Full wwPDB X-ray Structure Validation Report (i)

Dec 13, 2023 – 04:10 am GMT

PDB ID	:	2Y8C
Title	:	Plasmodium falciparum dUTPase in complex with a trityl ligand
Authors	:	Baragana, B.; McCarthy, O.; Sanchez, P.; Bosch, C.; Kaiser, M.; Brun, R.; Whittingham, J.L.; Roberts, S.; Zhou, X.-X.; Wilson, K.S.; Johansson, N.G.; Gonzalez-Pacanowska, D.; Gilbert, I.H.
Deposited on	:	2011-02-04
Resolution	:	2.10 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 (i) 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 (i)) were used in the production of this report:

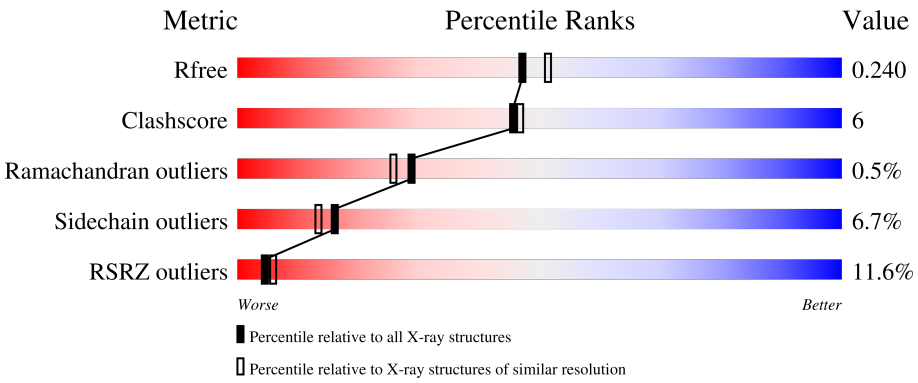
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.10 Å.

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



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	173	3% 66% 14% • • 16%
1	B	173	19% 56% 12% • • 29%
1	C	173	4% 62% 13% • • 21%

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
2	DUQ	A	400	X	-	-	-
2	DUQ	B	400	X	-	-	-
2	DUQ	C	400	X	-	-	-

2 Entry composition (i)

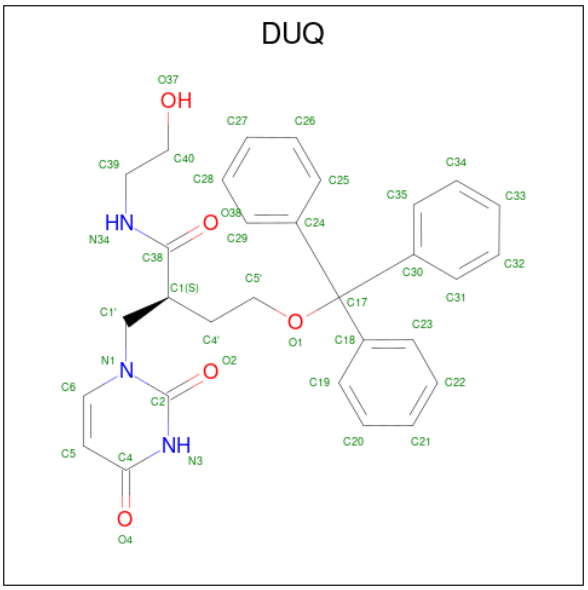
There are 4 unique types of molecules in this entry. The entry contains 3399 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 DEOXYURIDINE 5'-TRIPHOSPHATE NUCLEOTIDOHYDROLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	145	1132	719	191	219	3	67	0	0
1	B	123	961	616	159	183	3	69	0	0
1	C	136	1061	677	176	205	3	25	0	0

- Molecule 2 is (2S)-2-[(2,4-dioxopyrimidin-1-yl)methyl]-N-(2-hydroxyethyl)-4-trityloxy-butanamide (three-letter code: DUQ) (formula: $C_{30}H_{31}N_3O_5$).



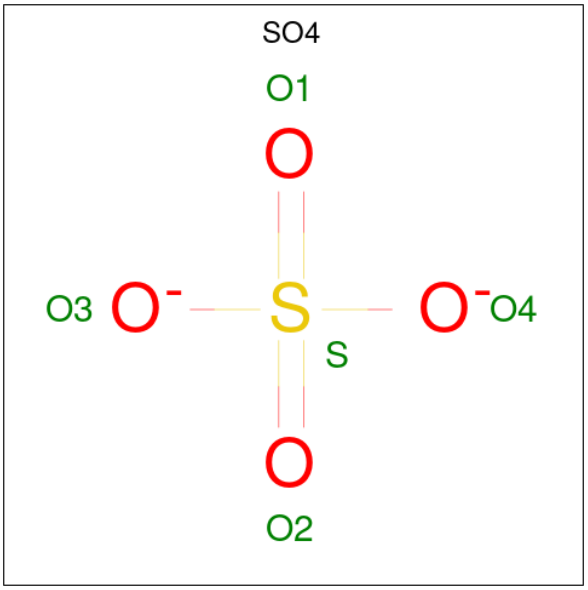
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	38	30	3	5	0	0
2	B	1	38	30	3	5	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	C	1	38	30	3	5	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	O	S			
3	A	1	5	4	1		0	0
3	C	1	Total	O	S		0	0
			5	4	1			

- Molecule 4 is water.

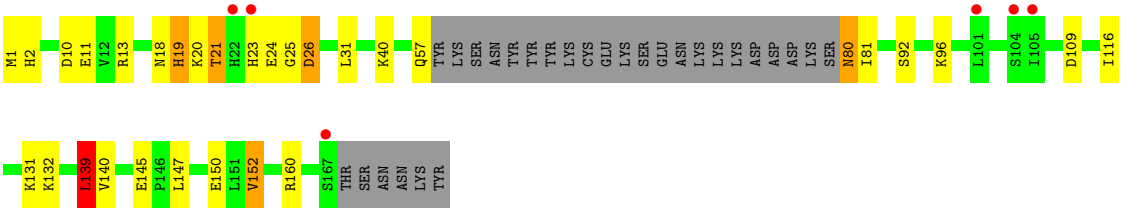
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	62	62	62	0	0
4	B	19	Total	O	0	0
			19	19		
4	C	40	Total	O	0	0
			40	40		

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.

