



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 6, 2022 – 10:08 PM EDT

PDB ID : 5SJM  
Title : Crystal Structure of human phosphodiesterase 10 in complex with 3-(2-phenylethyl)-5-pyridin-4-yl-1H-pyridin-2-one  
Authors : Joseph, C.; Benz, J.; Flohr, A.; Bleicher, K.; Rudolph, M.G.  
Deposited on : 2022-02-01  
Resolution : 2.30 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.31.2  
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) : 2.31.2

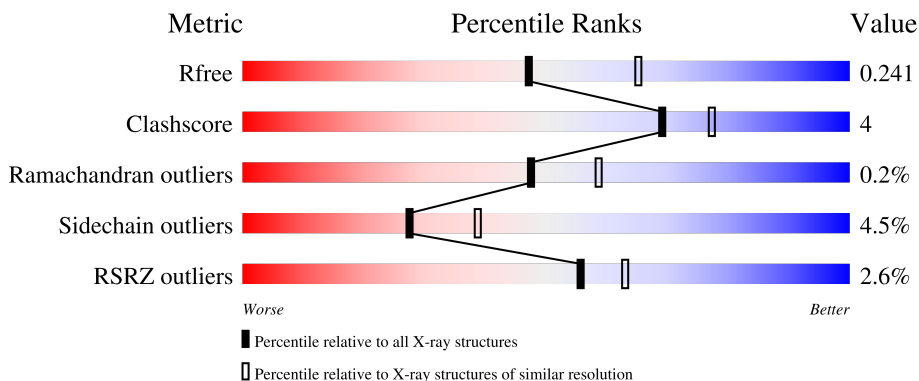
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	343	 81% 10% • 9%
1	B	343	 80% 11% • 8%
1	C	343	 80% 10% • 9%
1	D	343	 6% 78% 12% 10%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	K4L	A	803	X	-	-	-
4	K4L	B	803	X	-	-	-
4	K4L	C	803	X	-	-	-
4	K4L	D	803	X	-	-	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 10498 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 cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	313	Total 2541	C 1624	N 432	O 461	S 24	0	0	0
1	B	315	Total 2551	C 1630	N 434	O 463	S 24	0	0	0
1	C	313	Total 2541	C 1624	N 432	O 461	S 24	0	0	0
1	D	310	Total 2519	C 1612	N 429	O 454	S 24	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	447	GLY	-	expression tag	UNP Q9Y233
A	448	SER	-	expression tag	UNP Q9Y233
B	447	GLY	-	expression tag	UNP Q9Y233
B	448	SER	-	expression tag	UNP Q9Y233
C	447	GLY	-	expression tag	UNP Q9Y233
C	448	SER	-	expression tag	UNP Q9Y233
D	447	GLY	-	expression tag	UNP Q9Y233
D	448	SER	-	expression tag	UNP Q9Y233

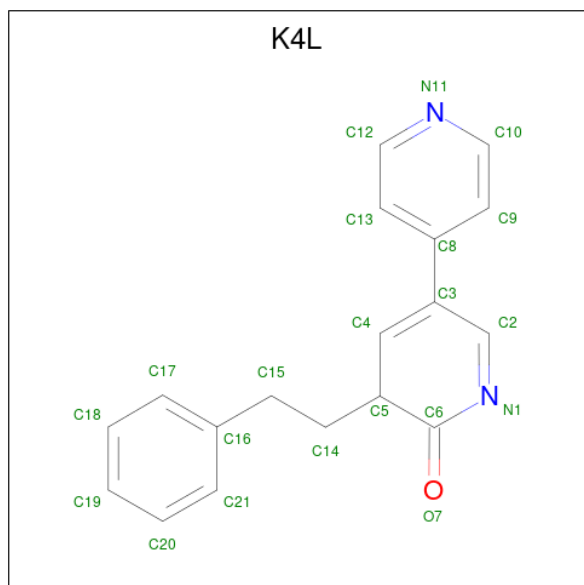
- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Zn 1	0	0
2	B	1	Total 1	Zn 1	0	0
2	C	1	Total 1	Zn 1	0	0
2	D	1	Total 1	Zn 1	0	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 4 is (5S)-5-(2-phenylethyl)[3,4'-bipyridin]-6(5H)-one (three-letter code: K4L) (formula: C<sub>18</sub>H<sub>16</sub>N<sub>2</sub>O) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O 21 18 2 1	0	0
4	B	1	Total C N O 21 18 2 1	0	0
4	C	1	Total C N O 21 18 2 1	0	0
4	D	1	Total C N O 21 18 2 1	0	0

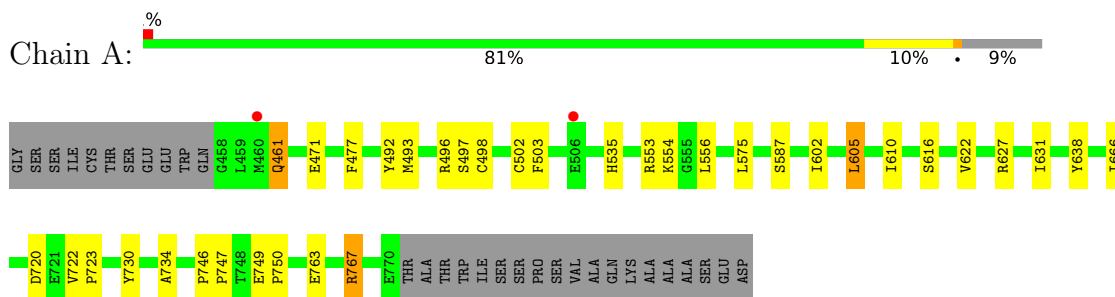
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	74	Total O 74 74	0	0
5	B	77	Total O 77 77	0	0
5	C	80	Total O 80 80	0	0
5	D	23	Total O 23 23	0	0

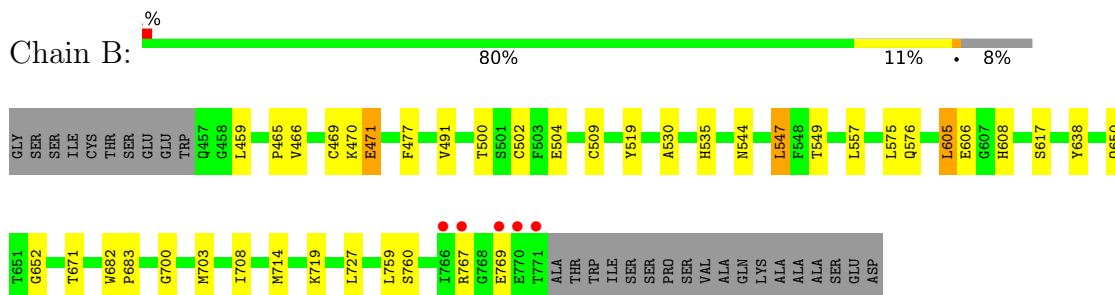
### 3 Residue-property plots i

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.

