wwPDB X-ray Structure Validation Summary Report (i)

Jan 20, 2024 – 03:53 pm GMT

PDB ID	:	6ZPL
Title	:	Inward-open structure of human glycine transporter 1 in complex with a benzoylisoindoline inhibitor, sybody Sb_GlyT1#7 and bound Na and Cl ions.
Authors	:	Shahsavar, A.; Stohler, P.; Bourenkov, G.; Zimmermann, I.; Siegrist, M.; Guba, W.; Pinard, E.; Sinning, S.; Seeger, M.A.; Schneider, T.R.; Dawson, R.J.P.; Nissen, P.
Deposited on	:	2020-07-08
Resolution	:	3.94 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 (1)) were used in the production of this report:

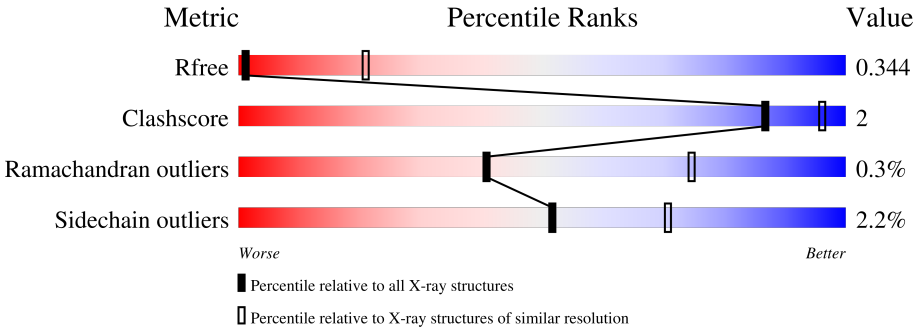
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
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 3.94 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	1036 (4.20-3.68)
Clashscore	141614	1009 (4.18-3.70)
Ramachandran outliers	138981	1057 (4.20-3.68)
Sidechain outliers	138945	1049 (4.20-3.68)

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%

Mol	Chain	Length	Quality of chain
1	A	576	84% 6% 10%
1	B	576	85% 6% 8%
2	C	273	95% . .
3	E	120	99% .
3	F	120	93% 7%

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
6	CL	B	703	-	-	X	-

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 24610 atoms, of which 12117 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 Sodium- and chloride-dependent glycine transporter 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	B	530	Total	C	H	N	O	S	0	0	0
			8372	2815	4162	655	706	34			
1	A	518	Total	C	H	N	O	S	0	0	0
			8186	2755	4071	636	690	34			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	153	ALA	LEU	engineered mutation	UNP P48067
B	?	-	ASN	deletion	UNP P48067
B	?	-	GLY	deletion	UNP P48067
B	?	-	SER	deletion	UNP P48067
B	?	-	ARG	deletion	UNP P48067
B	?	-	PRO	deletion	UNP P48067
B	?	-	ALA	deletion	UNP P48067
B	?	-	ALA	deletion	UNP P48067
B	?	-	LEU	deletion	UNP P48067
B	?	-	PRO	deletion	UNP P48067
B	?	-	SER	deletion	UNP P48067
B	?	-	ASN	deletion	UNP P48067
B	?	-	LEU	deletion	UNP P48067
B	?	-	SER	deletion	UNP P48067
B	?	-	HIS	deletion	UNP P48067
B	?	-	LEU	deletion	UNP P48067
B	?	-	LEU	deletion	UNP P48067
B	?	-	ASN	deletion	UNP P48067
B	297	ALA	SER	engineered mutation	UNP P48067
B	368	ALA	ILE	engineered mutation	UNP P48067
B	633	ALA	CYS	engineered mutation	UNP P48067
A	153	ALA	LEU	engineered mutation	UNP P48067
A	?	-	ASN	deletion	UNP P48067
A	?	-	GLY	deletion	UNP P48067
A	?	-	SER	deletion	UNP P48067

