



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 2, 2021 – 05:48 AM EDT

PDB ID : 3I1L  
Title : Structure of porcine torovirus Hemagglutinin-Esterase in complex with its receptor  
Authors : Zeng, Q.H.; Huizinga, E.G.  
Deposited on : 2009-06-26  
Resolution : 2.79 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](mailto: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 : 2.23.2  
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) : 2.23.2

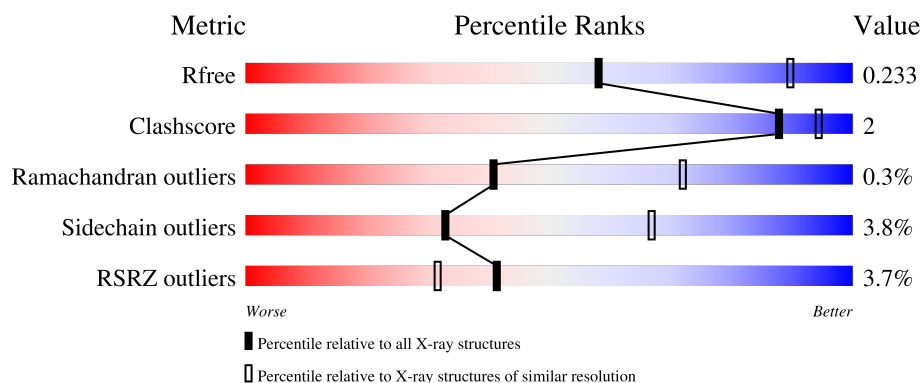
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.79 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	377	<div> <div>6%</div> <div> <div></div> <div>88%</div> <div>8%</div> <div></div> </div> <div>.</div> </div>
1	B	377	<div> <div>2%</div> <div> <div></div> <div>87%</div> <div>8%</div> <div></div> </div> <div>.</div> </div>
1	C	377	<div> <div>3%</div> <div> <div></div> <div>88%</div> <div>8%</div> <div></div> </div> <div>.</div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 8818 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 Hemagglutinin-esterase protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	362	Total	C	N	O	S	0	0	0
			2829	1799	460	542	28			
1	B	362	Total	C	N	O	S	0	0	0
			2829	1799	460	542	28			
1	C	362	Total	C	N	O	S	0	0	0
			2829	1799	460	542	28			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	46	ALA	SER	engineered mutation	UNP Q70KP4
A	394	SER	-	expression tag	UNP Q70KP4
A	395	ASP	-	expression tag	UNP Q70KP4
A	396	PRO	-	expression tag	UNP Q70KP4
A	397	LEU	-	expression tag	UNP Q70KP4
A	398	VAL	-	expression tag	UNP Q70KP4
A	399	PRO	-	expression tag	UNP Q70KP4
A	400	ARG	-	expression tag	UNP Q70KP4
B	46	ALA	SER	engineered mutation	UNP Q70KP4
B	394	SER	-	expression tag	UNP Q70KP4
B	395	ASP	-	expression tag	UNP Q70KP4
B	396	PRO	-	expression tag	UNP Q70KP4
B	397	LEU	-	expression tag	UNP Q70KP4
B	398	VAL	-	expression tag	UNP Q70KP4
B	399	PRO	-	expression tag	UNP Q70KP4
B	400	ARG	-	expression tag	UNP Q70KP4
C	46	ALA	SER	engineered mutation	UNP Q70KP4
C	394	SER	-	expression tag	UNP Q70KP4
C	395	ASP	-	expression tag	UNP Q70KP4
C	396	PRO	-	expression tag	UNP Q70KP4
C	397	LEU	-	expression tag	UNP Q70KP4
C	398	VAL	-	expression tag	UNP Q70KP4
C	399	PRO	-	expression tag	UNP Q70KP4

*Continued on next page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	400	ARG	-	expression tag	UNP Q70KP4

- # SIO

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total 28	C 16	N 1	O 11	0	0
2	B	1	Total 28	C 16	N 1	O 11	0	0
2	C	1	Total 28	C 16	N 1	O 11	0	0

- Molecule 3 is ACETIC ACID (three-letter code: ACY) (formula:  $\text{C}_2\text{H}_4\text{O}_2$ ).



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

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O 14 8 1 5	0	0
4	A	1	Total C N O 14 8 1 5	0	0
4	A	1	Total C N O 14 8 1 5	0	0
4	A	1	Total C N O 14 8 1 5	0	0
4	A	1	Total C N O 14 8 1 5	0	0
4	B	1	Total C N O 14 8 1 5	0	0
4	B	1	Total C N O 14 8 1 5	0	0
4	B	1	Total C N O 14 8 1 5	0	0
4	B	1	Total C N O 14 8 1 5	0	0
4	B	1	Total C N O 14 8 1 5	0	0
4	B	1	Total C N O 14 8 1 5	0	0
4	B	1	Total C N O 14 8 1 5	0	0
4	C	1	Total C N O 14 8 1 5	0	0
4	C	1	Total C N O 14 8 1 5	0	0
4	C	1	Total C N O 14 8 1 5	0	0
4	C	1	Total C N O 14 8 1 5	0	0
4	C	1	Total C N O 14 8 1 5	0	0

