



Full wwPDB X-ray Structure Validation Report ⓘ

Jun 27, 2023 – 04:15 PM EDT

PDB ID : 8SHJ
Title : Crystal structure of the WD-repeat domain of human WDR91 in complex with MR45279
Authors : Ahmad, H.; Zeng, H.; Dong, A.; Li, Y.; Hutchinson, A.; Seitova, A.; Xu, J.; Feng, J.W.; Brown, P.J.; Ackloo, S.; Arrowsmith, C.H.; Edwards, A.M.; Halabelian, L.; Structural Genomics Consortium (SGC)
Deposited on : 2023-04-14
Resolution : 2.21 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

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 : 2.33
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) : 2.33

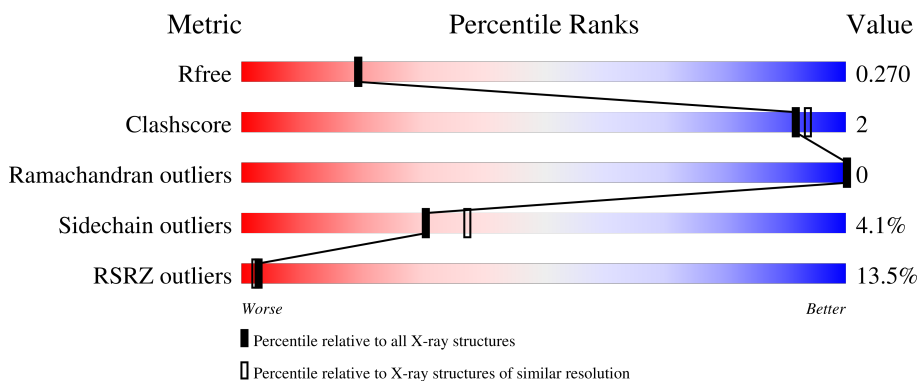
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.21 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	5912 (2.24-2.20)
Clashscore	141614	6646 (2.24-2.20)
Ramachandran outliers	138981	6543 (2.24-2.20)
Sidechain outliers	138945	6544 (2.24-2.20)
RSRZ outliers	127900	5797 (2.24-2.20)

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 $\leq 5\%$. 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	359	<div style="display: flex; align-items: center;"> <div style="width: 35%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 50%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 5%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 5%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">86% 6% • 7%</p>
1	B	359	<div style="display: flex; align-items: center;"> <div style="width: 35%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 45%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 5%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 5%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">85% 7% • 8%</p>
1	C	359	<div style="display: flex; align-items: center;"> <div style="width: 35%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 40%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 5%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 5%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">80% 5% 15%</p>

2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 7575 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 WD repeat-containing protein 91.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	334	2524	1591	432	479	22	0	4	0
1	B	331	2511	1584	428	478	21	0	3	0
1	C	306	2197	1384	375	418	20	0	1	0

There are 99 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	374	MET	-	initiating methionine	UNP A4D1P6
A	375	HIS	-	expression tag	UNP A4D1P6
A	376	HIS	-	expression tag	UNP A4D1P6
A	377	HIS	-	expression tag	UNP A4D1P6
A	378	HIS	-	expression tag	UNP A4D1P6
A	379	HIS	-	expression tag	UNP A4D1P6
A	380	HIS	-	expression tag	UNP A4D1P6
A	381	SER	-	expression tag	UNP A4D1P6
A	382	SER	-	expression tag	UNP A4D1P6
A	383	GLY	-	expression tag	UNP A4D1P6
A	384	ARG	-	expression tag	UNP A4D1P6
A	385	GLU	-	expression tag	UNP A4D1P6
A	386	ASN	-	expression tag	UNP A4D1P6
A	387	LEU	-	expression tag	UNP A4D1P6
A	388	TYR	-	expression tag	UNP A4D1P6
A	389	PHE	-	expression tag	UNP A4D1P6
A	390	GLN	-	expression tag	UNP A4D1P6
A	391	GLY	-	expression tag	UNP A4D1P6
A	?	-	VAL	deletion	UNP A4D1P6
A	?	-	ASP	deletion	UNP A4D1P6
A	?	-	PHE	deletion	UNP A4D1P6
A	?	-	SER	deletion	UNP A4D1P6
A	?	-	ALA	deletion	UNP A4D1P6

Continued on next page...

Continued from previous page...