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	ARG	deletion	UNP P48067
A	?	-	PRO	deletion	UNP P48067
A	?	-	ALA	deletion	UNP P48067
A	?	-	ALA	deletion	UNP P48067
A	?	-	LEU	deletion	UNP P48067
A	?	-	PRO	deletion	UNP P48067
A	?	-	SER	deletion	UNP P48067
A	?	-	ASN	deletion	UNP P48067
A	?	-	LEU	deletion	UNP P48067
A	?	-	SER	deletion	UNP P48067
A	?	-	HIS	deletion	UNP P48067
A	?	-	LEU	deletion	UNP P48067
A	?	-	LEU	deletion	UNP P48067
A	?	-	ASN	deletion	UNP P48067
A	297	ALA	SER	engineered mutation	UNP P48067
A	368	ALA	ILE	engineered mutation	UNP P48067
A	633	ALA	CYS	engineered mutation	UNP P48067

- Molecule 2 is a protein called Endoglucanase H.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	C	273	Total	C	H	N	O	S	13	0	0
			4286	1417	2076	378	411	4			

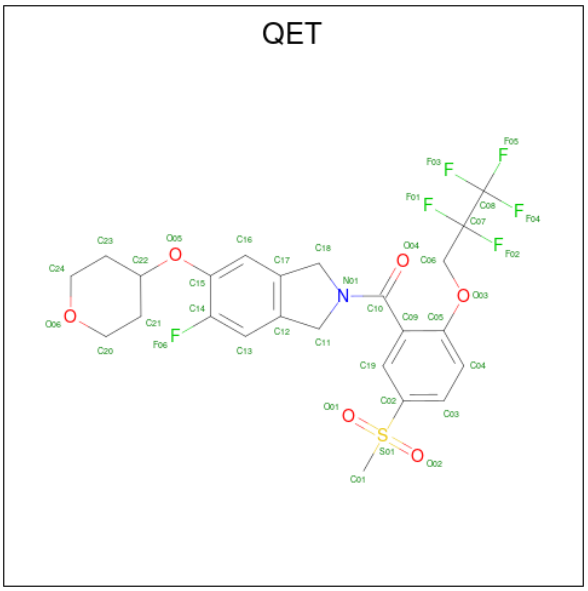
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	109	ALA	GLU	engineered mutation	UNP P16218

- Molecule 3 is a protein called Sybody Sb_GlyT1#7.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	E	120	Total	C	H	N	O	S	0	0	0
			1820	598	881	157	180	4			
3	F	120	Total	C	H	N	O	S	0	0	0
			1820	598	881	157	180	4			

- Molecule 4 is [5-fluoranyl-6-(oxan-4-yloxy)-1,3-dihydroisoindol-2-yl]-[5-methylsulfonyl-2-[2,2,3,3,3-pentakis(fluoranyl)propoxy]phenyl]methanone (three-letter code: QET) (formula: $C_{24}H_{23}F_6NO_6S$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total C F H N O S 61 24 6 23 1 6 1	0	0
4	A	1	Total C F H N O S 61 24 6 23 1 6 1	0	0

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total Na 1 1	0	0
5	A	1	Total Na 1 1	0	0

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	1	Total Cl 1 1	0	0
6	A	1	Total Cl 1 1	0	0

3 Residue-property plots (i)

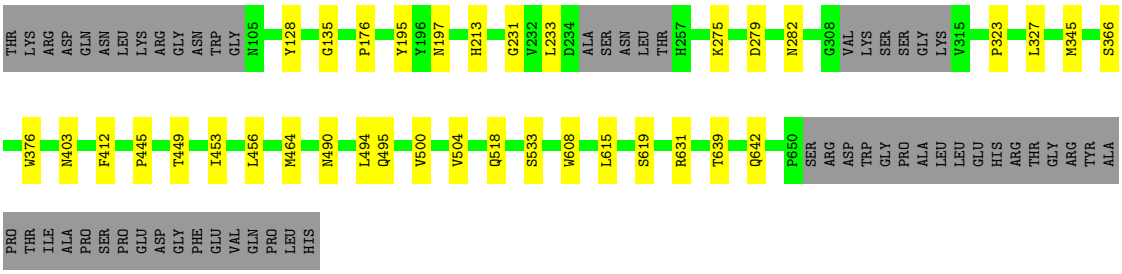
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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. 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: Sodium- and chloride-dependent glycine transporter 1