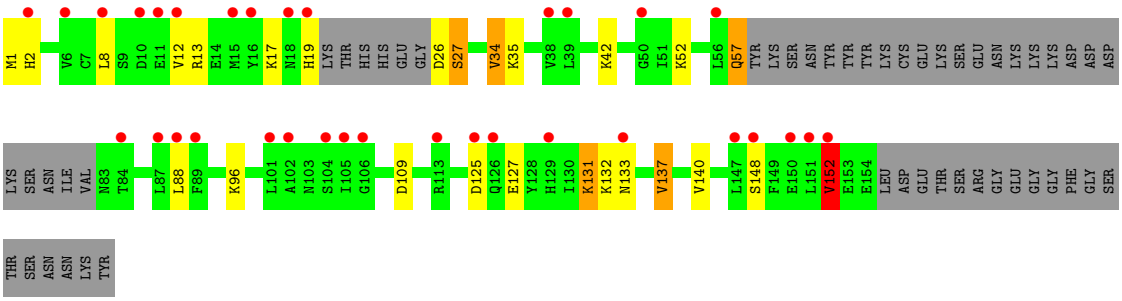
• Molecule 1: DEOXYURIDINE 5'-TRIPHOSPHATE NUCLEOTIDOHYDROLASE

Chain A:



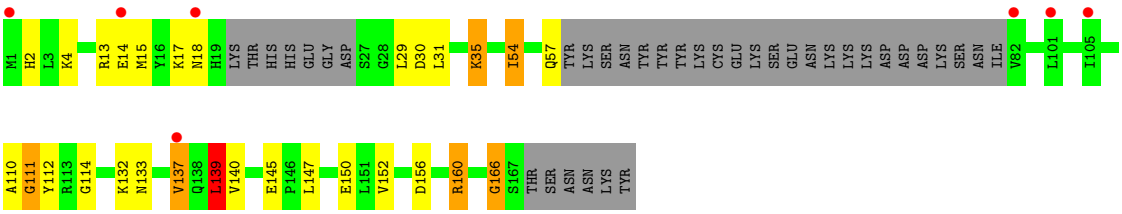
• Molecule 1: DEOXYURIDINE 5'-TRIPHOSPHATE NUCLEOTIDOHYDROLASE

Chain B:



• Molecule 1: DEOXYURIDINE 5'-TRIPHOSPHATE NUCLEOTIDOHYDROLASE

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, α, β, γ	77.16Å 77.16Å 110.98Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.90 – 2.10 38.90 – 2.10	Depositor EDS
% Data completeness (in resolution range)	100.0 (38.90-2.10) 100.0 (38.90-2.10)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.57 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.6.0119	Depositor
R, R_{free}	0.188 , 0.240 0.187 , 0.240	Depositor DCC
R_{free} test set	1902 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	35.7	Xtriage
Anisotropy	0.001	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 59.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.043 for h,-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	3399	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.41% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, DUQ

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	2.23	11/1149 (1.0%)	2.37	20/1545 (1.3%)
1	B	1.84	10/974 (1.0%)	1.66	12/1310 (0.9%)
1	C	1.41	3/1075 (0.3%)	1.20	8/1444 (0.6%)
All	All	1.87	24/3198 (0.8%)	1.83	40/4299 (0.9%)

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	A	1	3
1	B	1	1
1	C	0	2
All	All	2	6

All (24) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	21	THR	N-CA	-43.81	0.58	1.46
1	A	80	ASN	CB-CG	-27.04	0.88	1.51
1	C	57	GLN	C-O	-25.17	0.75	1.23
1	B	19	HIS	CA-C	-25.10	0.87	1.52
1	A	145	GLU	CB-CG	24.50	1.98	1.52
1	B	127	GLU	CG-CD	-24.09	1.15	1.51
1	A	25	GLY	C-N	-21.10	0.85	1.34
1	B	152	VAL	C-N	-18.50	0.91	1.34
1	B	57	GLN	CA-CB	-18.07	1.14	1.53
1	C	17	LYS	CB-CG	-16.25	1.08	1.52
1	A	21	THR	CA-CB	15.05	1.92	1.53
1	A	19	HIS	C-N	-14.72	1.00	1.34

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1	MET	CA-CB	-13.51	1.24	1.53
1	B	96	LYS	CG-CD	13.46	1.98	1.52
1	B	13	ARG	CA-CB	-12.92	1.25	1.53
1	A	26	ASP	CA-CB	-9.88	1.32	1.53
1	A	57	GLN	CB-CG	-9.79	1.26	1.52
1	A	131	LYS	CB-CG	-8.32	1.30	1.52
1	B	131	LYS	CG-CD	-8.26	1.24	1.52
1	A	40	LYS	CG-CD	-8.13	1.24	1.52
1	B	26	ASP	C-N	-7.61	1.16	1.34
1	C	145	GLU	CA-CB	-7.56	1.37	1.53
1	A	18	ASN	CA-CB	-5.54	1.38	1.53
1	B	17	LYS	CB-CG	-5.20	1.38	1.52

