wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 21, 2021 – 05:03 PM BST

PDB ID	:	7A5B
Title	:	Structure of DYRK1A in complex with complex 10
Authors	:	Dokurno, P.; Surgenor, A.E.; Hubbard, R.E.
Deposited on	:	2020-08-20
Resolution	:	2.60 Å(reported)

The following versions of software and data (see references ①) were used in the production of this report:

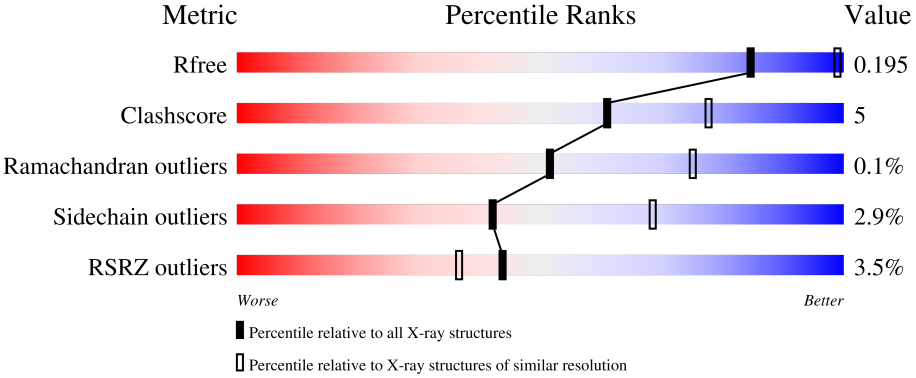
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.20
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.20

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

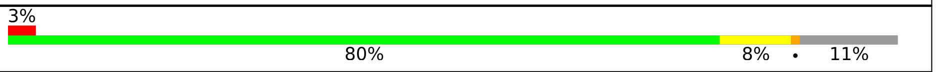
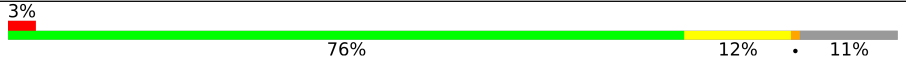
The reported resolution of this entry is 2.60 Å.

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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	382	
1	B	382	

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 5836 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 Dual specificity tyrosine-phosphorylation-regulated kinase 1A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	340	Total 2755	C 1771	N 469	O 497	P 1	S 17	0	0	0
1	B	341	Total 2778	C 1789	N 475	O 496	P 1	S 17	0	0	0

There are 46 discrepancies between the modelled and reference sequences:

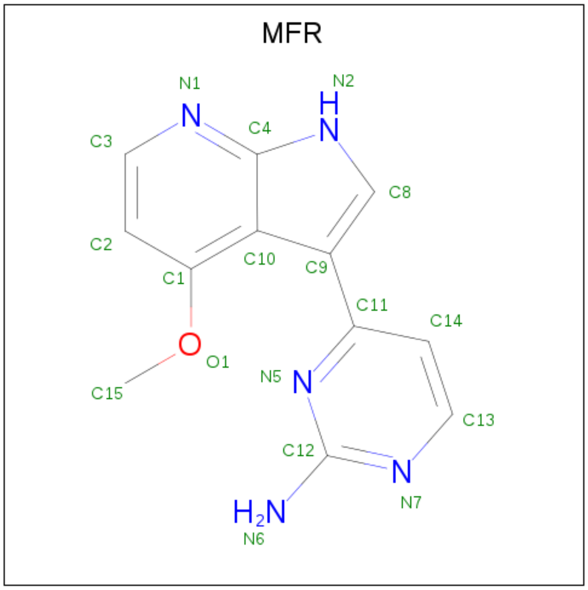
Chain	Residue	Modelled	Actual	Comment	Reference
A	104	MET	-	initiating methionine	UNP Q13627
A	105	HIS	-	expression tag	UNP Q13627
A	106	HIS	-	expression tag	UNP Q13627
A	107	HIS	-	expression tag	UNP Q13627
A	108	HIS	-	expression tag	UNP Q13627
A	109	HIS	-	expression tag	UNP Q13627
A	110	HIS	-	expression tag	UNP Q13627
A	111	SER	-	expression tag	UNP Q13627
A	112	SER	-	expression tag	UNP Q13627
A	113	GLY	-	expression tag	UNP Q13627
A	114	VAL	-	expression tag	UNP Q13627
A	115	ASP	-	expression tag	UNP Q13627
A	116	LEU	-	expression tag	UNP Q13627
A	117	GLY	-	expression tag	UNP Q13627
A	118	THR	-	expression tag	UNP Q13627
A	119	GLU	-	expression tag	UNP Q13627
A	120	ASN	-	expression tag	UNP Q13627
A	121	LEU	-	expression tag	UNP Q13627
A	122	TYR	-	expression tag	UNP Q13627
A	123	PHE	-	expression tag	UNP Q13627
A	124	GLN	-	expression tag	UNP Q13627
A	125	SER	-	expression tag	UNP Q13627
A	126	MET	-	expression tag	UNP Q13627
B	104	MET	-	initiating methionine	UNP Q13627
B	105	HIS	-	expression tag	UNP Q13627

