



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 6, 2022 – 07:51 PM EDT

PDB ID : 5SJU  
Title : Crystal Structure of human phosphodiesterase 10 in complex with 6-cyclopropyl-N-[4-(piperidin-1-ylmethyl)-1,3-thiazol-2-yl]-3-(pyrimidin-5-ylamino)pyrazine-2-carboxamide  
Authors : Joseph, C.; Benz, J.; Flohr, A.; Koerner, M.; Rudolph, M.G.  
Deposited on : 2022-02-01  
Resolution : 1.97 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)  
Xtrriage (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

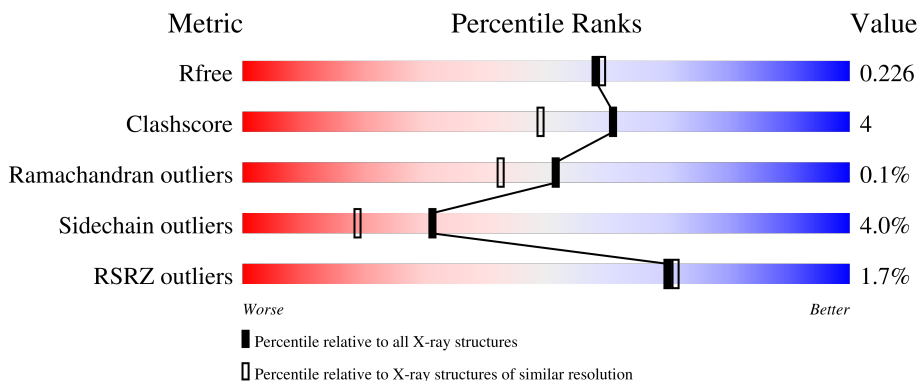
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.97 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	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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	 2% 82% 8% 7%
1	B	343	 % 82% 9% 7%
1	C	343	 2% 80% 10% 7%
1	D	343	 % 84% 6% 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:

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
4	K4U	A	803	X	-	-	-
4	K4U	B	803	X	-	-	-
4	K4U	C	803	X	-	-	-
4	K4U	D	803	X	-	-	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 11072 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	318	2590	1657	441	468	24	0	1	0
1	B	319	2595	1660	442	469	24	0	1	0
1	C	318	2598	1662	444	468	24	0	2	0
1	D	310	2519	1612	429	454	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

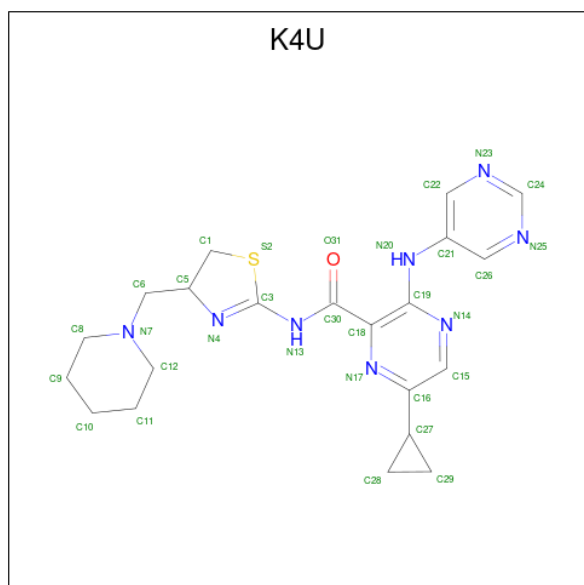
- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

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 6-cyclopropyl-N-{(4S)-4-[(piperidin-1-yl)methyl]-4,5-dihydro-1,3-thiazol-2-yl}-3-[(pyrimidin-5-yl)amino]pyrazine-2-carboxamide (three-letter code: K4U) (formula: C<sub>21</sub>H<sub>26</sub>N<sub>8</sub>OS) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O S 31 21 8 1 1	0	0
4	B	1	Total C N O S 31 21 8 1 1	0	0
4	C	1	Total C N O S 31 21 8 1 1	0	0
4	D	1	Total C N O S 31 21 8 1 1	0	0


- Molecule 5 is water.

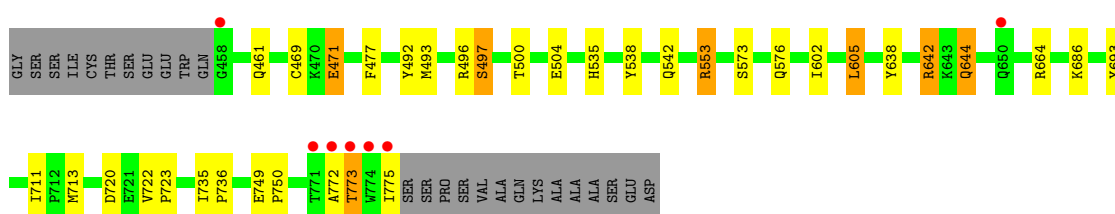
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
5	A	163	Total 163	O 163	0	0
5	B	177	Total 177	O 177	0	0
5	C	177	Total 177	O 177	0	0
5	D	121	Total 121	O 121	0	0

### 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.

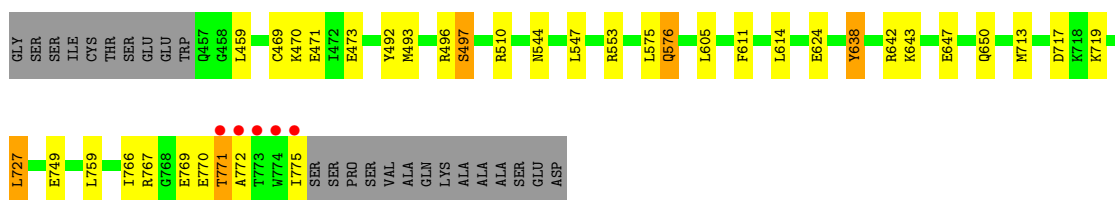
- Molecule 1: cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A

Chain A: 




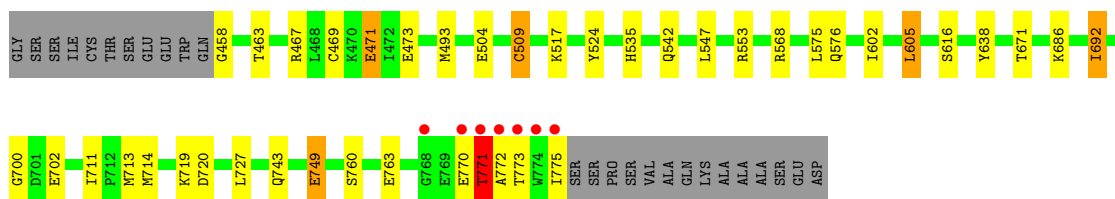
- Molecule 1: cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A

