wwPDB X-ray Structure Validation Summary Report (i)

Oct 31, 2023 – 10:30 AM JST

PDB ID	:	5BUN
Title	:	Crystal structure of an antigenic outer membrane protein ST50 from Salmonella Typhi
Authors	:	Yoshimura, M.; Chuankhayan, P.; Lin, C.C.; Chen, N.C.; Yang, M.C.; Fun, H.K.
Deposited on	:	2015-06-04
Resolution	:	2.98 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 (i)) were used in the production of this report:

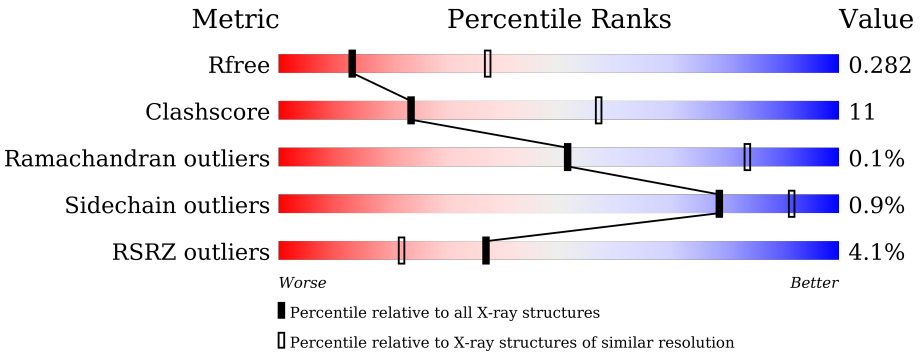
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 2.98 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	2754 (3.00-2.96)
Clashscore	141614	3103 (3.00-2.96)
Ramachandran outliers	138981	2993 (3.00-2.96)
Sidechain outliers	138945	2996 (3.00-2.96)
RSRZ outliers	127900	2644 (3.00-2.96)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=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	467	4% 71% 19% • 9%
1	B	467	4% 72% 19% 9%
1	C	467	2% 69% 22% • 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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	BOG	B	501	-	-	-	X

2 Entry composition (i)

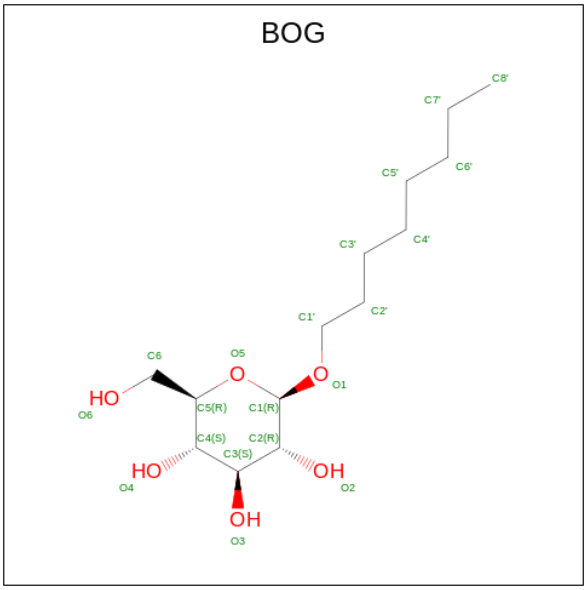
There are 3 unique types of molecules in this entry. The entry contains 10051 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Outer membrane protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	427	3313	2047	583	678	5	0	0	0
			Total	C	N	O	S			
1	B	425	3300	2039	581	675	5	0	0	0
			Total	C	N	O	S			
1	C	433	3356	2073	589	689	5	0	0	0

- Molecule 2 is octyl beta-D-glucopyranoside (three-letter code: BOG) (formula: $C_{14}H_{28}O_6$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	20	14	6	0	0
			Total	C	O		
2	A	1	20	14	6	0	0
			Total	C	O		
2	B	1	20	14	6	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	C	1	20	14	6	0	0

- Molecule 3 is water.