All (40) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	19	HIS	O-C-N	-71.92	7.62	122.70
1	B	152	VAL	O-C-N	-30.66	73.65	122.70
1	A	21	THR	N-CA-CB	26.51	160.67	110.30
1	B	152	VAL	CA-C-N	21.53	164.56	117.20
1	B	152	VAL	C-N-CA	14.92	158.99	121.70
1	A	21	THR	CA-CB-CG2	-14.27	92.42	112.40
1	B	96	LYS	CB-CG-CD	-12.55	78.96	111.60
1	A	21	THR	CB-CA-C	-12.19	78.70	111.60
1	B	1	MET	N-CA-CB	11.54	131.38	110.60
1	A	19	HIS	C-N-CA	10.44	147.80	121.70
1	B	13	ARG	N-CA-CB	9.67	128.00	110.60
1	B	131	LYS	CB-CG-CD	9.50	136.31	111.60
1	C	137	VAL	CG1-CB-CG2	9.30	125.78	110.90
1	A	80	ASN	CB-CG-OD1	-9.08	103.44	121.60
1	A	116	ILE	CG1-CB-CG2	-8.65	92.37	111.40
1	A	25	GLY	O-C-N	-8.50	109.10	122.70
1	C	18	ASN	CB-CA-C	-8.48	93.43	110.40
1	B	57	GLN	N-CA-CB	7.66	124.38	110.60
1	B	57	GLN	CB-CA-C	-7.32	95.76	110.40
1	A	145	GLU	CA-CB-CG	-7.30	97.34	113.40
1	A	80	ASN	CB-CG-ND2	7.08	133.70	116.70
1	A	21	THR	CA-CB-OG1	-6.90	94.51	109.00
1	C	17	LYS	CA-CB-CG	6.77	128.29	113.40
1	A	131	LYS	CA-CB-CG	6.71	128.17	113.40
1	A	25	GLY	C-N-CA	6.56	138.09	121.70
1	C	30	ASP	CB-CG-OD1	6.30	123.97	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	145	GLU	N-CA-CB	6.11	121.59	110.60
1	C	156	ASP	CB-CG-OD1	5.79	123.51	118.30
1	A	25	GLY	CA-C-N	5.75	129.85	117.20
1	A	140	VAL	CG1-CB-CG2	5.63	119.90	110.90
1	C	139	LEU	CB-CG-CD1	5.57	120.47	111.00
1	B	127	GLU	CG-CD-OE2	-5.47	107.35	118.30
1	C	166	GLY	O-C-N	-5.43	114.01	122.70
1	B	127	GLU	CG-CD-OE1	5.36	129.01	118.30
1	A	140	VAL	CA-CB-CG2	5.33	118.90	110.90
1	B	35	LYS	CA-CB-CG	5.31	125.08	113.40
1	A	21	THR	N-CA-C	-5.23	96.87	111.00
1	A	152	VAL	CA-CB-CG1	5.19	118.69	110.90
1	A	23	HIS	O-C-N	5.15	130.94	122.70
1	A	139	LEU	CB-CG-CD1	5.04	119.57	111.00

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	21	THR	CA
1	B	1	MET	CA

All (6) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	19	HIS	Mainchain,Peptide
1	A	20	LYS	Peptide
1	B	152	VAL	Mainchain
1	C	111	GLY	Peptide
1	C	166	GLY	Mainchain

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	1132	0	1144	10	0
1	B	961	0	981	8	0
1	C	1061	0	1085	16	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	38	0	31	0	0
2	B	38	0	31	3	0
2	C	38	0	30	6	0
3	A	5	0	0	0	0
3	C	5	0	0	1	0
4	A	62	0	0	2	0
4	B	19	0	0	0	0
4	C	40	0	0	0	0
All	All	3399	0	3302	40	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 6.

All (40) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:2:HIS:HD2	1:B:148:SER:OG	1.40	1.02
1:A:31:LEU:HD21	1:A:139:LEU:HD22	1.57	0.86
1:A:1:MET:HE2	1:A:147:LEU:HD21	1.58	0.84
1:A:1:MET:CE	1:A:147:LEU:HD21	2.08	0.84
1:C:29:LEU:HB2	1:C:147:LEU:HD11	1.61	0.81
1:B:2:HIS:CD2	1:B:148:SER:OG	2.31	0.78
1:A:11:GLU:OE2	1:A:132:LYS:HE3	1.90	0.72
1:B:8:LEU:HD21	1:B:52:LYS:HE3	1.71	0.71
1:C:15:MET:CE	1:C:132:LYS:HE3	2.23	0.68
1:C:2:HIS:NE2	1:C:150:GLU:HG2	2.09	0.68
2:C:400:DUQ:C35	2:C:400:DUQ:H4'1	2.26	0.65
1:B:88:LEU:HD11	1:B:137:VAL:HG13	1.82	0.61
1:C:35:LYS:NZ	1:C:35:LYS:HB2	2.16	0.60
2:B:400:DUQ:C25	2:B:400:DUQ:H4'1	2.32	0.60
1:B:12:VAL:HG22	1:B:34:VAL:HG22	1.84	0.59
1:C:15:MET:HE2	1:C:132:LYS:HE3	1.83	0.59
1:C:4:LYS:HB2	1:C:54:ILE:HG22	1.87	0.57
1:C:31:LEU:HD21	1:C:139:LEU:HD22	1.88	0.56
1:C:147:LEU:HD12	1:C:147:LEU:O	2.07	0.54
1:A:1:MET:HE3	1:A:147:LEU:HD21	1.89	0.54
2:C:400:DUQ:H4'1	2:C:400:DUQ:C30	2.40	0.51
1:C:147:LEU:HD12	1:C:147:LEU:C	2.33	0.49
1:B:42:LYS:HE2	1:B:125:ASP:OD1	2.14	0.48
1:A:13:ARG:NH1	4:A:2005:HOH:O	2.46	0.48
1:C:35:LYS:HB2	1:C:35:LYS:HZ2	1.78	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:400:DUQ:O38	2:B:400:DUQ:H401	2.14	0.46
1:C:110:ALA:O	1:C:111:GLY:C	2.53	0.45
1:A:109:ASP:OD1	1:B:27:SER:HB2	2.18	0.44
2:C:400:DUQ:H31	2:C:400:DUQ:C19	2.47	0.44
1:C:15:MET:HE1	1:C:132:LYS:HE3	2.01	0.43
1:A:92:SER:HB3	2:C:400:DUQ:H5'1	2.02	0.42
1:C:112:TYR:CZ	1:C:114:GLY:HA3	2.54	0.42
1:A:2:HIS:NE2	1:A:150:GLU:OE2	2.45	0.42
1:B:132:LYS:O	1:B:133:ASN:HB2	2.20	0.42
1:C:132:LYS:O	1:C:133:ASN:HB2	2.19	0.42
1:C:160:ARG:NH1	3:C:1168:SO4:O2	2.53	0.41
2:C:400:DUQ:C19	2:C:400:DUQ:C31	2.98	0.41
1:A:160:ARG:NH1	4:A:2043:HOH:O	2.43	0.40
2:B:400:DUQ:H23	2:B:400:DUQ:H392	2.02	0.40
1:C:112:TYR:HE1	2:C:400:DUQ:H392	1.86	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	141/173 (82%)	135 (96%)	5 (4%)	1 (1%)	22	18
1	B	117/173 (68%)	111 (95%)	5 (4%)	1 (1%)	17	12
1	C	130/173 (75%)	125 (96%)	5 (4%)	0	100	100
All	All	388/519 (75%)	371 (96%)	15 (4%)	2 (0%)	29	26

