Full wwPDB X-ray Structure Validation Report (i)

Oct 4, 2023 – 09:32 AM EDT

PDB ID	:	6OX0
Title	:	SETD3 in Complex with an Actin Peptide with Sinefungin Replacing SAH as Cofactor
Authors	:	Horton, J.R.; Dai, S.; Cheng, X.
Deposited on	:	2019-05-13
Resolution	:	1.75 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:

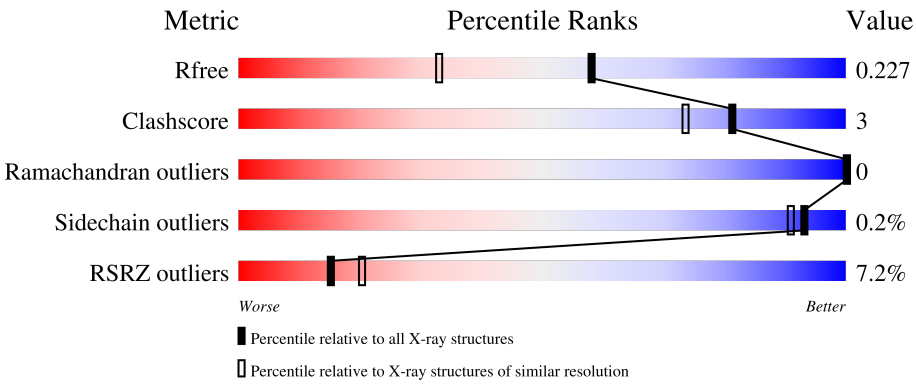
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 1.75 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	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (1.76-1.76)

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	Y	15	20% 93% 7%
1	Z	15	27% 93% 7%
2	A	599	5% 75% 6% 20%
2	B	599	6% 77% • 19%

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

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

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	EDO	A	1012	-	-	X	X
4	EDO	B	1007	-	-	X	-

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 8858 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 Actin Peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	Y	15	120	79	19	22	0	0	0
1	Z	15	120	79	19	22	0	0	0

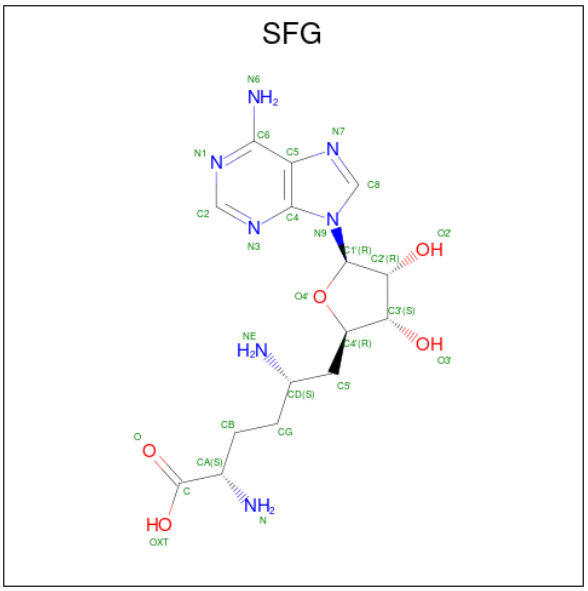
- Molecule 2 is a protein called Histone-lysine N-methyltransferase setd3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	481	3918	2513	655	730	20	0	13	0
2	B	483	3885	2488	650	729	18	0	7	0

There are 10 discrepancies between the modelled and reference sequences:

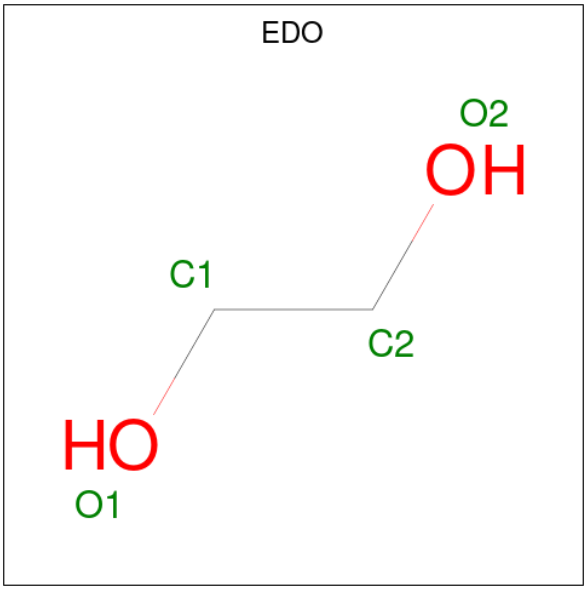
Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	GLY	-	expression tag	UNP Q86TU7
A	-4	PRO	-	expression tag	UNP Q86TU7
A	-3	LEU	-	expression tag	UNP Q86TU7
A	-2	GLY	-	expression tag	UNP Q86TU7
A	-1	SER	-	expression tag	UNP Q86TU7
B	-5	GLY	-	expression tag	UNP Q86TU7
B	-4	PRO	-	expression tag	UNP Q86TU7
B	-3	LEU	-	expression tag	UNP Q86TU7
B	-2	GLY	-	expression tag	UNP Q86TU7
B	-1	SER	-	expression tag	UNP Q86TU7