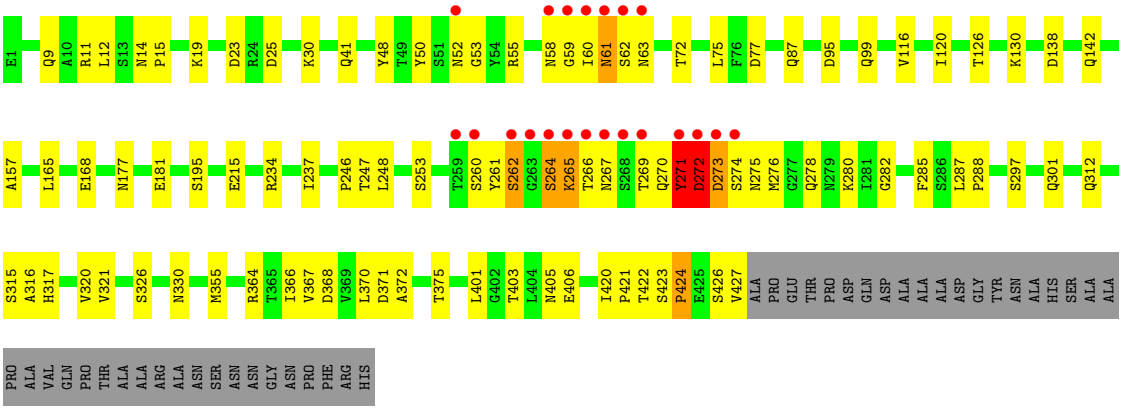
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	1	1	1	0	0
			Total	O		
3	C	1	1	1	0	0

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 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.

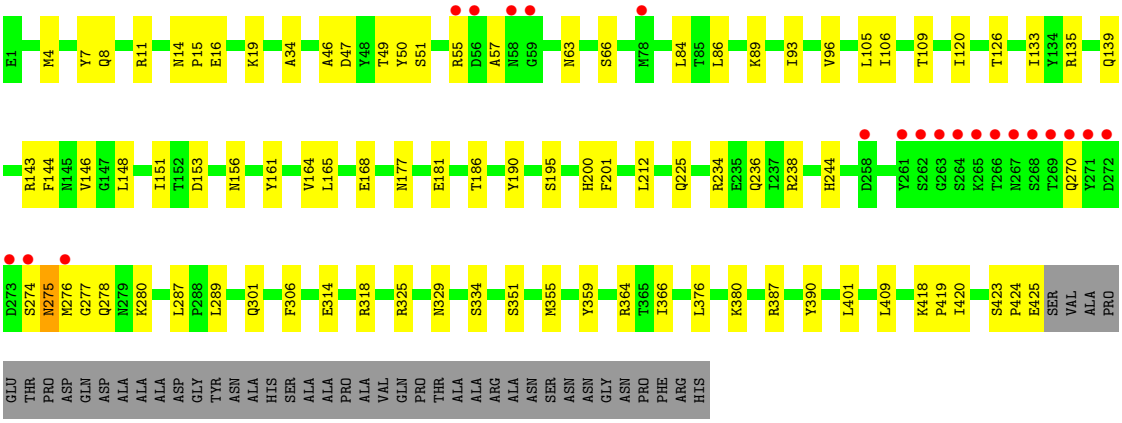
• Molecule 1: Outer membrane protein

Chain A:



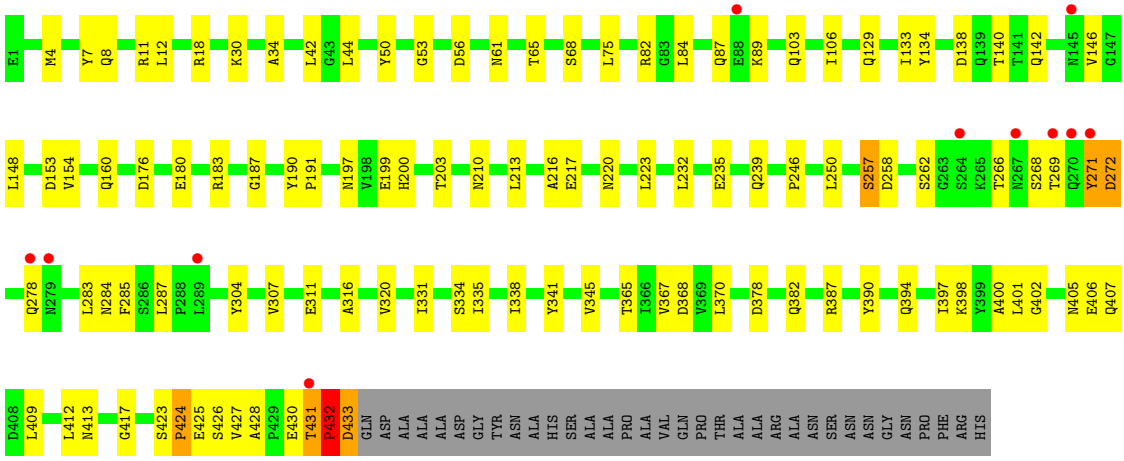
• Molecule 1: Outer membrane protein

Chain B:



• Molecule 1: Outer membrane protein

Chain C:


E1 M4 Y7 Q8 R11 L12 R18 K30 A34 L42 G43 L44 Y50 G53 D56 N61 T65 S68 L75 R82 G83 L84 Q87 E88 K89 Q103 I106 Q129 I133 Y134 D138 Q139 T140 T141 Q142 N145 V146 G147
L148 D153 V154 Q160 D176 E180 R183 G187 Y190 P191 N197 V198 E199 H200 T203 N210 L213 A216 E217 N220 L223 L232 E235 Q239 P246 L250 S257 D258 S262 G263 S264 K265 T266 N267 S268 T269 Q270 Y271 D272
Q278 N279 L283 N284 F285 S286 L287 P288 L289 Y304 V307 E311 A316 V320 I331 S334 I335 I338 Y341 V345 T365 I366 V367 D368 V369 L370 D378 Q382 R387 Y390 Q394 I397 K398 Y399 A400 L401 G402 N405 E406 Q407
D408 L409 L412 N413 G417 S423 P424 E425 S426 V427 A428 P429 E430 T431 P432 D433 GLN ASP ALA ALA ALA ASP GLY TYR ASN ALA HIS SER ALA ALA PRO ALA VAL GLN PRO THR ALA ALA ARG ALA ASN SER ASN ASN GLY ASN PRO PHE ARG HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 4	Depositor
Cell constants a, b, c, α, β, γ	246.85Å 246.85Å 65.68Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.93 – 2.98 29.93 – 2.98	Depositor EDS
% Data completeness (in resolution range)	85.2 (29.93-2.98) 85.2 (29.93-2.98)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	0.11	Depositor
$< I/\sigma(I) >$ [1]	2.24 (at 3.00Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.210 , 0.275 0.217 , 0.282	Depositor DCC
R_{free} test set	1756 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	56.1	Xtriage
Anisotropy	0.022	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 19.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.029 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	10051	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.35% 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: BOG

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	0.30	1/3356 (0.0%)	0.63	7/4559 (0.2%)
1	B	0.25	0/3343	0.48	0/4541
1	C	0.27	0/3401	0.61	7/4623 (0.2%)
All	All	0.27	1/10100 (0.0%)	0.57	14/13723 (0.1%)

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	3
1	C	0	3
All	All	0	6

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	421	PRO	N-CD	5.10	1.54	1.47

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	272	ASP	N-CA-C	12.14	143.79	111.00
1	C	432	PRO	C-N-CA	11.57	150.63	121.70
1	A	271	TYR	N-CA-C	10.44	139.18	111.00
1	C	428	ALA	N-CA-C	8.32	133.46	111.00
1	C	432	PRO	N-CA-C	8.27	133.59	112.10

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	264	SER	Peptide
1	A	265	LYS	Peptide
1	A	424	PRO	Peptide
1	C	423	SER	Peptide
1	C	68	SER	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	3313	0	3262	97	0
1	B	3300	0	3248	57	0
1	C	3356	0	3297	90	0
2	A	40	0	56	5	0
2	B	20	0	28	1	0
2	C	20	0	28	1	0
3	A	1	0	0	0	0
3	C	1	0	0	0	0
All	All	10051	0	9919	223	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 11.