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	106	HIS	-	expression tag	UNP Q13627
B	107	HIS	-	expression tag	UNP Q13627
B	108	HIS	-	expression tag	UNP Q13627
B	109	HIS	-	expression tag	UNP Q13627
B	110	HIS	-	expression tag	UNP Q13627
B	111	SER	-	expression tag	UNP Q13627
B	112	SER	-	expression tag	UNP Q13627
B	113	GLY	-	expression tag	UNP Q13627
B	114	VAL	-	expression tag	UNP Q13627
B	115	ASP	-	expression tag	UNP Q13627
B	116	LEU	-	expression tag	UNP Q13627
B	117	GLY	-	expression tag	UNP Q13627
B	118	THR	-	expression tag	UNP Q13627
B	119	GLU	-	expression tag	UNP Q13627
B	120	ASN	-	expression tag	UNP Q13627
B	121	LEU	-	expression tag	UNP Q13627
B	122	TYR	-	expression tag	UNP Q13627
B	123	PHE	-	expression tag	UNP Q13627
B	124	GLN	-	expression tag	UNP Q13627
B	125	SER	-	expression tag	UNP Q13627
B	126	MET	-	expression tag	UNP Q13627

- Molecule 2 is 4-(4-methoxy-1H-pyrrolo[2,3-b]pyridin-3-yl)pyrimidin-2-amine (three-letter code: MFR) (formula: $C_{12}H_{11}N_5O$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C N O 18 12 5 1	0	0
2	B	1	Total C N O 18 12 5 1	0	0

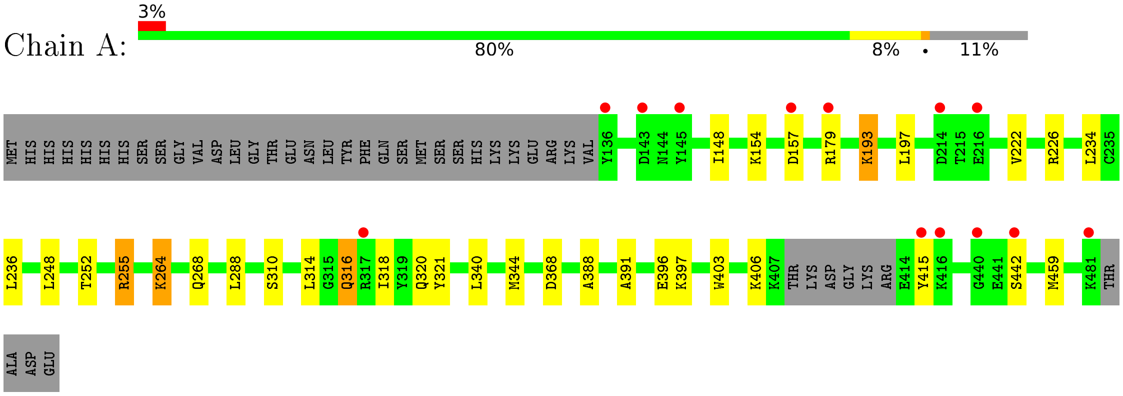
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	122	Total O 122 122	0	0
3	B	145	Total O 145 145	0	0

