



wwPDB X-ray Structure Validation Summary Report ⓘ

May 26, 2020 – 09:27 am BST

PDB ID : 6NTT
Title : X-ray Crystal Structure of Soybean Trypsin Inhibitor (Kunitz) Complexed with 1,5-Disulfonyl Naphthalene
Authors : McPherson, A.
Deposited on : 2019-01-30
Resolution : 2.40 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 ⓘ symbol.

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 : 2.11
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) : 2.11

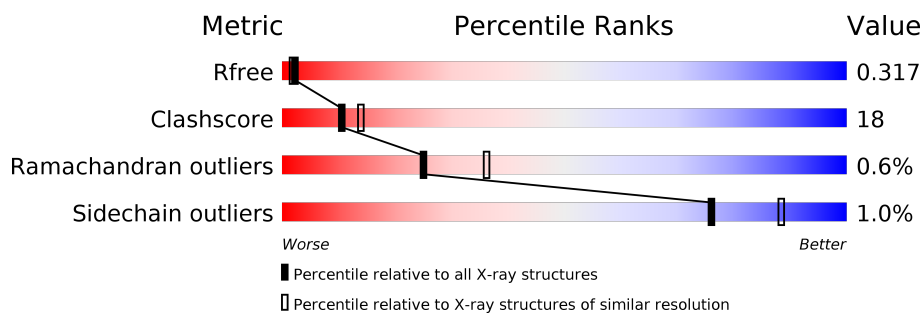
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	216	
1	B	216	

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	21D	A	201[A]	-	-	X	-
2	21D	A	201[B]	-	-	X	-
2	21D	B	201[A]	-	-	X	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	MES	A	202[A]	-	-	X	-

2 Entry composition [i](#)

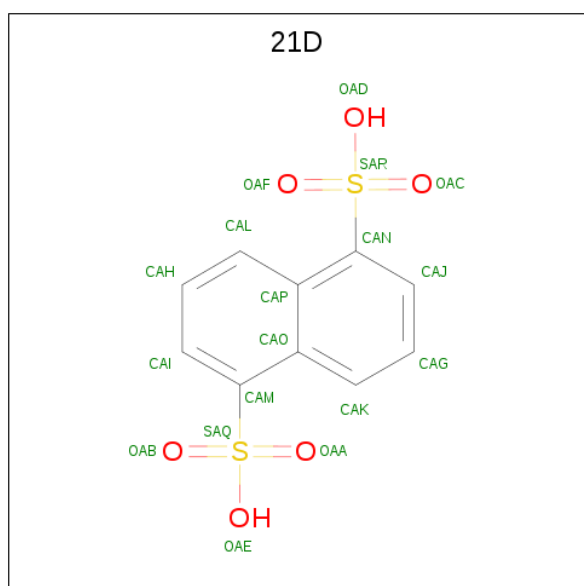
There are 4 unique types of molecules in this entry. The entry contains 3130 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Trypsin inhibitor A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	175	Total	C	N	O	S	0	2	0
			1376	871	232	267	6			
1	B	176	Total	C	N	O	S	0	0	0
			1373	868	232	267	6			

- Molecule 2 is naphthalene-1,5-disulfonic acid (three-letter code: 21D) (formula: C₁₀H₈O₆S₂).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	O	S	0	1
			36	20	12	4		
2	B	1	Total	C	O	S	0	1
			36	20	12	4		
2	B	1	Total	C	O	S	0	1
			36	20	12	4		

- Molecule 3 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C₆H₁₃NO₄S).




Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	1
			24	12	2	8	2		
3	B	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
3	B	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	123	Total	O	0	0
			123	123		
4	B	102	Total	O	0	0
			102	102		

These plots are drawn for all protein, RNA and DNA 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.

- Chain A: 

- Chain B:
-
- 63% 19% 19%
- MET LYS SER THR ILE PHE PHE LEU PHE LEU CYS ALA PHE ILE ALA D1 F2 V3 L11 L20 S21 D22 I23 R30 T34 L41 Q45 S46 R47 N48 I57 I64 R65 F66 I67 A68 H71 S74 F77 D78
- I89 T91 V95 G108 D115 D126 E127 F128 M129 M130 F135 Q138 E141 ASP D143 K144 T158 R159 M166 K167 V170 V171 Q172 F173 Q174 K175 L176 D177 LYS GLU SER LEU ALA LYS ASN HIS GLY LEU SER ARG SER GLU