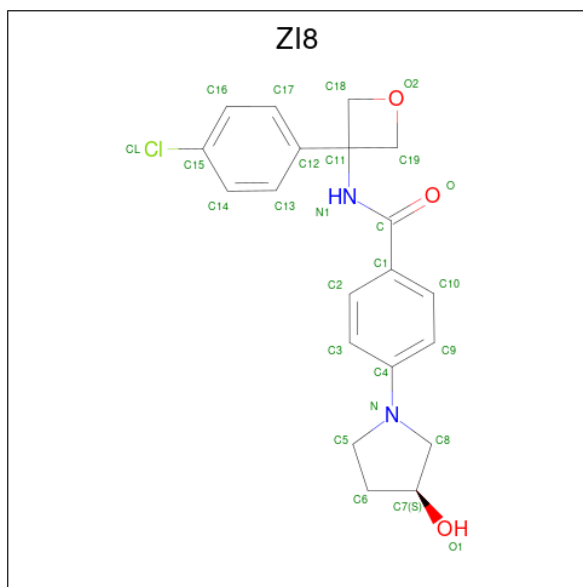
Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	PRO	deletion	UNP A4D1P6
A	?	-	ASP	deletion	UNP A4D1P6
A	?	-	ILE	deletion	UNP A4D1P6
A	?	-	GLY	deletion	UNP A4D1P6
A	?	-	SER	deletion	UNP A4D1P6
A	?	-	LYS	deletion	UNP A4D1P6
A	?	-	GLY	deletion	UNP A4D1P6
A	?	-	MET	deletion	UNP A4D1P6
A	?	-	ASN	deletion	UNP A4D1P6
A	?	-	GLN	deletion	UNP A4D1P6
B	374	MET	-	initiating methionine	UNP A4D1P6
B	375	HIS	-	expression tag	UNP A4D1P6
B	376	HIS	-	expression tag	UNP A4D1P6
B	377	HIS	-	expression tag	UNP A4D1P6
B	378	HIS	-	expression tag	UNP A4D1P6
B	379	HIS	-	expression tag	UNP A4D1P6
B	380	HIS	-	expression tag	UNP A4D1P6
B	381	SER	-	expression tag	UNP A4D1P6
B	382	SER	-	expression tag	UNP A4D1P6
B	383	GLY	-	expression tag	UNP A4D1P6
B	384	ARG	-	expression tag	UNP A4D1P6
B	385	GLU	-	expression tag	UNP A4D1P6
B	386	ASN	-	expression tag	UNP A4D1P6
B	387	LEU	-	expression tag	UNP A4D1P6
B	388	TYR	-	expression tag	UNP A4D1P6
B	389	PHE	-	expression tag	UNP A4D1P6
B	390	GLN	-	expression tag	UNP A4D1P6
B	391	GLY	-	expression tag	UNP A4D1P6
B	?	-	VAL	deletion	UNP A4D1P6
B	?	-	ASP	deletion	UNP A4D1P6
B	?	-	PHE	deletion	UNP A4D1P6
B	?	-	SER	deletion	UNP A4D1P6
B	?	-	ALA	deletion	UNP A4D1P6
B	?	-	PRO	deletion	UNP A4D1P6
B	?	-	ASP	deletion	UNP A4D1P6
B	?	-	ILE	deletion	UNP A4D1P6
B	?	-	GLY	deletion	UNP A4D1P6
B	?	-	SER	deletion	UNP A4D1P6
B	?	-	LYS	deletion	UNP A4D1P6
B	?	-	GLY	deletion	UNP A4D1P6
B	?	-	MET	deletion	UNP A4D1P6
B	?	-	ASN	deletion	UNP A4D1P6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	GLN	deletion	UNP A4D1P6
C	374	MET	-	initiating methionine	UNP A4D1P6
C	375	HIS	-	expression tag	UNP A4D1P6
C	376	HIS	-	expression tag	UNP A4D1P6
C	377	HIS	-	expression tag	UNP A4D1P6
C	378	HIS	-	expression tag	UNP A4D1P6
C	379	HIS	-	expression tag	UNP A4D1P6
C	380	HIS	-	expression tag	UNP A4D1P6
C	381	SER	-	expression tag	UNP A4D1P6
C	382	SER	-	expression tag	UNP A4D1P6
C	383	GLY	-	expression tag	UNP A4D1P6
C	384	ARG	-	expression tag	UNP A4D1P6
C	385	GLU	-	expression tag	UNP A4D1P6
C	386	ASN	-	expression tag	UNP A4D1P6
C	387	LEU	-	expression tag	UNP A4D1P6
C	388	TYR	-	expression tag	UNP A4D1P6
C	389	PHE	-	expression tag	UNP A4D1P6
C	390	GLN	-	expression tag	UNP A4D1P6
C	391	GLY	-	expression tag	UNP A4D1P6
C	?	-	VAL	deletion	UNP A4D1P6
C	?	-	ASP	deletion	UNP A4D1P6
C	?	-	PHE	deletion	UNP A4D1P6
C	?	-	SER	deletion	UNP A4D1P6
C	?	-	ALA	deletion	UNP A4D1P6
C	?	-	PRO	deletion	UNP A4D1P6
C	?	-	ASP	deletion	UNP A4D1P6
C	?	-	ILE	deletion	UNP A4D1P6
C	?	-	GLY	deletion	UNP A4D1P6
C	?	-	SER	deletion	UNP A4D1P6
C	?	-	LYS	deletion	UNP A4D1P6
C	?	-	GLY	deletion	UNP A4D1P6
C	?	-	MET	deletion	UNP A4D1P6
C	?	-	ASN	deletion	UNP A4D1P6
C	?	-	GLN	deletion	UNP A4D1P6

- Molecule 2 is N-[3-(4-chlorophenyl)oxetan-3-yl]-4-[(3S)-3-hydroxypyrrolidin-1-yl]benzamide (three-letter code: ZI8) (formula: C₂₀H₂₁ClN₂O₃) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
2	A	1	Total	C	Cl	N	O	0	0
			26	20	1	2	3		
2	B	1	Total	C	Cl	N	O	0	0
			26	20	1	2	3		
2	C	1	Total	C	Cl	N	O	0	0
			26	20	1	2	3		