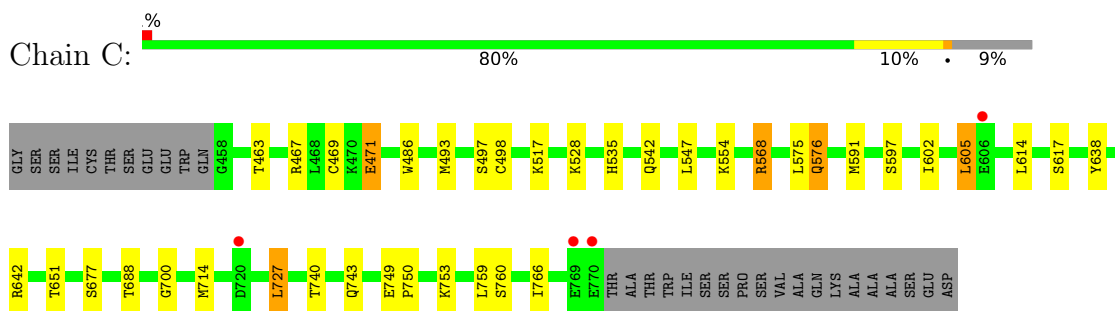
- Molecule 1: cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A



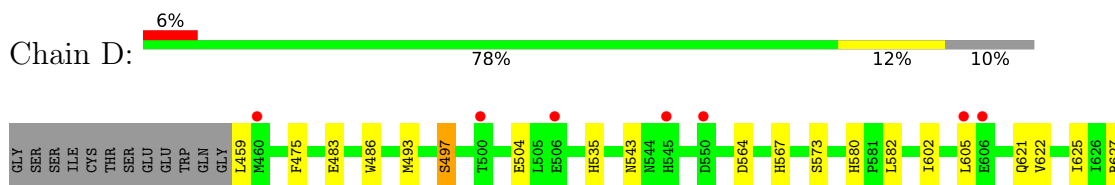
- Molecule 1: cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A

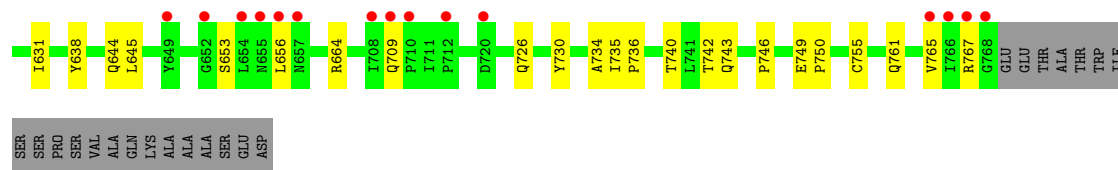


- Molecule 1: cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A



- Molecule 1: cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A







## 4 Data and refinement statistics i

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	134.80Å 134.80Å 232.62Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	43.22 – 2.30 43.22 – 2.30	Depositor EDS
% Data completeness (in resolution range)	95.2 (43.22-2.30) 95.2 (43.22-2.30)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.14 (at 2.29Å)	Xtrriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.184 , 0.242 0.188 , 0.241	Depositor DCC
$R_{free}$ test set	3481 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	41.4	Xtrriage
Anisotropy	0.302	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 42.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.027 for h,-h-k,-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	10498	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	55.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.82% 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: CME, MG, K4L, ZN

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.68	0/2592	0.81	0/3507
1	B	0.69	0/2602	0.82	0/3521
1	C	0.69	0/2592	0.81	3/3507 (0.1%)
1	D	0.70	0/2570	0.78	0/3478
All	All	0.69	0/10356	0.81	3/14013 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	568	ARG	NE-CZ-NH1	5.85	123.22	120.30
1	C	568	ARG	CG-CD-NE	5.64	123.65	111.80
1	C	568	ARG	NE-CZ-NH2	-5.20	117.70	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	2541	0	2513	19	0
1	B	2551	0	2515	17	0
1	C	2541	0	2511	19	0
1	D	2519	0	2498	21	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	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
3	D	1	0	0	0	0
4	A	21	0	0	1	0
4	B	21	0	0	0	0
4	C	21	0	0	0	0
4	D	21	0	0	2	0
5	A	74	0	0	0	0
5	B	77	0	0	0	0
5	C	80	0	0	3	0
5	D	23	0	0	0	0
All	All	10498	0	10037	75	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 4.

All (75) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:493:MET:O	1:D:497:SER:HB2	1.81	0.81
1:C:469:CYS:SG	5:C:976:HOH:O	2.43	0.75
1:A:461:GLN:HE21	1:A:461:GLN:HA	1.53	0.73
1:A:498:CYS:SG	1:A:554:LYS:HG3	2.32	0.70
1:D:493:MET:SD	1:D:535:HIS:HA	2.39	0.62
1:A:602:ILE:HA	1:A:605:LEU:HD22	1.81	0.61
1:A:498:CYS:HG	1:A:502:CYS:CB	2.14	0.60
1:B:491:VAL:HG21	1:B:509:CME:SD	2.41	0.60
1:A:492:TYR:CZ	1:A:496:ARG:HD2	2.37	0.60
1:D:726:GLN:HE22	4:D:803:K4L:C17	2.14	0.60
1:A:627:ARG:O	1:A:631:ILE:HG12	2.02	0.59
1:D:726:GLN:HE22	4:D:803:K4L:C16	2.17	0.58
1:C:740:THR:HA	1:C:743:GLN:HE21	1.69	0.57
1:A:477:PHE:HB3	1:A:535:HIS:CE1	2.40	0.57
1:C:467:ARG:O	1:C:471:GLU:HB2	2.07	0.55
1:C:493:MET:O	1:C:497:SER:HB2	2.07	0.55
1:C:498:CYS:SG	1:C:554:LYS:HG3	2.48	0.54