Chain B: 




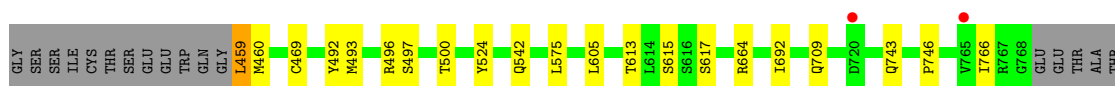
- Molecule 1: cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A

Chain C: 



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

Chain D: 



TRP  
ILE  
SER  
SER  
PRO  
SER  
VAL  
ALA  
GLN  
LYS  
ALA  
ALA  
ALA  
SER  
GLU  
ASP



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	135.63Å 135.63Å 235.65Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	43.63 – 1.97 43.63 – 1.97	Depositor EDS
% Data completeness (in resolution range)	96.2 (43.63-1.97) 96.2 (43.63-1.97)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.08 (at 1.97Å)	Xtrriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.176 , 0.220 0.185 , 0.226	Depositor DCC
$R_{free}$ test set	5705 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	33.0	Xtrriage
Anisotropy	0.074	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 42.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.026 for h,-h-k,-l	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	11072	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	41.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.65% 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: K4U, ZN, MG, CME

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	0.81	0/2646	0.87	4/3582 (0.1%)
1	B	0.78	1/2651 (0.0%)	0.87	1/3589 (0.0%)
1	C	0.79	2/2657 (0.1%)	0.86	0/3596
1	D	0.78	0/2570	0.84	1/3478 (0.0%)
All	All	0.79	3/10524 (0.0%)	0.86	6/14245 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	473	GLU	CD-OE1	-5.87	1.19	1.25
1	B	624	GLU	CD-OE2	-5.16	1.20	1.25
1	C	702	GLU	CD-OE2	-5.05	1.20	1.25

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	664	ARG	NE-CZ-NH1	6.03	123.32	120.30
1	B	510	ARG	NE-CZ-NH1	5.83	123.22	120.30
1	A	642	ARG	NE-CZ-NH2	-5.62	117.49	120.30
1	A	664	ARG	NE-CZ-NH2	-5.27	117.67	120.30
1	A	693	TYR	CB-CG-CD1	5.19	124.12	121.00
1	D	664	ARG	NE-CZ-NH1	5.05	122.83	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	771	THR	Peptide

## 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	2590	0	2564	32	0
1	B	2595	0	2566	20	0
1	C	2598	0	2577	25	0
1	D	2519	0	2496	12	0
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	31	0	0	3	0
4	B	31	0	0	1	0
4	C	31	0	0	1	0
4	D	31	0	0	2	0
5	A	163	0	0	6	0
5	B	177	0	0	3	0
5	C	177	0	0	6	0
5	D	121	0	0	4	0
All	All	11072	0	10203	92	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 (92) 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:C:772:ALA:HB3	1:C:775:ILE:HD12	1.35	1.07
1:C:469:CYS:SG	5:C:1056:HOH:O	2.10	1.06
1:D:469:CYS:SG	5:D:1013:HOH:O	2.17	1.02
1:A:469:CYS:SG	5:A:1047:HOH:O	2.20	0.98
1:B:576:GLN:HB2	5:B:1039:HOH:O	1.69	0.93
1:B:469:CYS:SG	5:B:1050:HOH:O	2.28	0.92
1:B:470:LYS:HE2	1:D:746:PRO:HG3	1.53	0.89
1:A:773:THR:HG21	1:B:717:ASP:OD1	1.72	0.89
1:A:461:GLN:HG3	5:A:970:HOH:O	1.75	0.87
1:A:471:GLU:HG2	5:A:1043:HOH:O	1.77	0.83
1:D:743:GLN:HG2	5:D:1016:HOH:O	1.79	0.83
1:A:772:ALA:O	1:A:773:THR:HG23	1.85	0.77
1:A:773:THR:CG2	1:B:717:ASP:OD1	2.34	0.75
1:A:553:ARG:HD3	5:A:941:HOH:O	1.88	0.73
1:C:568[A]:ARG:NH1	5:C:901:HOH:O	2.22	0.72
1:D:459:LEU:N	5:D:902:HOH:O	2.23	0.70
1:C:686:LYS:HE2	1:C:772:ALA:HA	1.75	0.69
1:A:553:ARG:NH1	5:A:901:HOH:O	2.22	0.68
1:B:772:ALA:HB3	1:B:775:ILE:HD12	1.76	0.67
1:C:772:ALA:HB3	1:C:775:ILE:CD1	2.19	0.64
1:C:686:LYS:CE	1:C:772:ALA:HA	2.29	0.61
1:A:772:ALA:HB2	1:A:775:ILE:HD12	1.84	0.59
1:C:524:TYR:CD2	1:C:692:ILE:HD13	2.38	0.59
1:A:602:ILE:HA	1:A:605:LEU:HD22	1.88	0.56
1:D:524:TYR:CE2	1:D:692:ILE:HD12	2.40	0.56
1:A:461:GLN:NE2	1:A:461:GLN:HA	2.21	0.55
1:A:686:LYS:HE3	1:A:772:ALA:HA	1.89	0.55
1:A:461:GLN:HE22	1:A:500:THR:CG2	2.20	0.55
1:D:493:MET:O	1:D:497:SER:HB3	2.06	0.55
1:A:644:GLN:HE21	1:A:644:GLN:HA	1.72	0.54
1:A:496:ARG:NH1	1:A:538:TYR:OH	2.39	0.54
1:D:496:ARG:NH1	5:D:904:HOH:O	2.40	0.54
4:A:803:K4U:S2	4:A:803:K4U:O31	2.65	0.54
1:A:576:GLN:HB3	5:A:1048:HOH:O	2.07	0.53
1:C:553:ARG:NH1	5:C:904:HOH:O	2.39	0.53
4:D:803:K4U:S2	4:D:803:K4U:O31	2.66	0.53
1:A:492:TYR:CZ	1:A:496:ARG:HD2	2.45	0.52
1:A:493:MET:O	1:A:497:SER:HB2	2.09	0.52
1:C:568[A]:ARG:HG2	1:C:568[A]:ARG:HH11	1.75	0.51
1:C:727:LEU:HD21	1:C:763:GLU:HG3	1.91	0.51
1:B:492:TYR:CZ	1:B:496:ARG:HD2	2.45	0.51
1:A:497:SER:O	1:A:553:ARG:HD2	2.10	0.51