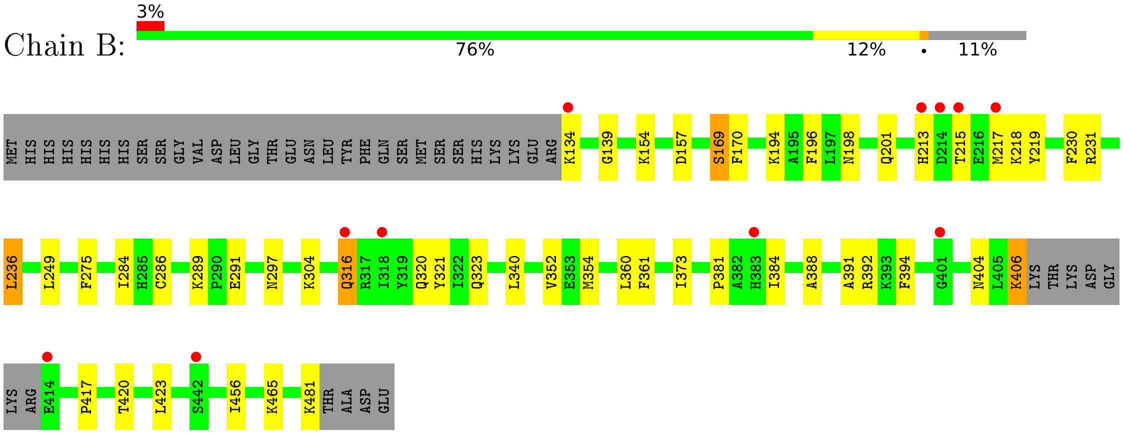
3 Residue-property plots ⓘ

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: Dual specificity tyrosine-phosphorylation-regulated kinase 1A



- Molecule 1: Dual specificity tyrosine-phosphorylation-regulated kinase 1A



4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	65.71Å 84.94Å 76.80Å 90.00° 108.50° 90.00°	Depositor
Resolution (Å)	25.00 – 2.60 24.72 – 2.60	Depositor EDS
% Data completeness (in resolution range)	98.3 (25.00-2.60) 98.4 (24.72-2.60)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.78 (at 2.60Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.154 , 0.192 0.161 , 0.195	Depositor DCC
R_{free} test set	1182 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å^2)	40.0	Xtriage
Anisotropy	0.054	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 43.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	5836	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.05% 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: MFR, PTR

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	0.66	0/2802	0.82	1/3783 (0.0%)
1	B	0.65	0/2825	0.81	0/3808
All	All	0.66	0/5627	0.82	1/7591 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	255	ARG	NE-CZ-NH2	-5.59	117.50	120.30

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	2755	0	2720	19	2
1	B	2778	0	2780	33	1
2	A	18	0	11	2	0
2	B	18	0	11	1	0
3	A	122	0	0	0	0
3	B	145	0	0	7	0
All	All	5836	0	5522	50	2

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 5.