Chain B:

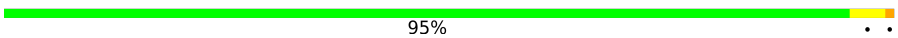
• Molecule 1: Sodium- and chloride-dependent glycine transporter 1

Chain A:



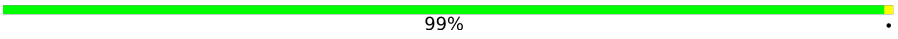
• Molecule 2: Endoglucanase H

Chain C:



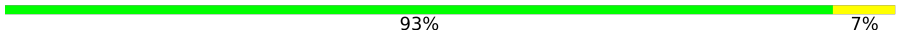
• Molecule 3: Sybody Sb_GlyT1#7

Chain E:



• Molecule 3: Sybody Sb_GlyT1#7

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	116.41Å 69.71Å 149.43Å 90.00° 92.86° 90.00°	Depositor
Resolution (Å)	29.07 – 3.94 29.89 – 3.60	Depositor EDS
% Data completeness (in resolution range)	100.0 (29.07-3.94) 98.6 (29.89-3.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.95 (at 3.56Å)	Xtriage
Refinement program	BUSTER 2.10.3 (6-FEB-2020)	Depositor
R, R_{free}	0.277 , 0.291 0.317 , 0.344	Depositor DCC
R_{free} test set	1391 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	111.6	Xtriage
Anisotropy	0.357	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 28.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.38$, $< L^2 > = 0.20$	Xtriage
Estimated twinning fraction	0.125 for h,-k,-l	Xtriage
F_o,F_c correlation	0.85	EDS
Total number of atoms	24610	wwPDB-VP
Average B, all atoms (Å^2)	106.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 30.80 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.2299e-03.*

[1] Intensities estimated from amplitudes.

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

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CL, NA, QET

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	0.39	0/4246	0.59	0/5793
1	B	0.39	0/4345	0.60	0/5928
2	C	0.36	0/2279	0.59	0/3110
3	E	0.38	0/966	0.64	0/1315
3	F	0.38	0/966	0.63	0/1315
All	All	0.38	0/12802	0.60	0/17461

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	4115	4071	4093	22	1
1	B	4210	4162	4184	16	0
2	C	2210	2076	2078	6	0
3	E	939	881	879	0	0
3	F	939	881	879	8	1
4	A	38	23	0	2	0
4	B	38	23	0	1	0
5	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	1	0	0	0	0
6	A	1	0	0	0	0
6	B	1	0	0	3	0
All	All	12493	12117	12113	52	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

The worst 5 of 52 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:323:PRO:O	1:A:327:LEU:HD23	1.59	1.02
1:A:490:ASN:O	1:A:494:LEU:HD23	1.88	0.74
3:F:6:GLU:OE2	3:F:96:CYS:HB3	1.87	0.73
3:F:27:PHE:CZ	3:F:29:VAL:HG12	2.25	0.72
1:A:176:PRO:HB2	1:A:631:ARG:NH2	2.06	0.70

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:345:MET:SD	3:F:11:LEU:O[2_656]	2.17	0.03

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	512/576 (89%)	491 (96%)	19 (4%)	2 (0%)	34 70
1	B	524/576 (91%)	499 (95%)	23 (4%)	2 (0%)	34 70
2	C	271/273 (99%)	254 (94%)	17 (6%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	E	118/120 (98%)	112 (95%)	6 (5%)	0	100	100
3	F	118/120 (98%)	110 (93%)	8 (7%)	0	100	100
All	All	1543/1665 (93%)	1466 (95%)	73 (5%)	4 (0%)	41	74

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	135	GLY
1	A	135	GLY
1	A	231	GLY
1	B	231	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	437/485 (90%)	429 (98%)	8 (2%)	59	77
1	B	446/485 (92%)	431 (97%)	15 (3%)	37	61
2	C	225/225 (100%)	221 (98%)	4 (2%)	59	77
3	E	94/94 (100%)	93 (99%)	1 (1%)	73	84
3	F	94/94 (100%)	94 (100%)	0	100	100
All	All	1296/1383 (94%)	1268 (98%)	28 (2%)	52	71