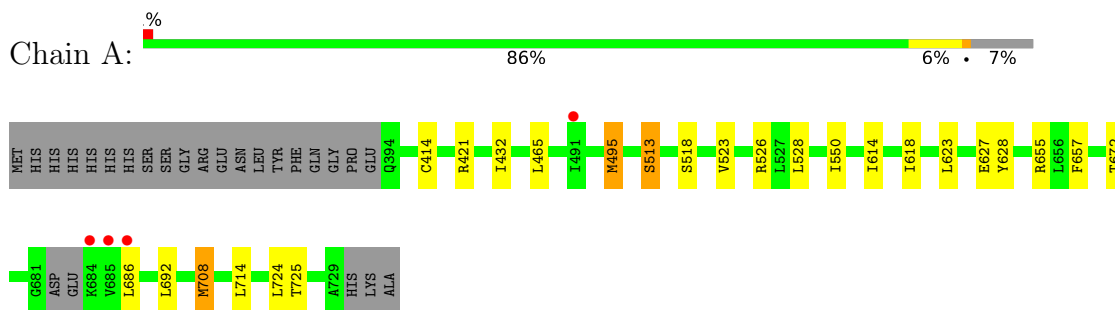
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	111	Total	O	0	1
			112	112		
3	B	113	Total	O	0	1
			114	114		
3	C	39	Total	O	0	0
			39	39		

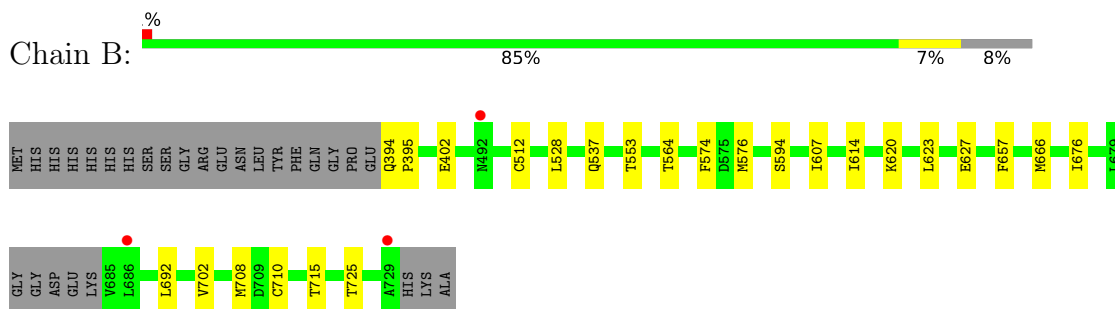
3 Residue-property plots [i](#)

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.

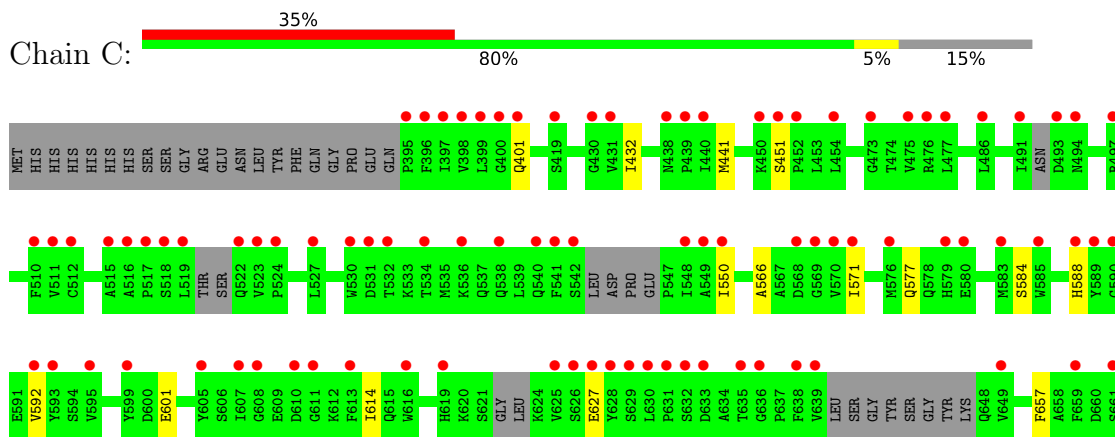
- Molecule 1: WD repeat-containing protein 91



- Molecule 1: WD repeat-containing protein 91



- Molecule 1: WD repeat-containing protein 91



E662	G663	N664	Y665	N666	L667	T668	C669	S670	A671	T672	G673	G674	V675	I676	Y677	K678	L679	GLY	GLY	ASP	GLU	GLU	LYS	VAL	LEU	GLU	SER	C689	L690	S691	L692	G693	G694	H695	R696	A697	P698	V699	Y700	W704	S705	T706	ALA	MET	D709	C710	T715	L724	T725	T726	L727	L728	ALA	HIS	LYS	ALA
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	------	------	------	------	------	------	------	------	-----	-----	-----	-----

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	76.86Å 121.37Å 131.80Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.25 – 2.21 46.25 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.4 (46.25-2.21) 99.4 (46.25-2.20)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.53 (at 2.20Å)	Xtrriage
Refinement program	BUSTER 2.10.3	Depositor
R, R_{free}	0.211 , 0.257 0.219 , 0.270	Depositor DCC
R_{free} test set	1224 reflections (1.95%)	wwPDB-VP
Wilson B-factor (Å ²)	36.0	Xtrriage
Anisotropy	0.184	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 54.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	7575	wwPDB-VP
Average B, all atoms (Å ²)	46.0	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 83.96 % 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 1.8905e-07. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZI8

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	0.51	0/2576	0.69	0/3497
1	B	0.50	0/2563	0.71	0/3480
1	C	0.52	0/2238	0.67	0/3041
All	All	0.51	0/7377	0.69	0/10018

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	2524	0	2422	8	0
1	B	2511	0	2419	10	0
1	C	2197	0	1970	5	0
2	A	26	0	0	0	0
2	B	26	0	0	0	0
2	C	26	0	0	0	0
3	A	112	0	0	0	0
3	B	114	0	0	0	0
3	C	39	0	0	0	0
All	All	7575	0	6811	23	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 2.