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	26	ASP
1	B	152	VAL

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	127/156 (81%)	119 (94%)	8 (6%)	18	15
1	B	109/156 (70%)	102 (94%)	7 (6%)	17	14
1	C	120/156 (77%)	111 (92%)	9 (8%)	13	10
All	All	356/468 (76%)	332 (93%)	24 (7%)	16	13

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	10	ASP
1	A	21	THR
1	A	24	GLU
1	A	80	ASN
1	A	81	ILE
1	A	96	LYS
1	A	139	LEU
1	A	152	VAL
1	B	27	SER
1	B	34	VAL
1	B	57	GLN
1	B	109	ASP
1	B	131	LYS
1	B	137	VAL
1	B	140	VAL
1	C	13	ARG
1	C	14	GLU
1	C	35	LYS
1	C	54	ILE
1	C	137	VAL
1	C	139	LEU
1	C	140	VAL
1	C	152	VAL
1	C	160	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such

sidechains are listed below:

Mol	Chain	Res	Type
1	B	2	HIS
1	B	133	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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)

5 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	DUQ	A	400	-	41,41,41	1.41	5 (12%)	49,55,55	2.58	12 (24%)
3	SO4	C	1168	-	4,4,4	0.50	0	6,6,6	0.54	0
2	DUQ	C	400	-	41,41,41	1.63	6 (14%)	49,55,55	3.16	19 (38%)
2	DUQ	B	400	-	41,41,41	1.31	6 (14%)	49,55,55	2.51	13 (26%)
3	SO4	A	1168	-	4,4,4	0.93	0	6,6,6	0.79	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	DUQ	B	400	-	1/1/3/4	12/37/37/37	0/4/4/4
2	DUQ	A	400	-	1/1/3/4	6/37/37/37	0/4/4/4
2	DUQ	C	400	-	1/1/3/4	7/37/37/37	0/4/4/4

All (17) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	400	DUQ	C1'-C1	3.93	1.58	1.53
2	C	400	DUQ	O2-C2	3.28	1.29	1.23
2	C	400	DUQ	C29-C24	3.18	1.44	1.39
2	C	400	DUQ	C6-C5	3.01	1.42	1.35
2	C	400	DUQ	C31-C30	2.95	1.44	1.39
2	A	400	DUQ	C20-C19	2.80	1.44	1.38
2	A	400	DUQ	C5-C4	-2.80	1.37	1.43
2	B	400	DUQ	C5-C4	-2.68	1.37	1.43
2	A	400	DUQ	C6-C5	2.59	1.41	1.35
2	B	400	DUQ	O2-C2	2.44	1.27	1.23
2	B	400	DUQ	C22-C23	2.26	1.43	1.38
2	C	400	DUQ	C33-C32	2.20	1.43	1.38
2	B	400	DUQ	C34-C35	2.17	1.43	1.38
2	B	400	DUQ	C6-C5	2.16	1.40	1.35
2	A	400	DUQ	C25-C24	2.06	1.42	1.39
2	B	400	DUQ	C29-C24	2.03	1.42	1.39
2	A	400	DUQ	C2-N1	2.02	1.40	1.37

All (44) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	400	DUQ	O38-C38-C1	-10.56	108.37	122.12
2	C	400	DUQ	C1-C38-N34	9.66	128.53	115.99
2	A	400	DUQ	O2-C2-N1	-8.47	115.41	122.85
2	B	400	DUQ	O2-C2-N1	-7.82	115.98	122.85
2	B	400	DUQ	C39-N34-C38	6.89	134.88	122.59
2	C	400	DUQ	C1'-N1-C2	-6.51	111.89	118.53
2	A	400	DUQ	C1-C38-N34	6.31	124.17	115.99
2	C	400	DUQ	C5-C4-N3	5.94	123.72	114.84
2	A	400	DUQ	N3-C2-N1	5.89	120.25	114.86
2	B	400	DUQ	N3-C2-N1	5.80	120.17	114.86
2	A	400	DUQ	O38-C38-C1	-5.78	114.60	122.12
2	C	400	DUQ	C4-N3-C2	-5.40	119.46	126.58
2	A	400	DUQ	C4-N3-C2	-5.13	119.81	126.58
2	B	400	DUQ	O38-C38-C1	-4.99	115.63	122.12

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	400	DUQ	C4-N3-C2	-4.93	120.08	126.58
2	A	400	DUQ	O4-C4-C5	-4.87	116.60	125.16
2	C	400	DUQ	N3-C2-N1	4.72	119.18	114.86
2	C	400	DUQ	C5-C6-N1	-4.62	118.11	122.44
2	C	400	DUQ	O2-C2-N1	-4.59	118.82	122.85
2	C	400	DUQ	C40-C39-N34	-4.53	101.45	111.64
2	A	400	DUQ	C5-C4-N3	4.12	121.00	114.84
2	B	400	DUQ	O4-C4-C5	-3.99	118.14	125.16
2	B	400	DUQ	C5-C4-N3	3.90	120.67	114.84
2	A	400	DUQ	C1'-C1-C38	3.88	117.06	108.66
2	B	400	DUQ	C5-C6-N1	-3.72	118.95	122.44
2	C	400	DUQ	C4'-C1-C38	3.67	115.76	109.53
2	A	400	DUQ	C40-C39-N34	-3.23	104.39	111.64
2	B	400	DUQ	C1'-N1-C2	-3.03	115.45	118.53
2	C	400	DUQ	C1'-C1-C38	2.98	115.11	108.66
2	C	400	DUQ	O37-C40-C39	-2.88	101.06	111.59
2	A	400	DUQ	C5-C6-N1	-2.83	119.79	122.44
2	C	400	DUQ	C20-C19-C18	-2.78	117.84	120.76
2	C	400	DUQ	O4-C4-N3	-2.73	115.29	119.31
2	B	400	DUQ	C1'-C1-C38	2.70	114.51	108.66
2	C	400	DUQ	O4-C4-C5	-2.37	120.99	125.16
2	C	400	DUQ	C6-N1-C2	2.33	123.10	121.04
2	C	400	DUQ	C18-C17-C24	-2.32	103.88	110.95
2	C	400	DUQ	C6-C5-C4	-2.31	116.35	119.52
2	A	400	DUQ	C34-C35-C30	2.25	123.12	120.76
2	B	400	DUQ	C20-C19-C18	2.15	123.01	120.76
2	A	400	DUQ	O4-C4-N3	2.11	122.40	119.31
2	B	400	DUQ	C18-C17-C30	-2.06	104.66	110.95
2	C	400	DUQ	C21-C22-C23	-2.03	117.09	120.19
2	B	400	DUQ	C34-C35-C30	2.02	122.88	120.76