5 of 28 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	654	TRP
3	E	45	ARG
1	A	197	ASN
2	C	69	TRP
1	A	195	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such

sidechains are listed below:

Mol	Chain	Res	Type
2	C	28	GLN
2	C	61	ASN
2	C	163	ASN
1	B	403	ASN
1	B	282	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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	QET	A	701	-	40,41,41	2.58	8 (20%)	59,63,63	2.89	14 (23%)
4	QET	B	701	-	40,41,41	2.62	8 (20%)	59,63,63	2.63	16 (27%)

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	QET	A	701	-	-	12/33/49/49	0/4/4/4
4	QET	B	701	-	-	15/33/49/49	0/4/4/4

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	701	QET	C02-S01	9.62	1.87	1.77
4	B	701	QET	C10-N01	8.87	1.54	1.34
4	A	701	QET	C10-N01	8.81	1.54	1.34
4	A	701	QET	C02-S01	8.52	1.86	1.77
4	A	701	QET	O05-C15	5.31	1.46	1.37

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	701	QET	O02-S01-O01	-10.12	100.50	117.92
4	B	701	QET	C12-C11-N01	8.76	106.60	102.46
4	A	701	QET	C12-C11-N01	8.47	106.47	102.46
4	A	701	QET	O02-S01-O01	-8.39	103.47	117.92
4	A	701	QET	C17-C18-N01	8.22	106.35	102.46

There are no chirality outliers.

5 of 27 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	701	QET	C06-C07-C08-F03
4	B	701	QET	C06-C07-C08-F04
4	B	701	QET	C06-C07-C08-F05
4	B	701	QET	F01-C07-C08-F03
4	B	701	QET	F02-C07-C08-F04

There are no ring outliers.

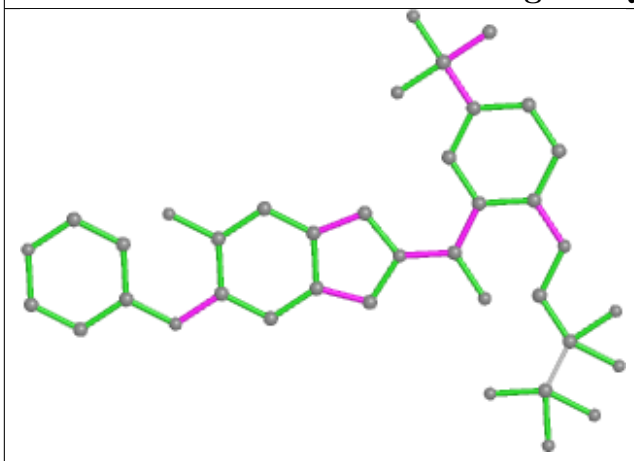
2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	701	QET	2	0
4	B	701	QET	1	0

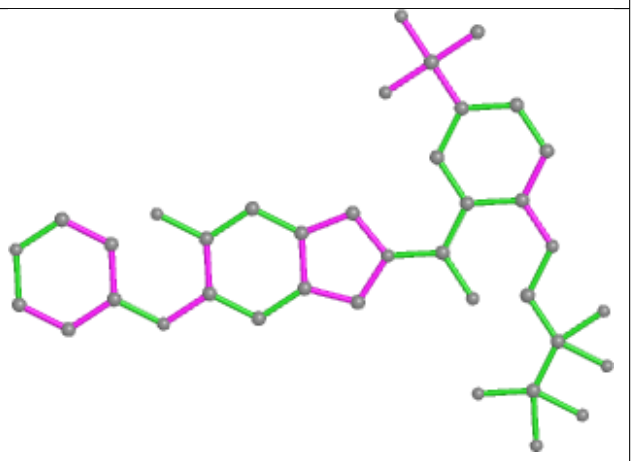
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less 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.

