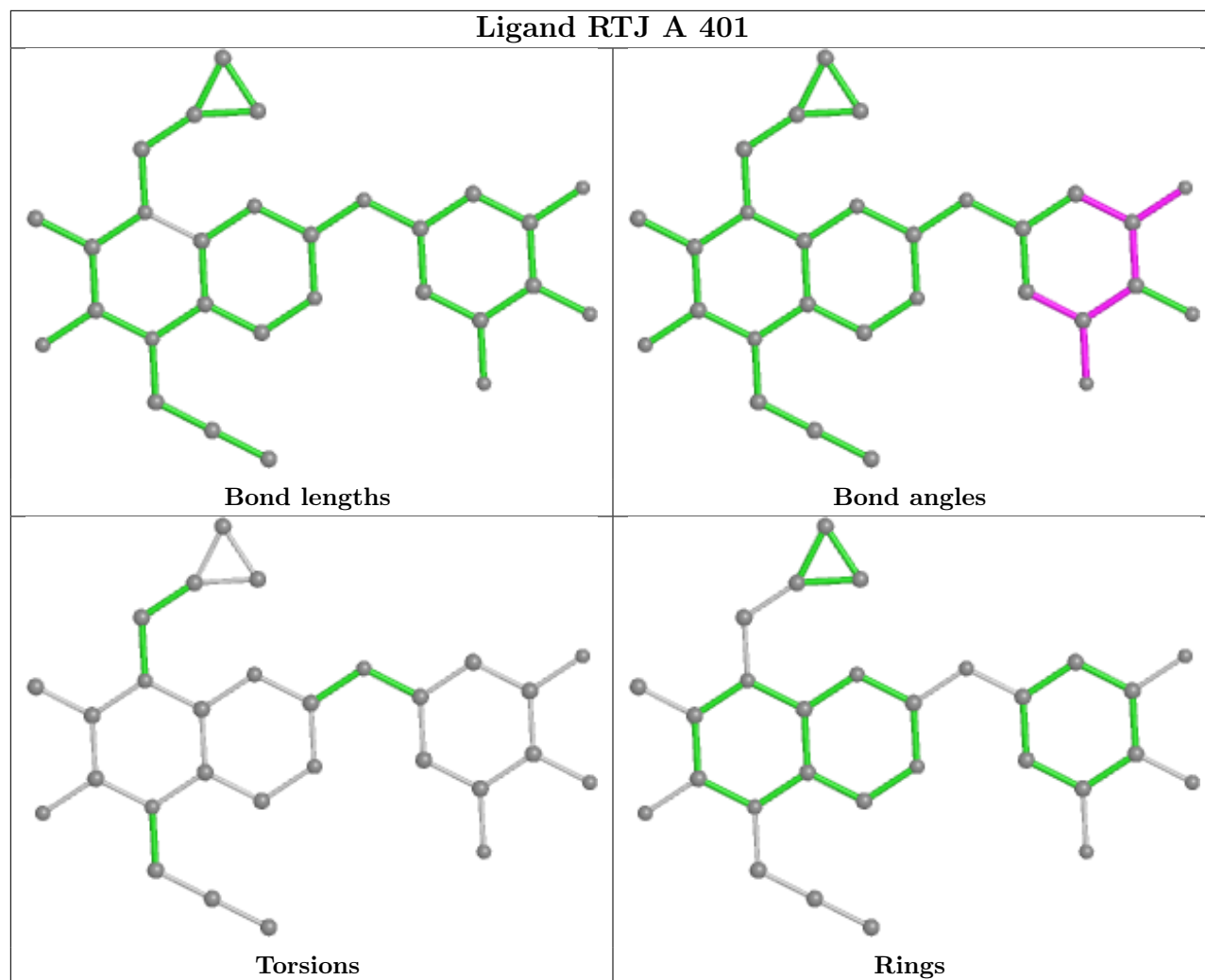
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	401	RTJ	1	0
2	A	401	RTJ	1	0
2	B	401	RTJ	2	0
4	C	401	VBD	1	0

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.