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	400	DUQ	C1
2	B	400	DUQ	C1
2	C	400	DUQ	C1

All (25) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	400	DUQ	C1-C1'-N1-C2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	A	400	DUQ	C4'-C1-C38-N34
2	A	400	DUQ	C4'-C1-C38-O38
2	B	400	DUQ	C1-C1'-N1-C2
2	B	400	DUQ	C4'-C1-C1'-N1
2	B	400	DUQ	C38-C1-C4'-C5'
2	C	400	DUQ	C1-C1'-N1-C2
2	C	400	DUQ	C4'-C1-C1'-N1
2	C	400	DUQ	C38-C1-C4'-C5'
2	B	400	DUQ	C40-C39-N34-C38
2	A	400	DUQ	C1-C4'-C5'-O1
2	B	400	DUQ	C1'-C1-C38-N34
2	B	400	DUQ	N34-C39-C40-O37
2	C	400	DUQ	C4'-C1-C38-O38
2	B	400	DUQ	C4'-C1-C38-N34
2	B	400	DUQ	C1-C4'-C5'-O1
2	B	400	DUQ	C38-C1-C1'-N1
2	B	400	DUQ	C4'-C1-C38-O38
2	C	400	DUQ	C4'-C1-C38-N34
2	C	400	DUQ	C4'-C5'-O1-C17
2	B	400	DUQ	C1'-C1-C38-O38
2	C	400	DUQ	C1-C4'-C5'-O1
2	A	400	DUQ	C38-C1-C4'-C5'
2	A	400	DUQ	C4'-C5'-O1-C17
2	B	400	DUQ	C4'-C5'-O1-C17

There are no ring outliers.

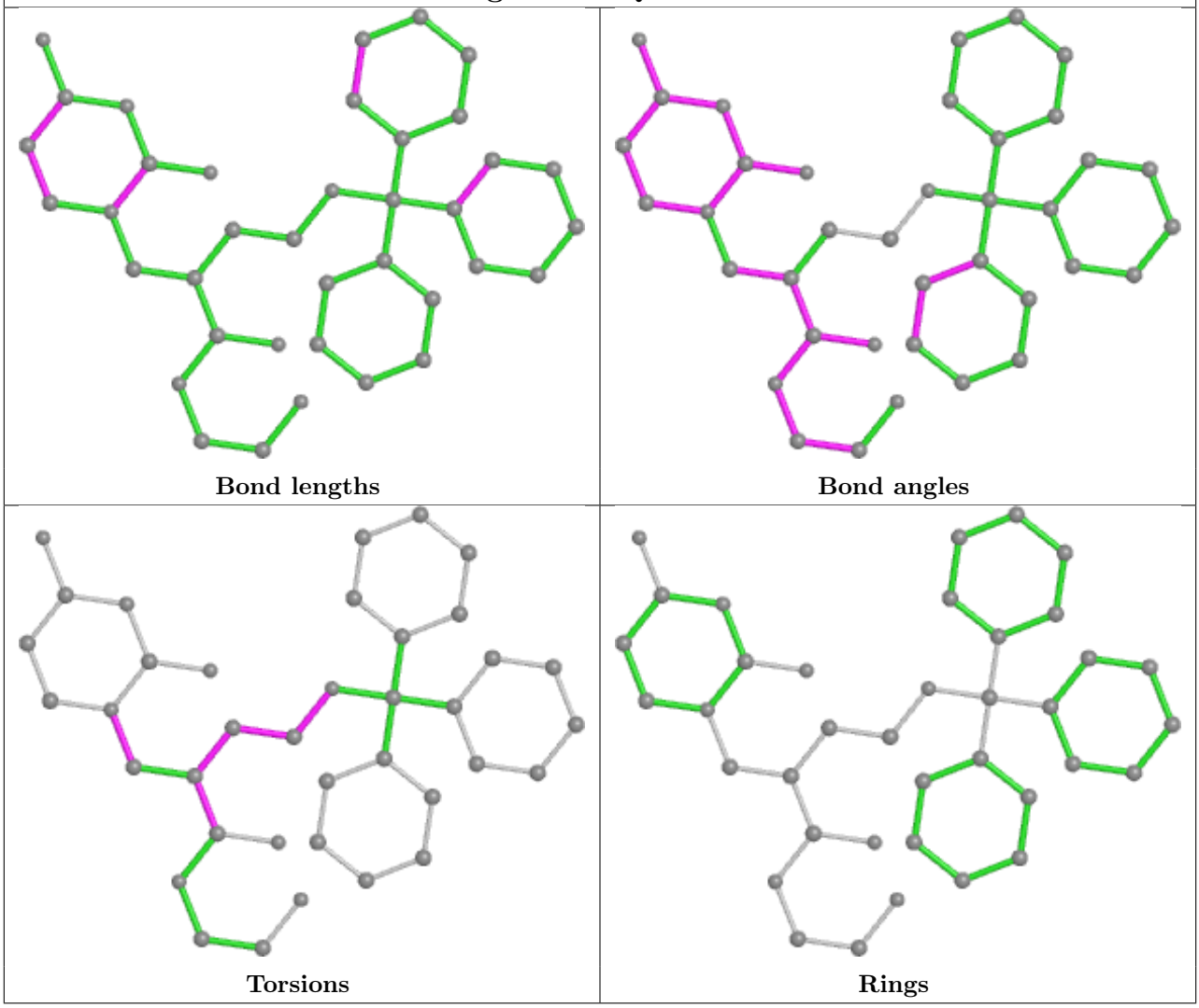
3 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	1168	SO4	1	0
2	C	400	DUQ	6	0
2	B	400	DUQ	3	0