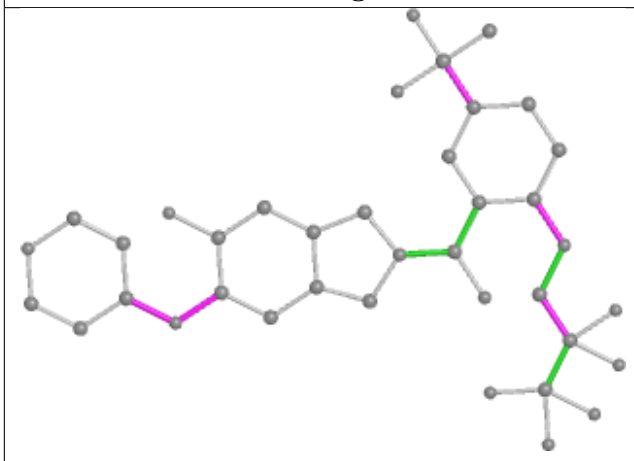
Ligand QET A 701



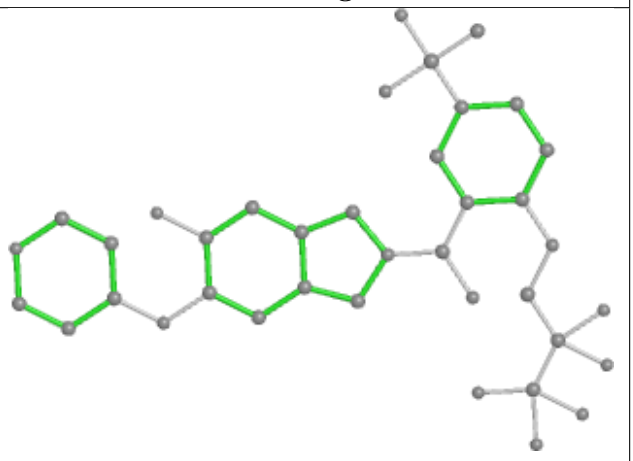
Bond lengths



Bond angles

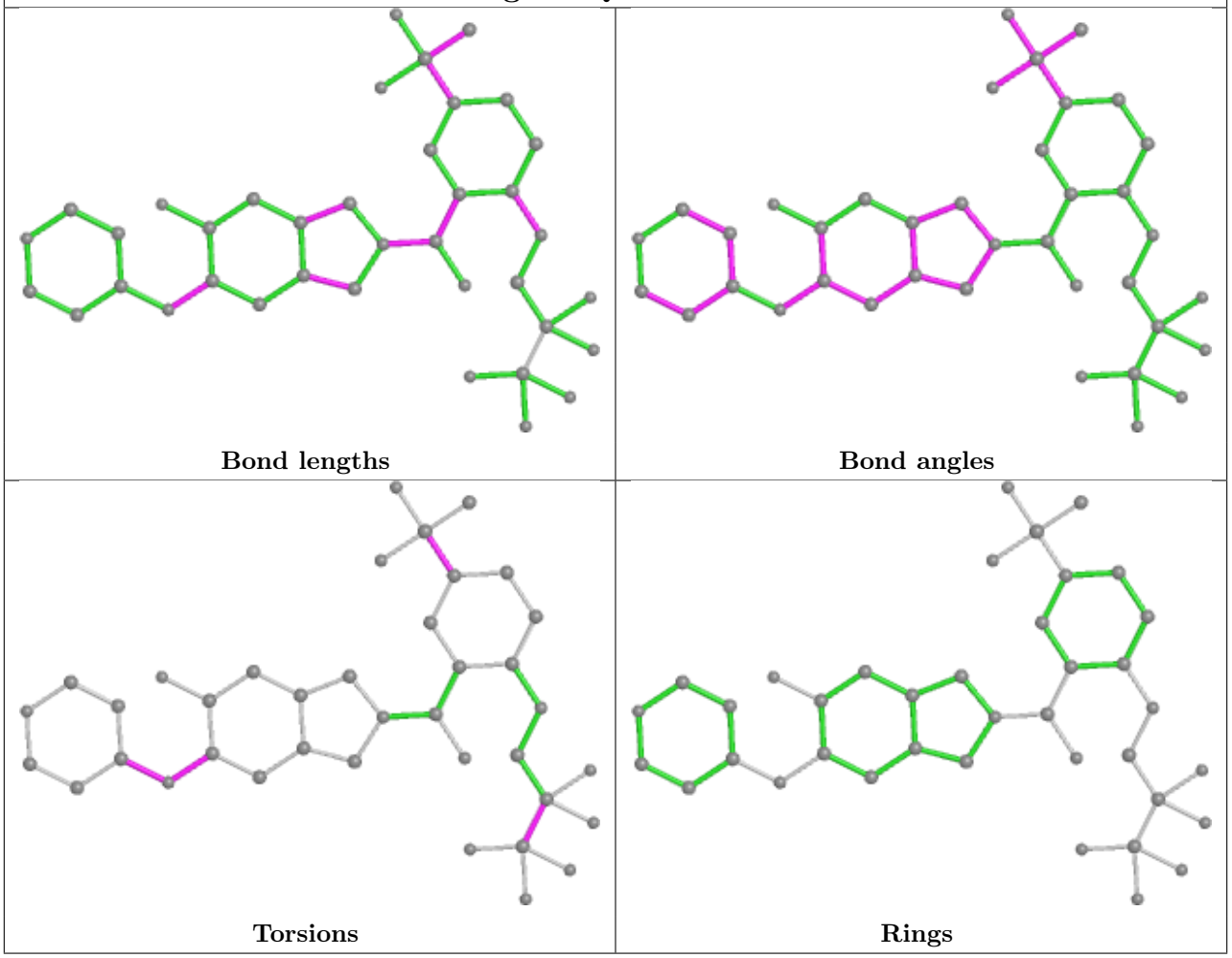


Torsions



Rings

Ligand QET B 701