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:740:THR:HA	1:D:743:GLN:HE21	1.72	0.53
1:C:727:LEU:HD12	1:C:766:ILE:CD1	2.38	0.53
1:B:703:MET:HE1	1:B:708:ILE:HG21	1.91	0.52
1:C:727:LEU:HD23	1:C:759:LEU:HD11	1.91	0.52
1:D:735:ILE:HD13	1:D:755:CYS:SG	2.49	0.52
1:B:477:PHE:HB3	1:B:535:HIS:CE1	2.45	0.52
1:A:503:PHE:CE1	1:A:610:ILE:HD12	2.46	0.51
1:C:493:MET:SD	1:C:535:HIS:HA	2.51	0.51
1:C:677:SER:HB2	1:C:688:THR:HG21	1.94	0.50
1:C:651:THR:HG23	5:C:972:HOH:O	2.10	0.49
1:B:470:LYS:HE2	1:D:746:PRO:HG3	1.95	0.49
1:A:493:MET:O	1:A:497:SER:HB2	2.13	0.48
1:D:730:TYR:HA	1:D:734:ALA:HB3	1.96	0.48
1:A:749:GLU:N	1:A:750:PRO:CD	2.77	0.48
1:D:580:HIS:CE1	1:D:582:LEU:HD12	2.48	0.47
1:C:576:GLN:HB2	5:C:970:HOH:O	2.14	0.47
1:C:602:ILE:HA	1:C:605:LEU:HD22	1.96	0.47
1:D:602:ILE:O	1:D:605:LEU:HB2	2.14	0.47
1:A:722:VAL:HB	1:A:723:PRO:HD3	1.97	0.47
1:A:493:MET:SD	1:A:535:HIS:HA	2.56	0.46
1:C:575:LEU:HD11	1:C:591:MET:SD	2.55	0.46
1:A:498:CYS:SG	1:A:502:CYS:SG	3.11	0.46
1:C:542:GLN:NE2	1:C:542:GLN:HA	2.31	0.46
1:D:475:PHE:CE2	1:D:543:ASN:ND2	2.84	0.46
1:B:700:GLY:HA3	1:B:714:MET:O	2.16	0.45
1:B:544:ASN:HB3	1:B:547:LEU:HD22	1.98	0.45
1:B:519:TYR:CZ	1:B:530:ALA:HB2	2.52	0.45
1:D:749:GLU:N	1:D:750:PRO:CD	2.80	0.45
1:A:556:LEU:HD21	1:A:666:ILE:CG2	2.47	0.45
1:B:682:TRP:HB3	1:B:683:PRO:HD3	1.98	0.44
1:D:622:VAL:HA	1:D:625:ILE:HD12	2.00	0.44
1:C:749:GLU:HB3	1:C:750:PRO:HD3	2.00	0.44
1:D:761:GLN:O	1:D:765:VAL:HG23	2.18	0.44
1:B:470:LYS:HD2	1:D:742:THR:HG22	2.00	0.44
1:D:645:LEU:HD11	1:D:664:ARG:O	2.18	0.43
1:B:727:LEU:HD23	1:B:759:LEU:CD1	2.47	0.43
1:B:767:ARG:NH2	1:B:769:GLU:OE1	2.51	0.43
1:B:638:TYR:CD1	1:B:671:THR:HG21	2.52	0.43
1:A:503:PHE:CD1	1:A:610:ILE:HB	2.54	0.43
1:A:730:TYR:HA	1:A:734:ALA:HB3	2.01	0.42
1:A:556:LEU:HD21	1:A:666:ILE:HG21	2.01	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:627:ARG:O	1:D:631:ILE:HG12	2.20	0.42
1:A:746:PRO:N	1:A:747:PRO:CD	2.82	0.42
1:D:735:ILE:O	1:D:736:PRO:C	2.58	0.42
1:C:727:LEU:HD23	1:C:759:LEU:CD1	2.49	0.42
1:B:470:LYS:HG3	1:B:471:GLU:N	2.33	0.41
1:D:735:ILE:HB	1:D:736:PRO:HD3	2.02	0.41
1:B:605:LEU:HB3	1:B:608:HIS:ND1	2.35	0.41
1:B:557:LEU:HD23	1:B:557:LEU:HA	1.95	0.41
1:C:638:TYR:OH	1:C:642:ARG:HD3	2.20	0.41
4:A:803:K4L:C16	4:A:803:K4L:C4	2.99	0.40
1:C:486:TRP:CH2	1:C:528:LYS:HB2	2.56	0.40
1:C:700:GLY:HA3	1:C:714:MET:O	2.21	0.40
1:B:465:PRO:O	1:B:466:VAL:C	2.59	0.40
1:D:483:GLU:HA	1:D:486:TRP:CE3	2.56	0.40
1:A:763:GLU:O	1:A:767:ARG:HG3	2.21	0.40
1:B:727:LEU:HD23	1:B:727:LEU:HA	1.88	0.40
1:D:564:ASP:O	1:D:567:HIS:HB2	2.21	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	310/343 (90%)	298 (96%)	11 (4%)	1 (0%)	41	50
1	B	312/343 (91%)	298 (96%)	13 (4%)	1 (0%)	41	50
1	C	310/343 (90%)	305 (98%)	5 (2%)	0	100	100
1	D	307/343 (90%)	286 (93%)	21 (7%)	0	100	100
All	All	1239/1372 (90%)	1187 (96%)	50 (4%)	2 (0%)	47	58

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	767	ARG
1	B	652	GLY

### 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	281/305 (92%)	271 (96%)	10 (4%)	35	49
1	B	281/305 (92%)	265 (94%)	16 (6%)	20	28
1	C	281/305 (92%)	268 (95%)	13 (5%)	27	38
1	D	279/305 (92%)	268 (96%)	11 (4%)	32	46
All	All	1122/1220 (92%)	1072 (96%)	50 (4%)	27	39