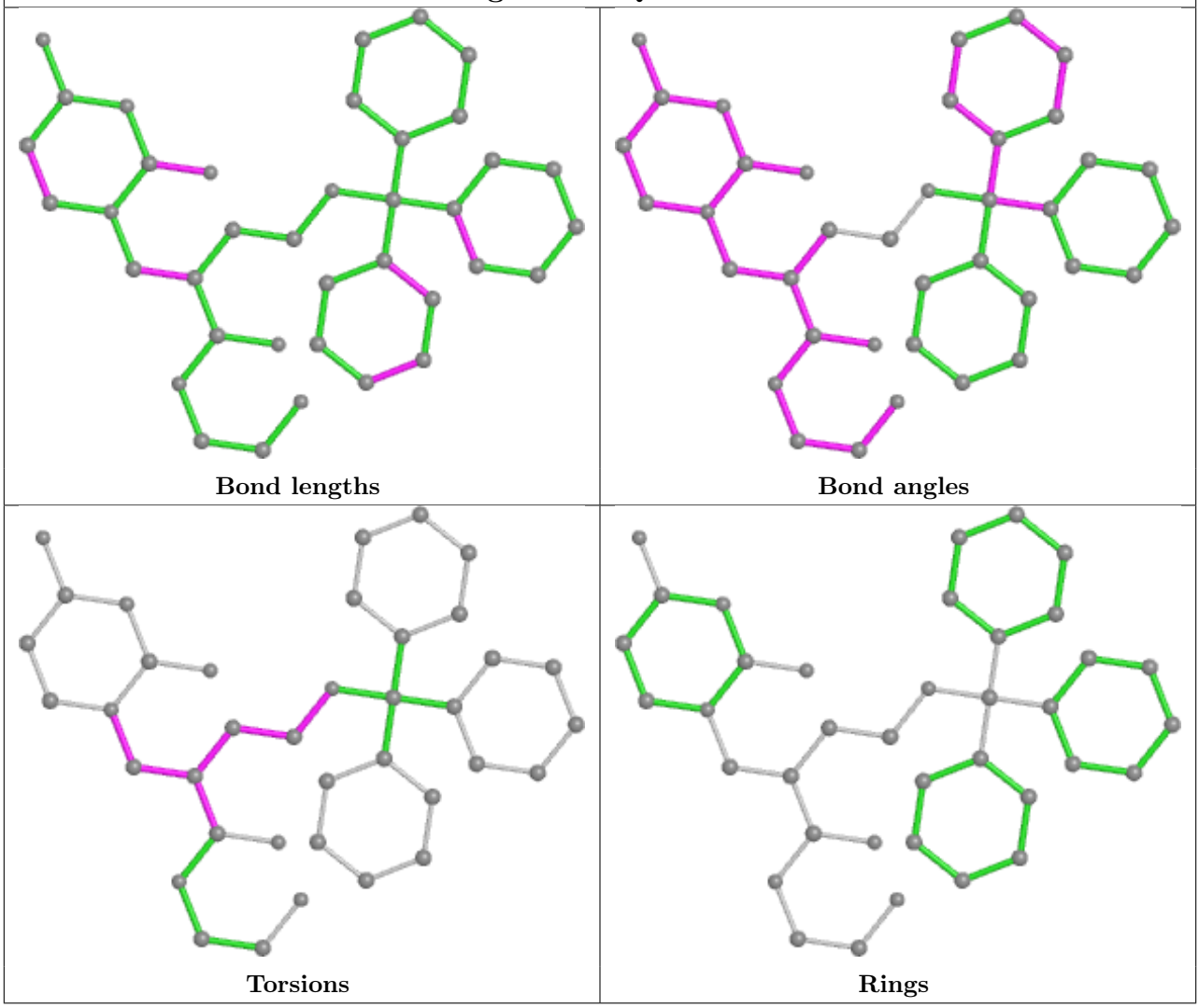
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and

