



# wwPDB X-ray Structure Validation Summary Report

Jun 26, 2024 – 06:49 PM EDT

PDB ID : 8V2F  
Title : Crystal structure of IRAK4 kinase domain with compound 9  
Authors : Weiss, M.M.; Zheng, X.; Browne, C.M.; Campbell, V.; Chen, D.; Enerson, B.;  
Fei, X.; Huang, X.; Klaus, C.R.; Li, H.; Mayo, M.; McDonald, A.A.; Paul, A.;  
Sharma, K.; Shi, Y.; Slavin, A.; Walter, D.M.; Yuan, K.; Zhang, Y.; Zhu, X.;  
Kelleher, J.; Ji, N.; Walker, D.; Mainolfi, N.  
Deposited on : 2023-11-22  
Resolution : 2.09 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](#) ①) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : 2.37.1  
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) : 2.37.1

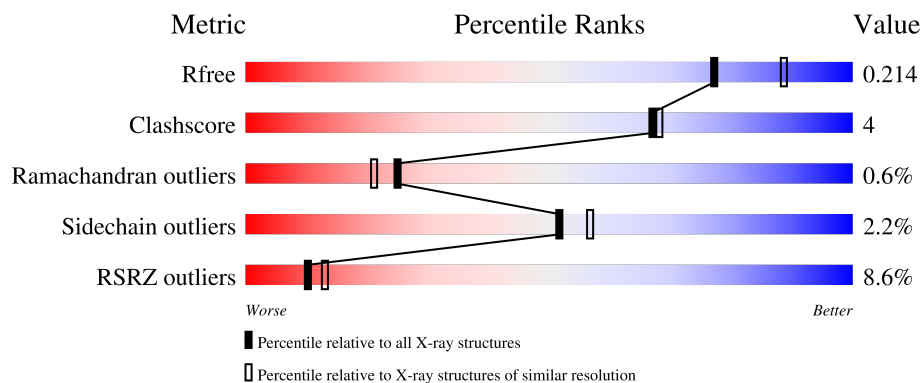
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.09 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	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  $\geq 3$ , 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	327	
1	B	327	
1	C	327	
1	D	327	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 10225 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 Interleukin-1 receptor-associated kinase 4.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
1	A	297	2353	1475	393	466	3	16	0	2	0
1	B	298	2358	1477	394	468	3	16	0	1	0
1	C	302	2392	1500	398	474	3	17	0	3	0
1	D	294	2328	1458	389	462	3	16	0	2	0

There are 104 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	134	MET	-	initiating methionine	UNP Q9NWZ3
A	135	SER	-	expression tag	UNP Q9NWZ3
A	136	TYR	-	expression tag	UNP Q9NWZ3
A	137	TYR	-	expression tag	UNP Q9NWZ3
A	138	HIS	-	expression tag	UNP Q9NWZ3
A	139	HIS	-	expression tag	UNP Q9NWZ3
A	140	HIS	-	expression tag	UNP Q9NWZ3
A	141	HIS	-	expression tag	UNP Q9NWZ3
A	142	HIS	-	expression tag	UNP Q9NWZ3
A	143	HIS	-	expression tag	UNP Q9NWZ3
A	144	ASP	-	expression tag	UNP Q9NWZ3
A	145	TYR	-	expression tag	UNP Q9NWZ3
A	146	ASP	-	expression tag	UNP Q9NWZ3
A	147	ILE	-	expression tag	UNP Q9NWZ3
A	148	PRO	-	expression tag	UNP Q9NWZ3
A	149	THR	-	expression tag	UNP Q9NWZ3
A	150	THR	-	expression tag	UNP Q9NWZ3
A	151	GLU	-	expression tag	UNP Q9NWZ3
A	152	ASN	-	expression tag	UNP Q9NWZ3
A	153	LEU	-	expression tag	UNP Q9NWZ3
A	154	TYR	-	expression tag	UNP Q9NWZ3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	155	PHE	-	expression tag	UNP Q9NWZ3
A	156	GLN	-	expression tag	UNP Q9NWZ3
A	157	SER	-	expression tag	UNP Q9NWZ3
A	158	ILE	-	expression tag	UNP Q9NWZ3
A	159	ALA	-	expression tag	UNP Q9NWZ3
B	134	MET	-	initiating methionine	UNP Q9NWZ3
B	135	SER	-	expression tag	UNP Q9NWZ3
B	136	TYR	-	expression tag	UNP Q9NWZ3
B	137	TYR	-	expression tag	UNP Q9NWZ3
B	138	HIS	-	expression tag	UNP Q9NWZ3
B	139	HIS	-	expression tag	UNP Q9NWZ3
B	140	HIS	-	expression tag	UNP Q9NWZ3
B	141	HIS	-	expression tag	UNP Q9NWZ3
B	142	HIS	-	expression tag	UNP Q9NWZ3
B	143	HIS	-	expression tag	UNP Q9NWZ3
B	144	ASP	-	expression tag	UNP Q9NWZ3
B	145	TYR	-	expression tag	UNP Q9NWZ3
B	146	ASP	-	expression tag	UNP Q9NWZ3
B	147	ILE	-	expression tag	UNP Q9NWZ3
B	148	PRO	-	expression tag	UNP Q9NWZ3
B	149	THR	-	expression tag	UNP Q9NWZ3
B	150	THR	-	expression tag	UNP Q9NWZ3
B	151	GLU	-	expression tag	UNP Q9NWZ3
B	152	ASN	-	expression tag	UNP Q9NWZ3
B	153	LEU	-	expression tag	UNP Q9NWZ3
B	154	TYR	-	expression tag	UNP Q9NWZ3
B	155	PHE	-	expression tag	UNP Q9NWZ3
B	156	GLN	-	expression tag	UNP Q9NWZ3
B	157	SER	-	expression tag	UNP Q9NWZ3
B	158	ILE	-	expression tag	UNP Q9NWZ3
B	159	ALA	-	expression tag	UNP Q9NWZ3
C	134	MET	-	initiating methionine	UNP Q9NWZ3
C	135	SER	-	expression tag	UNP Q9NWZ3
C	136	TYR	-	expression tag	UNP Q9NWZ3
C	137	TYR	-	expression tag	UNP Q9NWZ3
C	138	HIS	-	expression tag	UNP Q9NWZ3
C	139	HIS	-	expression tag	UNP Q9NWZ3
C	140	HIS	-	expression tag	UNP Q9NWZ3
C	141	HIS	-	expression tag	UNP Q9NWZ3
C	142	HIS	-	expression tag	UNP Q9NWZ3
C	143	HIS	-	expression tag	UNP Q9NWZ3
C	144	ASP	-	expression tag	UNP Q9NWZ3