All (50) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	461	GLN
1	A	471	GLU
1	A	553	ARG
1	A	575	LEU
1	A	587	SER
1	A	605	LEU
1	A	616	SER
1	A	622	VAL
1	A	638	TYR
1	A	720	ASP
1	B	459	LEU
1	B	469	CYS
1	B	471	GLU
1	B	500	THR
1	B	502	CYS
1	B	504	GLU
1	B	547	LEU
1	B	549	THR
1	B	575	LEU
1	B	576	GLN

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	B	605	LEU
1	B	606	GLU
1	B	617	SER
1	B	650	GLN
1	B	719	LYS
1	B	760	SER
1	C	463	THR
1	C	471	GLU
1	C	517	LYS
1	C	547	LEU
1	C	568	ARG
1	C	576	GLN
1	C	597	SER
1	C	605	LEU
1	C	614	LEU
1	C	617	SER
1	C	727	LEU
1	C	753	LYS
1	C	760	SER
1	D	459	LEU
1	D	497	SER
1	D	504	GLU
1	D	573	SER
1	D	621	GLN
1	D	638	TYR
1	D	644	GLN
1	D	653	SER
1	D	656	LEU
1	D	709	GLN
1	D	767	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (24) such sidechains are listed below:

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	A	461	GLN
1	A	604	GLN
1	A	644	GLN
1	A	726	GLN
1	A	743	GLN
1	B	484	ASN
1	B	576	GLN
1	B	604	GLN

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type
1	B	726	GLN
1	B	743	GLN
1	C	542	GLN
1	C	604	GLN
1	C	709	GLN
1	C	726	GLN
1	C	743	GLN
1	D	476	HIS
1	D	542	GLN
1	D	576	GLN
1	D	604	GLN
1	D	644	GLN
1	D	650	GLN
1	D	726	GLN
1	D	743	GLN
1	D	761	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](#)

4 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	CME	D	509	1	8,9,10	0.54	0	5,9,11	0.82	0
1	CME	C	509	1	8,9,10	0.57	0	5,9,11	0.40	0
1	CME	B	509	1	8,9,10	0.43	0	5,9,11	1.02	0
1	CME	A	509	1	8,9,10	0.44	0	5,9,11	0.71	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	CME	D	509	1	-	2/5/8/10	-
1	CME	C	509	1	-	2/5/8/10	-
1	CME	B	509	1	-	3/5/8/10	-
1	CME	A	509	1	-	2/5/8/10	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	B	509	CME	SD-CE-CZ-OH
1	C	509	CME	CE-SD-SG-CB
1	B	509	CME	CE-SD-SG-CB
1	A	509	CME	SD-CE-CZ-OH
1	A	509	CME	CZ-CE-SD-SG
1	B	509	CME	CZ-CE-SD-SG
1	C	509	CME	CZ-CE-SD-SG
1	D	509	CME	N-CA-CB-SG
1	D	509	CME	SD-CE-CZ-OH

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	509	CME	1	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 8 are monoatomic - leaving 4 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	K4L	C	803	-	21,23,23	1.82	4 (19%)	19,30,30	2.39	5 (26%)
4	K4L	B	803	-	21,23,23	1.76	3 (14%)	19,30,30	2.44	5 (26%)
4	K4L	D	803	-	21,23,23	1.74	4 (19%)	19,30,30	1.78	5 (26%)
4	K4L	A	803	-	21,23,23	1.78	3 (14%)	19,30,30	2.04	4 (21%)

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	K4L	C	803	-	1/1/3/5	1/9/22/22	0/3/3/3
4	K4L	B	803	-	1/1/3/5	1/9/22/22	0/3/3/3
4	K4L	D	803	-	1/1/3/5	1/9/22/22	0/3/3/3
4	K4L	A	803	-	1/1/3/5	1/9/22/22	0/3/3/3

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	803	K4L	C5-C4	-6.90	1.37	1.50
4	B	803	K4L	C5-C4	-6.31	1.38	1.50
4	A	803	K4L	C5-C4	-6.26	1.38	1.50
4	D	803	K4L	C5-C4	-6.25	1.38	1.50
4	D	803	K4L	C5-C6	-3.18	1.40	1.52
4	A	803	K4L	C5-C6	-2.98	1.40	1.52
4	B	803	K4L	C5-C6	-2.82	1.41	1.52
4	C	803	K4L	C5-C6	-2.66	1.42	1.52
4	D	803	K4L	C2-N1	2.42	1.36	1.32
4	A	803	K4L	C2-N1	2.38	1.36	1.32
4	B	803	K4L	C2-N1	2.19	1.36	1.32
4	C	803	K4L	C14-C5	-2.11	1.51	1.54
4	C	803	K4L	C2-N1	2.04	1.35	1.32
4	D	803	K4L	C14-C5	-2.02	1.51	1.54

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	803	K4L	C14-C5-C6	7.69	117.36	108.89
4	B	803	K4L	C14-C5-C6	7.56	117.21	108.89
4	A	803	K4L	C14-C5-C6	6.19	115.71	108.89
4	D	803	K4L	C14-C5-C6	5.29	114.72	108.89
4	C	803	K4L	C21-C16-C17	3.46	123.60	118.17
4	B	803	K4L	C9-C10-N11	-3.38	117.73	123.62
4	B	803	K4L	C21-C16-C17	3.26	123.29	118.17
4	A	803	K4L	C9-C10-N11	-3.16	118.11	123.62
4	B	803	K4L	C12-N11-C10	2.91	123.69	116.85
4	A	803	K4L	C12-N11-C10	2.90	123.66	116.85
4	D	803	K4L	C12-N11-C10	2.78	123.39	116.85
4	C	803	K4L	C12-N11-C10	2.71	123.22	116.85
4	B	803	K4L	C12-C13-C8	-2.68	116.12	119.05
4	C	803	K4L	C20-C21-C16	-2.52	116.77	120.63
4	C	803	K4L	C9-C10-N11	-2.24	119.71	123.62
4	A	803	K4L	C21-C16-C17	2.17	121.58	118.17
4	D	803	K4L	C9-C10-N11	-2.12	119.93	123.62
4	D	803	K4L	O7-C6-N1	2.10	122.16	119.91
4	D	803	K4L	C13-C12-N11	-2.07	120.02	123.62