- Molecule 3 is SINEFUNGIN (three-letter code: SFG) (formula: $C_{15}H_{23}N_7O_5$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	27	15	7	5	0	0
3	B	1	27	15	7	5	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O			
4	A	1	4	2	2		0	0
4	A	1	Total	C	O		0	0
			4	2	2			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O			
4	A	1	4	2	2		0	0
			Total	C	O			
4	A	1	4	2	2		0	0
			Total	C	O			
4	A	1	4	2	2		0	0
			Total	C	O			
4	A	1	4	2	2		0	0
			Total	C	O			
4	A	1	4	2	2		0	0
			Total	C	O			
4	A	1	4	2	2		0	0
			Total	C	O			
4	A	1	4	2	2		0	0
			Total	C	O			
4	A	1	4	2	2		0	0
			Total	C	O			
4	A	1	4	2	2		0	0
			Total	C	O			
4	A	1	4	2	2		0	0
			Total	C	O			
4	A	1	4	2	2		0	0
			Total	C	O			
4	A	1	4	2	2		0	0
			Total	C	O			
4	A	1	4	2	2		0	0
			Total	C	O			
4	A	1	4	2	2		0	0
			Total	C	O			
4	A	1	4	2	2		0	0
			Total	C	O			
4	A	1	4	2	2		0	0
			Total	C	O			
4	B	1	4	2	2		0	0
			Total	C	O			
4	B	1	4	2	2		0	0
			Total	C	O			
4	B	1	4	2	2		0	0
			Total	C	O			
4	B	1	4	2	2		0	0
			Total	C	O			
4	B	1	4	2	2		0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
4	B	1	4	2	2	0	0
			Total	C	O		
4	B	1	4	2	2	0	0
			Total	C	O		
4	B	1	4	2	2	0	0
			Total	C	O		
4	B	1	4	2	2	0	0
			Total	C	O		
4	B	1	4	2	2	0	0
			Total	C	O		
4	B	1	4	2	2	0	0

- Molecule 5 is water.

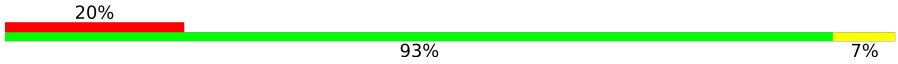
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	Y	11	11	11	0	0
			Total	O		
5	A	340	340	340	0	0
			Total	O		
5	Z	10	10	10	0	0
			Total	O		
5	B	284	284	284	0	0

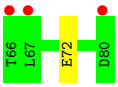
3 Residue-property plots (i)

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: Actin Peptide

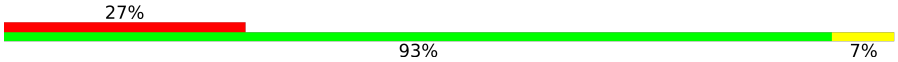
Chain Y:

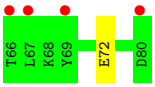




• Molecule 1: Actin Peptide

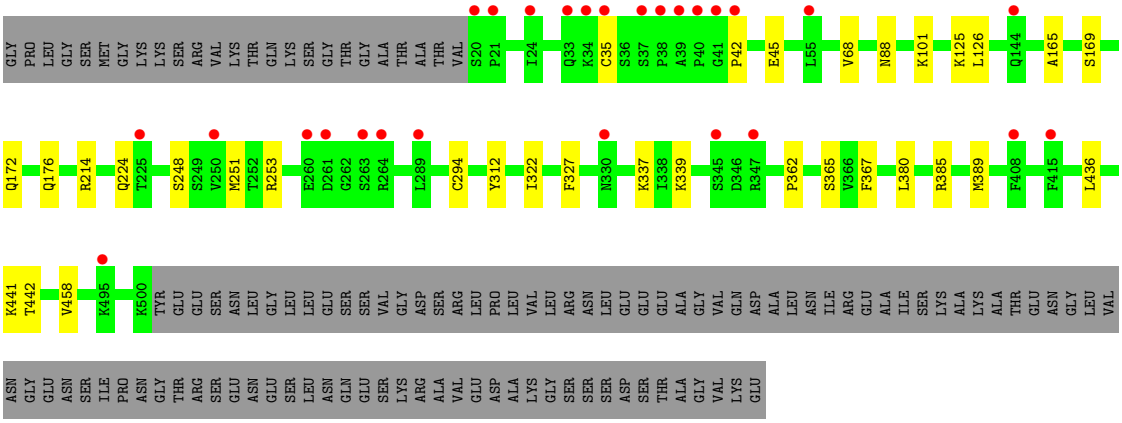
Chain Z:





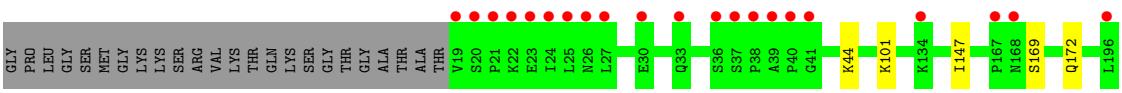
• Molecule 2: Histone-lysine N-methyltransferase setd3