The worst 5 of 223 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:266:THR:HB	1:C:271:TYR:CD2	1.51	1.44
1:A:272:ASP:CG	1:A:273:ASP:H	1.07	1.33
1:A:265:LYS:CD	1:A:271:TYR:OH	1.84	1.24
1:C:271:TYR:C	1:C:271:TYR:HD1	1.41	1.23
1:C:266:THR:CB	1:C:271:TYR:HD2	1.54	1.19

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	425/467 (91%)	392 (92%)	32 (8%)	1 (0%)	47	80
1	B	423/467 (91%)	393 (93%)	30 (7%)	0	100	100
1	C	431/467 (92%)	402 (93%)	29 (7%)	0	100	100
All	All	1279/1401 (91%)	1187 (93%)	91 (7%)	1 (0%)	51	83

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	272	ASP

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	362/389 (93%)	359 (99%)	3 (1%)	81	92
1	B	360/389 (92%)	358 (99%)	2 (1%)	86	94
1	C	367/389 (94%)	362 (99%)	5 (1%)	67	86
All	All	1089/1167 (93%)	1079 (99%)	10 (1%)	78	91

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	427	VAL
1	C	431	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	433	ASP
1	B	270	GLN
1	B	275	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	127	GLN
1	C	197	ASN
1	C	413	ASN
1	C	312	GLN
1	B	278	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 (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

4 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	BOG	C	501	-	20,20,20	1.16	1 (5%)	25,25,25	0.97	1 (4%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	BOG	B	501	-	20,20,20	1.19	1 (5%)	25,25,25	0.82	0
2	BOG	A	1001	-	20,20,20	1.10	2 (10%)	25,25,25	1.71	5 (20%)
2	BOG	A	1002	-	20,20,20	1.17	1 (5%)	25,25,25	0.94	1 (4%)

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	BOG	C	501	-	-	8/11/31/31	0/1/1/1
2	BOG	B	501	-	-	5/11/31/31	0/1/1/1
2	BOG	A	1001	-	-	7/11/31/31	0/1/1/1
2	BOG	A	1002	-	-	9/11/31/31	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	501	BOG	O5-C1	2.91	1.49	1.41
2	C	501	BOG	O5-C1	2.80	1.49	1.41
2	A	1002	BOG	O5-C1	2.77	1.48	1.41
2	A	1001	BOG	O5-C1	2.77	1.48	1.41
2	A	1001	BOG	O2-C2	2.00	1.47	1.43

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1001	BOG	C1-O5-C5	-4.05	105.75	113.69
2	A	1001	BOG	C1-C2-C3	3.45	117.19	110.00
2	A	1001	BOG	O1-C1-C2	3.31	113.48	108.30
2	A	1001	BOG	C4-C3-C2	2.85	115.80	110.82
2	A	1002	BOG	O5-C5-C4	2.41	114.07	109.69

There are no chirality outliers.

5 of 29 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	501	BOG	C2-C1-O1-C1'
2	B	501	BOG	O5-C1-O1-C1'
2	B	501	BOG	C2'-C1'-O1-C1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	C	501	BOG	C2'-C1'-O1-C1
2	C	501	BOG	O5-C5-C6-O6

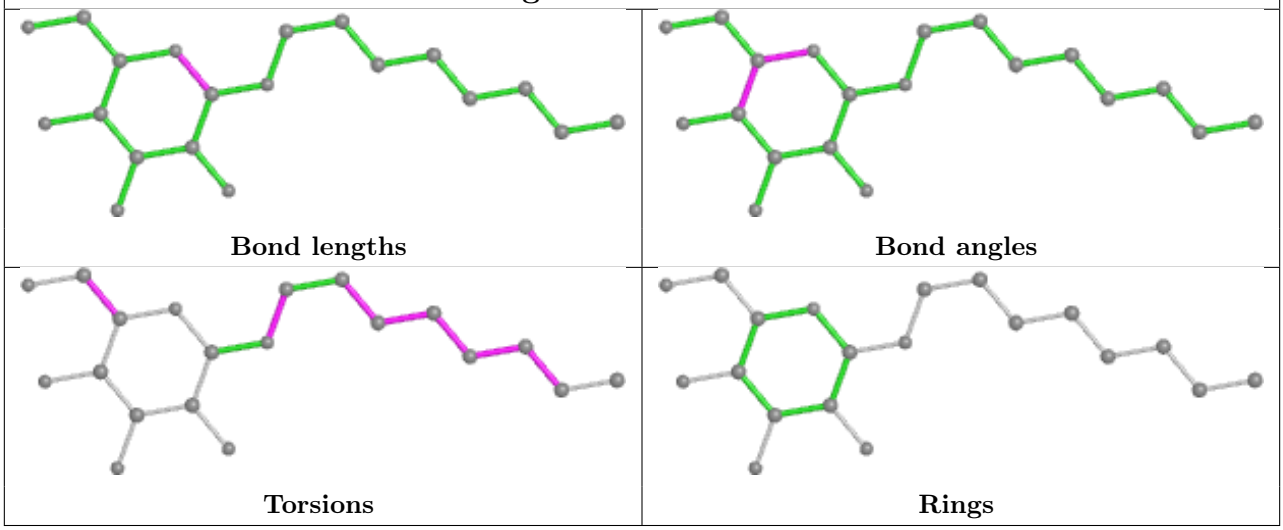
There are no ring outliers.