All (23) 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:A:414[A]:CYS:SG	1:A:714:LEU:HD13	2.41	0.60
1:C:614:ILE:HG12	1:C:627:GLU:HG3	1.86	0.57
1:A:614:ILE:HG12	1:A:627:GLU:HG3	1.89	0.53
1:B:394:GLN:N	1:B:395:PRO:HD2	2.26	0.51
1:B:553:THR:HG22	1:B:564:THR:HG22	1.93	0.50
1:B:574:PHE:HB3	1:B:576:MET:CE	2.43	0.49
1:B:614:ILE:HG12	1:B:627:GLU:HG3	1.96	0.47
1:C:550:ILE:HD13	1:C:566:ALA:HB2	1.97	0.47
1:C:401:GLN:HE22	1:C:694:GLY:HA3	1.79	0.47
1:C:588:HIS:HD2	1:C:592:VAL:HG22	1.80	0.47
1:A:692:LEU:HD22	1:A:725:THR:HG21	1.98	0.46
1:A:513:SER:HB2	1:A:550:ILE:HB	1.98	0.45
1:A:421:ARG:HH12	1:A:708:MET:HG3	1.84	0.43
1:A:495:MET:HG3	1:A:526:ARG:HE	1.84	0.43
1:C:571:ILE:O	1:C:584:SER:HA	2.19	0.42
1:B:394:GLN:N	1:B:395:PRO:CD	2.82	0.42
1:A:628:TYR:CD1	1:A:686:LEU:HB2	2.54	0.42
1:A:618:ILE:HD12	1:A:618:ILE:HA	1.93	0.42
1:B:574:PHE:HB3	1:B:576:MET:HE1	2.02	0.41
1:B:666:MET:O	1:B:676:ILE:HA	2.20	0.41
1:B:702:VAL:HG12	1:B:715:THR:HG22	2.03	0.41
1:B:594:SER:HB2	1:B:607:ILE:HG22	2.03	0.40
1:B:692:LEU:HD22	1:B:725:THR:HG21	2.03	0.40

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	334/359 (93%)	325 (97%)	9 (3%)	0	100	100
1	B	330/359 (92%)	322 (98%)	8 (2%)	0	100	100
1	C	291/359 (81%)	277 (95%)	14 (5%)	0	100	100
All	All	955/1077 (89%)	924 (97%)	31 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	266/305 (87%)	252 (95%)	14 (5%)	22	26
1	B	267/305 (88%)	258 (97%)	9 (3%)	37	46
1	C	210/305 (69%)	202 (96%)	8 (4%)	33	41
All	All	743/915 (81%)	712 (96%)	31 (4%)	30	36

All (31) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	432	ILE
1	A	465	LEU
1	A	495	MET
1	A	513	SER
1	A	518	SER
1	A	523	VAL
1	A	528	LEU
1	A	623	LEU
1	A	655[A]	ARG
1	A	655[B]	ARG
1	A	657	PHE
1	A	672	THR
1	A	708	MET
1	A	724	LEU
1	B	402	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	512	CYS
1	B	528	LEU
1	B	537	GLN
1	B	620	LYS
1	B	623	LEU
1	B	657	PHE
1	B	708	MET
1	B	710	CYS
1	C	432	ILE
1	C	441	MET
1	C	451	SER
1	C	577	GLN
1	C	601	GLU
1	C	657	PHE
1	C	670	SER
1	C	691	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	401	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

3 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	ZI8	C	801	-	25,29,29	0.38	0	36,42,42	0.56	0
2	ZI8	A	801	-	25,29,29	0.34	0	36,42,42	0.46	0
2	ZI8	B	801	-	25,29,29	0.34	0	36,42,42	0.46	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	ZI8	C	801	-	-	1/19/36/36	0/4/4/4
2	ZI8	A	801	-	-	1/19/36/36	0/4/4/4
2	ZI8	B	801	-	-	1/19/36/36	0/4/4/4