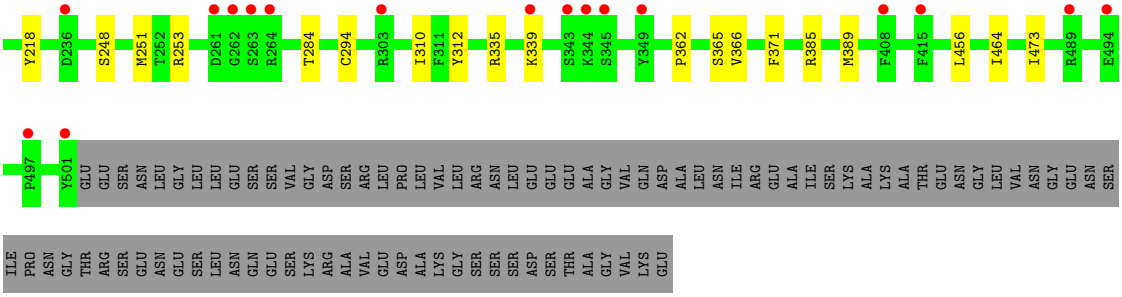
Chain A:



• Molecule 2: Histone-lysine N-methyltransferase setd3

Chain B:




Y218
D236
S248
M251
T252
R253
D261
G262
S263
R264
T284
C294
R303
I310
F311
Y312
R335
K339
S343
K344
S345
Y349
P362
S365
V366
F371
R385
M389
F408
F415
L456
I464
I473
R489
E494
P497
Y501
GLU GLU SER ASN LEU GLY LEU LEU GLU SER SER VAL GLY ASP SER ARG LEU PRO LEU VAL LEU ARG ASN LEU GLU GLU GLU ALA GLY VAL GLN ASP ALA LEU ASN ILE ARG GLU ALA ILE SER LYS ALA LYS ALA THR GLU ASN GLY LEU VAL ASN GLY GLU ASN SER
ILE PRO ASN GLY THR ARG SER GLU ASN GLU SER LEU ASN GLN GLU SER LYS ARG ALA VAL GLU ASP ALA LYS GLY SER SER SER ASP SER THR ALA GLY VAL LYS GLU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	60.27Å 175.19Å 66.03Å 90.00° 92.80° 90.00°	Depositor
Resolution (Å)	36.49 – 1.75 36.49 – 1.75	Depositor EDS
% Data completeness (in resolution range)	97.1 (36.49-1.75) 97.1 (36.49-1.75)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.68 (at 1.76Å)	Xtriage
Refinement program	PHENIX (1.15.2_3472)	Depositor
R, R_{free}	0.206 , 0.225 0.208 , 0.227	Depositor DCC
R_{free} test set	1991 reflections (1.51%)	wwPDB-VP
Wilson B-factor (Å^2)	23.7	Xtriage
Anisotropy	0.764	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 42.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.034 for h,-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	8858	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.24% 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: SFG, EDO

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	Y	0.21	0/124	0.36	0/171
1	Z	0.21	0/124	0.36	0/171
2	A	0.24	0/4037	0.40	0/5466
2	B	0.24	0/3983	0.40	0/5403
All	All	0.24	0/8268	0.40	0/11211

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 (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	Y	120	0	104	1	0
1	Z	120	0	104	1	0
2	A	3918	0	3852	23	0
2	B	3885	0	3765	19	0
3	A	27	0	22	1	0
3	B	27	0	21	1	0
4	A	72	0	108	9	0
4	B	44	0	66	5	0
5	A	340	0	0	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	284	0	0	0	0
5	Y	11	0	0	0	0
5	Z	10	0	0	0	0
All	All	8858	0	8042	42	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 3.