4 monomers are involved in 7 short contacts:

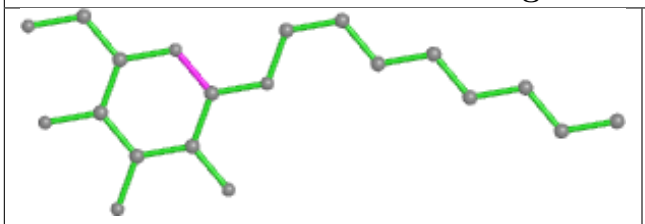
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	501	BOG	1	0
2	B	501	BOG	1	0
2	A	1001	BOG	4	0
2	A	1002	BOG	1	0

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.

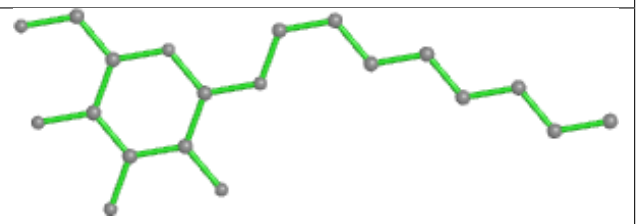
Ligand BOG C 501



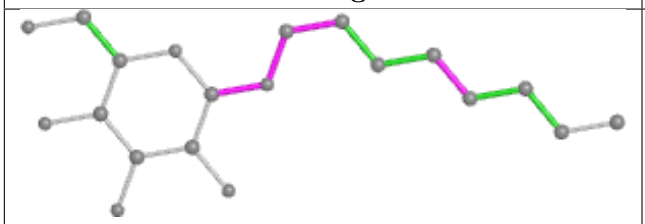
Ligand BOG B 501



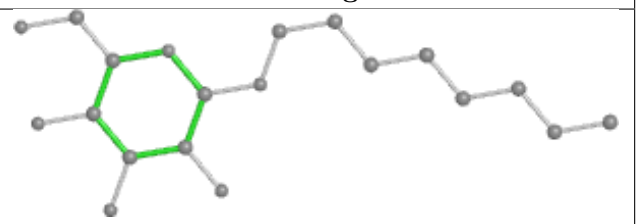
Bond lengths



Bond angles

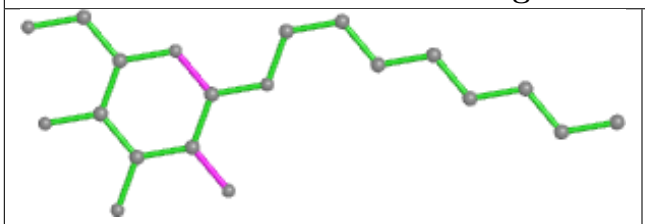


Torsions

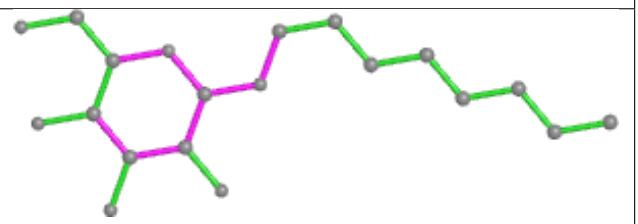


Rings

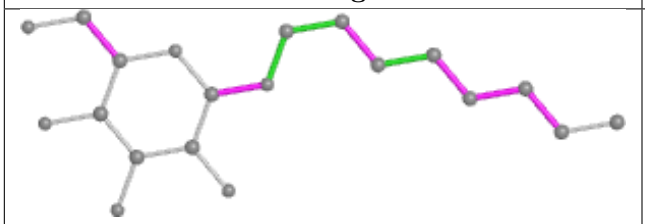
Ligand BOG A 1001



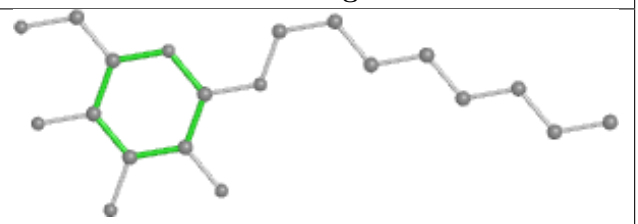
Bond lengths



Bond angles

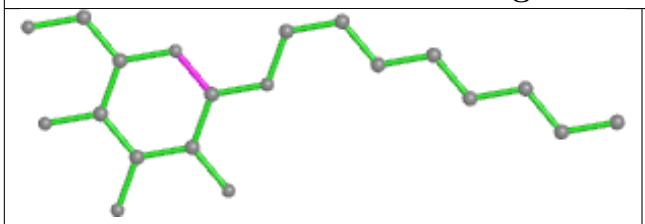