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

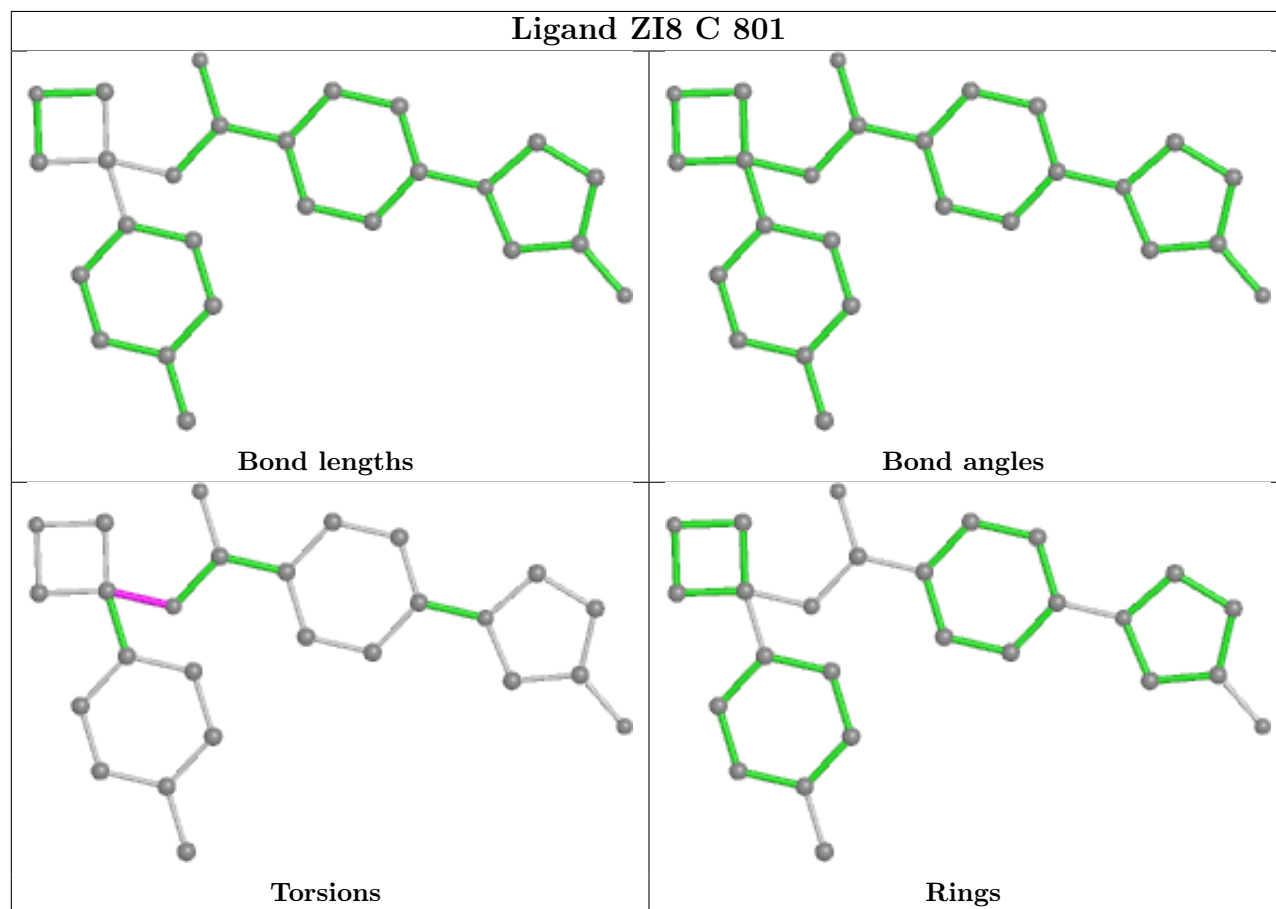
Mol	Chain	Res	Type	Atoms
2	C	801	ZI8	C19-C11-N1-C
2	A	801	ZI8	C19-C11-N1-C
2	B	801	ZI8	C19-C11-N1-C

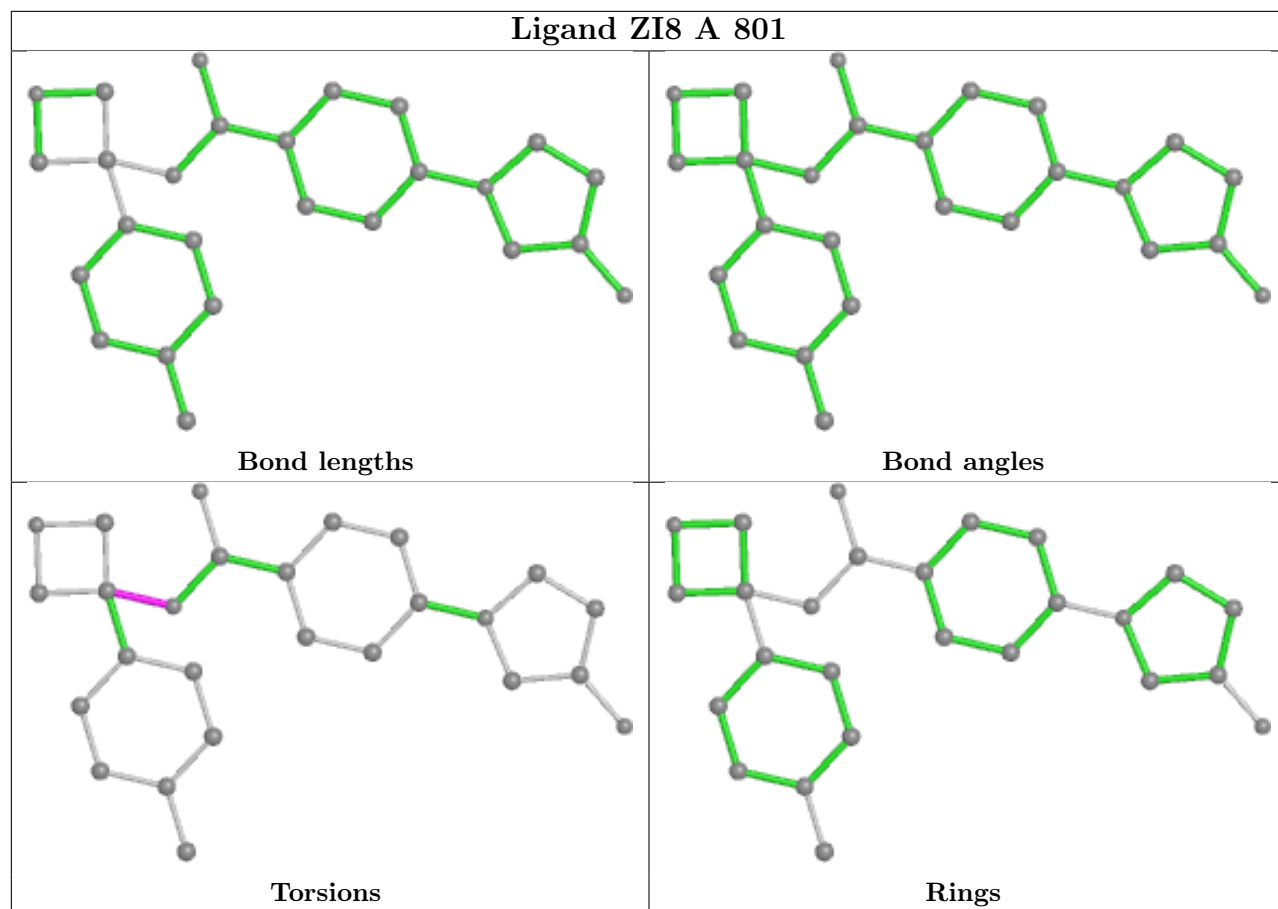
There are no ring outliers.