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	41.88Å 64.68Å 69.08Å 90.00° 97.52° 90.00°	Depositor
Resolution (Å)	32.61 – 2.40 32.61 – 2.00	Depositor EDS
% Data completeness (in resolution range)	93.2 (32.61-2.40) 88.4 (32.61-2.00)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	0.12	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.15 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.209 , 0.306 0.215 , 0.317	Depositor DCC
R_{free} test set	1153 reflections (4.78%)	wwPDB-VP
Wilson B-factor (Å ²)	42.1	Xtriage
Anisotropy	0.384	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 56.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	3130	wwPDB-VP
Average B, all atoms (Å ²)	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.14% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: 21D, MES

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	0.65	0/1408	0.76	0/1905
1	B	0.65	0/1400	0.76	0/1896
All	All	0.65	0/2808	0.76	0/3801

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	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	88	GLY	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	1376	0	1348	37	0
1	B	1373	0	1340	46	0
2	A	36	0	12	21	0
2	B	72	0	27	17	0
3	A	24	0	26	9	0
3	B	24	0	26	2	0
4	A	123	0	0	7	0
4	B	102	0	0	8	0
All	All	3130	0	2779	105	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 18.

The worst 5 of 105 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:22:ASP:CB	1:B:172:GLN:HE22	1.62	1.12
2:A:201[A]:21D:H4	4:A:308:HOH:O	1.49	1.11
1:A:159:ARG:HH22	3:A:202[A]:MES:C2	1.72	1.01
1:B:30:ARG:CZ	2:B:201[A]:21D:OAD	2.09	1.01
2:A:201[B]:21D:OAF	4:A:305:HOH:O	1.80	1.00

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	171/216 (79%)	157 (92%)	13 (8%)	1 (1%)	25	36
1	B	172/216 (80%)	158 (92%)	13 (8%)	1 (1%)	25	36
All	All	343/432 (79%)	315 (92%)	26 (8%)	2 (1%)	25	36

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	126	ASP
1	A	13	ASN

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	153/187 (82%)	150 (98%)	3 (2%)	55	74
1	B	152/187 (81%)	151 (99%)	1 (1%)	84	92
All	All	305/374 (82%)	301 (99%)	4 (1%)	76	84

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	110[A]	ASN
1	A	110[B]	ASN
1	A	142	ASP
1	B	74	SER

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

Mol	Chain	Res	Type
1	B	45	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

10 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	21D	B	201[A]	-	17,19,19	0.22	0	24,30,30	0.15	0
2	21D	A	201[B]	-	17,19,19	0.42	0	24,30,30	0.23	0
2	21D	B	201[B]	-	17,19,19	0.44	0	24,30,30	0.18	0
3	MES	B	204	-	12,12,12	0.75	0	14,16,16	0.36	0
2	21D	B	202[B]	-	17,19,19	0.33	0	24,30,30	0.22	0
2	21D	A	201[A]	-	17,19,19	0.31	0	24,30,30	0.25	0
3	MES	A	202[A]	-	12,12,12	0.70	0	14,16,16	0.41	0
2	21D	B	202[A]	-	17,19,19	0.33	0	24,30,30	0.15	0
3	MES	B	203	-	12,12,12	0.74	0	14,16,16	0.39	0
3	MES	A	202[B]	-	12,12,12	0.76	0	14,16,16	0.44	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	21D	B	201[A]	-	-	5/12/12/12	0/2/2/2
2	21D	A	201[B]	-	-	11/12/12/12	0/2/2/2
2	21D	B	201[B]	-	-	6/12/12/12	0/2/2/2
3	MES	B	204	-	-	3/6/14/14	0/1/1/1
2	21D	B	202[B]	-	-	6/12/12/12	0/2/2/2
2	21D	A	201[A]	-	-	12/12/12/12	0/2/2/2
3	MES	A	202[A]	-	-	5/6/14/14	0/1/1/1
2	21D	B	202[A]	-	-	0/12/12/12	0/2/2/2

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	MES	B	203	-	-	3/6/14/14	0/1/1/1
3	MES	A	202[B]	-	-	6/6/14/14	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 57 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	201[A]	21D	CAI-CAM-SAQ-OAE
2	B	201[A]	21D	CAI-CAM-SAQ-OAA
2	B	201[A]	21D	CAO-CAM-SAQ-OAA
2	B	201[A]	21D	CAI-CAM-SAQ-OAB
2	B	201[A]	21D	CAO-CAM-SAQ-OAB