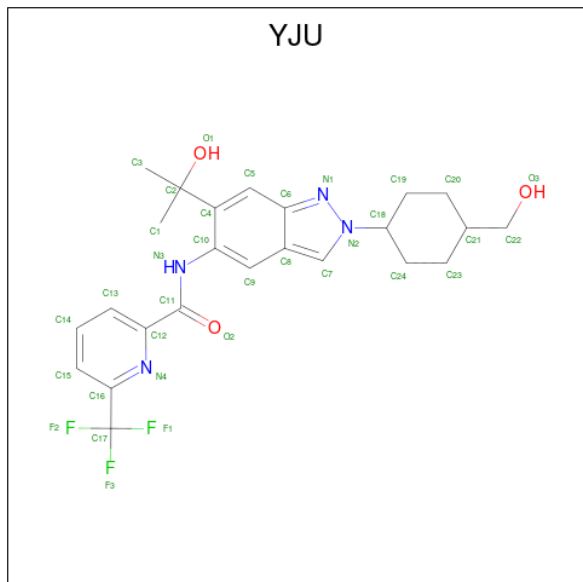
*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	145	TYR	-	expression tag	UNP Q9NWZ3
C	146	ASP	-	expression tag	UNP Q9NWZ3
C	147	ILE	-	expression tag	UNP Q9NWZ3
C	148	PRO	-	expression tag	UNP Q9NWZ3
C	149	THR	-	expression tag	UNP Q9NWZ3
C	150	THR	-	expression tag	UNP Q9NWZ3
C	151	GLU	-	expression tag	UNP Q9NWZ3
C	152	ASN	-	expression tag	UNP Q9NWZ3
C	153	LEU	-	expression tag	UNP Q9NWZ3
C	154	TYR	-	expression tag	UNP Q9NWZ3
C	155	PHE	-	expression tag	UNP Q9NWZ3
C	156	GLN	-	expression tag	UNP Q9NWZ3
C	157	SER	-	expression tag	UNP Q9NWZ3
C	158	ILE	-	expression tag	UNP Q9NWZ3
C	159	ALA	-	expression tag	UNP Q9NWZ3
D	134	MET	-	initiating methionine	UNP Q9NWZ3
D	135	SER	-	expression tag	UNP Q9NWZ3
D	136	TYR	-	expression tag	UNP Q9NWZ3
D	137	TYR	-	expression tag	UNP Q9NWZ3
D	138	HIS	-	expression tag	UNP Q9NWZ3
D	139	HIS	-	expression tag	UNP Q9NWZ3
D	140	HIS	-	expression tag	UNP Q9NWZ3
D	141	HIS	-	expression tag	UNP Q9NWZ3
D	142	HIS	-	expression tag	UNP Q9NWZ3
D	143	HIS	-	expression tag	UNP Q9NWZ3
D	144	ASP	-	expression tag	UNP Q9NWZ3
D	145	TYR	-	expression tag	UNP Q9NWZ3
D	146	ASP	-	expression tag	UNP Q9NWZ3
D	147	ILE	-	expression tag	UNP Q9NWZ3
D	148	PRO	-	expression tag	UNP Q9NWZ3
D	149	THR	-	expression tag	UNP Q9NWZ3
D	150	THR	-	expression tag	UNP Q9NWZ3
D	151	GLU	-	expression tag	UNP Q9NWZ3
D	152	ASN	-	expression tag	UNP Q9NWZ3
D	153	LEU	-	expression tag	UNP Q9NWZ3
D	154	TYR	-	expression tag	UNP Q9NWZ3
D	155	PHE	-	expression tag	UNP Q9NWZ3
D	156	GLN	-	expression tag	UNP Q9NWZ3
D	157	SER	-	expression tag	UNP Q9NWZ3
D	158	ILE	-	expression tag	UNP Q9NWZ3
D	159	ALA	-	expression tag	UNP Q9NWZ3

- Molecule 2 is N-{2-[(1r,4r)-4-(hydroxymethyl)cyclohexyl]-6-(2-hydroxypropan-2-yl)-2H-in

dazol-5-yl}-6-(trifluoromethyl)pyridine-2-carboxamide (three-letter code: YJU) (formula:  $C_{24}H_{27}F_3N_4O_3$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	F	N	O		
2	A	1	Total 34	C 24	F 3	N 4	O 3	0	0
2	B	1	Total 34	C 24	F 3	N 4	O 3	0	0
2	C	1	Total 34	C 24	F 3	N 4	O 3	0	0
2	D	1	Total 34	C 24	F 3	N 4	O 3	0	0

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
3	A	1	Total 1	Cl 1	0	0
3	B	1	Total 1	Cl 1	0	0
3	C	1	Total 1	Cl 1	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



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

- Molecule 5 is water.

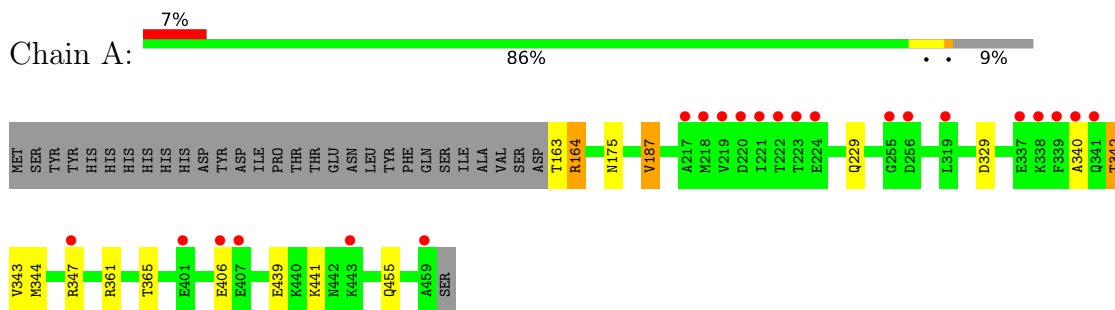
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	176	Total O 176 176	0	0
5	B	139	Total O 139 139	0	0
5	C	149	Total O 149 149	0	0
5	D	155	Total O 155 155	0	0