The worst 5 of 50 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:198:ASN:ND2	3:B:601:HOH:O	2.13	0.81
1:A:316:GLN:HB3	1:A:318:ILE:HG13	1.66	0.77
1:B:297:ASN:HB3	3:B:686:HOH:O	1.84	0.76
1:B:381:PRO:HG2	1:B:384:ILE:HD12	1.70	0.73
1:B:219:TYR:O	1:B:304:LYS:HD3	1.95	0.67

All (2) 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:A:148:ILE:O	1:A:255:ARG:NH2[2_545]	2.01	0.19
1:A:179:ARG:NH1	1:B:297:ASN:OD1[1_455]	2.16	0.04

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	335/382 (88%)	324 (97%)	10 (3%)	1 (0%)	41	64
1	B	336/382 (88%)	320 (95%)	16 (5%)	0	100	100
All	All	671/764 (88%)	644 (96%)	26 (4%)	1 (0%)	51	75

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	316	GLN

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	294/339 (87%)	288 (98%)	6 (2%)	55	78
1	B	299/339 (88%)	288 (96%)	11 (4%)	34	60
All	All	593/678 (88%)	576 (97%)	17 (3%)	42	68

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	316	GLN
1	B	465	LYS
1	B	154	LYS
1	B	157	ASP
1	B	169	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	198	ASN
1	B	211	ASN
1	B	387	GLN
1	B	213	HIS
1	A	211	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 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	PTR	B	321	1	15,16,17	0.41	0	19,22,24	1.12	2 (10%)
1	PTR	A	321	1	15,16,17	0.75	1 (6%)	19,22,24	0.86	1 (5%)

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	PTR	B	321	1	-	0/10/11/13	0/1/1/1
1	PTR	A	321	1	-	3/10/11/13	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	321	PTR	P-OH	-2.18	1.55	1.59

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	321	PTR	OH-P-O1P	-2.77	98.85	109.31
1	A	321	PTR	CG-CB-CA	2.43	119.02	114.10
1	B	321	PTR	O3P-P-OH	2.30	112.43	105.24

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	321	PTR	O-C-CA-CB
1	A	321	PTR	CZ-OH-P-O1P
1	A	321	PTR	CZ-OH-P-O3P

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

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	MFR	B	501	-	19,20,20	1.02	3 (15%)	20,28,28	1.70	4 (20%)
2	MFR	A	501	-	19,20,20	0.95	2 (10%)	20,28,28	1.62	4 (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	MFR	B	501	-	-	0/2/6/6	0/3/3/3
2	MFR	A	501	-	-	0/2/6/6	0/3/3/3

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	501	MFR	C4-N1	-2.48	1.33	1.37
2	B	501	MFR	C9-C11	-2.46	1.46	1.49
2	A	501	MFR	C9-C10	2.34	1.45	1.41
2	B	501	MFR	C9-C10	2.32	1.45	1.41
2	A	501	MFR	C4-N1	-2.12	1.34	1.37

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	501	MFR	C14-C11-C9	3.66	125.79	121.04
2	B	501	MFR	C11-N5-C12	3.60	118.39	116.34

Continued on next page...

Continued from previous page...

Mol	**Chain**	**Res**	**Type**	**Atoms**	**Z**	**Observed(°)**	**Ideal(°)**
2	A	501	MFR	C11-N5-C12	3.36	118.25	116.34
2	B	501	MFR	C14-C11-C9	3.28	125.30	121.04
2	B	501	MFR	O1-C1-C10	2.98	120.27	115.89

There are no chirality outliers.

There are no torsion outliers.