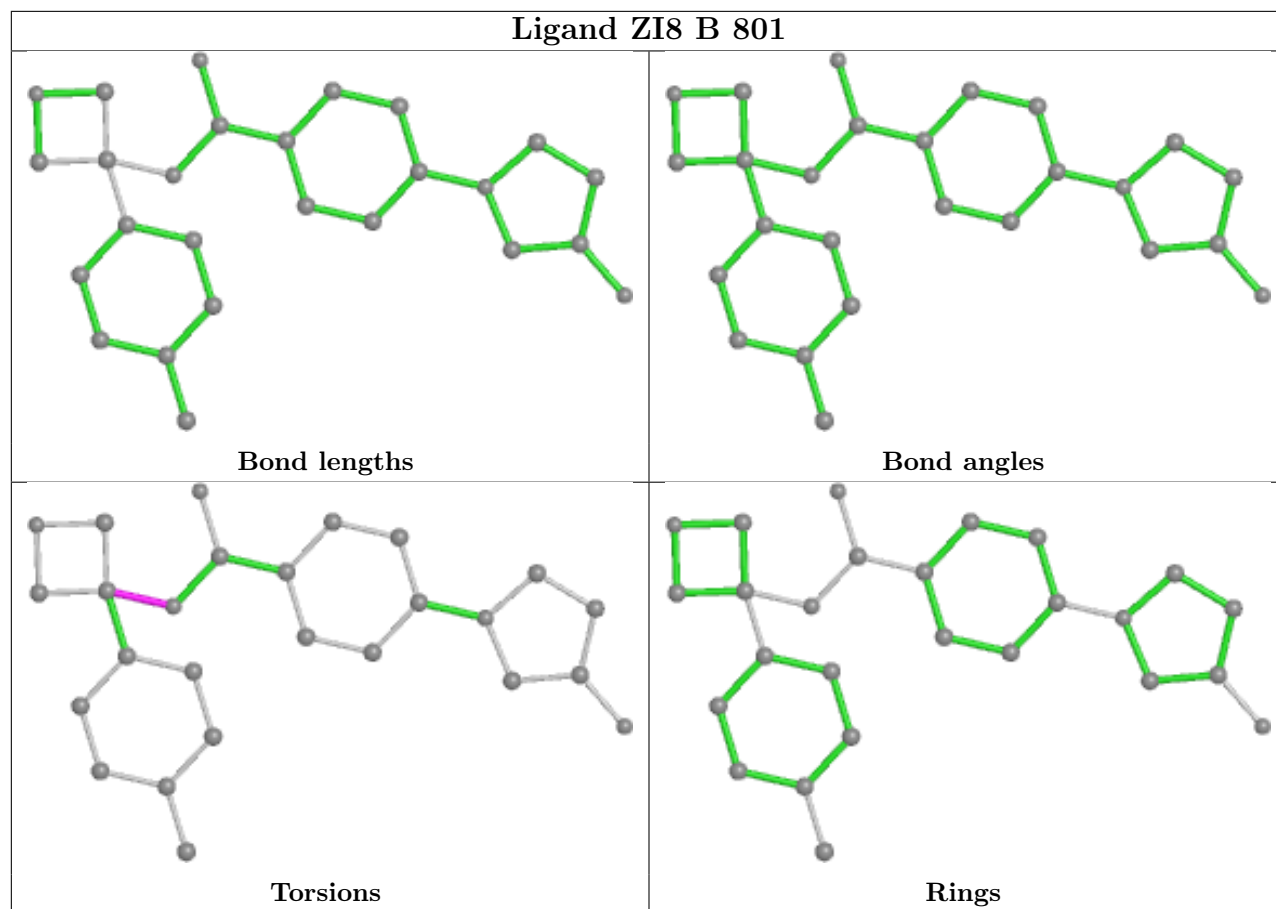
No monomer is involved in short contacts.

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

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	334/359 (93%)	-0.07	4 (1%) 79 77	22, 35, 60, 92	0
1	B	331/359 (92%)	-0.23	3 (0%) 84 83	24, 38, 61, 94	0
1	C	306/359 (85%)	1.80	124 (40%) 0 0	39, 64, 90, 108	0
All	All	971/1077 (90%)	0.46	131 (13%) 3 2	22, 44, 82, 108	0