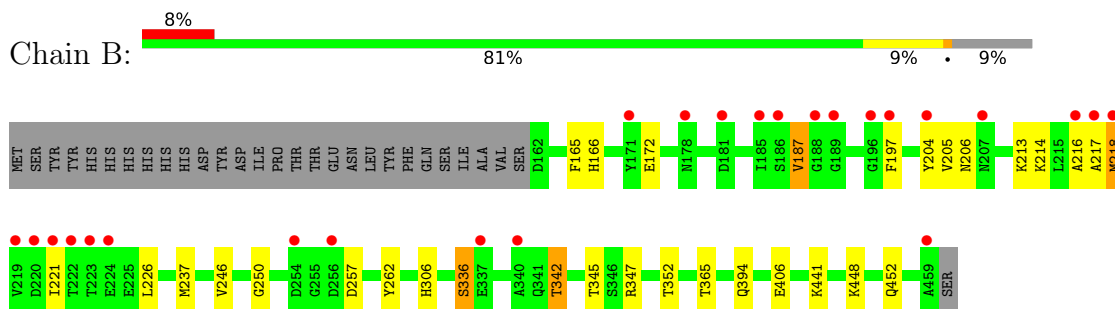
### 3 Residue-property plots [i](#)

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.

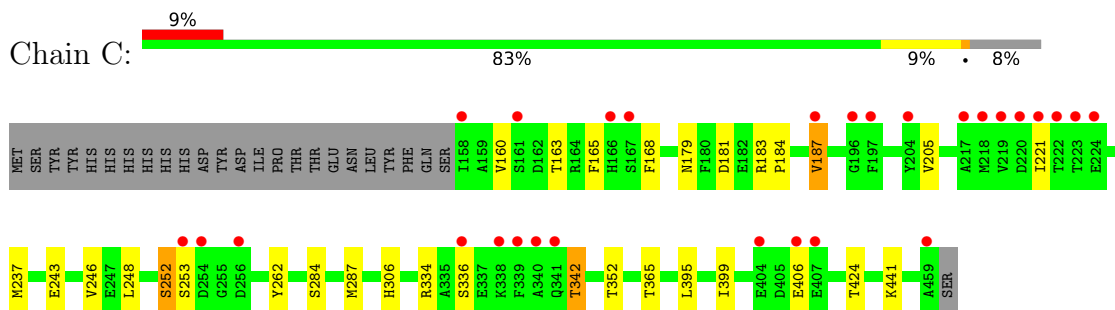
- Molecule 1: Interleukin-1 receptor-associated kinase 4



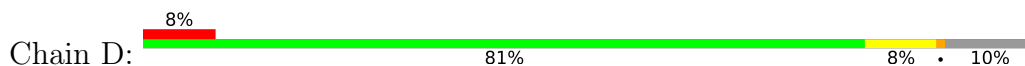
- Molecule 1: Interleukin-1 receptor-associated kinase 4

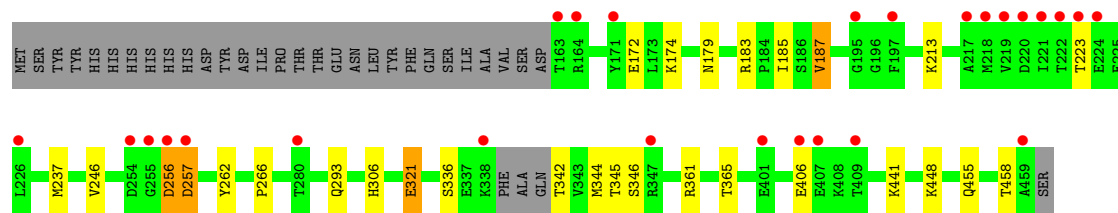


- Molecule 1: Interleukin-1 receptor-associated kinase 4



- Molecule 1: Interleukin-1 receptor-associated kinase 4





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	142.87Å 140.68Å 89.02Å 90.00° 124.00° 90.00°	Depositor
Resolution (Å)	49.14 – 2.09 49.09 – 2.09	Depositor EDS
% Data completeness (in resolution range)	99.9 (49.14-2.09) 99.9 (49.09-2.09)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.02 (at 2.08Å)	Xtrriage
Refinement program	REFMAC 5.8.0238	Depositor
R, $R_{free}$	0.168 , 0.208 0.178 , 0.214	Depositor DCC
$R_{free}$ test set	4334 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	36.0	Xtrriage
Anisotropy	0.052	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 51.9	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	0.97	EDS
Total number of atoms	10225	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 19.92 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 9.7314e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: YJU, SEP, TPO, CL, GOL

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	0.71	0/2364	0.82	0/3183
1	B	0.71	0/2367	0.81	0/3189
1	C	0.71	0/2407	0.81	1/3243 (0.0%)
1	D	0.71	0/2339	0.83	0/3151
All	All	0.71	0/9477	0.81	1/12766 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	334	ARG	NE-CZ-NH1	-6.91	116.85	120.30

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	2353	0	2320	12	0
1	B	2358	0	2319	25	0
1	C	2392	0	2363	17	1
1	D	2328	0	2299	21	2
2	A	34	0	0	0	0
2	B	34	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	34	0	0	0	0
2	D	34	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
4	A	12	0	16	1	0
4	B	12	0	16	1	0
4	C	6	0	8	2	0
4	D	6	0	8	0	0
5	A	176	0	0	4	2
5	B	139	0	0	4	0
5	C	149	0	0	3	1
5	D	155	0	0	4	0
All	All	10225	0	9349	74	3

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 4.