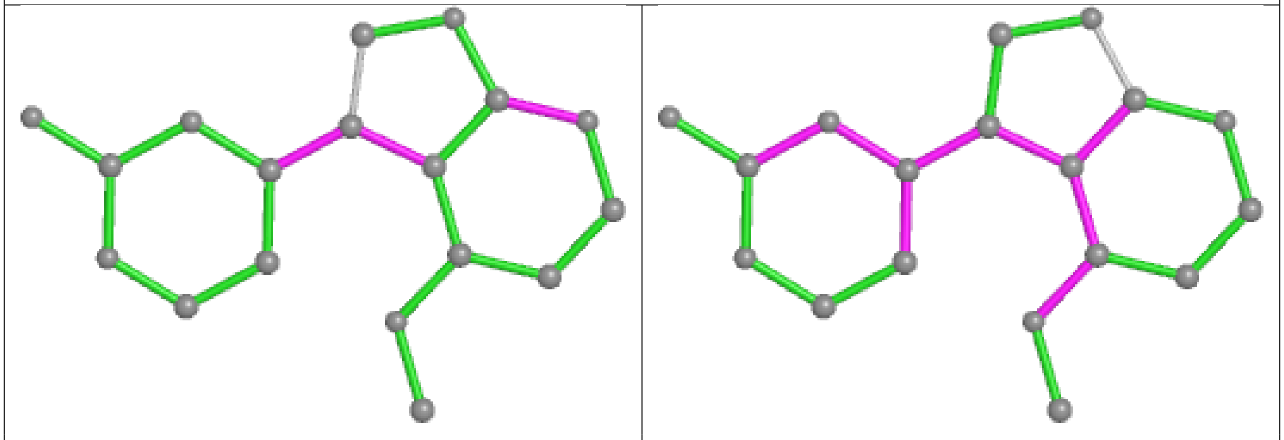
There are no ring outliers.

2 monomers are involved in 3 short contacts:

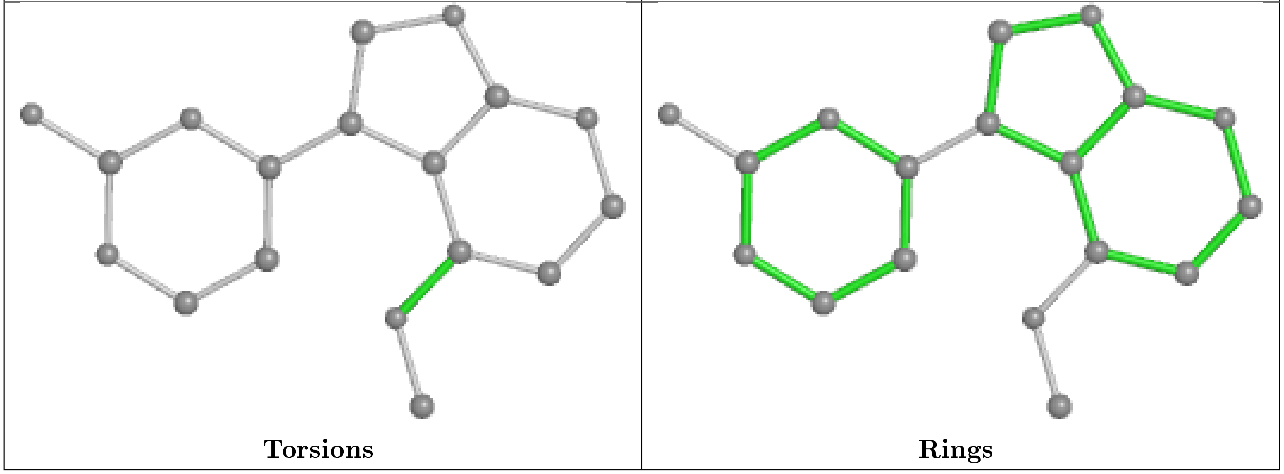
Mol	**Chain**	**Res**	**Type**	**Clashes**	**Symm-Clashes**
2	B	501	MFR	1	0
2	A	501	MFR	2	0

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 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 MFR B 501