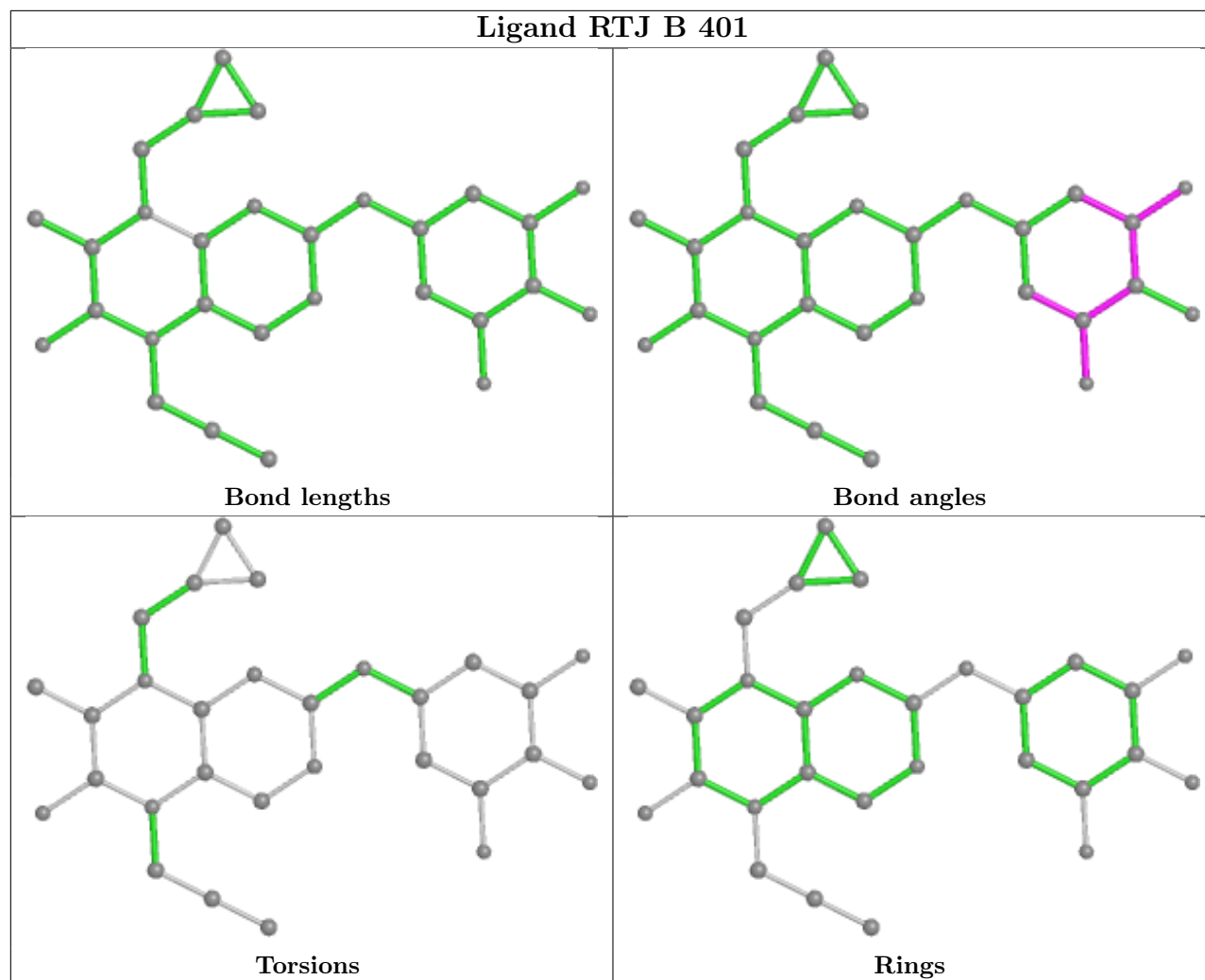
## Ligand RTJ D 401

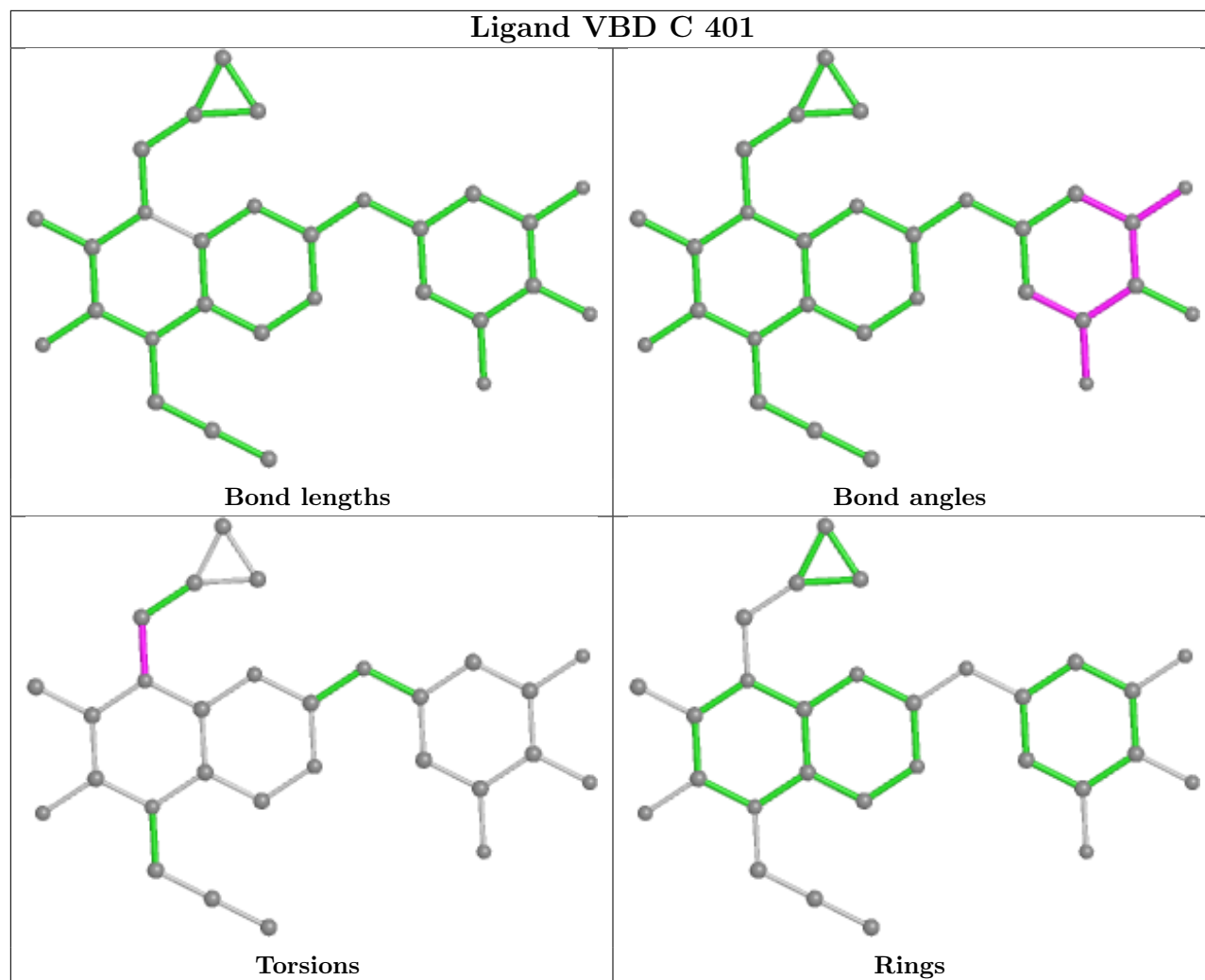


## Ligand RTJ A 401



## Ligand RTJ B 401





## 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

### 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, 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	320/364 (87%)	0.12	6 (1%) 66 65	28, 39, 60, 90	0
1	B	305/364 (83%)	0.54	37 (12%) 4 3	26, 45, 86, 109	0
1	C	311/364 (85%)	0.08	1 (0%) 94 93	26, 39, 60, 79	0
1	D	315/364 (86%)	0.45	28 (8%) 9 8	25, 40, 76, 98	0
All	All	1251/1456 (85%)	0.30	72 (5%) 23 22	25, 40, 75, 109	0