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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

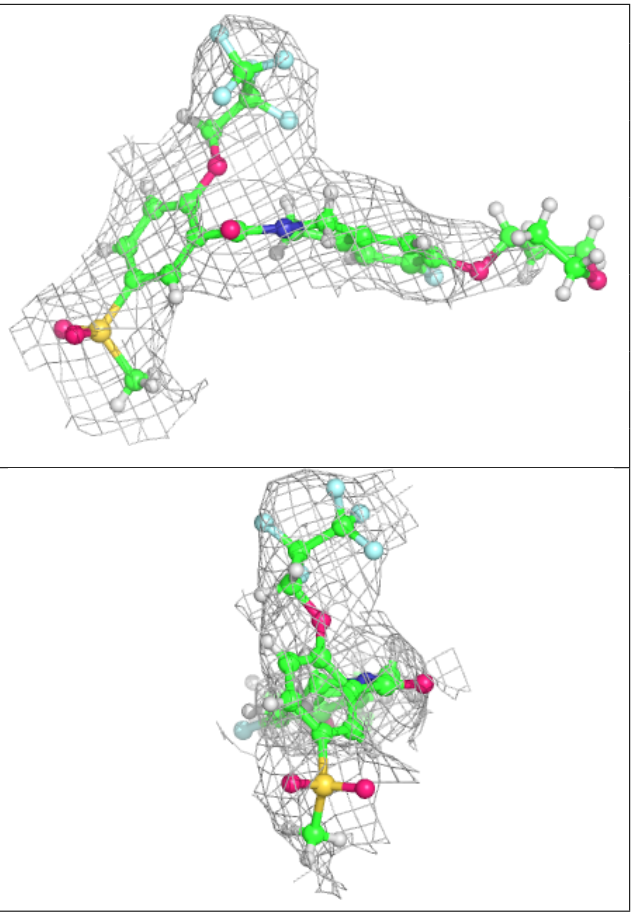
6.4 Ligands

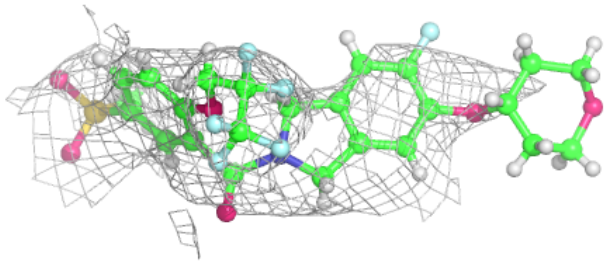
Unable to reproduce the depositors R factor - this section is therefore empty.