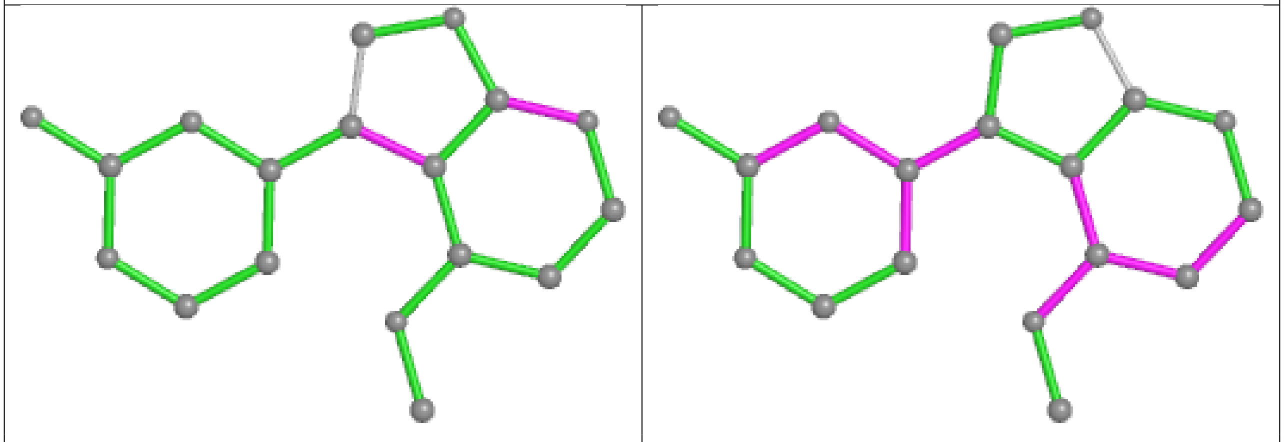


Bond lengths **Bond angles**

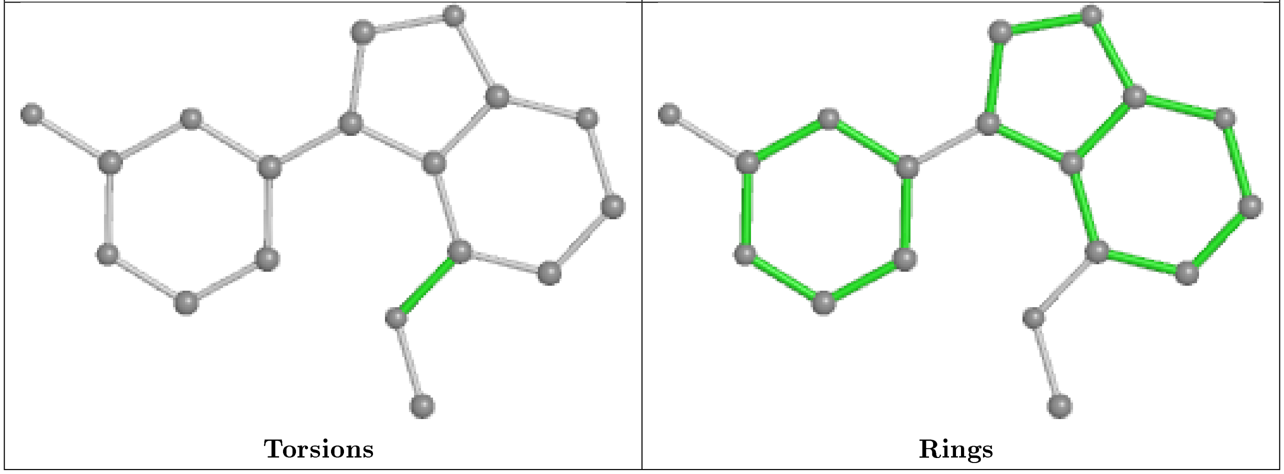


Torsions **Rings**

Ligand MFR A 501



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($Å^2$)	Q<0.9
1	A	339/382 (88%)	-0.34	13 (3%)	40	33	29, 42, 73, 105	2 (0%)
1	B	340/382 (89%)	-0.30	11 (3%)	47	40	28, 42, 76, 105	0
All	All	679/764 (88%)	-0.32	24 (3%)	44	36	28, 42, 74, 105	2 (0%)

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	214	ASP	4.3
1	A	481	LYS	4.1
1	A	214	ASP	3.6
1	A	317	ARG	3.2
1	A	440	GLY	3.2

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

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($Å^2$)	Q<0.9
1	PTR	B	321	16/17	0.97	0.11	45,55,63,63	0
1	PTR	A	321	16/17	0.98	0.09	27,37,43,44	0