Torsions

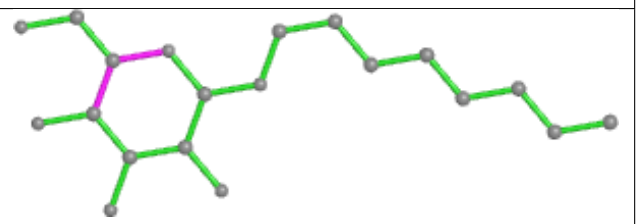


Rings

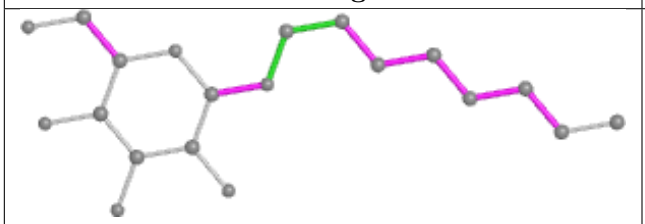
Ligand BOG A 1002



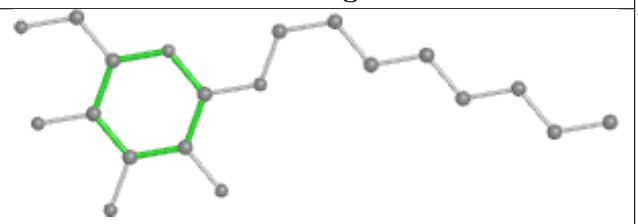
Bond lengths



Bond angles



Torsions



Rings

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	A	427/467 (91%)	0.19	21 (4%)	29	17	22, 36, 90, 136	0
1	B	425/467 (91%)	0.16	21 (4%)	29	17	18, 39, 92, 138	0
1	C	433/467 (92%)	0.03	11 (2%)	57	38	26, 47, 81, 98	0
All	All	1285/1401 (91%)	0.12	53 (4%)	37	22	18, 42, 87, 138	0

The worst 5 of 53 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	271	TYR	13.3
1	B	263	GLY	8.1
1	A	263	GLY	7.8
1	B	272	ASP	7.5
1	B	270	GLN	7.3

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)