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 QET B 701:

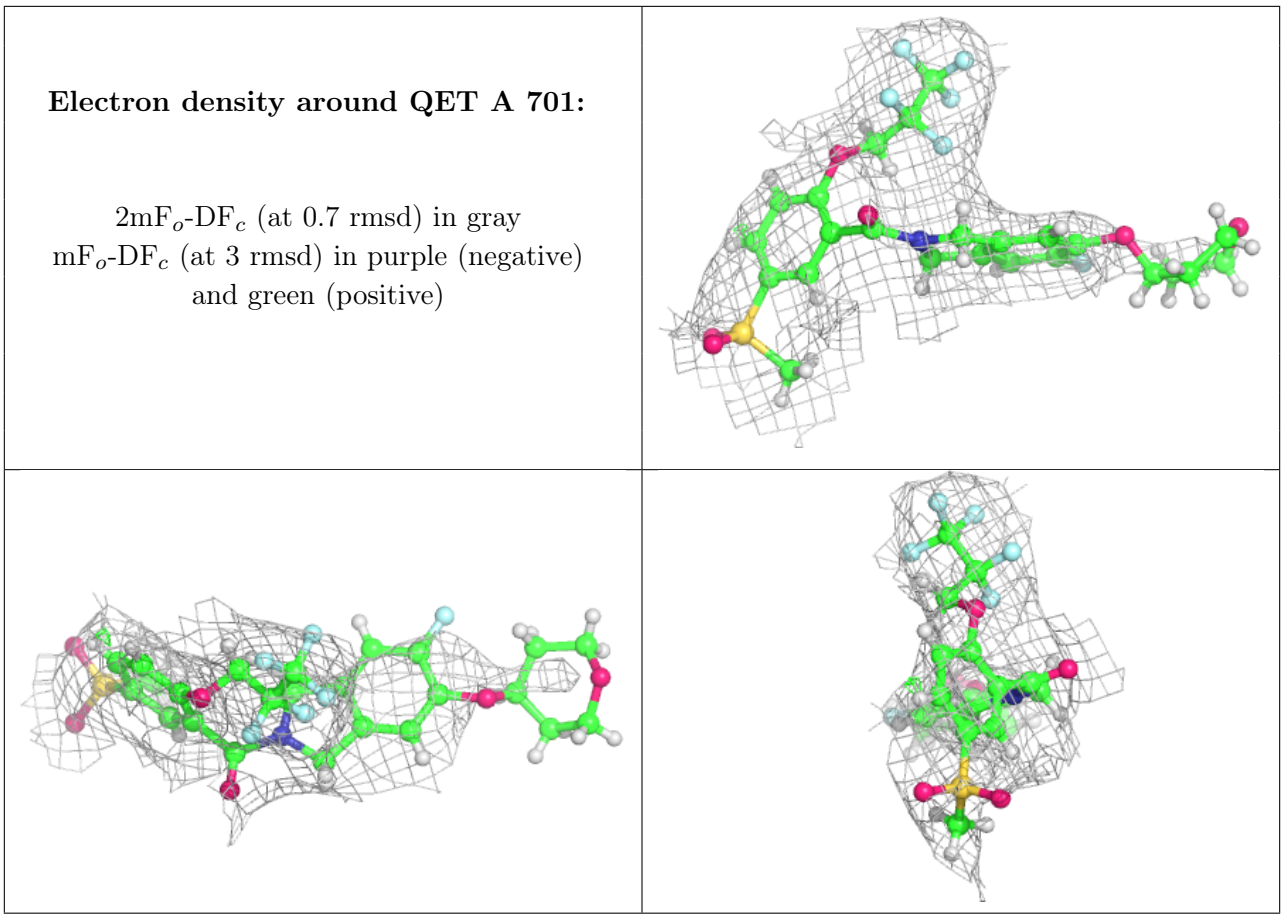
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)





Electron density around QET A 701:

$2mF_o$-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

Unable to reproduce the depositors R factor - this section is therefore empty.