6.3 Carbohydrates ⓘ

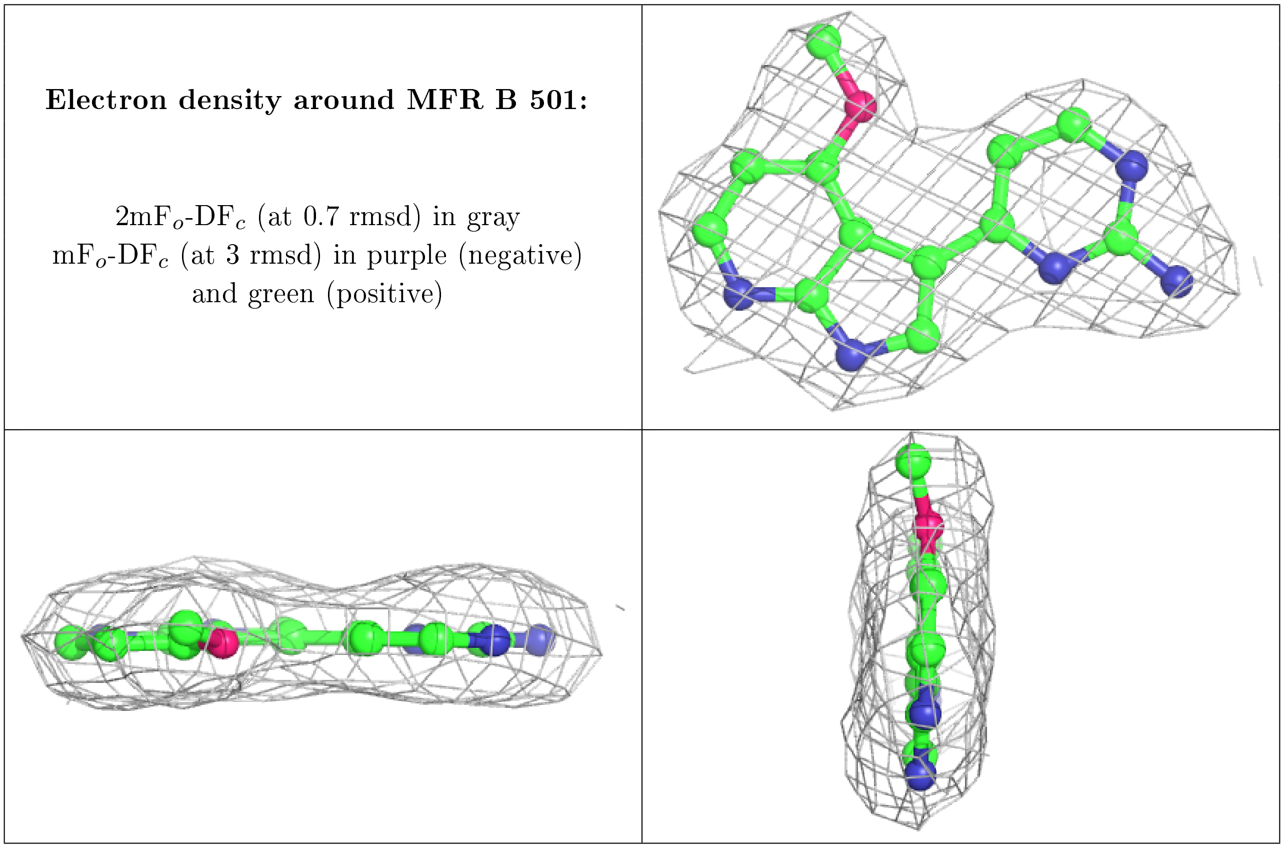
There are no monosaccharides in this entry.

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.