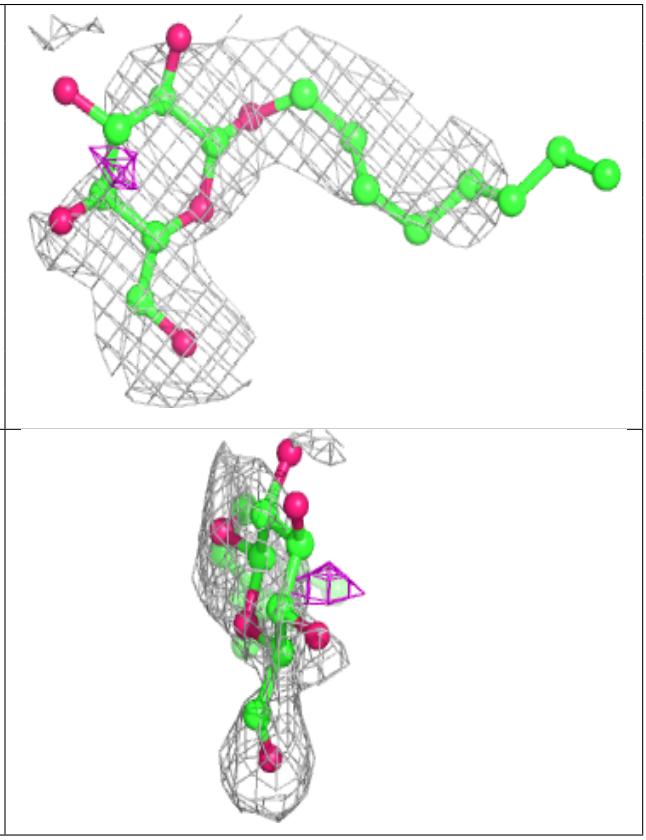
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	BOG	B	501	20/20	0.64	0.45	48,86,98,102	0
2	BOG	C	501	20/20	0.68	0.35	55,83,100,101	0
2	BOG	A	1002	20/20	0.74	0.29	53,74,90,92	0
2	BOG	A	1001	20/20	0.76	0.36	56,65,73,75	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.

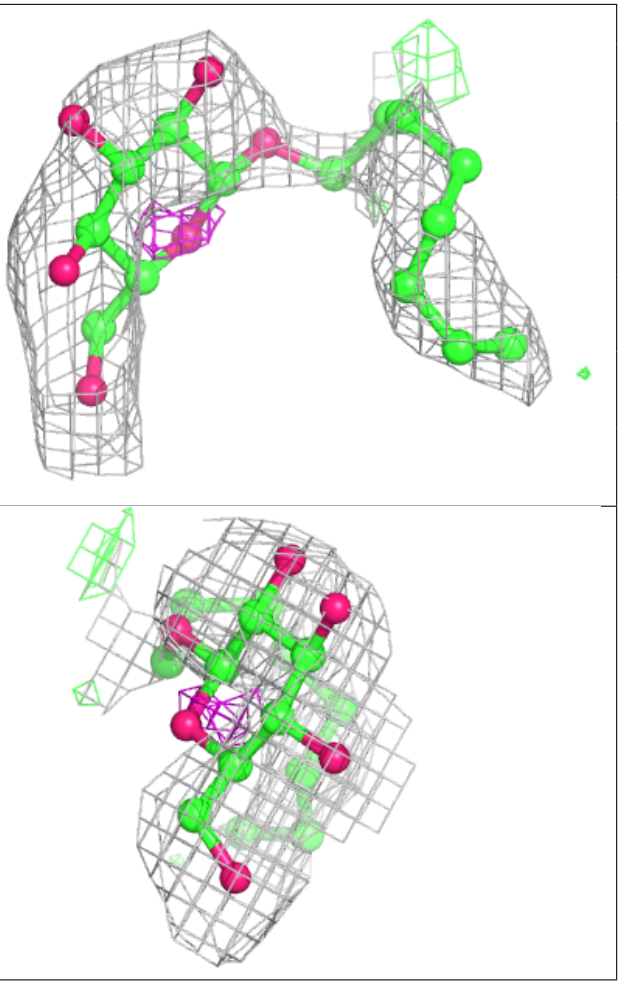
Electron density around BOG B 501:

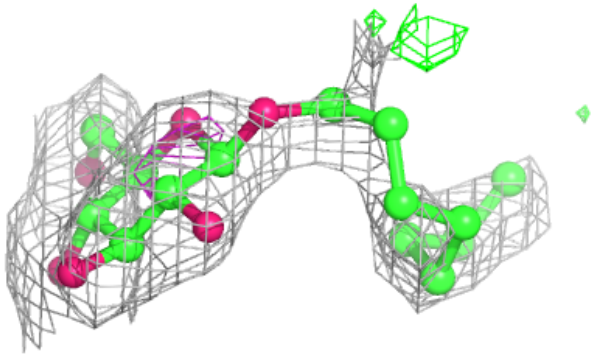
2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative) and green (positive)