All (42) 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)
2:A:101[B]:LYS:HE2	2:B:101:LYS:HE2	1.72	0.71
2:B:284:THR:H	4:B:1007:EDO:H12	1.60	0.67
2:A:176:GLN:NE2	5:A:1101:HOH:O	2.29	0.65
4:A:1012:EDO:H21	5:A:1416:HOH:O	2.01	0.59
2:B:339:LYS:HG2	2:B:366:VAL:HG22	1.83	0.59
2:B:253:ARG:HG2	3:B:1001:SFG:HNE2	1.68	0.57
2:B:335:ARG:NH2	4:B:1008:EDO:O1	2.40	0.54
2:B:362:PRO:HG2	2:B:365:SER:HB2	1.90	0.54
2:A:101[A]:LYS:HE3	2:B:101:LYS:HG3	1.90	0.54
2:A:458:VAL:HG11	4:A:1012:EDO:H11	1.91	0.52
2:B:284:THR:H	4:B:1007:EDO:C1	2.22	0.51
2:A:169:SER:HB3	2:A:172:GLN:HG2	1.92	0.50
2:A:322:ILE:HD12	4:A:1003:EDO:H11	1.93	0.50
2:A:362:PRO:HG2	2:A:365:SER:HB2	1.93	0.49
2:A:441:LYS:HG3	2:A:442:THR:HG23	1.96	0.48
2:A:68:VAL:HG21	2:A:165:ALA:HA	1.96	0.48
2:A:253:ARG:HG2	3:A:1001:SFG:HNE2	1.79	0.47
2:A:88:ASN:HB3	2:A:125:LYS:HD2	1.96	0.46
2:B:169:SER:HB3	2:B:172:GLN:HG2	1.96	0.46
2:B:310:ILE:HB	4:B:1007:EDO:H22	1.98	0.46
2:A:327:PHE:HB3	4:A:1009:EDO:H11	1.97	0.46
2:B:248:SER:HA	2:B:251:MET:HG2	1.99	0.44
2:A:42:PRO:HA	2:A:45:GLU:HG3	1.98	0.44
4:A:1012:EDO:O2	2:B:101:LYS:NZ	2.47	0.44
4:A:1012:EDO:C2	5:A:1416:HOH:O	2.64	0.44
2:B:44:LYS:HA	2:B:44:LYS:HE2	2.00	0.44
2:B:385:ARG:O	2:B:389[B]:MET:HG3	2.18	0.44
2:A:248:SER:HA	2:A:251:MET:HG2	2.00	0.44
2:A:436:LEU:HD21	4:A:1018:EDO:H12	1.99	0.44
1:Y:72:GLU:HG3	2:A:312:TYR:O	2.18	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:337:LYS:HE3	4:A:1004:EDO:H21	2.00	0.42
1:Z:72:GLU:HG3	2:B:312:TYR:O	2.19	0.42
2:A:224:GLN:NE2	5:A:1123:HOH:O	2.48	0.42
2:A:339:LYS:HB3	4:A:1018:EDO:H22	2.01	0.41
2:B:310:ILE:HB	4:B:1007:EDO:H11	2.01	0.41
2:B:147:ILE:HD12	2:B:218:TYR:CE2	2.56	0.41
2:A:367:PHE:CG	2:A:380:LEU:HD13	2.56	0.41
2:B:371:PHE:HB2	2:B:473:ILE:HG12	2.01	0.41
2:B:456:LEU:HD13	2:B:464:ILE:HD12	2.02	0.41
2:A:385:ARG:O	2:A:389[A]:MET:HG3	2.20	0.40
2:A:35:CYS:HB3	2:A:214:ARG:HB2	2.03	0.40
2:A:88:ASN:HB2	2:A:126:LEU:HD21	2.03	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	Y	13/15 (87%)	13 (100%)	0	0	100	100
1	Z	13/15 (87%)	13 (100%)	0	0	100	100
2	A	492/599 (82%)	483 (98%)	9 (2%)	0	100	100
2	B	488/599 (82%)	480 (98%)	8 (2%)	0	100	100
All	All	1006/1228 (82%)	989 (98%)	17 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	Y	12/14 (86%)	12 (100%)	0	100	100
1	Z	12/14 (86%)	12 (100%)	0	100	100
2	A	423/521 (81%)	421 (100%)	2 (0%)	88	83
2	B	414/521 (80%)	412 (100%)	2 (0%)	88	83
All	All	861/1070 (80%)	857 (100%)	4 (0%)	93	83

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

Mol	Chain	Res	Type
2	A	294[A]	CYS
2	A	294[B]	CYS
2	B	294[A]	CYS
2	B	294[B]	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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)