All (72) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	24	ALA	6.2
1	D	23	PHE	5.2
1	A	220	CYS	5.0
1	B	43	ILE	4.8
1	B	62	VAL	4.7
1	B	23	PHE	4.7
1	B	50	CYS	4.5
1	D	24	ALA	4.3
1	D	26	GLY	4.2
1	D	126	TYR	4.1
1	B	61	SER	4.0
1	B	187	TYR	3.8
1	D	43	ILE	3.8
1	B	57	ASN	3.8
1	B	56	MET	3.8
1	B	120	ASP	3.8
1	B	63	GLY	3.7
1	D	63	GLY	3.7
1	B	126	TYR	3.6
1	D	59	SER	3.6
1	B	25	VAL	3.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	26	GLY	3.5
1	B	66	ALA	3.3
1	B	64	SER	3.2
1	B	42	PRO	3.1
1	D	120	ASP	3.1
1	D	62	VAL	3.1
1	B	52	TYR	3.0
1	A	48	PHE	2.9
1	B	35	ALA	2.9
1	D	52	TYR	2.9
1	B	53	LEU	2.9
1	B	67	PRO	2.9
1	B	29	ILE	2.8
1	B	41	LEU	2.8
1	D	28	ILE	2.8
1	D	57	ASN	2.8
1	D	76	ASP	2.8
1	B	136	SER	2.7
1	D	77	ASN	2.7
1	B	68	CYS	2.7
1	A	221	HIS	2.7
1	D	22	GLN	2.6
1	B	28	ILE	2.6
1	B	30	THR	2.6
1	D	25	VAL	2.5
1	D	40	GLY	2.5
1	A	217	PRO	2.5
1	A	44	GLY	2.5
1	D	187	TYR	2.4
1	B	39	VAL	2.4
1	D	58	SER	2.4
1	D	61	SER	2.4
1	A	325	LEU	2.3
1	D	50	CYS	2.3
1	B	58	SER	2.3
1	D	56	MET	2.2
1	B	55	ASP	2.2
1	D	53	LEU	2.2
1	B	40	GLY	2.2
1	B	34	ALA	2.2
1	B	293	ALA	2.2
1	B	44	GLY	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	101	ARG	2.2
1	D	66	ALA	2.1
1	D	60	GLU	2.1
1	D	36	ALA	2.1
1	D	122	ASN	2.1
1	B	322	ASP	2.1
1	C	217	PRO	2.1
1	B	54	ALA	2.0
1	D	41	LEU	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 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(Å <sup>2</sup> )	Q<0.9
3	SO4	A	403	5/5	0.80	0.24	94,94,98,100	0
4	VBD	C	401	29/29	0.81	0.27	52,59,65,67	0
3	SO4	D	405	5/5	0.84	0.16	96,98,100,100	0
3	SO4	D	407	5/5	0.86	0.16	93,94,95,96	0
3	SO4	D	404	5/5	0.87	0.20	72,72,78,79	0
3	SO4	C	404	5/5	0.87	0.20	86,89,93,93	0
3	SO4	C	403	5/5	0.88	0.30	70,77,81,84	0
2	RTJ	A	401	29/29	0.88	0.18	39,44,49,50	0
3	SO4	D	406	5/5	0.89	0.21	116,117,118,118	0
3	SO4	A	402	5/5	0.89	0.26	79,85,87,88	0
3	SO4	B	404	5/5	0.89	0.22	96,98,99,102	0
2	RTJ	B	401	29/29	0.90	0.26	45,58,65,67	0
3	SO4	D	402	5/5	0.91	0.16	73,73,78,81	0
5	GOL	C	405	6/6	0.91	0.15	38,46,46,47	0