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	A	803	K4L	C5
4	B	803	K4L	C5
4	C	803	K4L	C5
4	D	803	K4L	C5

All (4) torsion outliers are listed below:

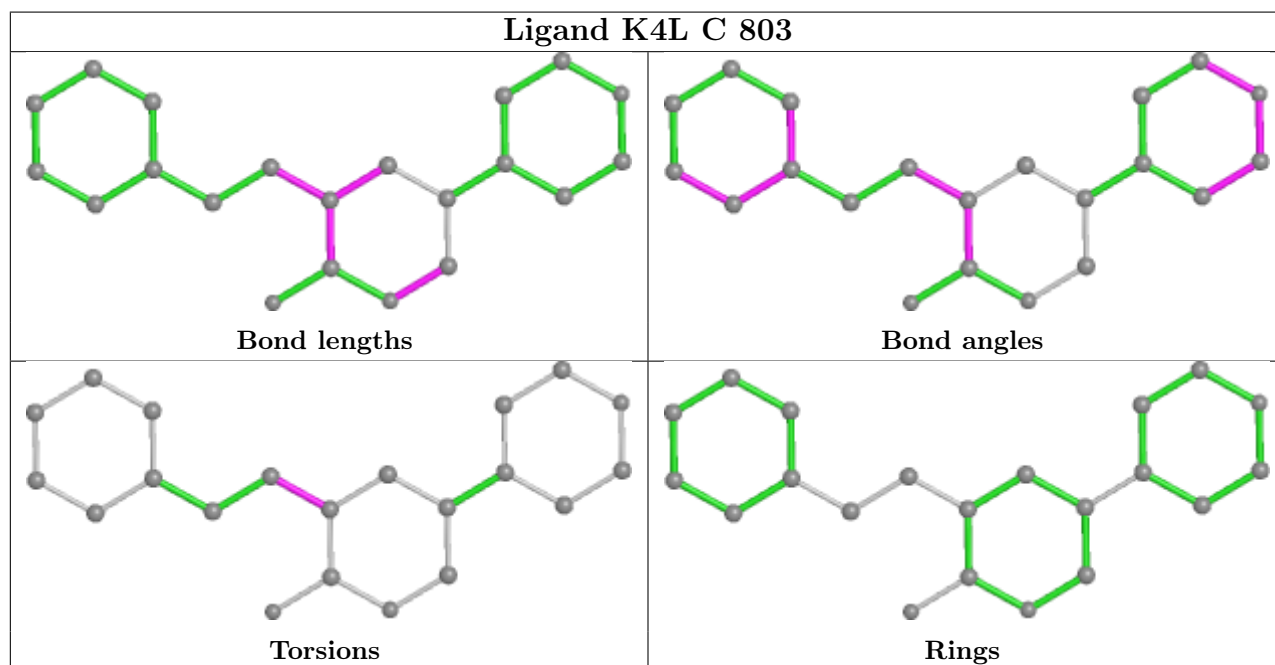
Mol	Chain	Res	Type	Atoms
4	A	803	K4L	C15-C14-C5-C4
4	B	803	K4L	C15-C14-C5-C4
4	C	803	K4L	C15-C14-C5-C4
4	D	803	K4L	C15-C14-C5-C4