31 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	$\#\|Z\| > 2$
4	EDO	B	1008	-	3,3,3	0.47	0	2,2,2	0.28	0
4	EDO	B	1007	-	3,3,3	0.38	0	2,2,2	0.33	0
4	EDO	A	1010	-	3,3,3	0.46	0	2,2,2	0.34	0
4	EDO	A	1011	-	3,3,3	0.45	0	2,2,2	0.35	0
4	EDO	A	1007	-	3,3,3	0.46	0	2,2,2	0.34	0
4	EDO	B	1010	-	3,3,3	0.45	0	2,2,2	0.35	0
4	EDO	A	1008	-	3,3,3	0.47	0	2,2,2	0.27	0
4	EDO	B	1004	-	3,3,3	0.46	0	2,2,2	0.31	0
4	EDO	A	1018	-	3,3,3	0.46	0	2,2,2	0.35	0
4	EDO	B	1012	-	3,3,3	0.45	0	2,2,2	0.34	0
4	EDO	B	1006	-	3,3,3	0.45	0	2,2,2	0.34	0
4	EDO	B	1009	-	3,3,3	0.46	0	2,2,2	0.34	0
4	EDO	A	1014	-	3,3,3	0.45	0	2,2,2	0.36	0
4	EDO	A	1019	-	3,3,3	0.47	0	2,2,2	0.30	0
4	EDO	B	1003	-	3,3,3	0.45	0	2,2,2	0.34	0
4	EDO	A	1015	-	3,3,3	0.45	0	2,2,2	0.33	0
3	SFG	B	1001	-	25,29,29	3.23	2 (8%)	23,42,42	1.25	2 (8%)
4	EDO	A	1012	-	3,3,3	0.46	0	2,2,2	0.31	0
3	SFG	A	1001	-	25,29,29	3.21	2 (8%)	23,42,42	1.21	2 (8%)
4	EDO	B	1011	-	3,3,3	0.45	0	2,2,2	0.35	0
4	EDO	A	1003	-	3,3,3	0.46	0	2,2,2	0.32	0
4	EDO	A	1004	-	3,3,3	0.47	0	2,2,2	0.28	0
4	EDO	B	1002	-	3,3,3	0.45	0	2,2,2	0.34	0
4	EDO	B	1005	-	3,3,3	0.45	0	2,2,2	0.34	0
4	EDO	A	1017	-	3,3,3	0.46	0	2,2,2	0.33	0
4	EDO	A	1016	-	3,3,3	0.46	0	2,2,2	0.35	0
4	EDO	A	1006	-	3,3,3	0.44	0	2,2,2	0.40	0
4	EDO	A	1009	-	3,3,3	0.46	0	2,2,2	0.30	0
4	EDO	A	1013	-	3,3,3	0.45	0	2,2,2	0.34	0
4	EDO	A	1002	-	3,3,3	0.46	0	2,2,2	0.32	0
4	EDO	A	1005	-	3,3,3	0.47	0	2,2,2	0.32	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
4	EDO	B	1008	-	-	0/1/1/1	-
4	EDO	B	1007	-	-	1/1/1/1	-
4	EDO	A	1010	-	-	0/1/1/1	-
4	EDO	A	1011	-	-	0/1/1/1	-
4	EDO	A	1007	-	-	0/1/1/1	-
4	EDO	B	1010	-	-	0/1/1/1	-
4	EDO	A	1008	-	-	0/1/1/1	-
4	EDO	B	1004	-	-	0/1/1/1	-
4	EDO	A	1018	-	-	1/1/1/1	-
4	EDO	B	1012	-	-	0/1/1/1	-
4	EDO	B	1006	-	-	0/1/1/1	-
4	EDO	B	1009	-	-	0/1/1/1	-
4	EDO	A	1014	-	-	0/1/1/1	-
4	EDO	A	1019	-	-	0/1/1/1	-
4	EDO	B	1003	-	-	0/1/1/1	-
4	EDO	A	1015	-	-	0/1/1/1	-
3	SFG	B	1001	-	-	5/13/33/33	0/3/3/3
4	EDO	A	1012	-	-	0/1/1/1	-
3	SFG	A	1001	-	-	4/13/33/33	0/3/3/3
4	EDO	B	1011	-	-	0/1/1/1	-
4	EDO	A	1003	-	-	0/1/1/1	-
4	EDO	A	1004	-	-	0/1/1/1	-
4	EDO	B	1002	-	-	0/1/1/1	-
4	EDO	B	1005	-	-	0/1/1/1	-
4	EDO	A	1017	-	-	0/1/1/1	-
4	EDO	A	1016	-	-	0/1/1/1	-
4	EDO	A	1006	-	-	0/1/1/1	-
4	EDO	A	1009	-	-	0/1/1/1	-
4	EDO	A	1013	-	-	0/1/1/1	-
4	EDO	A	1002	-	-	0/1/1/1	-
4	EDO	A	1005	-	-	0/1/1/1	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1001	SFG	C2'-C1'	-15.49	1.30	1.53
3	A	1001	SFG	C2'-C1'	-15.42	1.30	1.53
3	A	1001	SFG	O-C	2.04	1.28	1.22
3	B	1001	SFG	O-C	2.02	1.28	1.22

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	1001	SFG	C3'-C2'-C1'	3.79	106.69	100.98
3	A	1001	SFG	C3'-C2'-C1'	3.45	106.17	100.98
3	B	1001	SFG	O4'-C1'-C2'	2.52	110.61	106.93
3	A	1001	SFG	O4'-C1'-C2'	2.50	110.58	106.93

There are no chirality outliers.

All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1001	SFG	C5'-CD-CG-CB
3	B	1001	SFG	C5'-CD-CG-CB
4	B	1007	EDO	O1-C1-C2-O2
3	A	1001	SFG	NE-CD-CG-CB
3	B	1001	SFG	NE-CD-CG-CB
4	A	1018	EDO	O1-C1-C2-O2
3	B	1001	SFG	C-CA-CB-CG
3	A	1001	SFG	C4'-C5'-CD-NE
3	B	1001	SFG	C4'-C5'-CD-NE
3	A	1001	SFG	OXT-C-CA-CB
3	B	1001	SFG	OXT-C-CA-CB

There are no ring outliers.

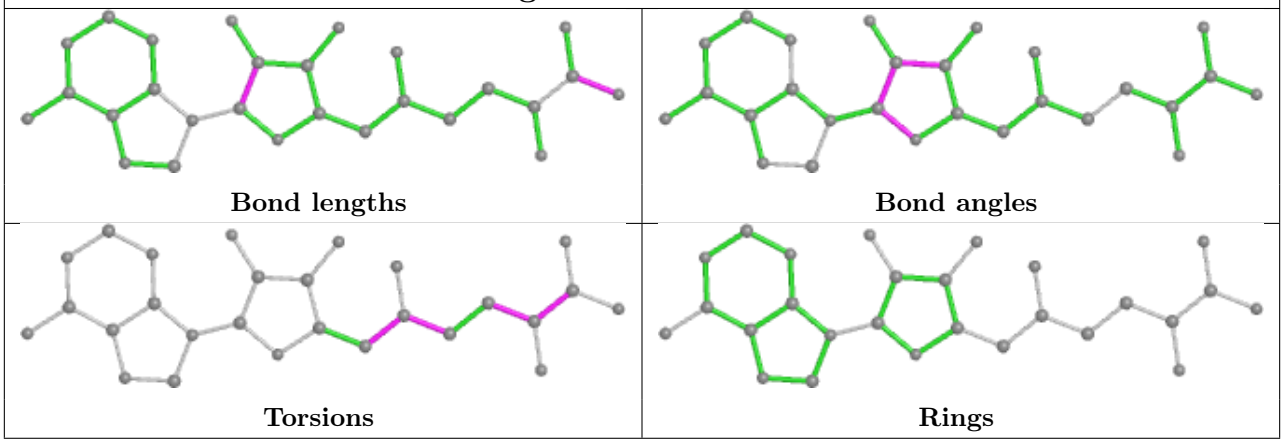
9 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	1008	EDO	1	0
4	B	1007	EDO	4	0
4	A	1018	EDO	2	0
3	B	1001	SFG	1	0
4	A	1012	EDO	4	0
3	A	1001	SFG	1	0
4	A	1003	EDO	1	0
4	A	1004	EDO	1	0
4	A	1009	EDO	1	0