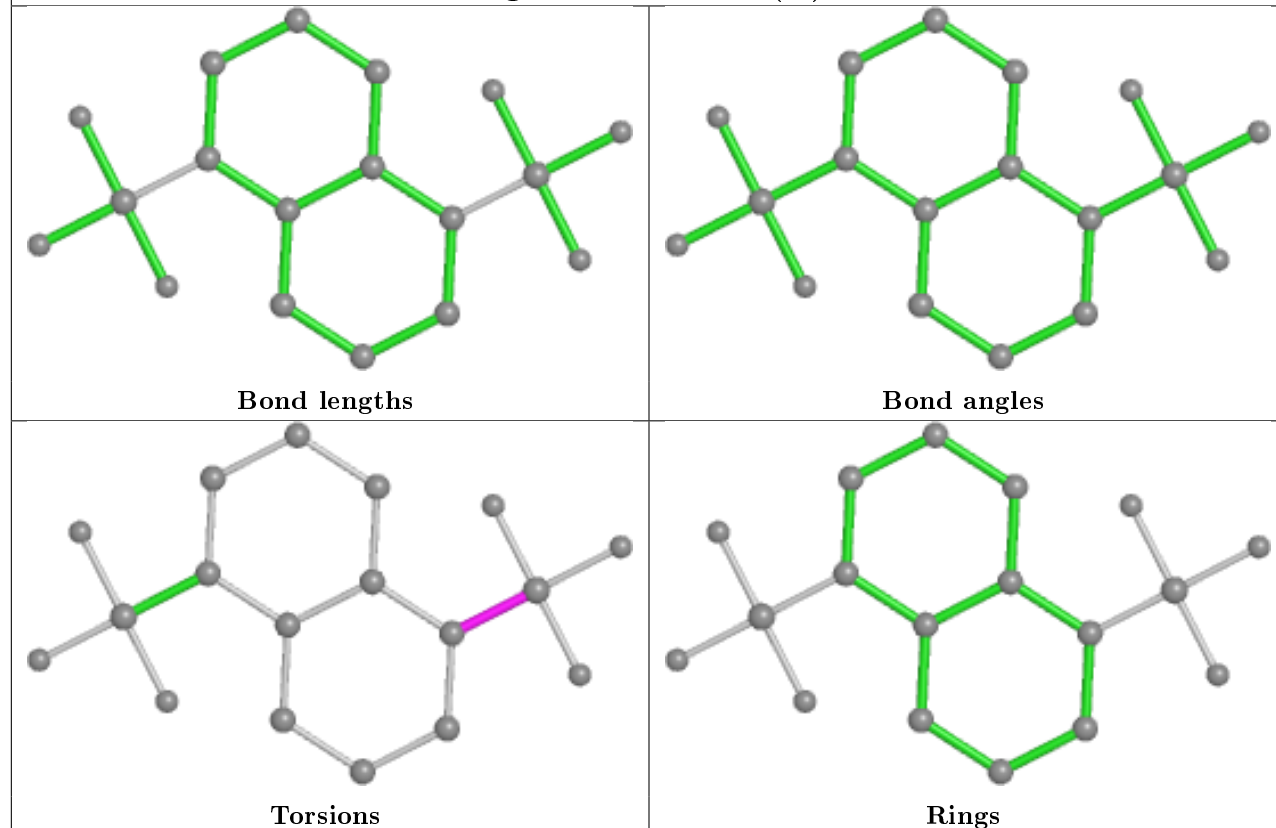
There are no ring outliers.

9 monomers are involved in 46 short contacts:

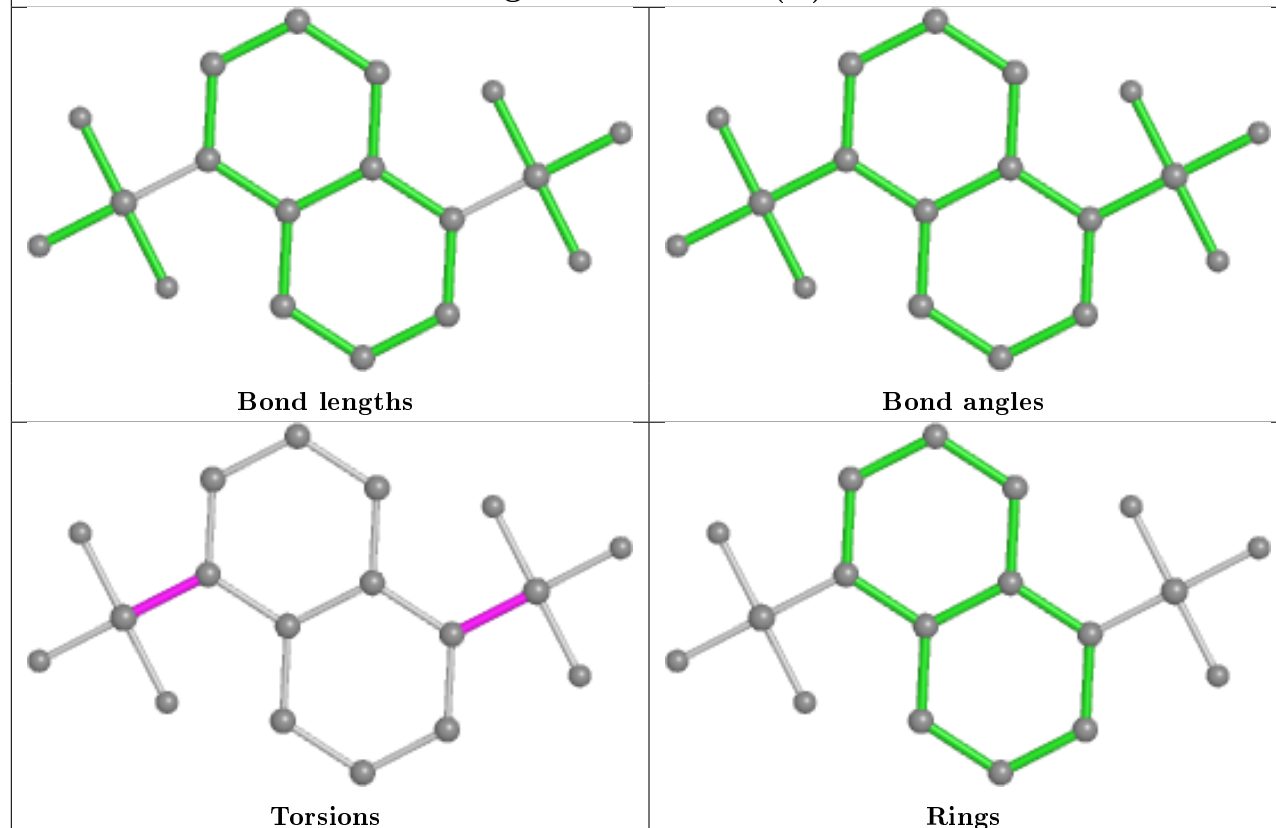
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	201[A]	21D	7	0
2	A	201[B]	21D	7	0
2	B	201[B]	21D	5	0
2	B	202[B]	21D	4	0
2	A	201[A]	21D	14	0
3	A	202[A]	MES	7	0
2	B	202[A]	21D	1	0
3	B	203	MES	2	0
3	A	202[B]	MES	2	0

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.

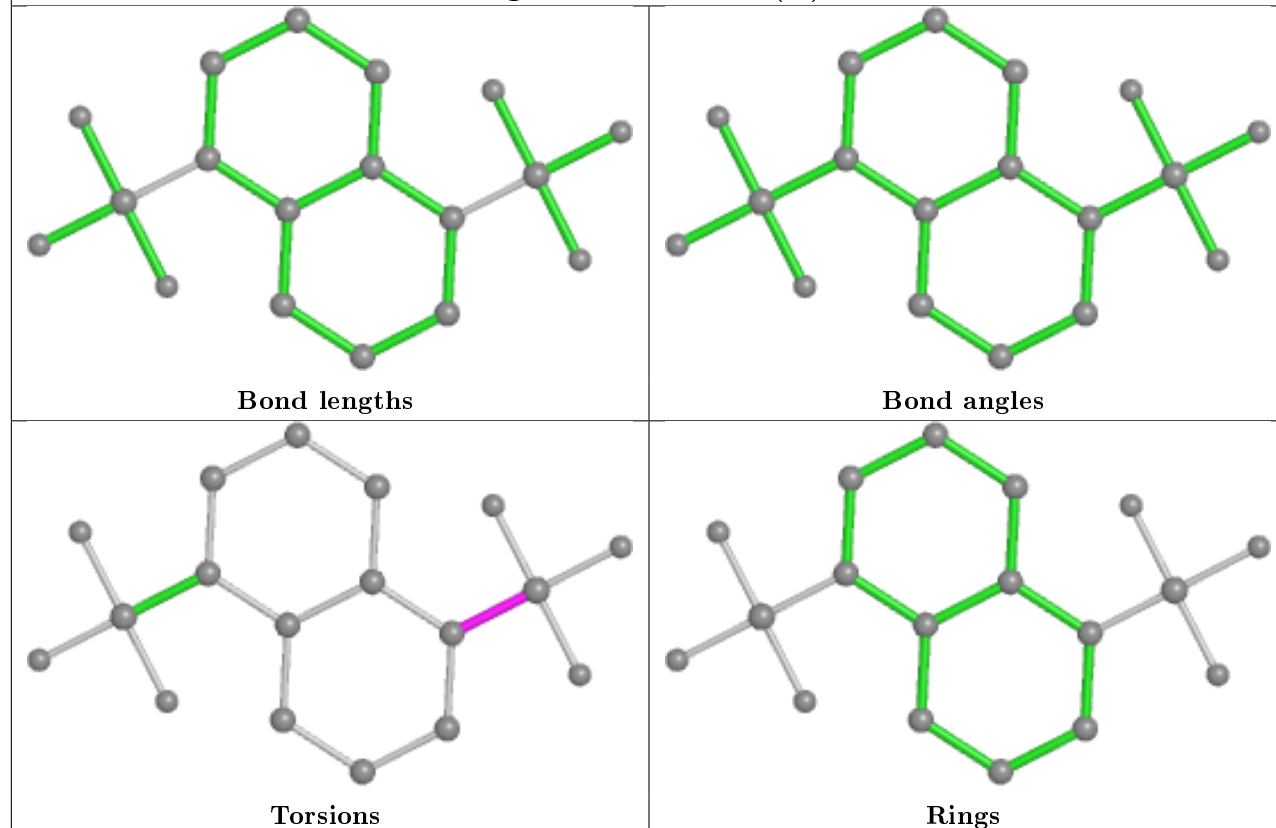
Ligand 21D B 201 (A)