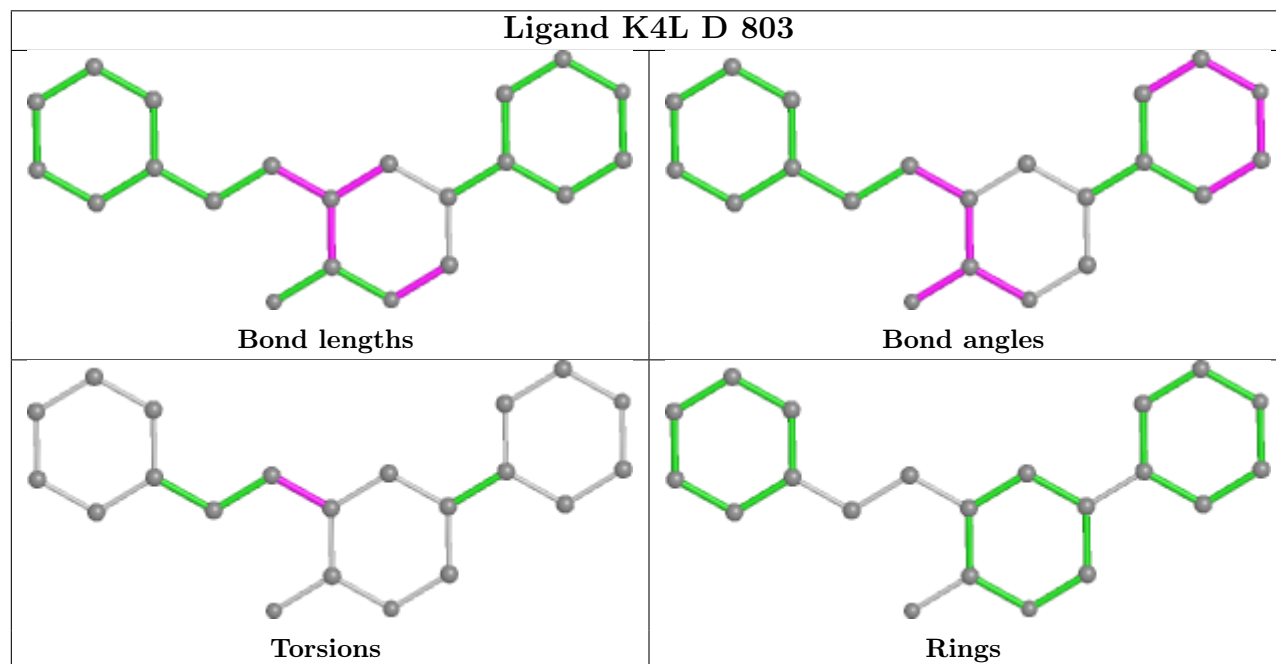
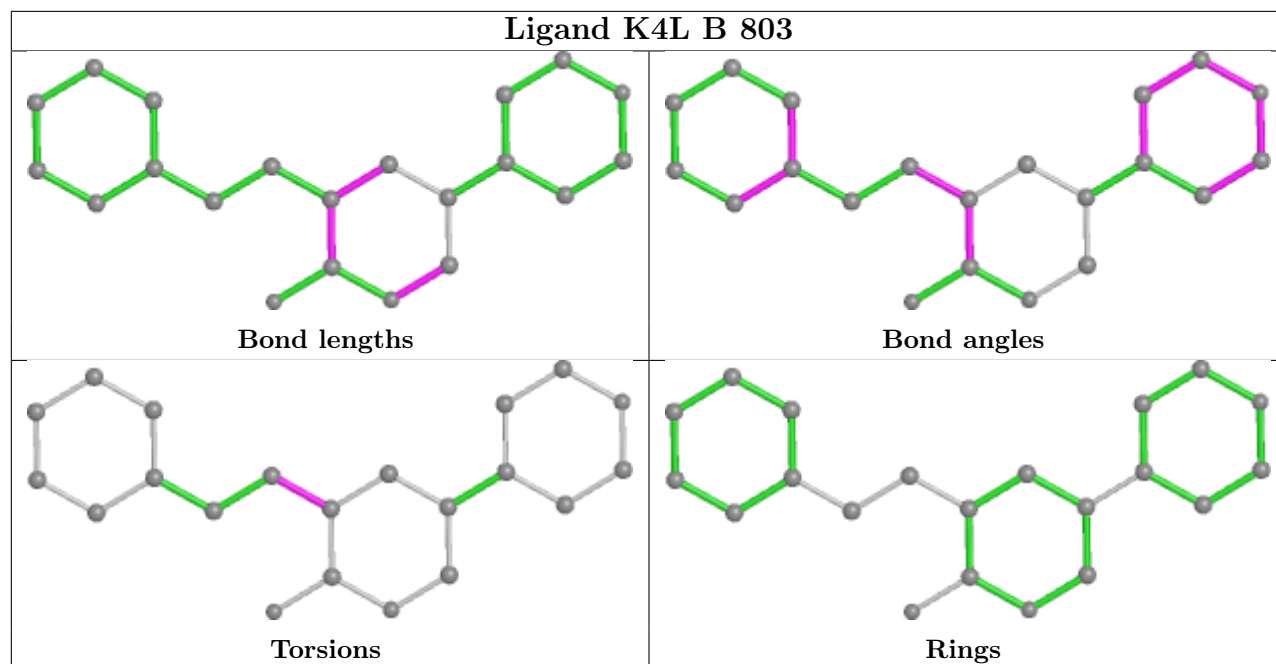
There are no ring outliers.

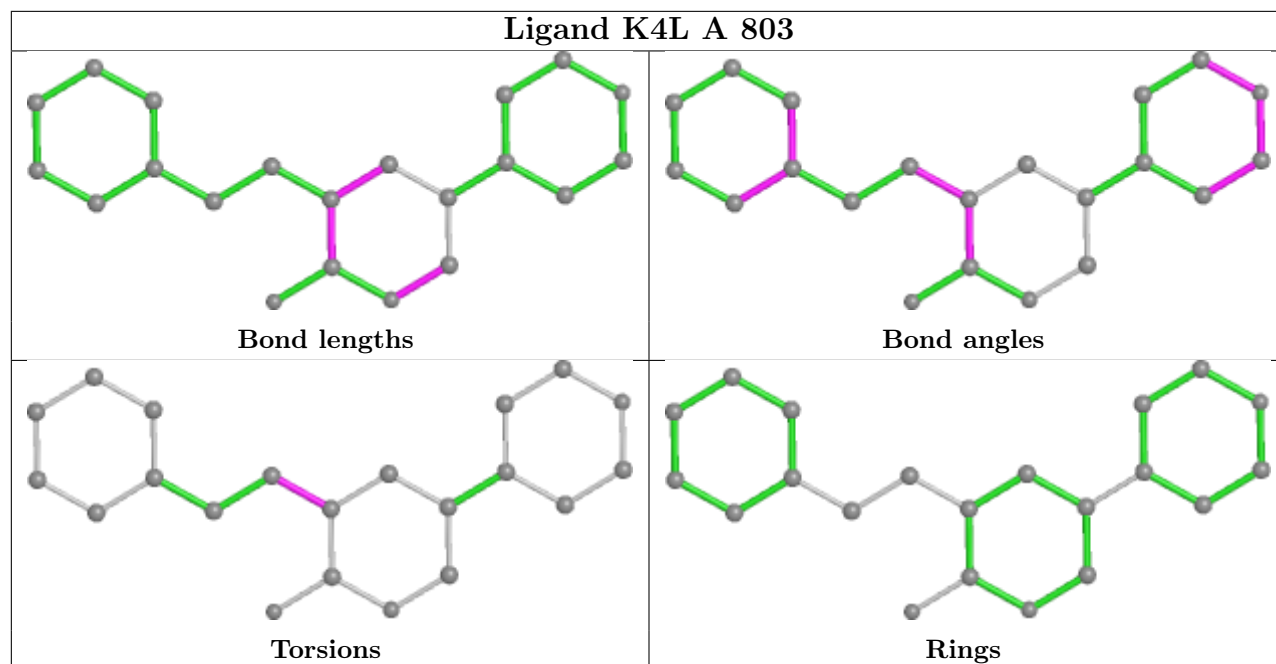
2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	803	K4L	2	0
4	A	803	K4L	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

### 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	312/343 (90%)	-0.16	2 (0%) 89 92	32, 45, 78, 112	0
1	B	314/343 (91%)	-0.13	5 (1%) 72 77	30, 47, 81, 110	0
1	C	312/343 (90%)	-0.23	4 (1%) 77 81	28, 45, 72, 117	0
1	D	309/343 (90%)	0.23	22 (7%) 16 21	45, 69, 99, 128	0
All	All	1247/1372 (90%)	-0.07	33 (2%) 56 63	28, 52, 89, 128	0