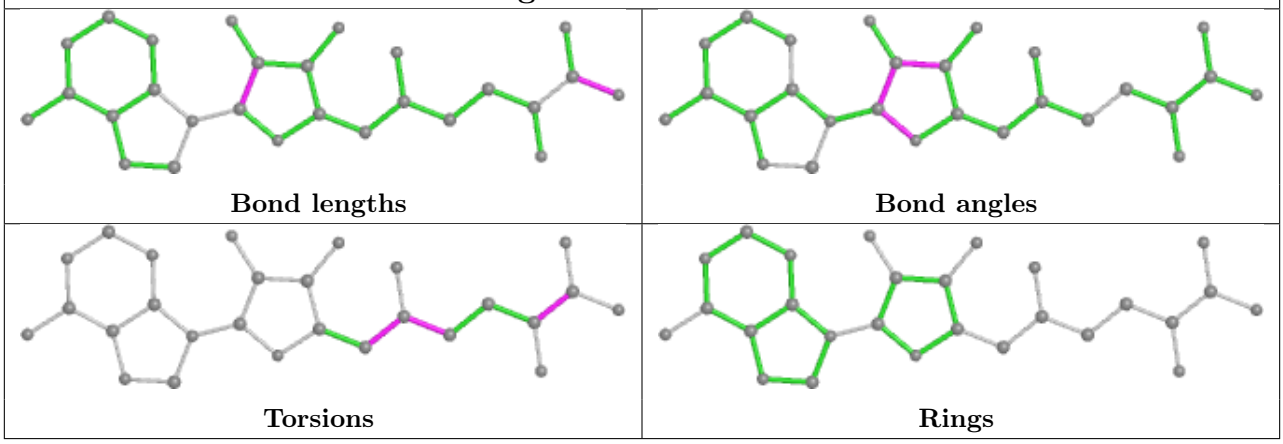
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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 SFG B 1001



Ligand SFG A 1001



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 (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	Y	15/15 (100%)	0.75	3 (20%)	1	1	19, 34, 70, 72	0
1	Z	15/15 (100%)	0.99	4 (26%)	0	0	28, 41, 69, 79	0
2	A	481/599 (80%)	0.43	27 (5%)	24	30	18, 31, 56, 88	0
2	B	483/599 (80%)	0.52	38 (7%)	12	16	19, 36, 63, 77	0
All	All	994/1228 (80%)	0.49	72 (7%)	15	20	18, 34, 62, 88	0

All (72) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	19	VAL	7.5
2	A	39	ALA	7.0
2	A	20	SER	6.3
2	B	21	PRO	5.9
2	A	42	PRO	5.9
2	B	261	ASP	5.7
1	Z	66	THR	5.5
2	B	39	ALA	5.4
2	B	38	PRO	4.9
2	A	408	PHE	4.6
2	A	264	ARG	4.6
2	B	20	SER	4.6
2	A	21	PRO	4.4
1	Y	66	THR	4.4
2	A	261	ASP	4.4
2	B	408	PHE	4.2
2	B	24	ILE	4.2
2	B	37	SER	4.2
2	A	38	PRO	4.0
2	B	501	TYR	4.0
2	A	40	PRO	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	A	330	ASN	3.9
2	B	41	GLY	3.6
2	B	344	LYS	3.5
2	B	33	GLN	3.4
2	A	347	ARG	3.3
2	A	263	SER	3.3
2	B	349	TYR	3.2
2	B	40	PRO	3.2
2	A	41	GLY	3.2
2	A	37	SER	3.1
2	B	415	PHE	3.0
2	B	262	GLY	2.9
2	A	35	CYS	2.8
2	A	33	GLN	2.8
2	B	23	GLU	2.8
1	Y	67	LEU	2.7
2	B	263	SER	2.7
2	B	345	SER	2.7
2	A	34	LYS	2.7
2	A	260	GLU	2.6
1	Z	80	ASP	2.6
2	B	36	SER	2.6
1	Z	69	TYR	2.6
2	B	236	ASP	2.5
2	A	24	ILE	2.5
1	Z	67	LEU	2.5
2	B	167	PRO	2.4
2	B	494	GLU	2.4
2	B	26	ASN	2.3
2	A	415	PHE	2.3
2	A	289	LEU	2.3
1	Y	80	ASP	2.2
2	A	495	LYS	2.2
2	B	30	GLU	2.2
2	A	250	VAL	2.2
2	B	134	LYS	2.2
2	A	55	LEU	2.2
2	B	339	LYS	2.2
2	A	345	SER	2.1
2	B	303	ARG	2.1
2	A	225	THR	2.1
2	B	264	ARG	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	489	ARG	2.1
2	B	25	LEU	2.1
2	B	196	LEU	2.1
2	B	168	ASN	2.0
2	A	144	GLN	2.0
2	B	497	PRO	2.0
2	B	22	LYS	2.0
2	B	343	SER	2.0
2	B	27	LEU	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.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	EDO	B	1005	4/4	0.33	0.20	76,76,76,76	0
4	EDO	A	1019	4/4	0.49	0.28	49,50,50,50	0
4	EDO	B	1012	4/4	0.49	0.24	54,54,54,54	0
4	EDO	A	1011	4/4	0.53	0.30	52,53,53,53	0
4	EDO	A	1004	4/4	0.57	0.28	52,52,52,52	0
4	EDO	A	1012	4/4	0.59	0.42	71,71,71,71	0
4	EDO	B	1006	4/4	0.60	0.27	93,93,93,93	0
4	EDO	A	1003	4/4	0.60	0.27	60,60,60,60	0
4	EDO	A	1010	4/4	0.68	0.35	49,49,49,49	0
4	EDO	A	1009	4/4	0.69	0.23	63,63,63,63	0
4	EDO	A	1013	4/4	0.70	0.25	45,45,45,45	0
4	EDO	A	1018	4/4	0.70	0.29	57,57,57,57	0
4	EDO	B	1011	4/4	0.71	0.23	54,54,54,55	0
4	EDO	A	1014	4/4	0.73	0.23	64,64,64,64	0