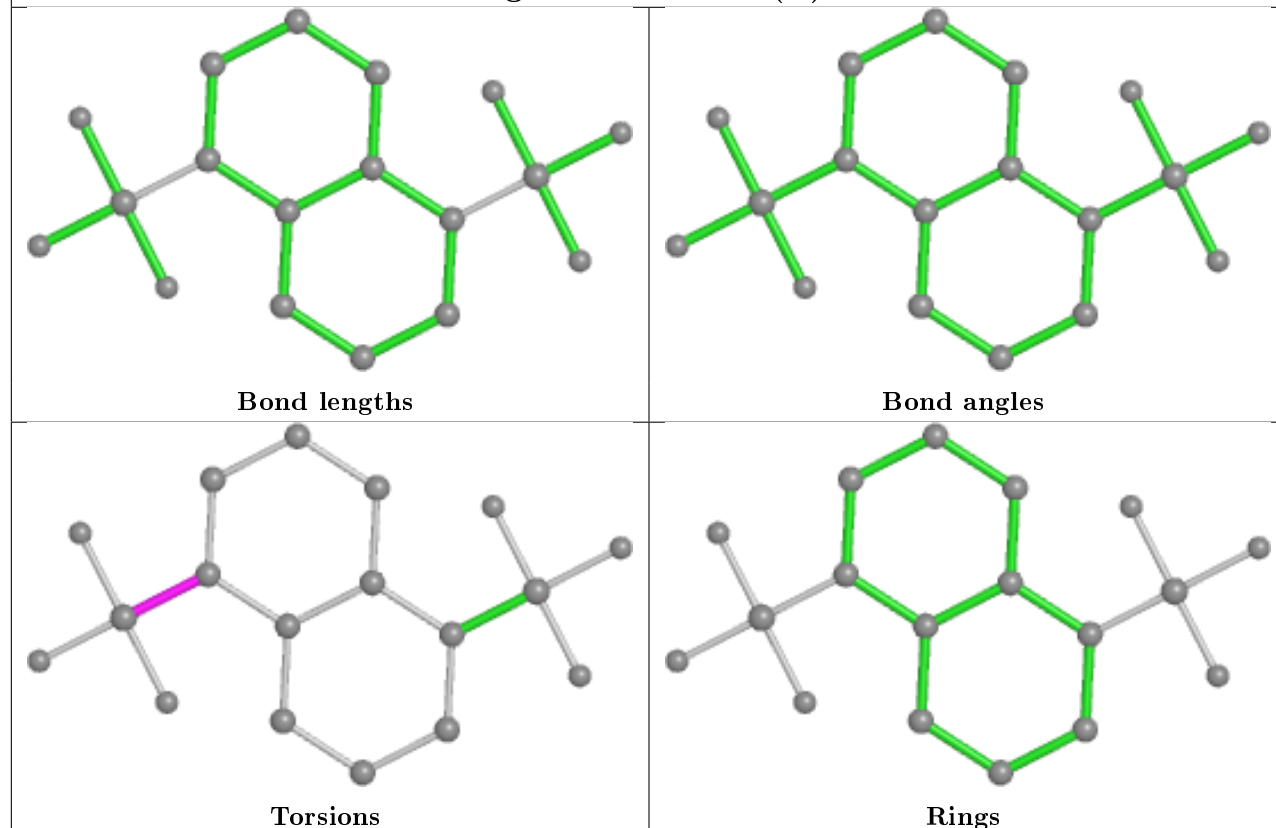
Ligand 21D A 201 (B)