The worst 5 of 74 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:217:ALA:HB2	1:B:226:LEU:HD22	1.43	1.01
1:C:284:SER:H	1:C:287[A]:MET:HE3	1.31	0.95
1:D:361:ARG:NH1	5:D:601:HOH:O	2.07	0.85
1:D:293:GLN:HE22	1:D:458:THR:HG21	1.44	0.82
1:C:179:ASN:HB2	5:C:606:HOH:O	1.81	0.81

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:424:THR:OG1	1:D:256:ASP:OD2[2_556]	1.94	0.26
1:D:172:GLU:OE1	5:A:601:HOH:O[3_555]	2.18	0.02
5:A:696:HOH:O	5:C:732:HOH:O[4_446]	2.18	0.02

## 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	294/327 (90%)	283 (96%)	9 (3%)	2 (1%)	22	18
1	B	294/327 (90%)	285 (97%)	7 (2%)	2 (1%)	22	18
1	C	300/327 (92%)	291 (97%)	8 (3%)	1 (0%)	41	41
1	D	290/327 (89%)	279 (96%)	9 (3%)	2 (1%)	22	18
All	All	1178/1308 (90%)	1138 (97%)	33 (3%)	7 (1%)	25	21

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	406	GLU
1	B	406	GLU
1	D	406	GLU
1	B	218	MET
1	C	406	GLU

### 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	257/284 (90%)	253 (98%)	4 (2%)	62	69
1	B	257/284 (90%)	252 (98%)	5 (2%)	57	63
1	C	262/284 (92%)	255 (97%)	7 (3%)	44	48
1	D	255/284 (90%)	248 (97%)	7 (3%)	44	48
All	All	1031/1136 (91%)	1008 (98%)	23 (2%)	52	57

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	365	THR
1	D	187	VAL
1	D	185	ILE
1	D	223	THR
1	B	213	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	206	ASN
1	D	175	ASN
1	D	293	GLN
1	D	179	ASN
1	B	206	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

12 non-standard protein/DNA/RNA residues 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	$\# Z  > 2$
1	TPO	C	345	1	8,10,11	0.93	0	10,14,16	0.90	0
1	TPO	D	342	1	8,10,11	0.64	0	10,14,16	0.91	0
1	TPO	A	342	1	8,10,11	0.50	0	10,14,16	1.07	1 (10%)
1	SEP	B	346	1	8,9,10	0.68	0	8,12,14	0.82	0
1	TPO	D	345	1	8,10,11	1.00	1 (12%)	10,14,16	0.94	0
1	SEP	D	346	1	8,9,10	0.46	0	8,12,14	1.12	1 (12%)
1	TPO	B	342	1	8,10,11	0.97	1 (12%)	10,14,16	1.14	1 (10%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	SEP	C	346	1	8,9,10	0.54	0	8,12,14	0.80	0
1	TPO	B	345	1	8,10,11	0.86	0	10,14,16	1.02	1 (10%)
1	TPO	C	342	1	8,10,11	0.63	0	10,14,16	1.20	1 (10%)
1	TPO	A	345	1	8,10,11	0.97	0	10,14,16	0.94	0
1	SEP	A	346	1	8,9,10	0.55	0	8,12,14	0.81	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
1	TPO	C	345	1	-	4/9/11/13	-
1	TPO	D	342	1	-	1/9/11/13	-
1	TPO	A	342	1	-	1/9/11/13	-
1	SEP	B	346	1	-	0/5/8/10	-
1	TPO	D	345	1	-	3/9/11/13	-
1	SEP	D	346	1	-	3/5/8/10	-
1	TPO	B	342	1	-	1/9/11/13	-
1	SEP	C	346	1	-	0/5/8/10	-
1	TPO	B	345	1	-	3/9/11/13	-
1	TPO	C	342	1	-	0/9/11/13	-
1	TPO	A	345	1	-	3/9/11/13	-
1	SEP	A	346	1	-	0/5/8/10	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	342	TPO	P-OG1	2.36	1.63	1.59
1	D	345	TPO	P-OG1	2.19	1.63	1.59

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	342	TPO	P-OG1-CB	-2.51	115.63	123.21
1	A	342	TPO	O3P-P-O2P	2.29	116.39	107.64
1	B	345	TPO	O3P-P-O2P	2.25	116.23	107.64
1	D	346	SEP	OG-P-O1P	-2.24	100.19	106.47
1	C	342	TPO	OG1-P-O1P	-2.08	101.37	109.39



There are no chirality outliers.

5 of 19 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	345	TPO	N-CA-CB-OG1
1	A	345	TPO	O-C-CA-CB
1	B	342	TPO	O-C-CA-CB
1	B	345	TPO	N-CA-CB-OG1
1	B	345	TPO	O-C-CA-CB

There are no ring outliers.

4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	D	342	TPO	1	0
1	A	342	TPO	1	0
1	B	342	TPO	1	0
1	C	342	TPO	1	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 13 ligands modelled in this entry, 3 are monoatomic - leaving 10 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	$\# Z  > 2$
4	GOL	A	504	-	5,5,5	0.11	0	5,5,5	0.37	0
2	YJU	B	501	-	35,37,37	1.51	5 (14%)	48,56,56	1.79	6 (12%)
4	GOL	A	503	-	5,5,5	0.15	0	5,5,5	0.27	0
2	YJU	A	501	-	35,37,37	1.44	7 (20%)	48,56,56	1.53	8 (16%)
4	GOL	C	503	-	5,5,5	0.12	0	5,5,5	0.29	0
4	GOL	B	504	-	5,5,5	0.15	0	5,5,5	0.37	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	GOL	B	503	-	5,5,5	0.13	0	5,5,5	0.29	0
2	YJU	D	501	-	35,37,37	1.69	8 (22%)	48,56,56	2.21	9 (18%)
2	YJU	C	501	-	35,37,37	1.94	9 (25%)	48,56,56	2.55	12 (25%)
4	GOL	D	502	-	5,5,5	0.10	0	5,5,5	0.30	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
4	GOL	A	504	-	-	3/4/4/4	-
2	YJU	B	501	-	-	2/22/36/36	0/4/4/4
4	GOL	A	503	-	-	3/4/4/4	-
2	YJU	A	501	-	-	2/22/36/36	0/4/4/4
4	GOL	C	503	-	-	4/4/4/4	-
4	GOL	B	504	-	-	2/4/4/4	-
4	GOL	B	503	-	-	2/4/4/4	-
2	YJU	D	501	-	-	2/22/36/36	0/4/4/4
2	YJU	C	501	-	-	2/22/36/36	0/4/4/4
4	GOL	D	502	-	-	2/4/4/4	-

The worst 5 of 29 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	501	YJU	C6-N1	-5.55	1.24	1.37
2	C	501	YJU	C2-C4	-4.99	1.50	1.54
2	B	501	YJU	C16-N4	-4.74	1.28	1.34
2	D	501	YJU	C3-C2	-4.44	1.47	1.52
2	D	501	YJU	C6-N1	-4.21	1.27	1.37

The worst 5 of 35 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	501	YJU	C7-N2-C18	12.95	136.96	125.48
2	D	501	YJU	C7-N2-C18	8.35	132.88	125.48
2	B	501	YJU	C7-N2-C18	8.32	132.85	125.48
2	D	501	YJU	C3-C2-C1	-6.29	102.80	110.49
2	D	501	YJU	C19-C20-C21	-4.51	104.96	112.42

There are no chirality outliers.