Continued on next page...

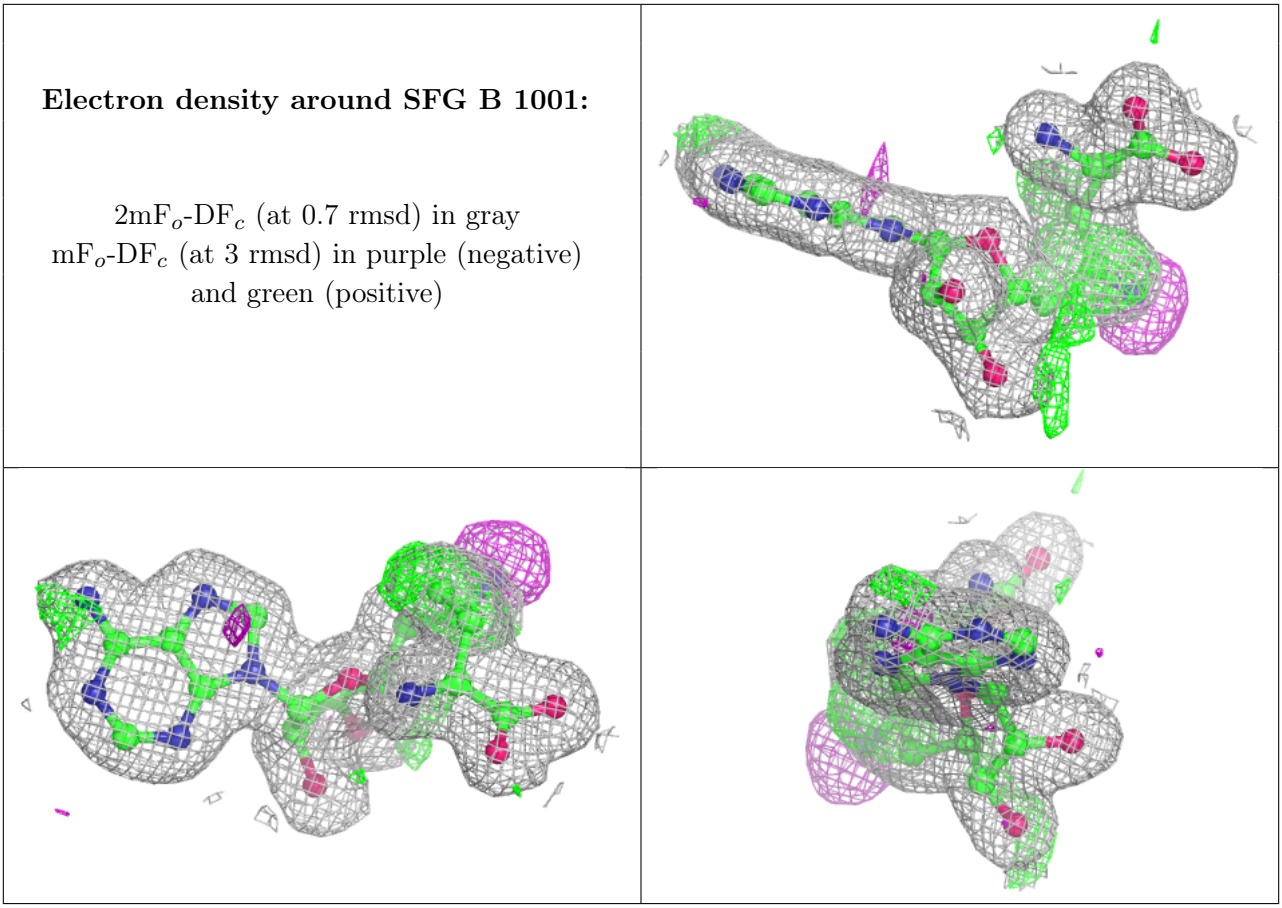
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	EDO	B	1010	4/4	0.74	0.25	53,53,53,53	0
4	EDO	B	1008	4/4	0.77	0.39	54,54,54,54	0
4	EDO	A	1015	4/4	0.78	0.21	63,63,63,63	0
4	EDO	A	1005	4/4	0.79	0.29	60,60,60,60	0
4	EDO	A	1017	4/4	0.80	0.20	47,47,47,47	0
4	EDO	A	1008	4/4	0.81	0.17	43,43,43,43	0
4	EDO	A	1006	4/4	0.82	0.22	45,45,45,45	0
4	EDO	B	1002	4/4	0.82	0.16	62,62,62,62	0
4	EDO	B	1003	4/4	0.83	0.32	68,68,68,68	0
4	EDO	A	1002	4/4	0.84	0.43	49,49,49,49	0
4	EDO	B	1009	4/4	0.85	0.28	52,52,52,52	0
3	SFG	B	1001	27/27	0.86	0.13	18,19,21,22	0
3	SFG	A	1001	27/27	0.87	0.13	16,18,21,21	0
4	EDO	B	1007	4/4	0.91	0.37	30,30,30,30	0
4	EDO	B	1004	4/4	0.91	0.20	47,47,47,47	0
4	EDO	A	1007	4/4	0.93	0.14	41,41,42,42	0
4	EDO	A	1016	4/4	0.95	0.08	58,58,58,58	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.