All (33) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	460	MET	6.7
1	C	770	GLU	4.5
1	D	768	GLY	4.1
1	D	606	GLU	4.0
1	B	769	GLU	3.9
1	D	657	ASN	3.7
1	D	656	LEU	3.6
1	D	766	ILE	3.4
1	D	654	LEU	3.4
1	C	720	ASP	3.3
1	D	652	GLY	3.1
1	B	771	THR	3.1
1	D	550	ASP	2.9
1	D	712	PRO	2.8
1	D	767	ARG	2.8
1	D	765	VAL	2.8
1	A	506	GLU	2.8
1	A	460	MET	2.6
1	D	720	ASP	2.6
1	B	770	GLU	2.6
1	D	655	ASN	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	766	ILE	2.5
1	C	769	GLU	2.4
1	D	500	THR	2.3
1	D	710	PRO	2.3
1	D	605	LEU	2.3
1	D	545	HIS	2.3
1	B	767	ARG	2.2
1	C	606	GLU	2.2
1	D	708	ILE	2.2
1	D	709	GLN	2.1
1	D	649	TYR	2.1
1	D	506	GLU	2.0

## 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	CME	C	509	10/11	0.89	0.14	46,57,90,94	0
1	CME	B	509	10/11	0.90	0.17	50,62,102,108	0
1	CME	D	509	10/11	0.91	0.25	69,99,132,139	0
1	CME	A	509	10/11	0.92	0.12	50,59,88,95	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(Å <sup>2</sup> )	Q<0.9
4	K4L	B	803	21/21	0.92	0.19	46,57,64,70	0
4	K4L	C	803	21/21	0.94	0.17	48,55,69,76	0

*Continued on next page...*



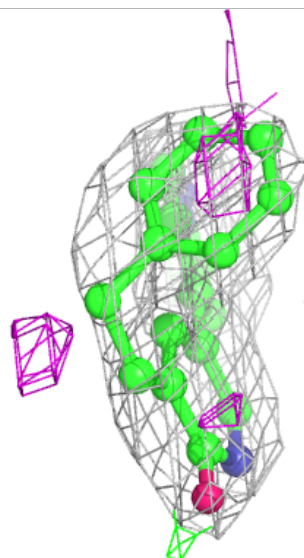
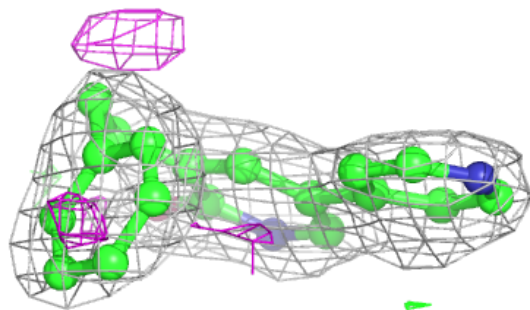
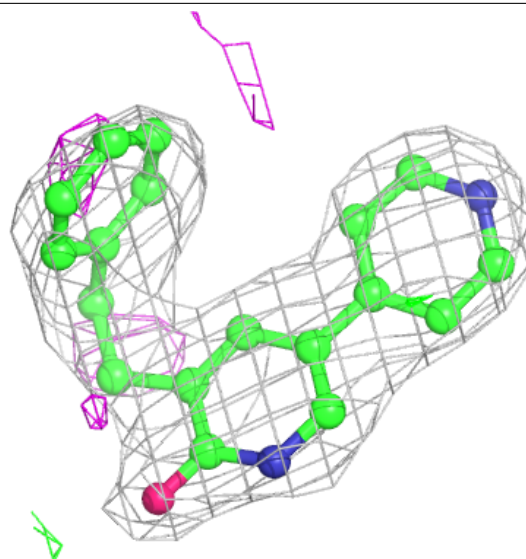
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	K4L	D	803	21/21	0.94	0.17	60,67,72,75	0
4	K4L	A	803	21/21	0.95	0.14	44,53,68,71	0
3	MG	B	802	1/1	0.97	0.08	28,28,28,28	0
3	MG	D	802	1/1	0.98	0.04	50,50,50,50	0
3	MG	C	802	1/1	0.99	0.12	32,32,32,32	0
3	MG	A	802	1/1	0.99	0.05	28,28,28,28	0
2	ZN	D	801	1/1	0.99	0.10	55,55,55,55	0
2	ZN	A	801	1/1	1.00	0.11	37,37,37,37	0
2	ZN	B	801	1/1	1.00	0.10	34,34,34,34	0
2	ZN	C	801	1/1	1.00	0.12	37,37,37,37	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.

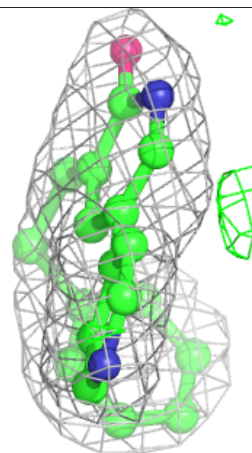
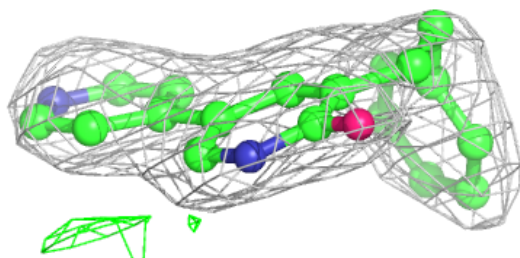
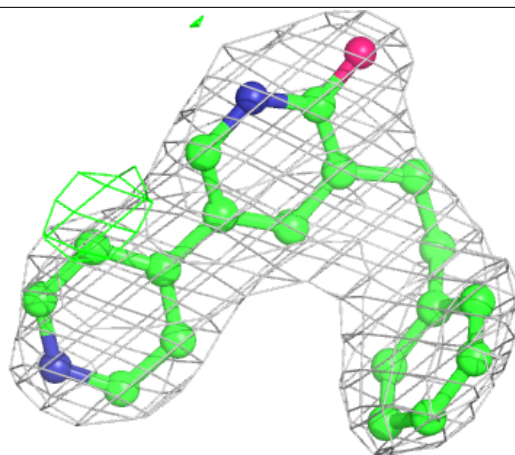
**Electron density around K4L B 803:**