Electron density around BOG C 501:

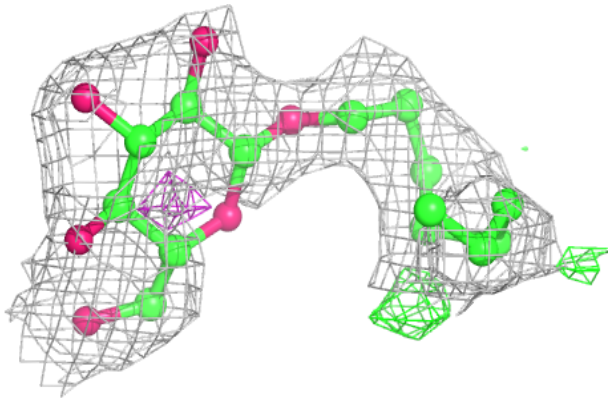
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

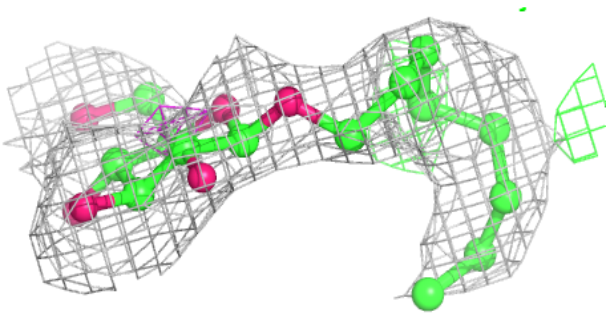


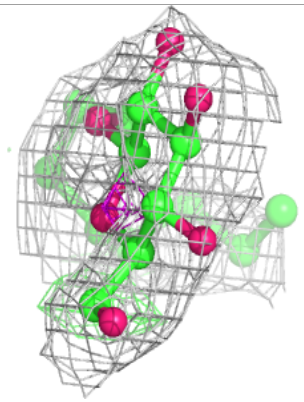


Electron density around BOG A 1002:

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

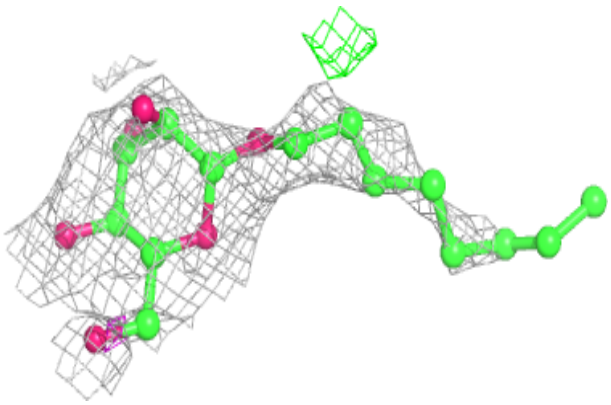


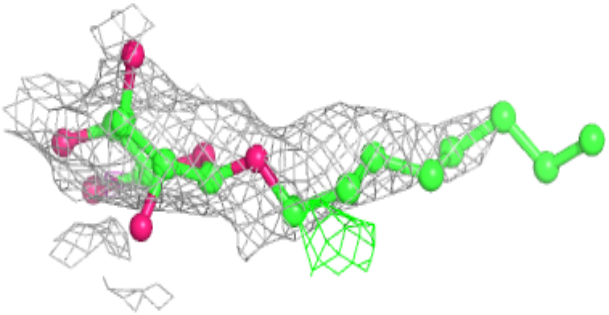


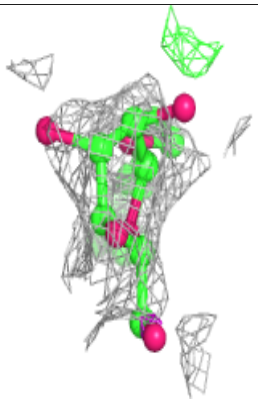


Electron density around BOG A 1001:

$2mF_o$-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.