All (131) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	518	SER	8.7
1	C	523	VAL	6.7
1	C	689	CYS	6.4
1	C	590	GLY	6.1
1	C	398	VAL	6.1
1	C	628	TYR	6.1
1	C	516	ALA	6.1
1	C	399	LEU	5.8
1	C	530	TRP	5.4
1	C	692	LEU	5.4
1	C	579	HIS	5.4
1	B	686	LEU	5.3
1	C	630	LEU	5.3
1	C	677	TYR	5.1
1	C	517	PRO	4.8
1	C	534	THR	4.7
1	C	694	GLY	4.7
1	C	397	ILE	4.7
1	C	669	CYS	4.6
1	C	691	SER	4.6
1	C	494	ASN	4.5
1	C	627	GLU	4.5
1	C	625	VAL	4.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	635	THR	4.4
1	C	674	GLY	4.4
1	C	633	ASP	4.4
1	C	512	CYS	4.4
1	C	726	THR	4.4
1	C	616	TRP	4.4
1	C	727	LEU	4.3
1	C	475	VAL	4.3
1	C	439	PRO	4.2
1	C	649	VAL	4.2
1	C	629	SER	4.2
1	C	396	PHE	4.1
1	C	670	SER	4.1
1	C	613	PHE	4.0
1	C	400	GLY	4.0
1	C	693	GLY	3.9
1	C	704	TRP	3.8
1	C	710	CYS	3.8
1	C	690	LEU	3.8
1	C	454	LEU	3.7
1	C	631	PRO	3.7
1	C	610	ASP	3.6
1	C	473	GLY	3.6
1	C	542	SER	3.5
1	C	585	TRP	3.5
1	C	709	ASP	3.5
1	C	519	LEU	3.5
1	C	715	THR	3.5
1	C	678	LYS	3.4
1	C	395	PRO	3.3
1	C	605	TYR	3.3
1	C	452	PRO	3.3
1	C	705	SER	3.3
1	C	431	VAL	3.3
1	C	589	TYR	3.2
1	C	510	PHE	3.2
1	C	608	GLY	3.2
1	C	725	THR	3.1
1	C	524	PRO	3.1
1	C	511	VAL	3.1
1	C	611	GLY	3.1
1	C	476	ARG	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	592	VAL	3.0
1	C	728	LEU	3.0
1	C	588	HIS	3.0
1	C	675	VAL	3.0
1	C	700	VAL	3.0
1	C	541	PHE	2.9
1	C	665	TYR	2.9
1	C	663	GLY	2.8
1	C	532	THR	2.8
1	B	492	ASN	2.8
1	C	672	THR	2.8
1	A	491	ILE	2.7
1	C	626	SER	2.7
1	A	686	LEU	2.7
1	C	583	MET	2.7
1	C	570	VAL	2.7
1	C	667	LEU	2.7
1	C	697	ALA	2.6
1	C	538	GLN	2.6
1	C	531	ASP	2.6
1	C	549	ALA	2.6
1	C	522	GLN	2.6
1	A	684	LYS	2.6
1	C	632	SER	2.5
1	C	430	GLY	2.5
1	C	451	SER	2.5
1	C	619	HIS	2.5
1	C	536	LYS	2.5
1	C	438	ASN	2.5
1	C	440	ILE	2.5
1	C	491	ILE	2.5
1	C	401	GLN	2.5
1	C	450	LYS	2.4
1	C	724	LEU	2.4
1	C	419	SER	2.4
1	C	661	SER	2.4
1	C	676	ILE	2.4
1	C	515	ALA	2.4
1	A	685	VAL	2.4
1	C	486	LEU	2.3
1	C	550	ILE	2.3
1	C	696	ARG	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	599	TYR	2.3
1	C	706	THR	2.3
1	C	497	ARG	2.3
1	C	607	ILE	2.3
1	C	576	MET	2.3
1	C	595	VAL	2.3
1	C	662	GLU	2.2
1	C	568	ASP	2.2
1	C	593	TYR	2.2
1	C	659	PHE	2.2
1	C	548	ILE	2.2
1	C	540	GLN	2.2
1	C	639	VAL	2.2
1	C	636	GLY	2.2
1	B	729	ALA	2.1
1	C	477	LEU	2.1
1	C	569	GLY	2.1
1	C	493	ASP	2.1
1	C	571	ILE	2.1
1	C	580	GLU	2.1
1	C	664	ASN	2.1
1	C	638	PHE	2.1
1	C	698	PRO	2.0
1	C	527	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](#)