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



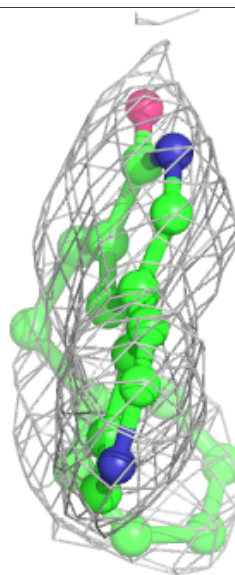
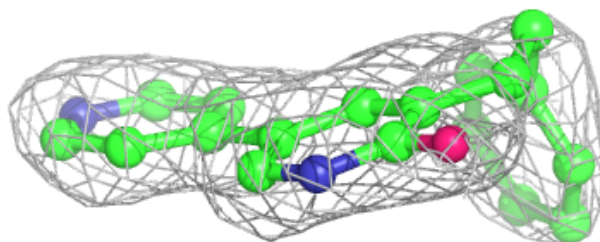
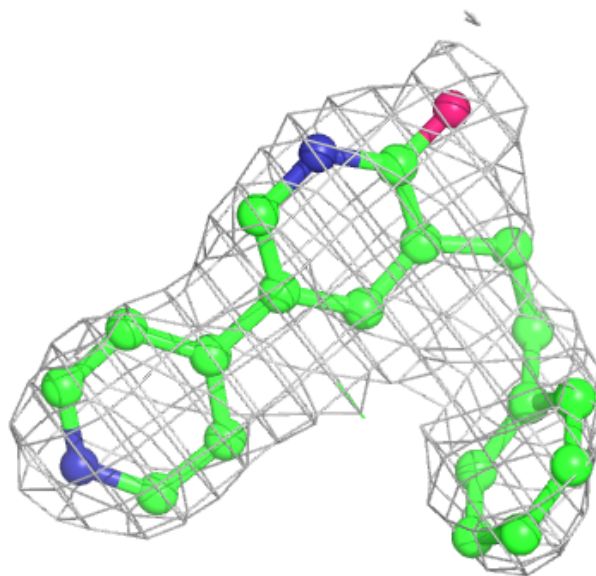
**Electron density around K4L C 803:**

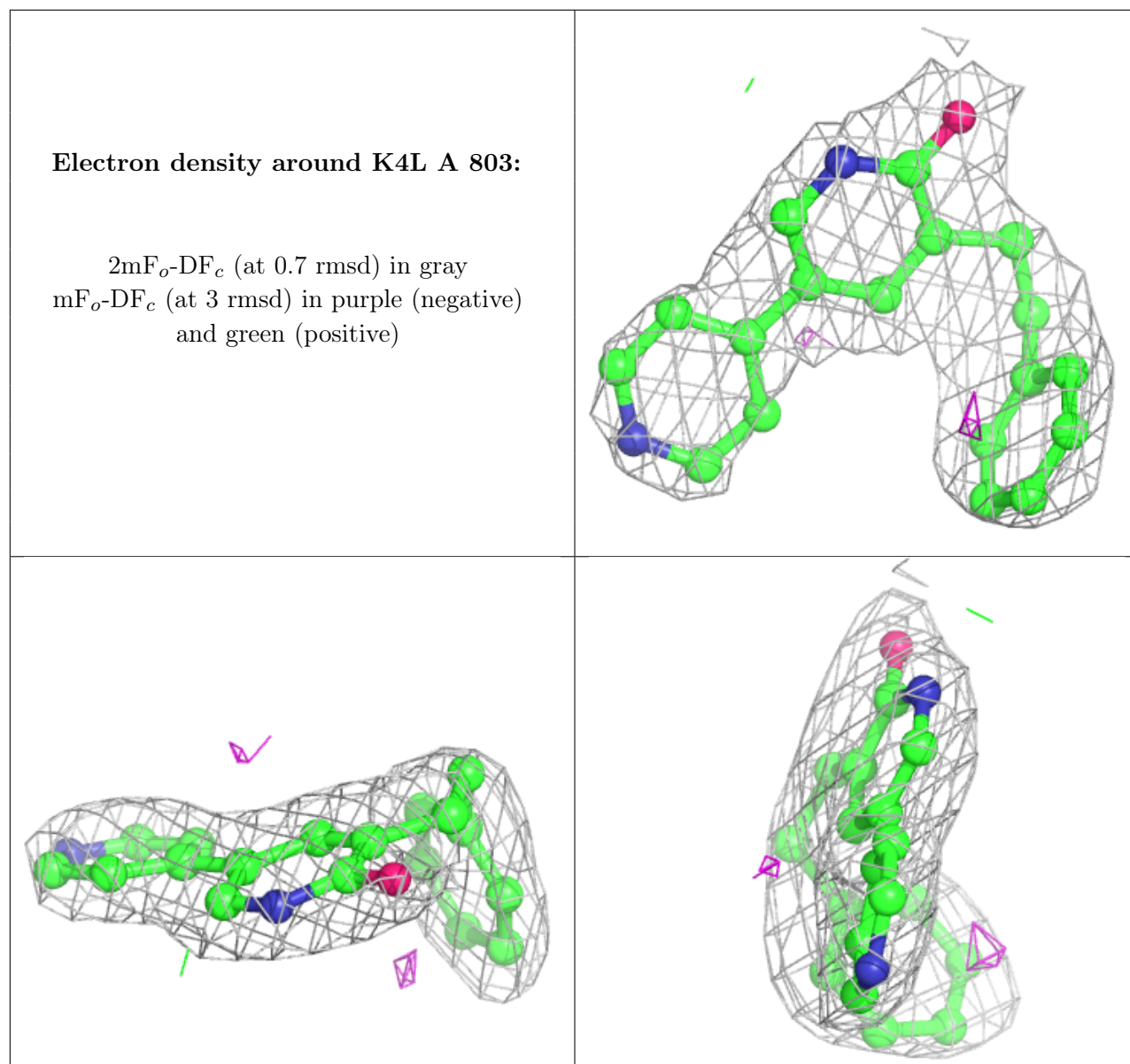
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around K4L D 803:**

$2mF_o-DF_c$  (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.