5 of 24 torsion outliers are listed below:

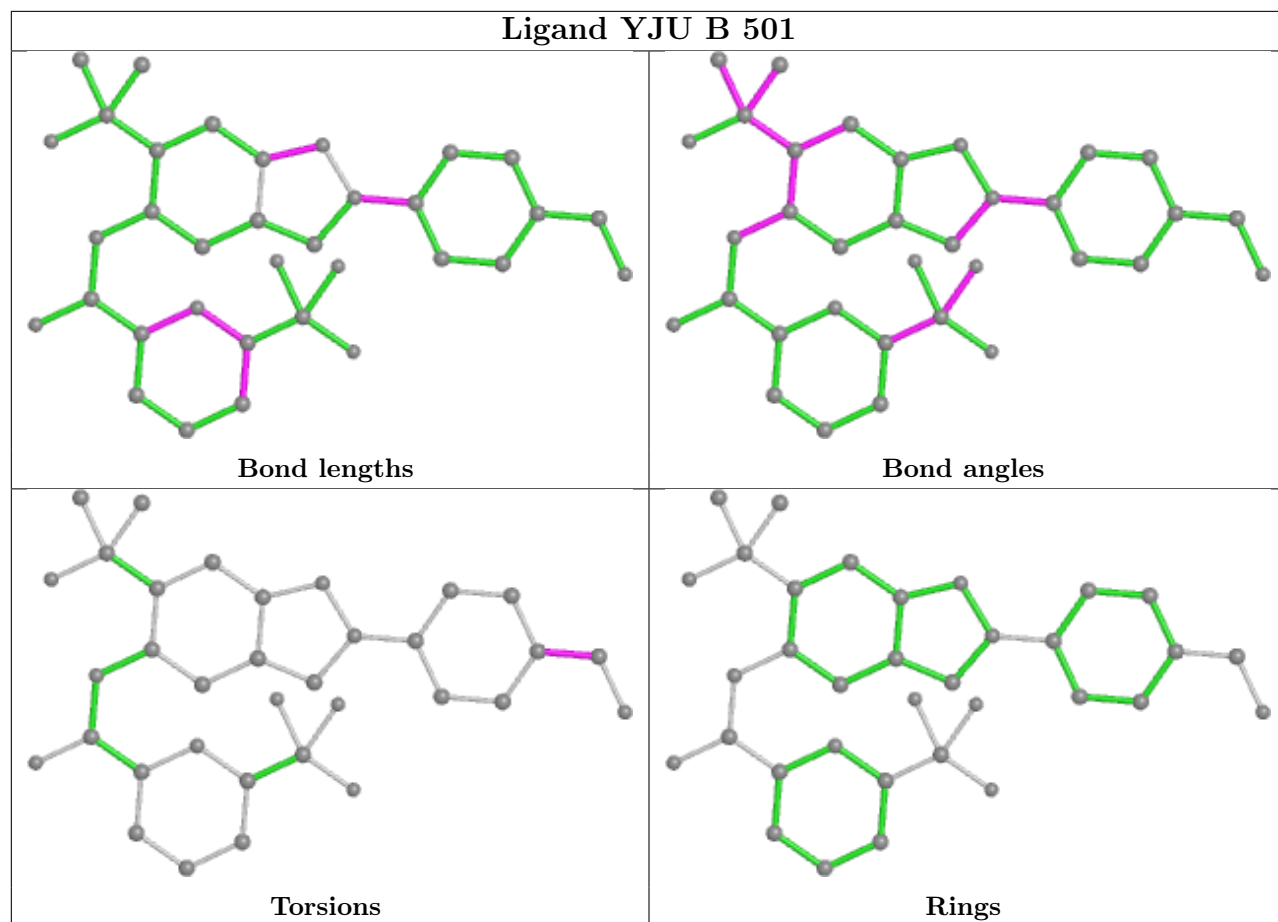
Mol	Chain	Res	Type	Atoms
2	B	501	YJU	C20-C21-C22-O3
2	C	501	YJU	C23-C21-C22-O3
2	D	501	YJU	C20-C21-C22-O3
2	D	501	YJU	C23-C21-C22-O3
4	A	503	GOL	C1-C2-C3-O3

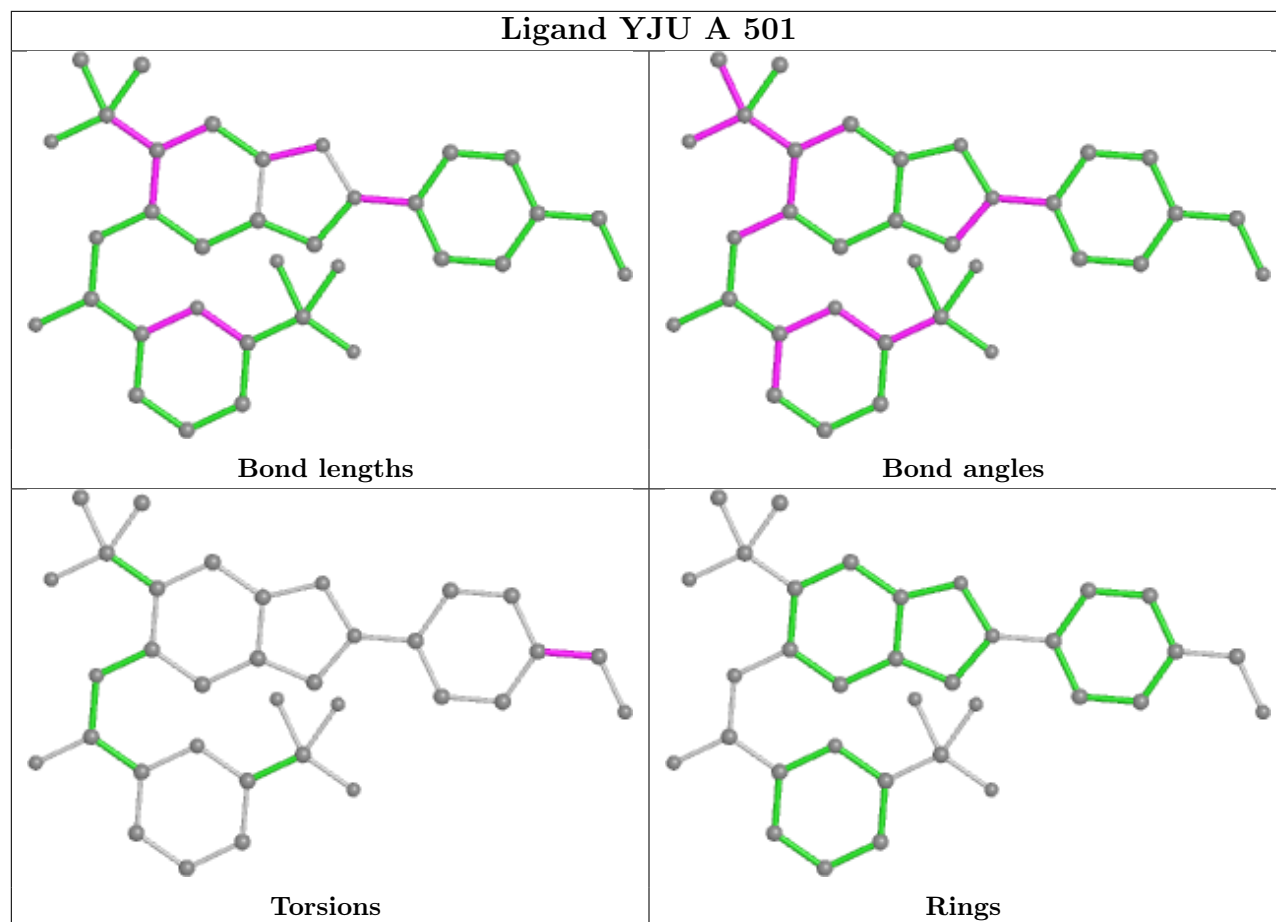
There are no ring outliers.