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:458:GLY:N	5:C:907:HOH:O	2.44	0.51
1:C:720:ASP:N	1:C:720:ASP:OD1	2.45	0.50
1:B:767:ARG:NH2	1:B:769:GLU:OE1	2.45	0.49
1:C:719:LYS:O	1:C:775:ILE:HG23	2.12	0.49
1:C:770:GLU:HG3	1:C:771:THR:N	2.27	0.49
1:D:497:SER:HA	1:D:542:GLN:NE2	2.27	0.49
1:C:749:GLU:HG3	5:C:1043:HOH:O	2.12	0.49
1:B:719:LYS:O	1:B:775:ILE:HG23	2.14	0.48
1:C:700:GLY:HA3	1:C:714:MET:O	2.15	0.47
1:B:643:LYS:HE2	1:B:647:GLU:OE2	2.14	0.47
1:D:492:TYR:CZ	1:D:496:ARG:HD2	2.50	0.47
1:A:461:GLN:NE2	1:A:500:THR:HG21	2.30	0.47
1:A:772:ALA:O	1:A:773:THR:CG2	2.60	0.46
1:A:772:ALA:CB	1:A:775:ILE:HD12	2.45	0.46
1:A:542:GLN:NE2	1:A:542:GLN:HA	2.31	0.46
1:A:722:VAL:HB	1:A:723:PRO:HD3	1.99	0.45
4:D:803:K4U:N14	4:D:803:K4U:C26	2.79	0.45
1:A:735:ILE:N	1:A:736:PRO:HD2	2.31	0.45
1:B:727:LEU:HD23	1:B:759:LEU:HD11	1.99	0.45
1:B:493:MET:O	1:B:497:SER:HB2	2.17	0.45
1:A:749:GLU:N	1:A:750:PRO:CD	2.80	0.45
1:A:713:MET:HE2	4:A:803:K4U:C6	2.47	0.44
1:B:611:PHE:HB3	1:B:614:LEU:HD22	2.00	0.44
1:B:770:GLU:HG3	1:B:771:THR:N	2.33	0.44
1:C:509:CME:HB2	5:C:1048:HOH:O	2.17	0.44
1:A:461:GLN:HE22	1:A:500:THR:HG22	1.82	0.44
1:B:727:LEU:HD12	1:B:766:ILE:HD13	2.00	0.43
1:A:461:GLN:HE22	1:A:500:THR:HG21	1.82	0.43
1:A:711:ILE:CD1	1:A:713:MET:HE3	2.48	0.42
1:C:524:TYR:CG	1:C:692:ILE:HD13	2.54	0.42
1:C:467:ARG:O	1:C:471:GLU:HB2	2.19	0.42
4:A:803:K4U:C26	4:A:803:K4U:N14	2.83	0.42
1:B:544:ASN:HB3	1:B:547:LEU:HD22	2.02	0.42
1:C:602:ILE:HA	1:C:605:LEU:HD22	2.02	0.42
1:A:638:TYR:OH	1:A:642:ARG:HD3	2.20	0.41
1:B:638:TYR:OH	1:B:642:ARG:HD3	2.20	0.41
1:C:638:TYR:CD1	1:C:671:THR:HG21	2.55	0.41
4:C:803:K4U:O31	4:C:803:K4U:S2	2.79	0.41
1:D:766:ILE:HG22	1:D:766:ILE:O	2.19	0.41
1:C:542:GLN:NE2	1:C:542:GLN:HA	2.35	0.41
1:C:568[A]:ARG:NH1	1:C:568[A]:ARG:HG2	2.35	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:497:SER:HA	1:D:542:GLN:HE22	1.86	0.41
1:B:473:GLU:OE2	1:B:496:ARG:NH1	2.54	0.41
1:A:477:PHE:HB3	1:A:535:HIS:CE1	2.55	0.41
1:B:496:ARG:NH1	5:B:901:HOH:O	2.42	0.41
4:B:803:K4U:C26	4:B:803:K4U:N14	2.83	0.41
1:C:743:GLN:HE21	1:C:743:GLN:HB2	1.66	0.41
1:D:766:ILE:O	1:D:766:ILE:CG2	2.69	0.40
1:B:497:SER:O	1:B:553:ARG:HD2	2.21	0.40
1:C:493:MET:SD	1:C:535:HIS:HA	2.61	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	316/343 (92%)	312 (99%)	4 (1%)	0	100	100
1	B	317/343 (92%)	311 (98%)	6 (2%)	0	100	100
1	C	317/343 (92%)	312 (98%)	5 (2%)	0	100	100
1	D	307/343 (90%)	299 (97%)	7 (2%)	1 (0%)	41	29
All	All	1257/1372 (92%)	1234 (98%)	22 (2%)	1 (0%)	51	42

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	615	SER

### 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	286/305 (94%)	277 (97%)	9 (3%)	40	28
1	B	286/305 (94%)	274 (96%)	12 (4%)	30	17
1	C	287/305 (94%)	271 (94%)	16 (6%)	21	9
1	D	279/305 (92%)	271 (97%)	8 (3%)	42	31
All	All	1138/1220 (93%)	1093 (96%)	45 (4%)	31	19

All (45) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	471	GLU
1	A	497	SER
1	A	504	GLU
1	A	553	ARG
1	A	573	SER
1	A	605	LEU
1	A	644	GLN
1	A	720	ASP
1	A	773	THR
1	B	459	LEU
1	B	471	GLU
1	B	497	SER
1	B	575	LEU
1	B	576	GLN
1	B	605	LEU
1	B	638	TYR
1	B	650	GLN
1	B	713	MET
1	B	727	LEU
1	B	749	GLU
1	B	771	THR
1	C	463	THR
1	C	471	GLU
1	C	504	GLU
1	C	517	LYS
1	C	547	LEU
1	C	575	LEU
1	C	576	GLN

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	C	605	LEU
1	C	616	SER
1	C	692	ILE
1	C	711	ILE
1	C	713	MET
1	C	749	GLU
1	C	760	SER
1	C	771	THR
1	C	773	THR
1	D	459	LEU
1	D	460	MET
1	D	500	THR
1	D	575	LEU
1	D	605	LEU
1	D	613	THR
1	D	617	SER
1	D	709	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (18) such sidechains are listed below:

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	A	461	GLN
1	A	542	GLN
1	A	604	GLN
1	A	644	GLN
1	A	743	GLN
1	B	542	GLN
1	B	576	GLN
1	B	604	GLN
1	C	542	GLN
1	C	576	GLN
1	C	604	GLN
1	C	743	GLN
1	D	484	ASN
1	D	542	GLN
1	D	604	GLN
1	D	709	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	B	509	1	8,9,10	0.52	0	5,9,11	0.60	0
1	CME	A	509	1	8,9,10	0.76	0	5,9,11	1.00	0
1	CME	D	509	1	8,9,10	0.56	0	5,9,11	0.67	0
1	CME	C	509	1	8,9,10	0.82	1 (12%)	5,9,11	1.09	1 (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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	CME	B	509	1	-	1/5/8/10	-
1	CME	A	509	1	-	2/5/8/10	-
1	CME	D	509	1	-	2/5/8/10	-
1	CME	C	509	1	-	2/5/8/10	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	509	CME	O-C	2.13	1.28	1.19

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	509	CME	CB-SG-SD	-2.07	98.46	103.82

There are no chirality outliers.