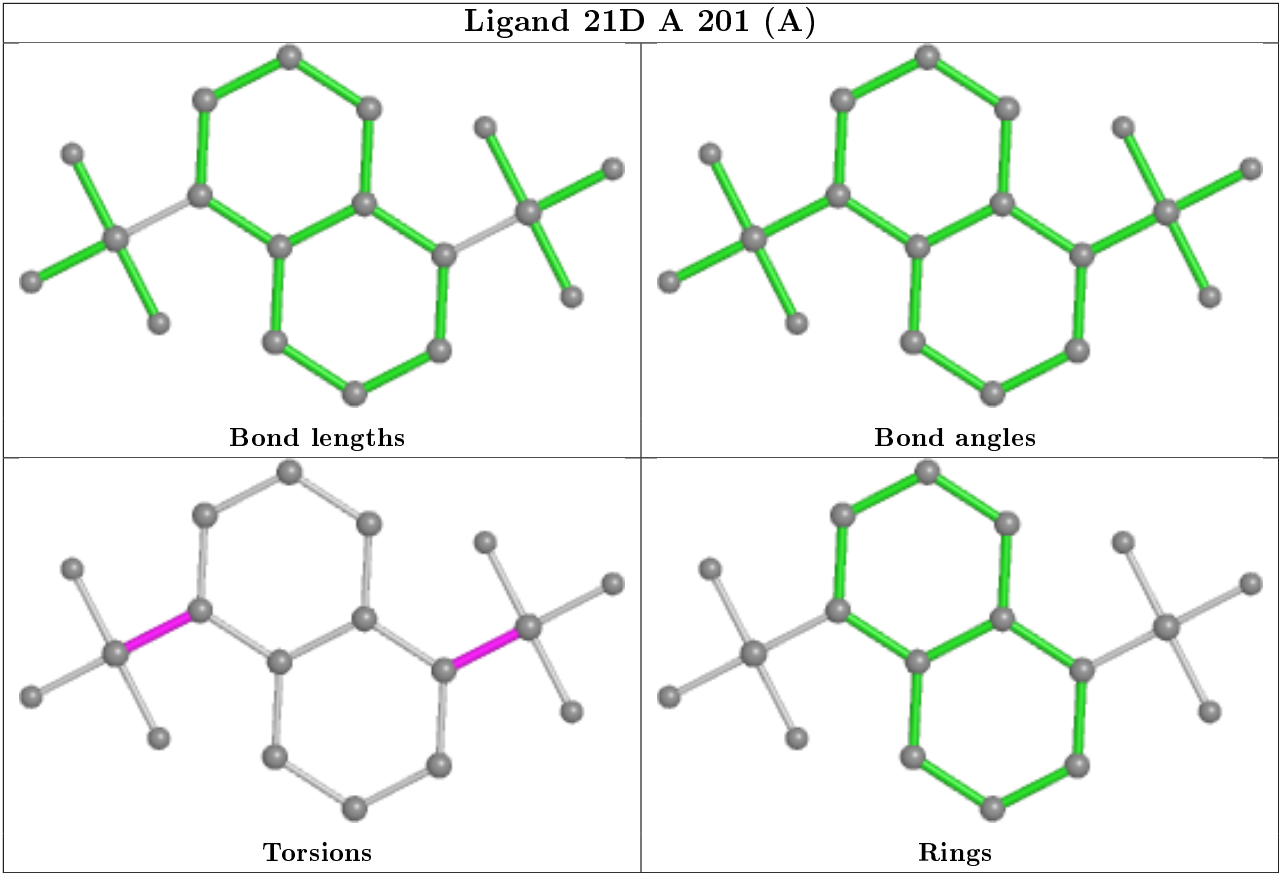


Ligand 21D B 201 (B)



Ligand 21D B 202 (B)





5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	138:GLN	C	139:GLN	N	13.79

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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