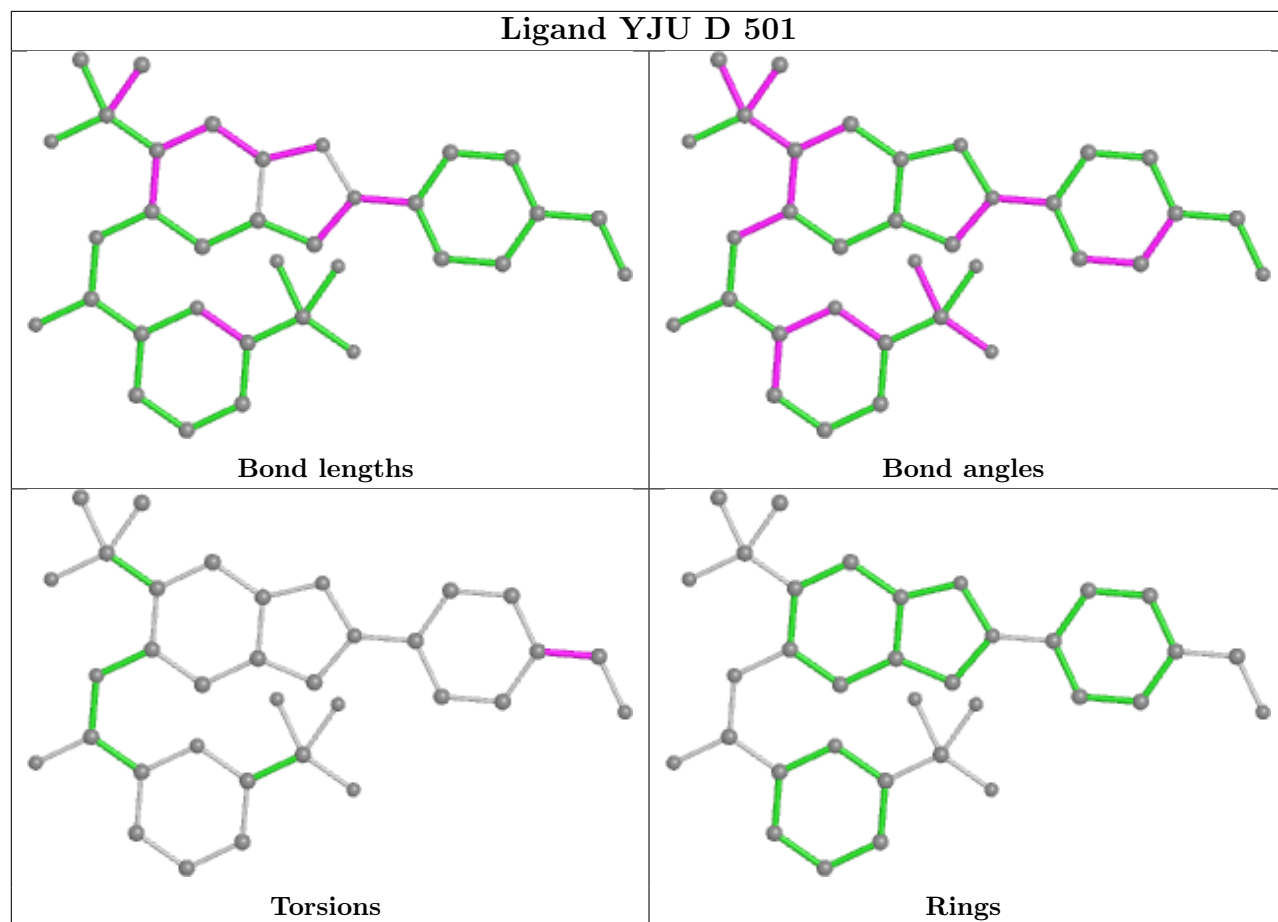
3 monomers are involved in 4 short contacts:

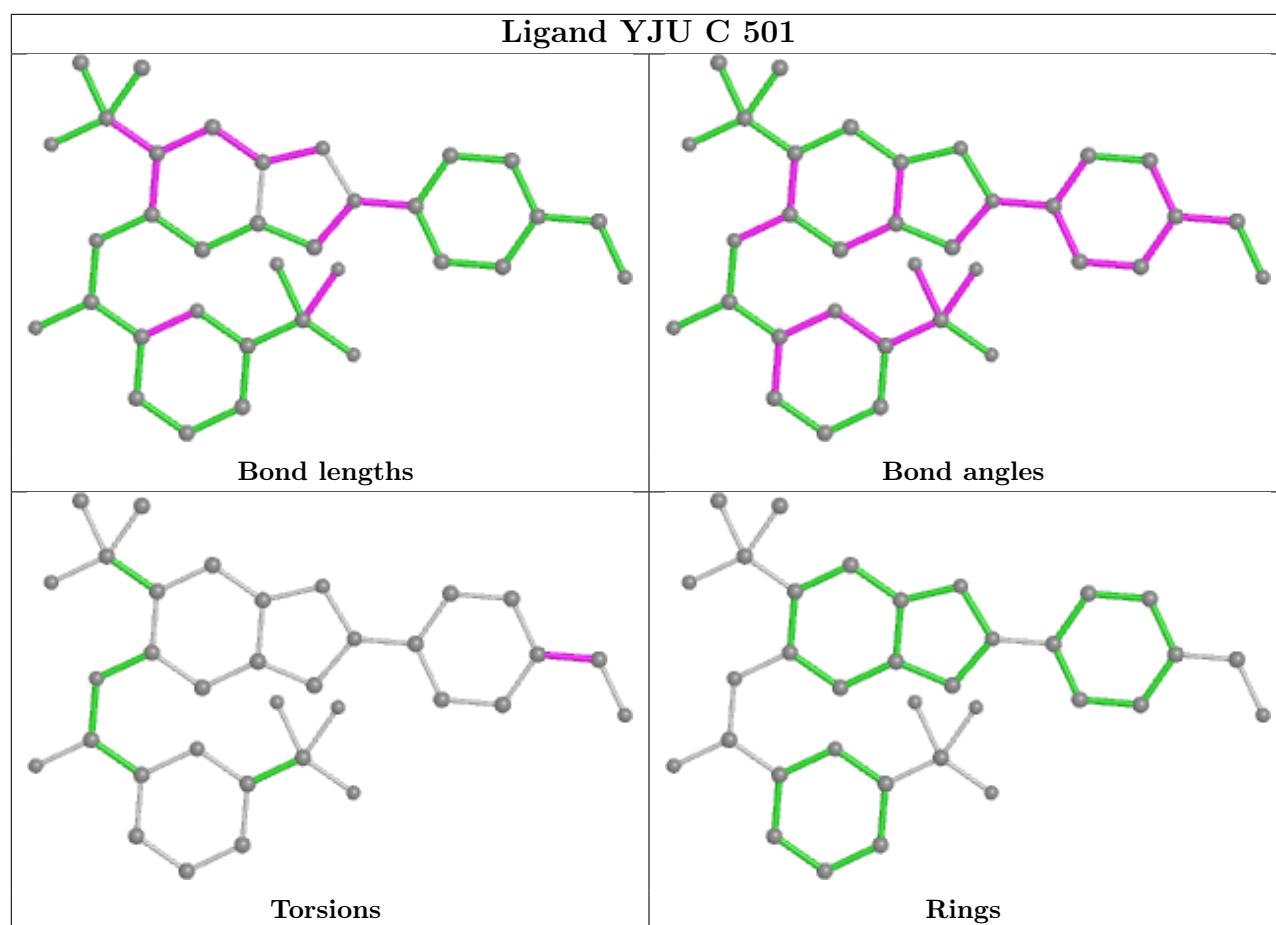
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	504	GOL	1	0
4	C	503	GOL	2	0
4	B	504	GOL	1	0

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	294/327 (89%)	0.28	22 (7%) 14 18	25, 36, 78, 165	0
1	B	295/327 (90%)	0.31	25 (8%) 10 13	25, 43, 88, 140	0
1	C	299/327 (91%)	0.48	28 (9%) 8 11	25, 43, 90, 147	0
1	D	291/327 (88%)	0.29	26 (8%) 9 12	26, 41, 86, 136	0
All	All	1179/1308 (90%)	0.34	101 (8%) 10 13	25, 40, 86, 165	0

The worst 5 of 101 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	339	PHE	19.3
1	A	340	ALA	14.5
1	C	339	PHE	13.3
1	B	217	ALA	9.1
1	B	218	MET	8.4

### 6.2 Non-standard residues in protein, DNA, RNA chains [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
1	TPO	D	342	11/12	0.88	0.12	70,74,90,93	0
1	SEP	B	346	10/11	0.89	0.11	46,60,83,88	0
1	SEP	D	346	10/11	0.90	0.12	52,70,93,100	0
1	TPO	B	342	11/12	0.92	0.12	56,62,74,78	0
1	SEP	C	346	10/11	0.92	0.11	49,61,89,91	0
1	SEP	A	346	10/11	0.93	0.15	42,58,78,85	0
1	TPO	A	342	11/12	0.94	0.14	54,65,78,83	0

*Continued on next page...*



Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
1	TPO	C	342	11/12	0.95	0.08	54,64,75,83	0
1	TPO	C	345	11/12	0.97	0.08	42,46,51,62	0
1	TPO	B	345	11/12	0.97	0.10	44,47,52,62	0
1	TPO	D	345	11/12	0.98	0.07	41,47,51,56	0
1	TPO	A	345	11/12	0.98	0.08	38,43,48,53	0