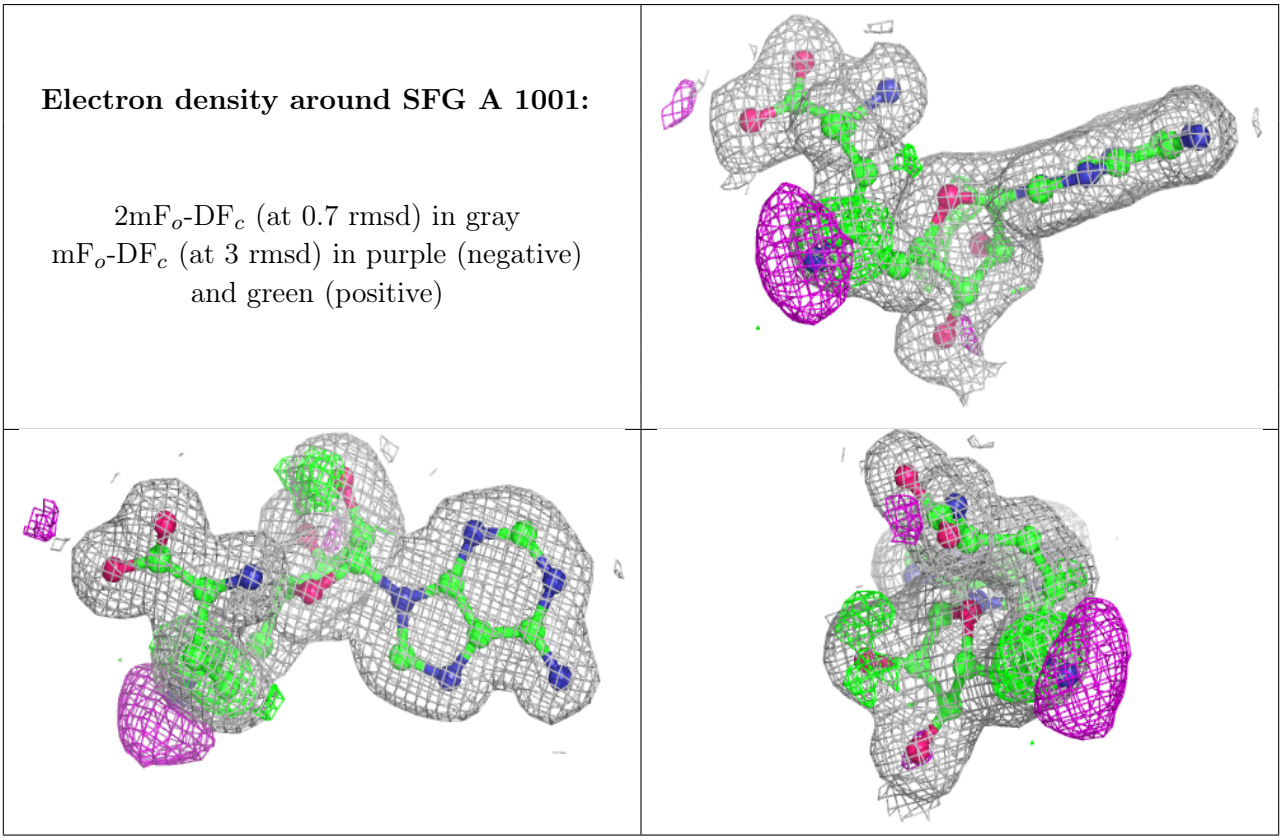
Electron density around SFG B 1001:

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



Electron density around SFG A 1001:

$2mF_o$-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.