any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

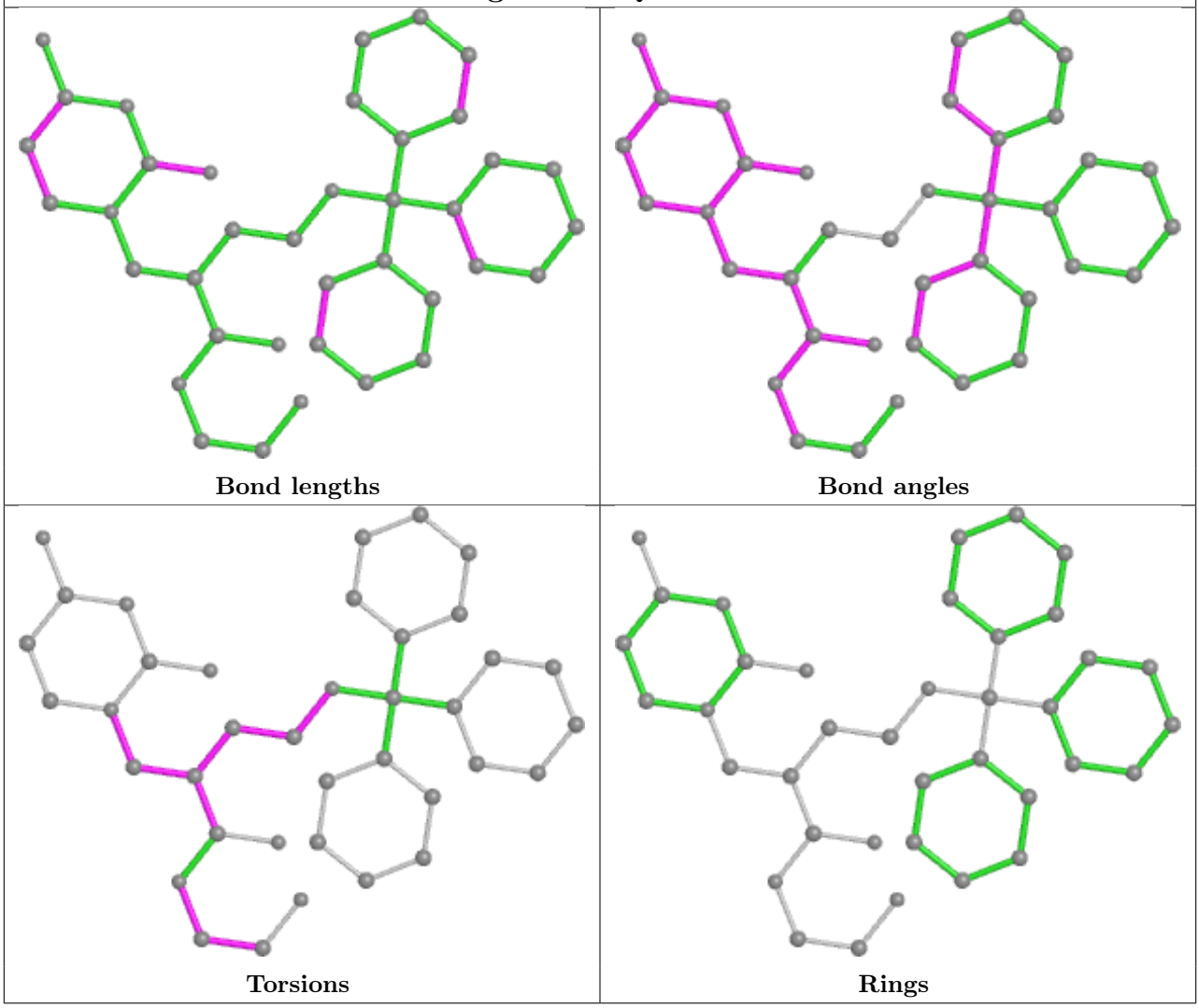
Ligand DUQ A 400



Ligand DUQ C 400



Ligand DUQ B 400



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	2
1	A	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	26:ASP	C	27:SER	N	1.16
1	A	19:HIS	C	20:LYS	N	1.00
1	B	152:VAL	C	153:GLU	N	0.91
1	A	25:GLY	C	26:ASP	N	0.85

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	142/173 (82%)	0.15	6 (4%)	36	42	21, 33, 56, 85	14 (9%)
1	B	120/173 (69%)	1.00	33 (27%)	0	0	25, 47, 72, 90	13 (10%)
1	C	135/173 (78%)	0.25	7 (5%)	27	32	21, 33, 59, 75	6 (4%)
All	All	397/519 (76%)	0.44	46 (11%)	4	6	21, 36, 67, 90	33 (8%)

All (46) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	105	ILE	4.3
1	B	56	LEU	4.1
1	B	8	LEU	4.0
1	B	38	VAL	3.8
1	A	105	ILE	3.8
1	B	129	HIS	3.5
1	A	104	SER	3.2
1	B	18	ASN	3.2
1	B	133	ASN	3.1
1	C	105	ILE	3.1
1	B	152	VAL	3.1
1	B	88	LEU	3.1
1	B	16	TYR	3.1
1	C	101	LEU	3.0
1	B	6	VAL	3.0
1	B	150	GLU	3.0
1	A	23	HIS	2.8
1	A	22	HIS	2.8
1	C	14	GLU	2.8
1	C	82	VAL	2.8
1	B	10	ASP	2.7
1	C	18	ASN	2.7
1	B	126	GLN	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	104	SER	2.7
1	B	148	SER	2.6
1	A	101	LEU	2.6
1	B	15	MET	2.6
1	B	106	GLY	2.6
1	B	12	VAL	2.6
1	B	101	LEU	2.5
1	B	151	LEU	2.4
1	B	2	HIS	2.4
1	B	19	HIS	2.4
1	B	11	GLU	2.4
1	B	113	ARG	2.4
1	B	89	PHE	2.4
1	B	39	LEU	2.3
1	C	1	MET	2.2
1	A	167	SER	2.2
1	B	84	THR	2.2
1	B	50	GLY	2.2
1	B	125	ASP	2.1
1	B	87	LEU	2.1
1	B	147	LEU	2.1
1	B	102	ALA	2.1
1	C	137	VAL	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.

Continued on next page...

Continued from previous page...

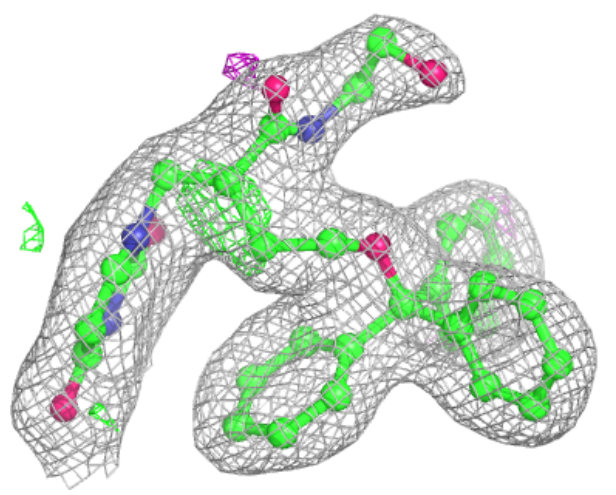
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9

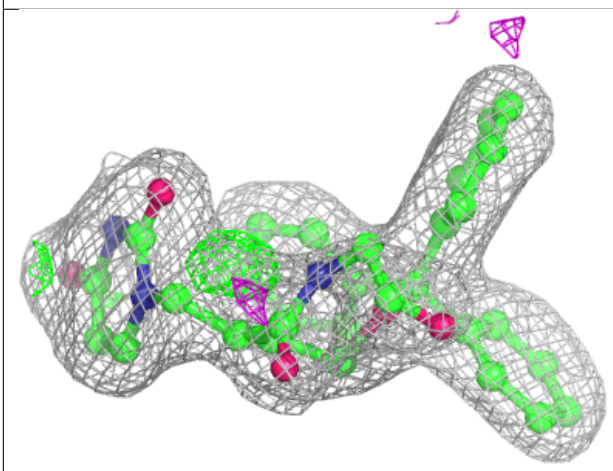
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	DUQ	B	400	38/38	0.92	0.15	34,40,61,69	0
2	DUQ	A	400	38/38	0.94	0.13	28,39,56,65	0
3	SO4	A	1168	5/5	0.94	0.17	31,33,37,38	5
2	DUQ	C	400	38/38	0.95	0.12	22,31,50,67	0
3	SO4	C	1168	5/5	0.95	0.10	46,47,49,49	5