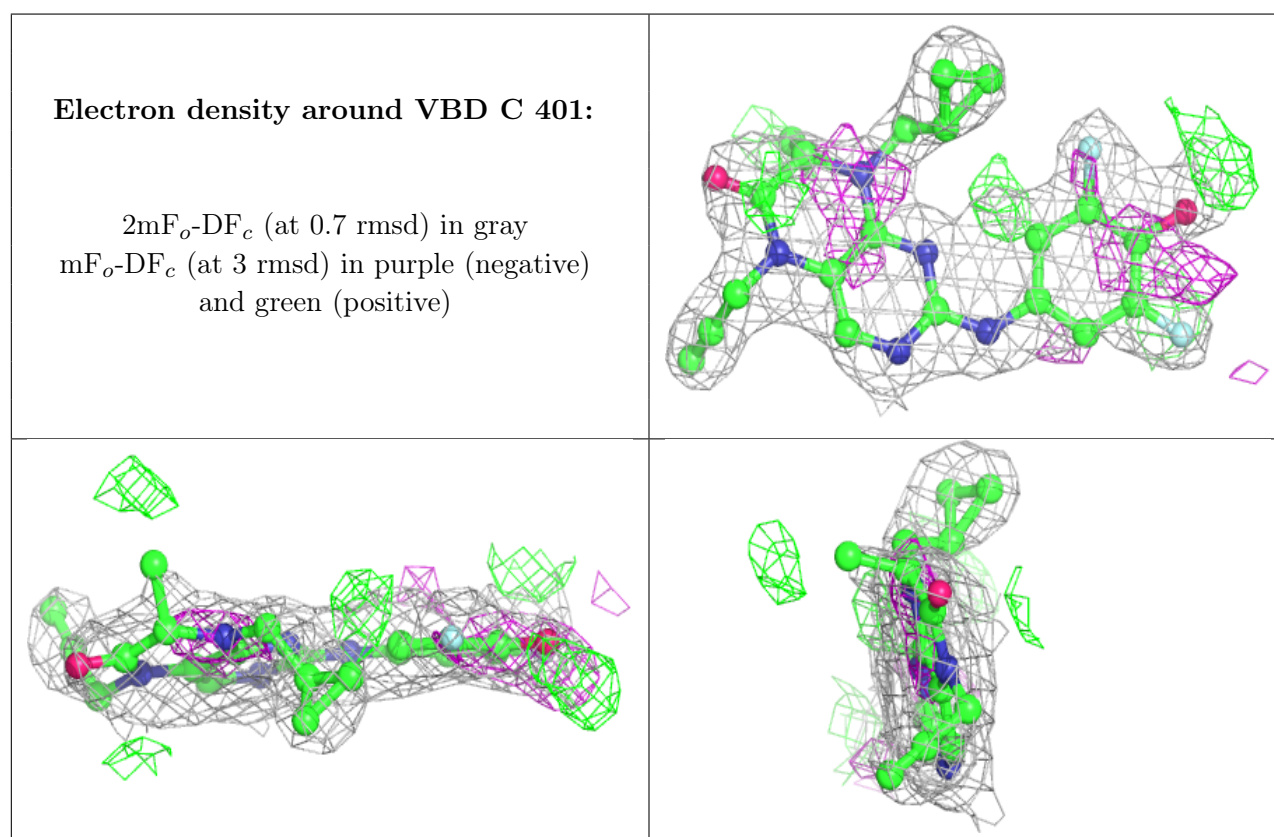
*Continued on next page...*



*Continued from previous page...*

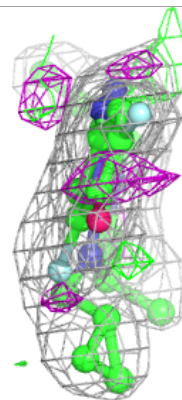
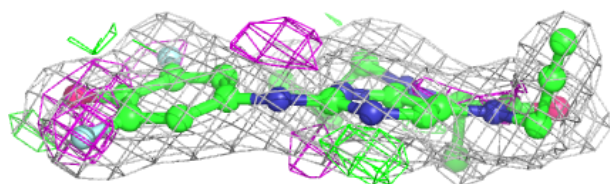
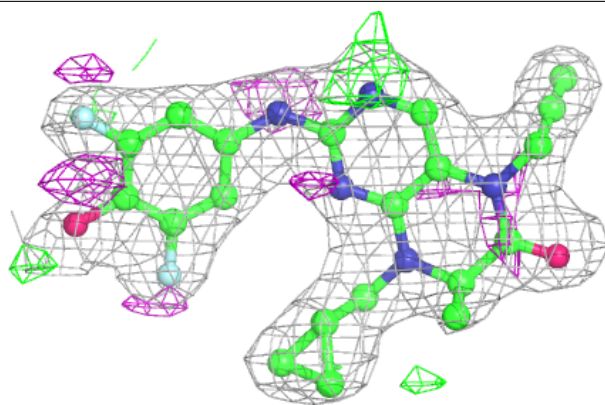
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	RTJ	D	401	29/29	0.92	0.18	37,45,50,52	0
3	SO4	B	402	5/5	0.93	0.13	77,80,85,86	0
3	SO4	B	403	5/5	0.94	0.13	92,92,93,95	0
3	SO4	D	403	5/5	0.95	0.12	63,65,68,69	0
3	SO4	C	402	5/5	0.97	0.13	61,64,65,66	0