All (7) torsion outliers are listed below:

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

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	C	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	K4U	B	803	-	34,35,35	2.58	11 (32%)	40,48,48	2.77	17 (42%)
4	K4U	C	803	-	34,35,35	2.21	4 (11%)	40,48,48	3.49	14 (35%)
4	K4U	D	803	-	34,35,35	2.57	8 (23%)	40,48,48	2.74	18 (45%)
4	K4U	A	803	-	34,35,35	2.30	9 (26%)	40,48,48	3.00	12 (30%)

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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	K4U	B	803	-	1/1/5/7	0/20/39/39	0/5/5/5
4	K4U	C	803	-	1/1/5/7	5/20/39/39	0/5/5/5
4	K4U	D	803	-	1/1/5/7	3/20/39/39	0/5/5/5
4	K4U	A	803	-	1/1/5/7	2/20/39/39	0/5/5/5

All (32) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	803	K4U	C1-C5	-8.83	1.35	1.53
4	D	803	K4U	C1-S2	-7.96	1.64	1.81
4	B	803	K4U	C1-C5	-7.40	1.38	1.53
4	C	803	K4U	C5-N4	-7.37	1.37	1.48
4	A	803	K4U	C1-C5	-7.25	1.39	1.53
4	C	803	K4U	C1-C5	-7.07	1.39	1.53
4	B	803	K4U	C5-N4	-6.39	1.39	1.48
4	B	803	K4U	C1-S2	-6.14	1.68	1.81
4	A	803	K4U	C5-N4	-5.83	1.40	1.48
4	C	803	K4U	C1-S2	-5.23	1.70	1.81
4	A	803	K4U	C1-S2	-5.04	1.70	1.81
4	D	803	K4U	C5-N4	-4.97	1.41	1.48
4	B	803	K4U	C3-S2	-4.12	1.70	1.76
4	B	803	K4U	C22-C21	4.08	1.46	1.39
4	A	803	K4U	C3-N4	3.82	1.38	1.29
4	D	803	K4U	C3-S2	-3.76	1.71	1.76
4	B	803	K4U	C3-N4	3.22	1.37	1.29
4	C	803	K4U	C3-N4	3.07	1.36	1.29
4	D	803	K4U	C22-C21	3.02	1.44	1.39
4	B	803	K4U	C24-N25	2.85	1.39	1.33
4	B	803	K4U	C15-C16	2.76	1.46	1.40
4	B	803	K4U	C19-N14	2.66	1.39	1.35
4	A	803	K4U	C3-S2	-2.65	1.72	1.76
4	A	803	K4U	C26-C21	2.61	1.44	1.39
4	B	803	K4U	C26-C21	2.59	1.43	1.39
4	B	803	K4U	C24-N23	2.43	1.38	1.33
4	A	803	K4U	C16-C27	-2.24	1.45	1.50
4	D	803	K4U	C15-C16	2.17	1.45	1.40
4	D	803	K4U	C3-N4	2.16	1.34	1.29
4	A	803	K4U	C24-N25	2.16	1.37	1.33
4	A	803	K4U	O31-C30	2.15	1.27	1.23

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	803	K4U	C18-C19	2.15	1.47	1.41

All (61) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	803	K4U	C6-C5-C1	17.36	131.01	113.86
4	A	803	K4U	C6-C5-C1	14.99	128.67	113.86
4	B	803	K4U	C6-C5-C1	10.84	124.57	113.86
4	D	803	K4U	C6-C5-C1	9.32	123.08	113.86
4	D	803	K4U	C28-C27-C16	5.29	126.55	120.33
4	D	803	K4U	C11-C12-N7	5.06	119.45	111.33
4	C	803	K4U	N20-C19-N14	4.94	124.25	118.40
4	A	803	K4U	C5-C1-S2	4.81	113.14	105.55
4	B	803	K4U	C29-C27-C16	-4.79	114.70	120.33
4	C	803	K4U	C18-C19-N14	-4.67	117.40	121.97
4	D	803	K4U	C5-C1-S2	4.56	112.75	105.55
4	C	803	K4U	C5-C1-S2	4.46	112.59	105.55
4	D	803	K4U	C6-C5-N4	4.45	123.66	110.55
4	B	803	K4U	C28-C27-C16	4.27	125.34	120.33
4	D	803	K4U	C30-C18-N17	-4.20	112.07	115.99
4	C	803	K4U	C29-C27-C16	-4.18	115.42	120.33
4	A	803	K4U	O31-C30-C18	4.10	126.71	120.59
4	C	803	K4U	C11-C12-N7	4.08	117.88	111.33
4	B	803	K4U	C1-C5-N4	4.07	115.69	104.60
4	B	803	K4U	O31-C30-C18	4.04	126.62	120.59
4	B	803	K4U	C10-C9-C8	3.99	118.91	111.19
4	C	803	K4U	C1-C5-N4	3.91	115.25	104.60
4	B	803	K4U	C11-C12-N7	3.83	117.47	111.33
4	D	803	K4U	C12-N7-C6	3.73	122.03	111.20
4	B	803	K4U	C26-N25-C24	3.35	120.08	115.80
4	A	803	K4U	C5-C6-N7	3.27	117.55	113.36
4	D	803	K4U	C1-C5-N4	3.13	113.14	104.60
4	B	803	K4U	C6-C5-N4	3.10	119.69	110.55
4	A	803	K4U	C18-C19-N14	-3.08	118.96	121.97
4	B	803	K4U	C5-C1-S2	3.02	110.32	105.55
4	A	803	K4U	C1-C5-N4	2.98	112.72	104.60
4	D	803	K4U	C29-C27-C16	-2.95	116.87	120.33
4	B	803	K4U	N20-C19-N14	2.92	121.85	118.40
4	A	803	K4U	C12-N7-C6	2.90	119.61	111.20
4	A	803	K4U	C28-C27-C16	-2.75	117.10	120.33
4	A	803	K4U	C6-C5-N4	2.72	118.57	110.55
4	D	803	K4U	C26-N25-C24	2.66	119.20	115.80