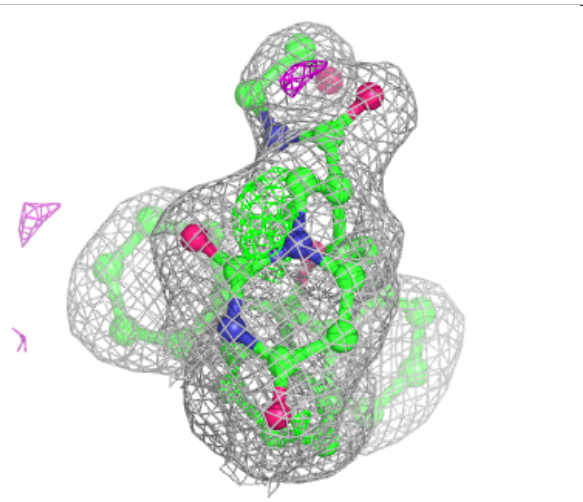
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around DUQ B 400:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)

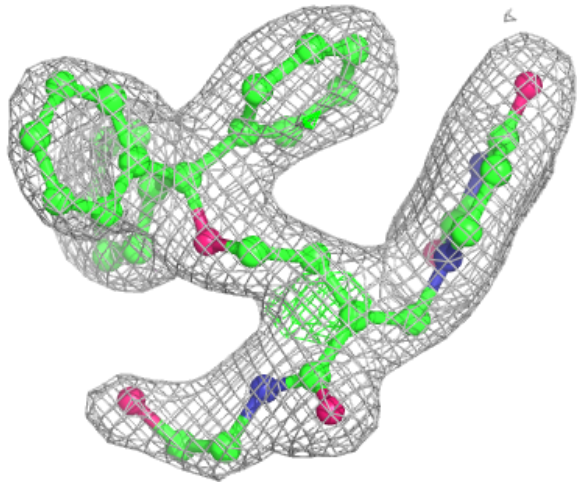


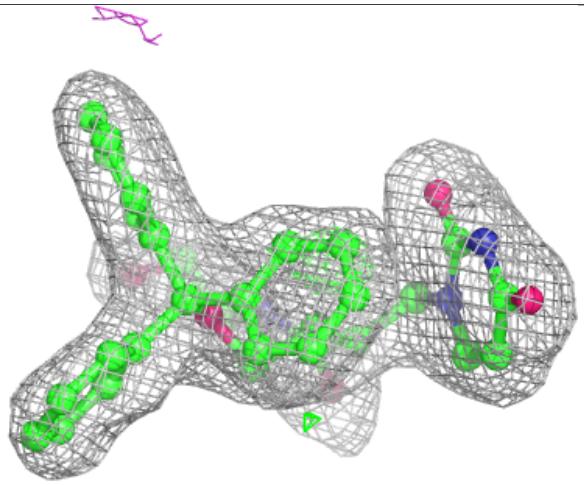


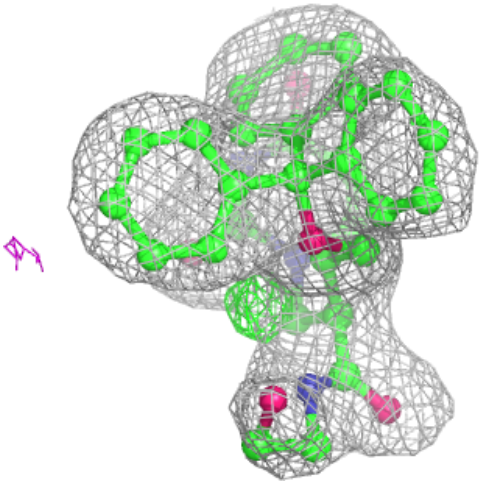


Electron density around DUQ A 400:

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

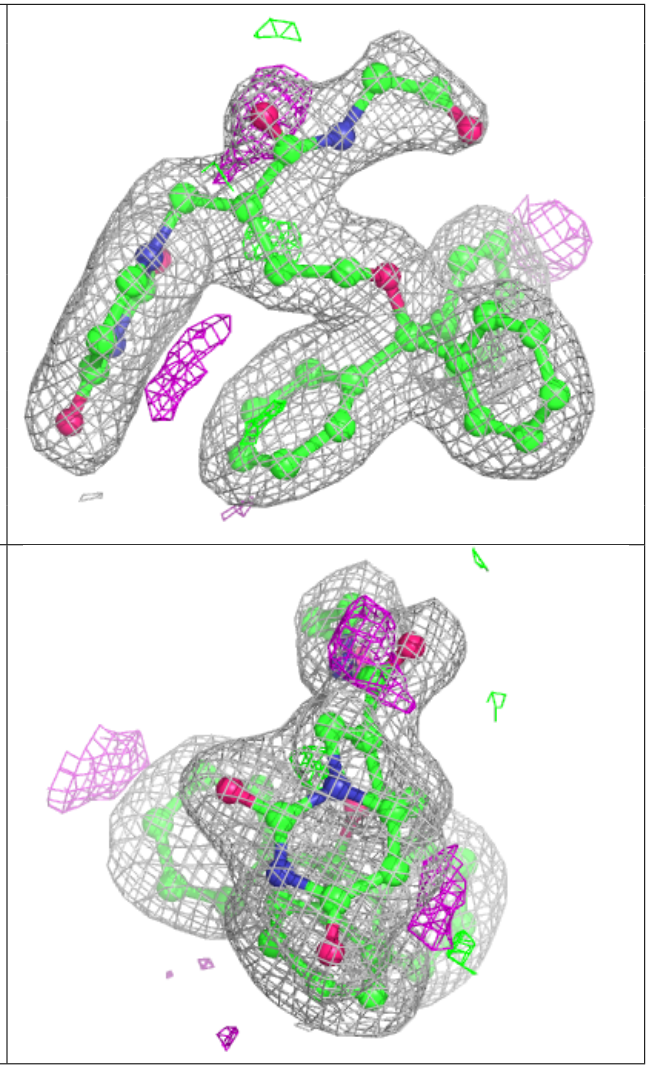






Electron density around DUQ C 400:

$2mF_o$-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.