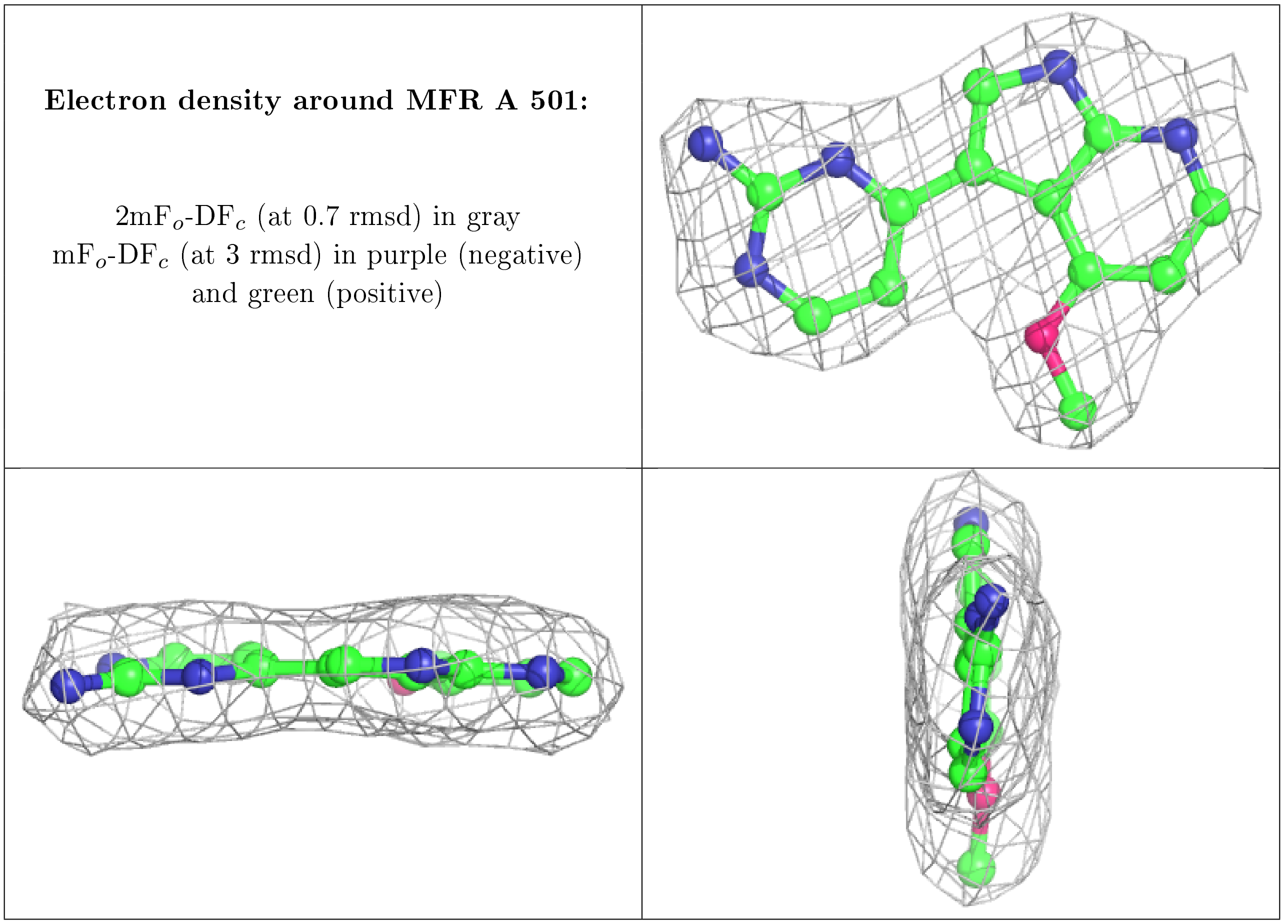
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
2	MFR	B	501	18/18	0.96	0.15	27,30,32,34	0
2	MFR	A	501	18/18	0.98	0.13	32,38,41,41	0

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 MFR B 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 MFR 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 ⓘ

There are no such residues in this entry.