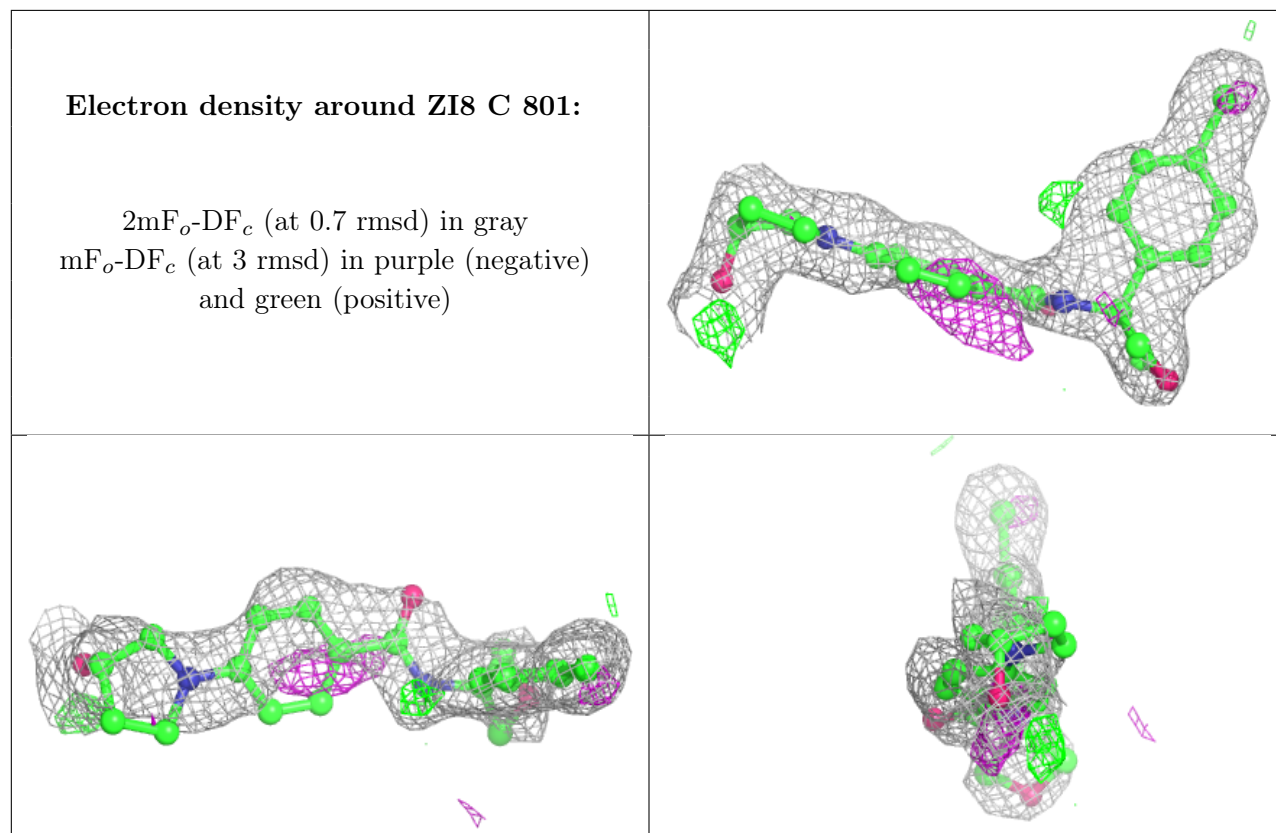
There are no monosaccharides in this entry.

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, 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.

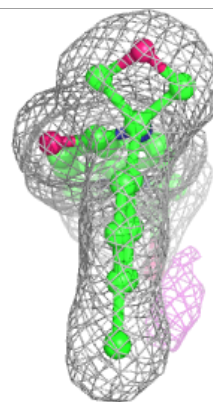
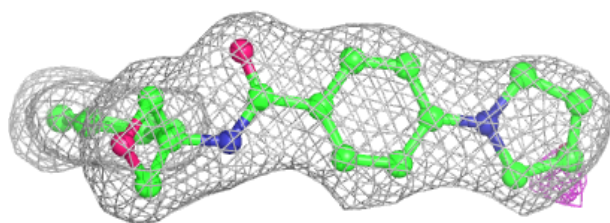
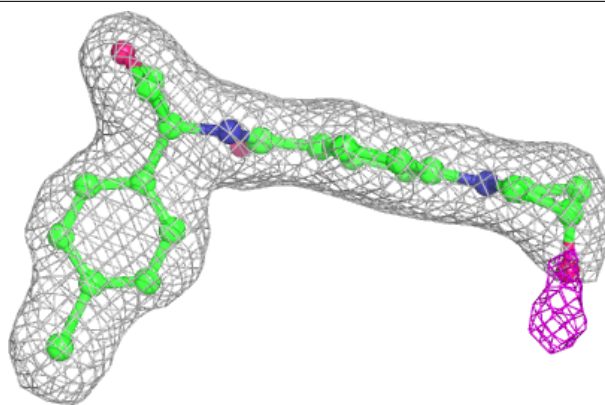
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	ZI8	C	801	26/26	0.81	0.27	50,69,79,80	0
2	ZI8	B	801	26/26	0.96	0.12	32,41,54,56	0
2	ZI8	A	801	26/26	0.96	0.13	32,39,52,58	0

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.

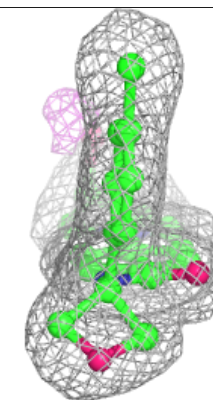
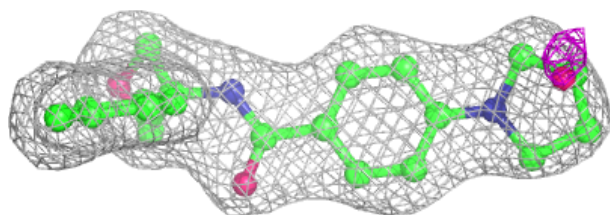
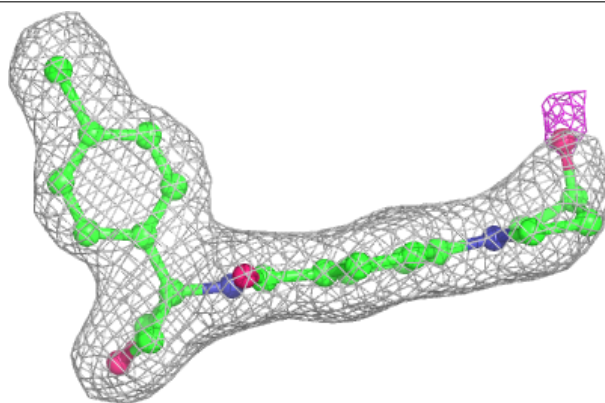


Electron density around ZI8 B 801:

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

**Electron density around ZI8 A 801:**

$2mF_o-DF_c$ (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.