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	803	K4U	O31-C30-C18	2.65	124.55	120.59
4	D	803	K4U	O31-C30-C18	2.60	124.48	120.59
4	C	803	K4U	C10-C11-C12	2.55	116.11	111.19
4	A	803	K4U	N20-C19-N14	2.54	121.40	118.40
4	C	803	K4U	C9-C8-N7	-2.51	107.30	111.33
4	C	803	K4U	C18-C30-N13	-2.51	111.10	113.80
4	D	803	K4U	C9-C8-N7	-2.47	107.35	111.33
4	A	803	K4U	C30-C18-N17	2.46	118.29	115.99
4	D	803	K4U	N20-C19-N14	2.44	121.29	118.40
4	B	803	K4U	C18-C30-N13	-2.40	111.22	113.80
4	A	803	K4U	C28-C29-C27	-2.39	58.57	60.56
4	D	803	K4U	N13-C3-N4	2.36	125.14	121.18
4	D	803	K4U	C10-C9-C8	2.35	115.73	111.19
4	D	803	K4U	C12-N7-C8	2.34	114.10	108.83
4	C	803	K4U	C12-N7-C8	2.31	114.03	108.83
4	B	803	K4U	C12-N7-C8	2.28	113.97	108.83
4	B	803	K4U	C21-C26-N25	-2.24	118.70	123.13
4	C	803	K4U	C11-C10-C9	-2.22	104.33	111.18
4	B	803	K4U	C11-C10-C9	2.22	118.01	111.18
4	B	803	K4U	C22-N23-C24	-2.13	113.07	115.80
4	C	803	K4U	C10-C9-C8	2.09	115.23	111.19
4	D	803	K4U	C18-C19-N14	-2.05	119.96	121.97
4	B	803	K4U	C18-C19-N14	-2.03	119.98	121.97
4	D	803	K4U	C21-C26-N25	-2.01	119.15	123.13

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	A	803	K4U	C5
4	B	803	K4U	C5
4	C	803	K4U	C5
4	D	803	K4U	C5

All (10) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	C	803	K4U	C15-C16-C27-C28
4	D	803	K4U	N17-C16-C27-C29
4	C	803	K4U	C15-C16-C27-C29
4	C	803	K4U	N17-C16-C27-C28
4	C	803	K4U	N17-C16-C27-C29
4	D	803	K4U	C15-C16-C27-C28

*Continued on next page...*

*Continued from previous page...*

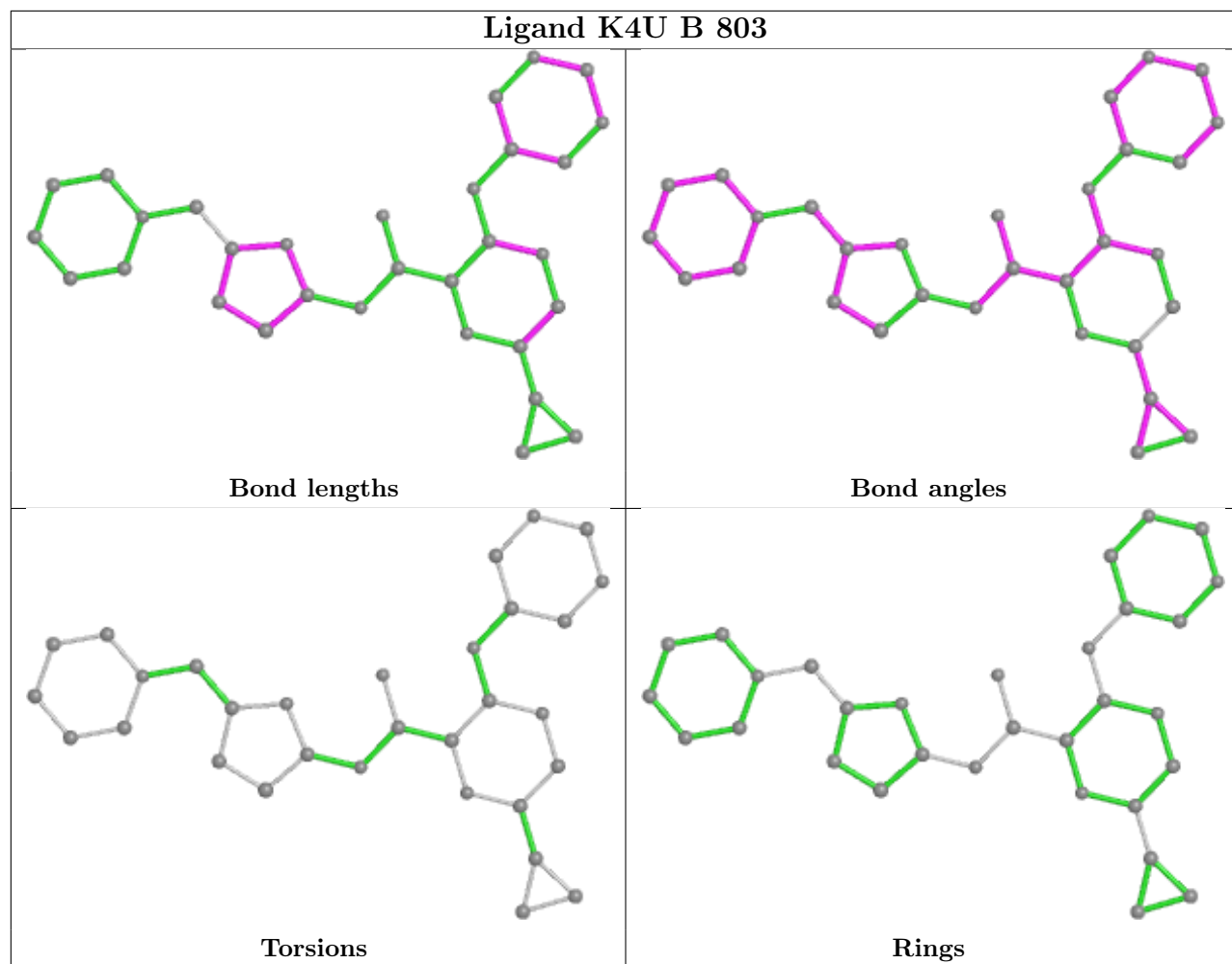
Mol	Chain	Res	Type	Atoms
4	A	803	K4U	C1-C5-C6-N7
4	D	803	K4U	C1-C5-C6-N7
4	C	803	K4U	N4-C5-C6-N7
4	A	803	K4U	C15-C16-C27-C28