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

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

6.3 Carbohydrates [i](#)

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

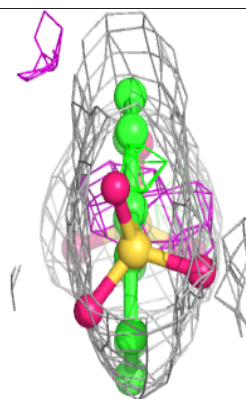
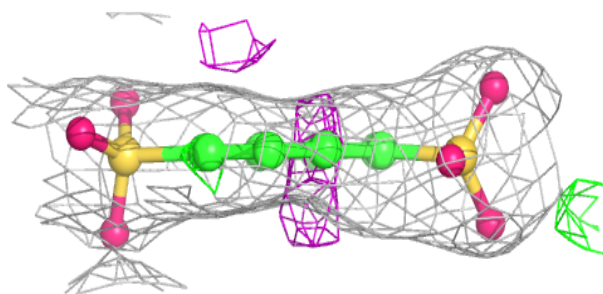
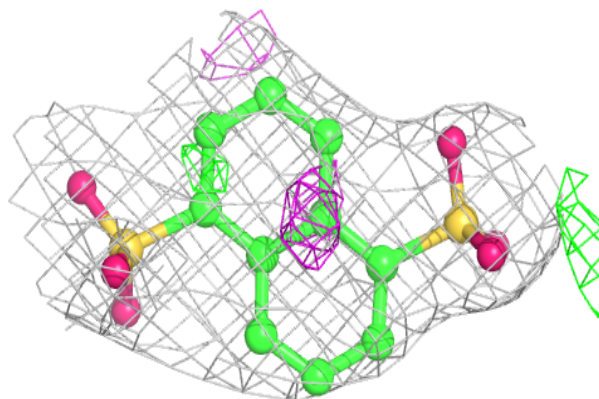
6.4 Ligands [i](#)

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.

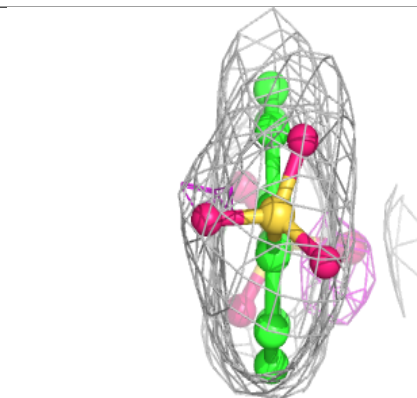
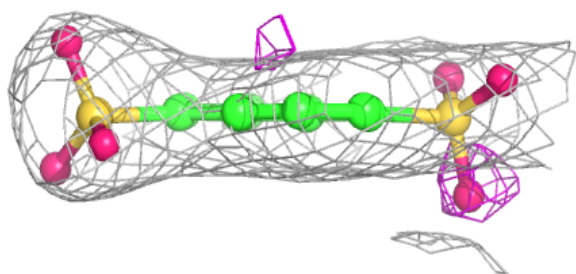
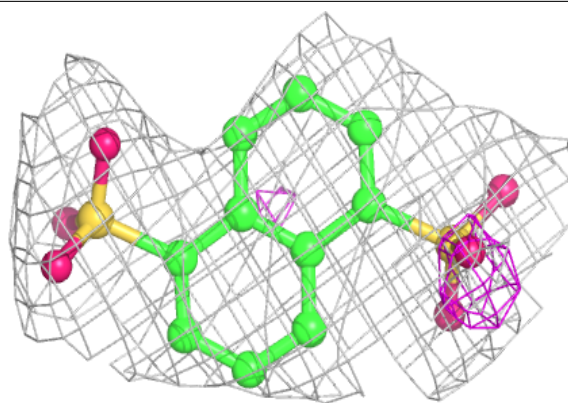
Electron density around 21D B 201 (A):

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

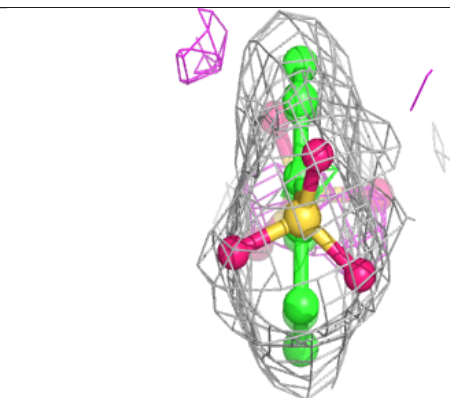
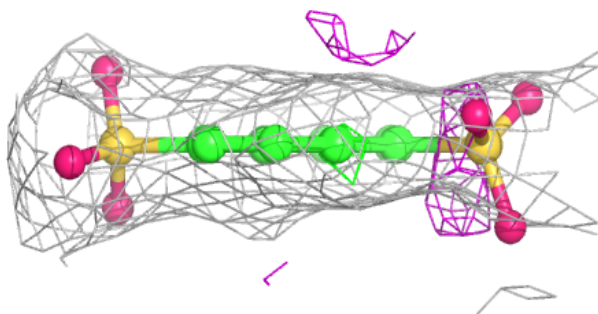
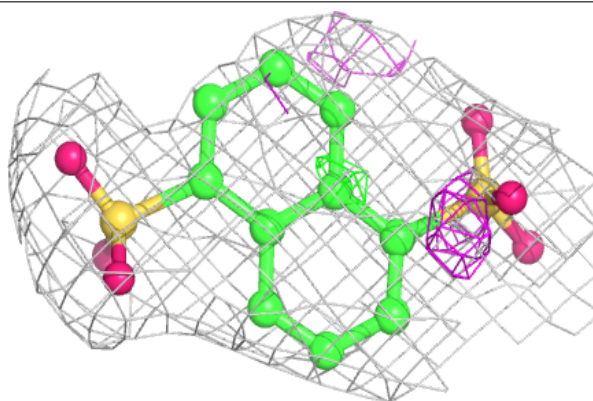