### 6.3 Carbohydrates [i](#)

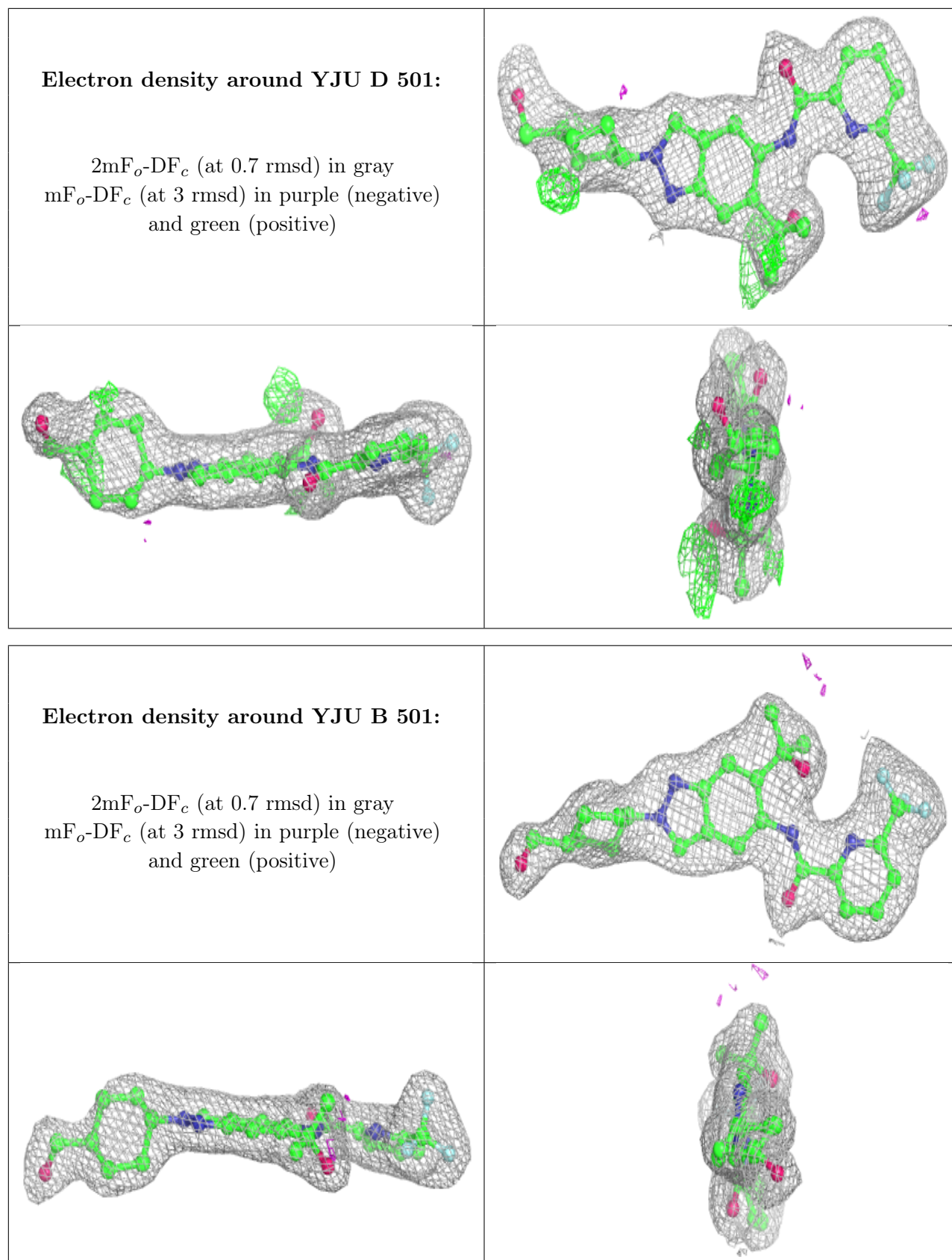
There are no monosaccharides in this entry.

### 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.

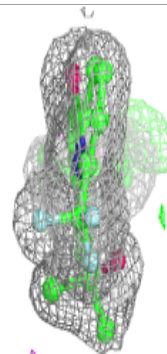
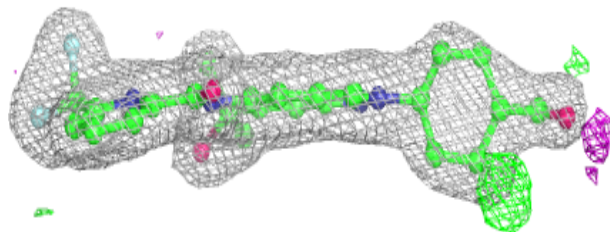
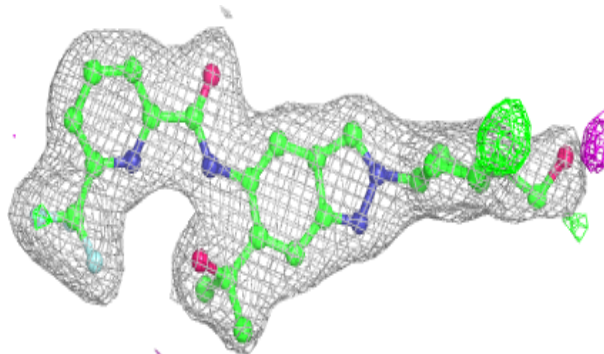
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	CL	B	502	1/1	0.74	0.10	84,84,84,84	0
4	GOL	B	504	6/6	0.86	0.13	45,51,57,58	0
3	CL	C	502	1/1	0.92	0.10	83,83,83,83	0
4	GOL	A	503	6/6	0.93	0.29	56,63,68,75	0
4	GOL	C	503	6/6	0.93	0.26	64,71,74,81	0
4	GOL	D	502	6/6	0.94	0.19	55,58,62,66	0
2	YJU	D	501	34/34	0.95	0.13	27,34,56,61	0
4	GOL	A	504	6/6	0.95	0.34	56,61,67,68	0
4	GOL	B	503	6/6	0.95	0.14	57,62,68,77	0
3	CL	A	502	1/1	0.96	0.10	69,69,69,69	0
2	YJU	B	501	34/34	0.97	0.09	29,34,50,51	0
2	YJU	C	501	34/34	0.97	0.10	26,30,48,50	0
2	YJU	A	501	34/34	0.97	0.14	26,29,49,57	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.

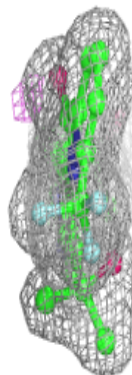
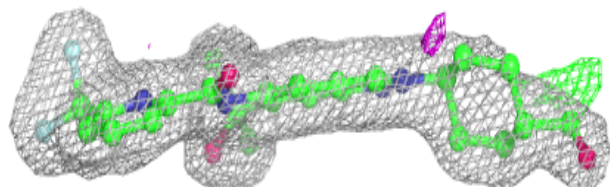
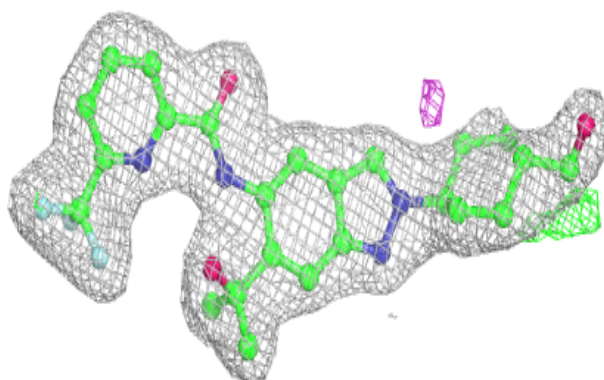


**Electron density around YJU C 501:**

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

**Electron density around YJU A 501:**

$2mF_o-DF_c$  (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.