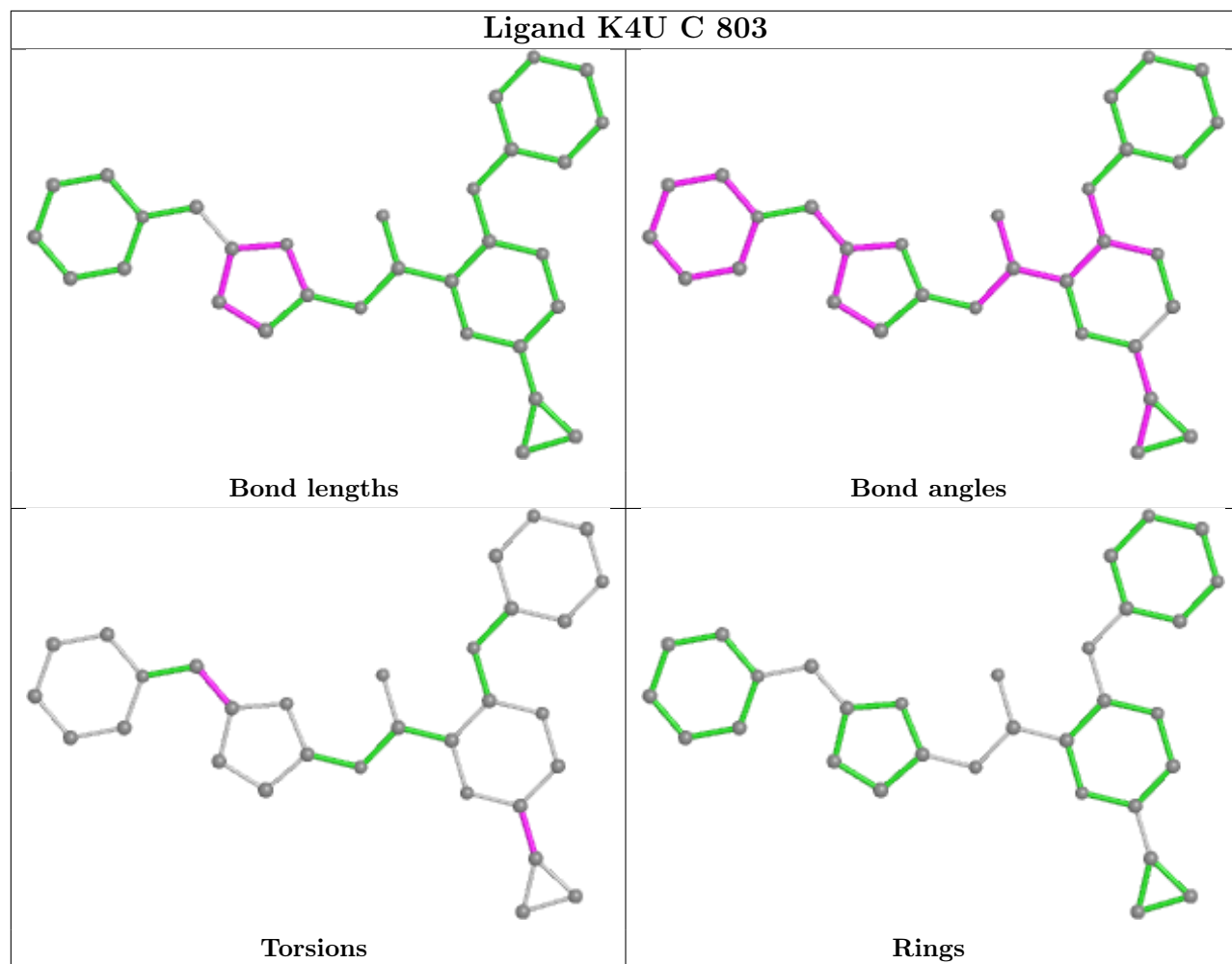
There are no ring outliers.

4 monomers are involved in 7 short contacts:

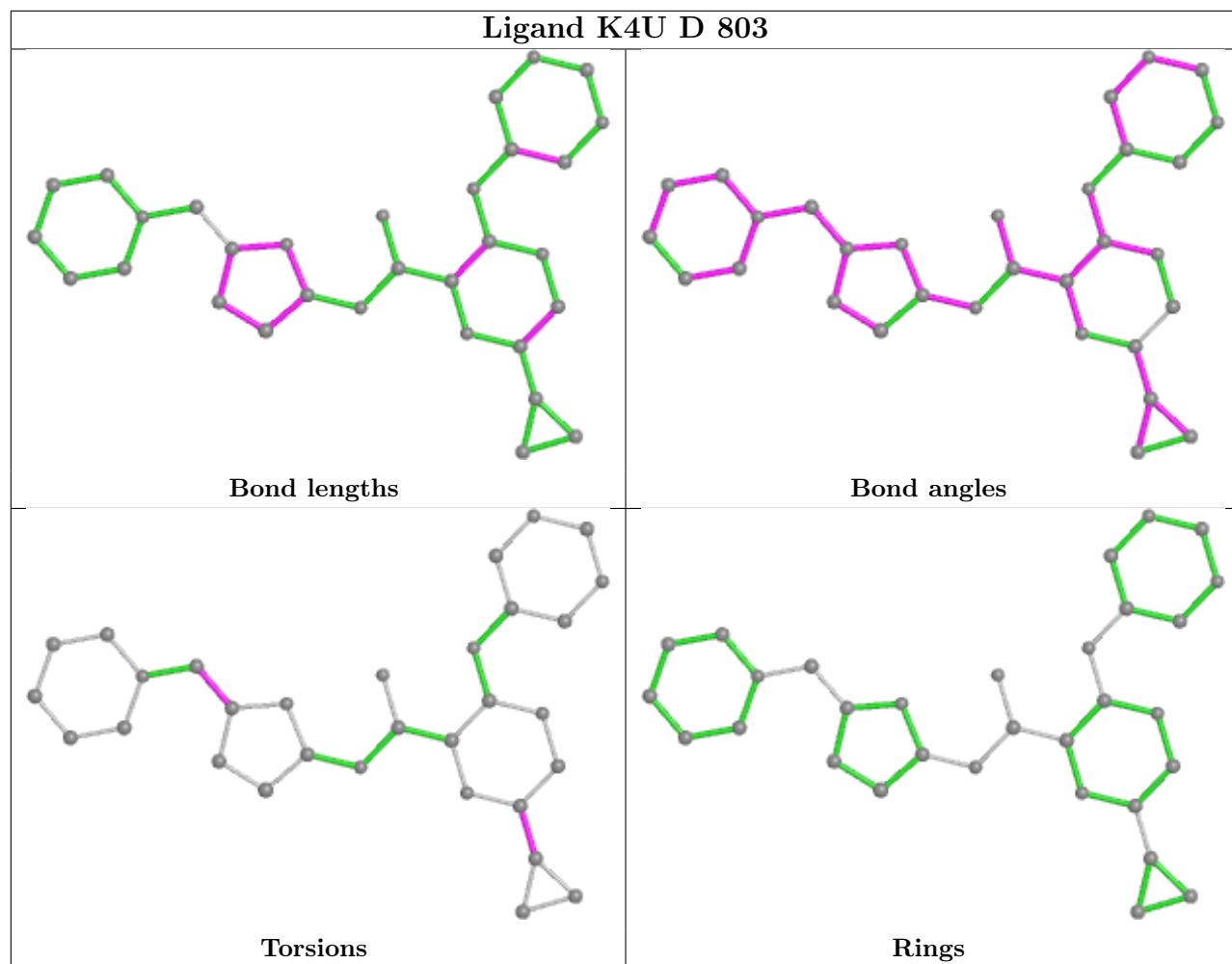
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	803	K4U	1	0
4	C	803	K4U	1	0
4	D	803	K4U	2	0
4	A	803	K4U	3	0