Electron density around 21D A 201 (B):

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

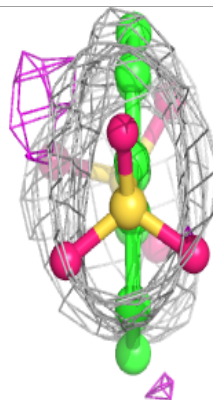
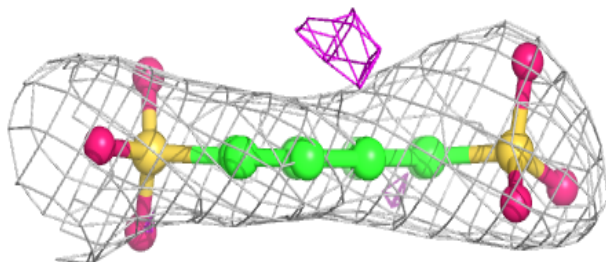
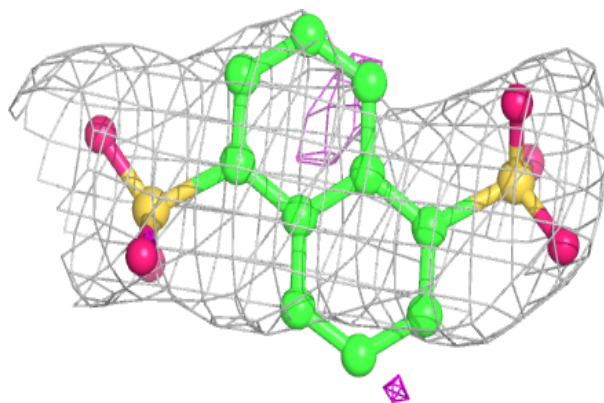
**Electron density around 21D B 201 (B):**

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

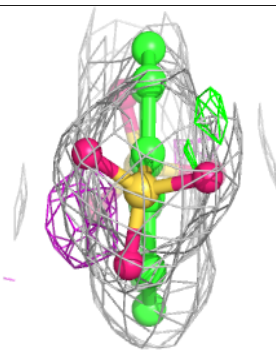
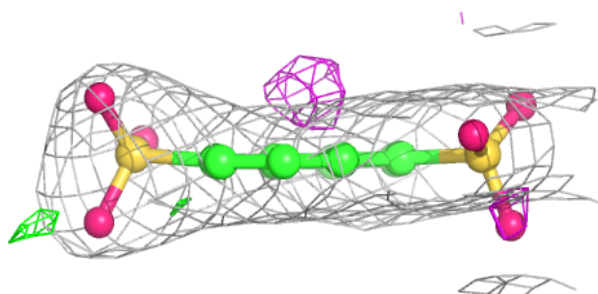
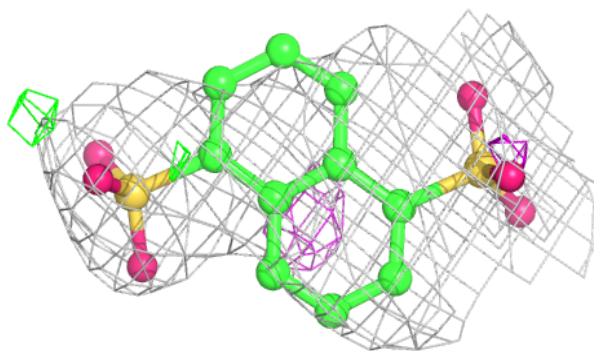


Electron density around 21D B 202 (B):

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

**Electron density around 21D A 201 (A):**

$2mF_o-DF_c$ (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.