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.

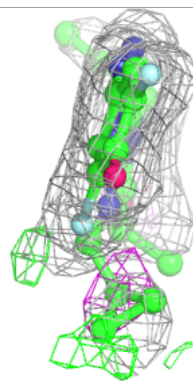
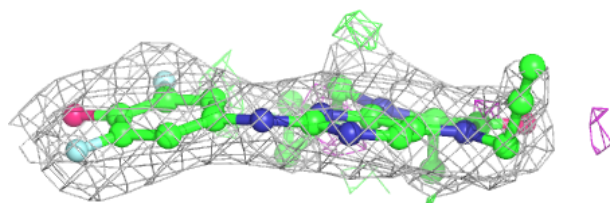
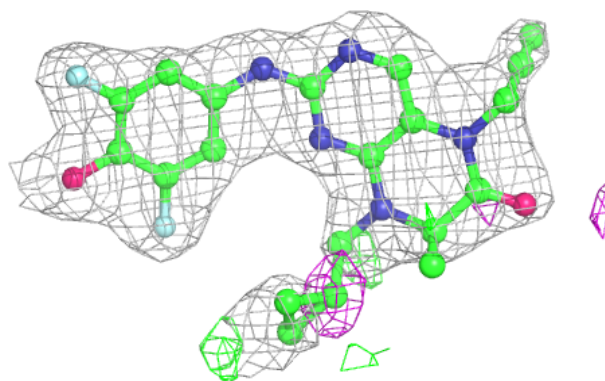


**Electron density around RTJ A 401:**

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

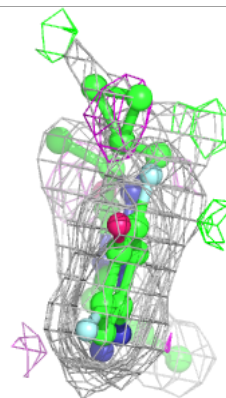
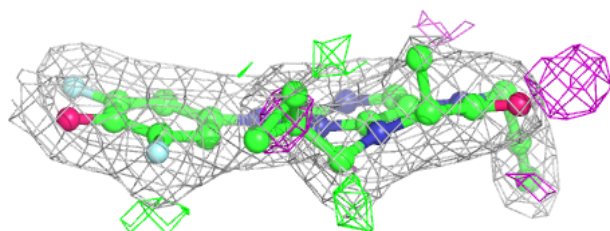
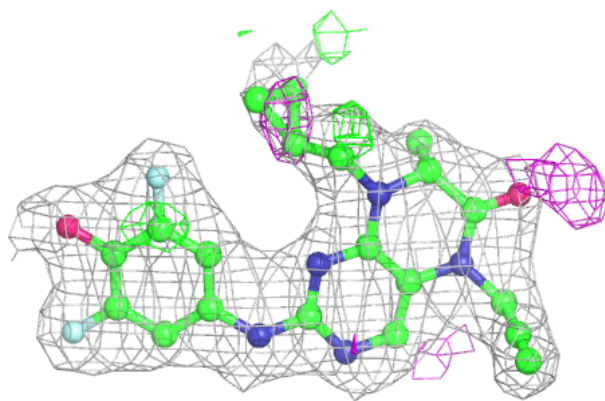
**Electron density around RTJ B 401:**

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



**Electron density around RTJ D 401:**

$2mF_o-DF_c$  (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.