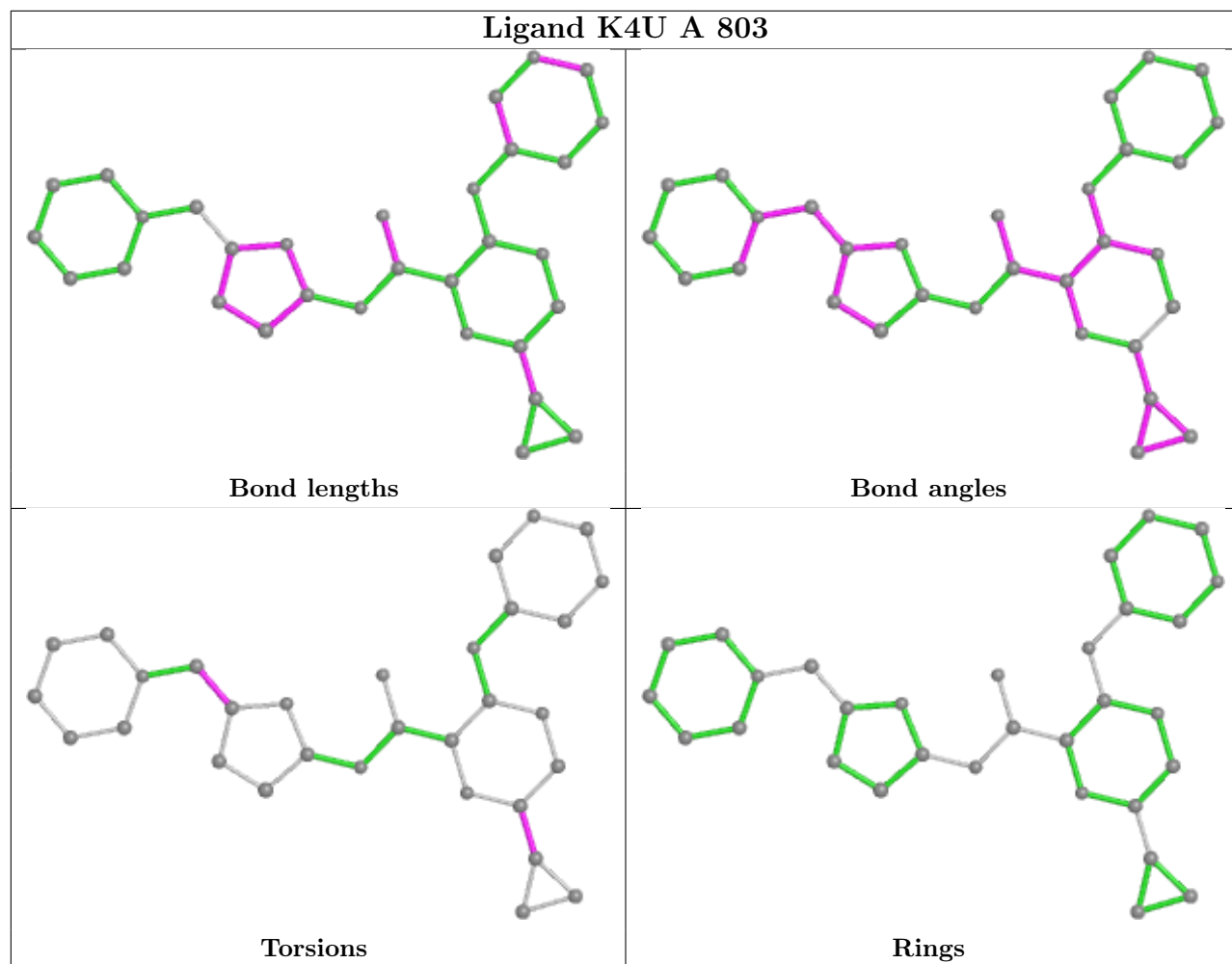
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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	317/343 (92%)	-0.16	7 (2%) 62 63	25, 37, 64, 90	0
1	B	318/343 (92%)	-0.07	5 (1%) 72 73	24, 35, 65, 89	0
1	C	317/343 (92%)	-0.30	7 (2%) 62 63	25, 37, 62, 91	0
1	D	309/343 (90%)	-0.31	2 (0%) 89 90	28, 40, 61, 94	0
All	All	1261/1372 (91%)	-0.21	21 (1%) 70 71	24, 38, 63, 94	0

All (21) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	774	TRP	6.4
1	C	772	ALA	5.8
1	B	773	THR	5.7
1	C	771	THR	5.7
1	A	773	THR	5.2
1	B	771	THR	4.5
1	C	773	THR	4.2
1	B	775	ILE	4.0
1	A	771	THR	4.0
1	A	774	TRP	3.9
1	B	774	TRP	3.9
1	C	775	ILE	3.7
1	B	772	ALA	3.4
1	A	775	ILE	3.0
1	C	770	GLU	2.5
1	A	458	GLY	2.4
1	C	768	GLY	2.3
1	A	772	ALA	2.3
1	D	720	ASP	2.3
1	A	650	GLN	2.2
1	D	765	VAL	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	A	509	10/11	0.92	0.10	36,44,81,85	0
1	CME	D	509	10/11	0.93	0.11	34,52,97,100	0
1	CME	C	509	10/11	0.94	0.10	37,45,84,91	0
1	CME	B	509	10/11	0.94	0.10	32,43,83,85	0

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

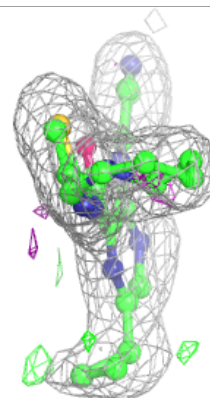
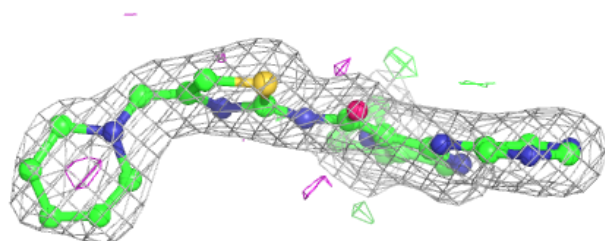
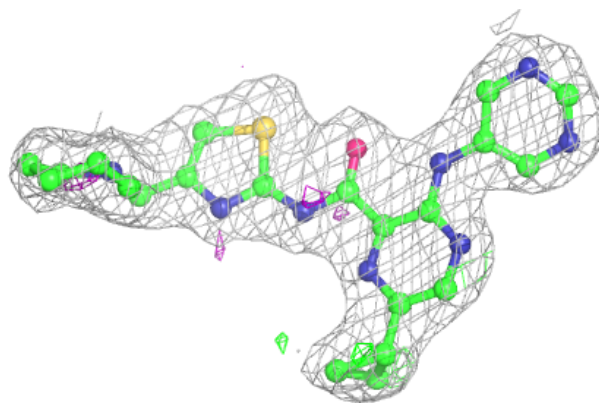
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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	K4U	C	803	31/31	0.95	0.11	33,39,53,55	0
4	K4U	B	803	31/31	0.96	0.10	30,37,54,62	0
4	K4U	D	803	31/31	0.96	0.10	36,41,65,69	0
4	K4U	A	803	31/31	0.97	0.09	30,39,56,59	0
3	MG	A	802	1/1	0.99	0.06	25,25,25,25	0
3	MG	B	802	1/1	0.99	0.10	24,24,24,24	0
3	MG	C	802	1/1	0.99	0.04	25,25,25,25	0
3	MG	D	802	1/1	0.99	0.05	34,34,34,34	0
2	ZN	A	801	1/1	1.00	0.10	31,31,31,31	0
2	ZN	B	801	1/1	1.00	0.09	30,30,30,30	0
2	ZN	C	801	1/1	1.00	0.08	31,31,31,31	0
2	ZN	D	801	1/1	1.00	0.06	35,35,35,35	0

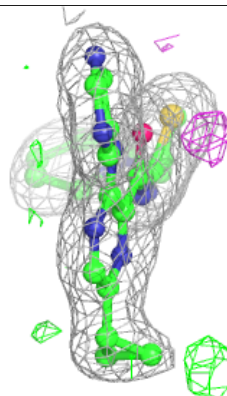
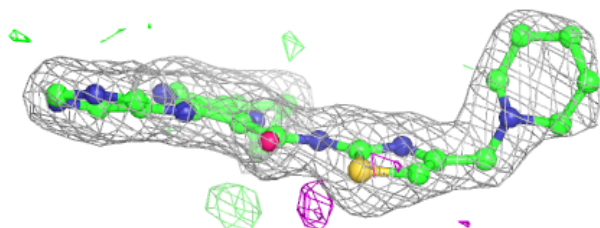
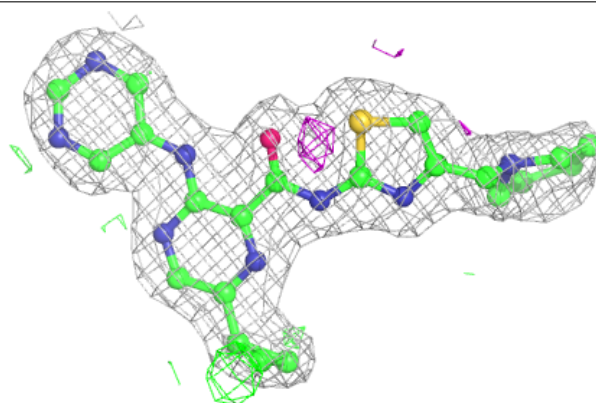
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around K4U 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)

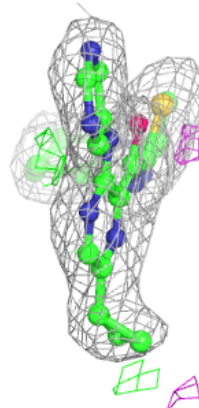
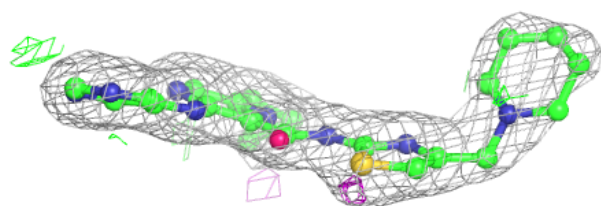
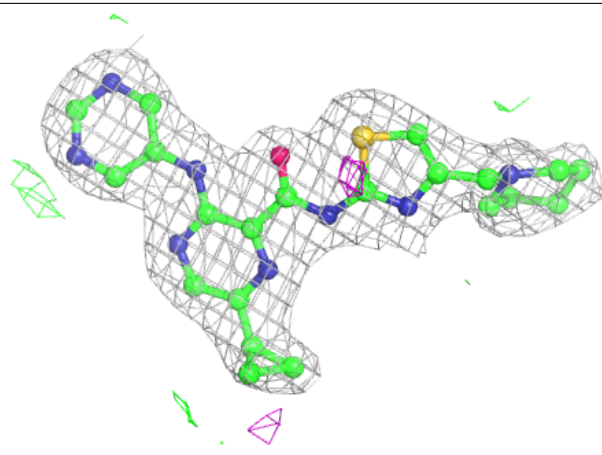
**Electron density around K4U B 803:**

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

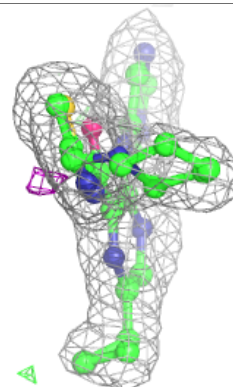
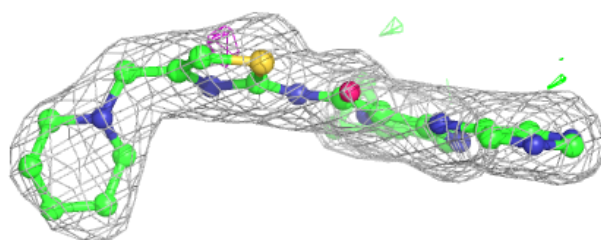
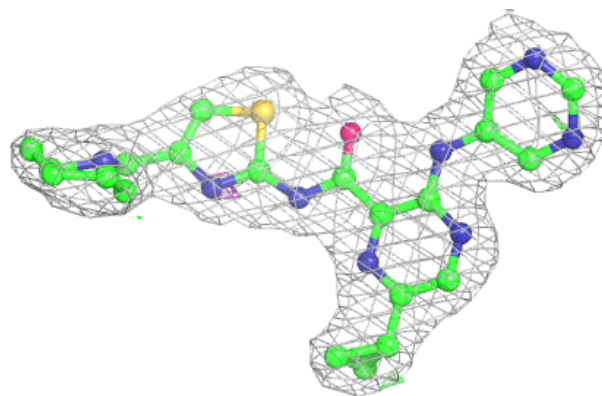


**Electron density around K4U D 803:**

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

**Electron density around K4U A 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.