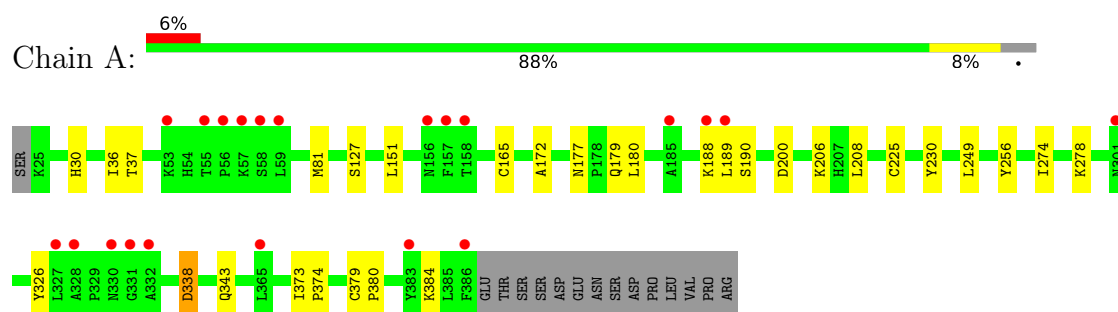
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	3	Total O 3 3	0	0
5	B	3	Total O 3 3	0	0
5	C	5	Total O 5 5	0	0

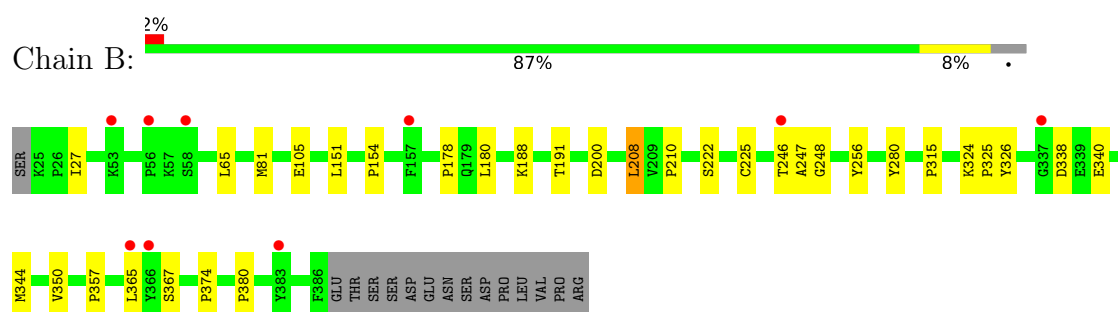
### 3 Residue-property plots [i](#)

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.

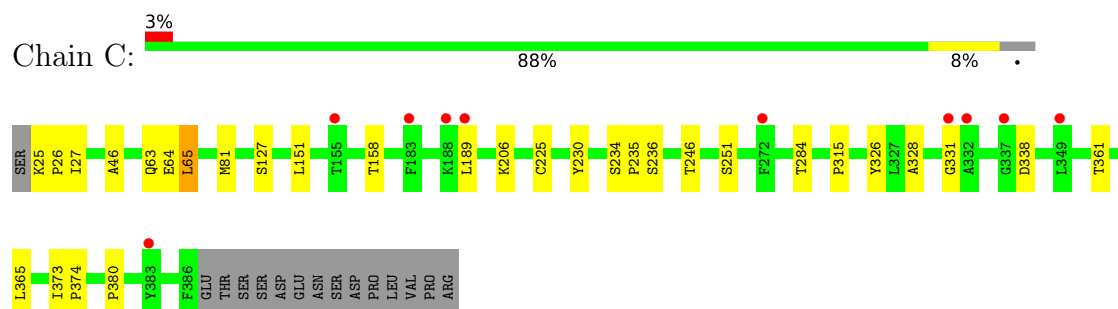
#### • Molecule 1: Hemagglutinin-esterase protein



#### • Molecule 1: Hemagglutinin-esterase protein



#### • Molecule 1: Hemagglutinin-esterase protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	156.03Å 103.69Å 97.07Å 90.00° 96.02° 90.00°	Depositor
Resolution (Å)	46.27 – 2.79 46.28 – 2.79	Depositor EDS
% Data completeness (in resolution range)	93.8 (46.27-2.79) 93.8 (46.28-2.79)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.79 (at 2.77Å)	Xtriage
Refinement program	REFMAC 5.5.0072	Depositor
R, $R_{free}$	0.192 , 0.243 0.190 , 0.233	Depositor DCC
$R_{free}$ test set	1817 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	51.4	Xtriage
Anisotropy	0.070	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 51.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	8818	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.18% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: ACY, SIO, NAG

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	0.42	0/2920	0.54	0/3970
1	B	0.43	0/2920	0.55	0/3970
1	C	0.41	0/2920	0.54	0/3970
All	All	0.42	0/8760	0.54	0/11910

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	A	2829	0	2622	10	0
1	B	2829	0	2621	15	0
1	C	2829	0	2622	11	0
2	A	28	0	24	1	0
2	B	28	0	24	1	0
2	C	28	0	24	0	0
3	A	4	0	3	0	0
3	B	4	0	3	0	0
3	C	4	0	3	1	0
4	A	70	0	65	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	84	0	78	0	0
4	C	70	0	65	0	0
5	A	3	0	0	0	0
5	B	3	0	0	0	0
5	C	5	0	0	0	0
All	All	8818	0	8154	34	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 2.

All (34) 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:27:ILE:HD11	1:C:380:PRO:HG3	1.76	0.68
1:B:208:LEU:HB3	1:B:280:TYR:HB2	1.83	0.60
1:B:367:SER:HB2	1:B:380:PRO:HB2	1.81	0.60
1:B:326:TYR:CZ	1:B:338:ASP:HB2	2.39	0.57
1:B:315:PRO:HG2	1:C:315:PRO:HG2	1.88	0.56
1:C:326:TYR:CZ	1:C:338:ASP:HB2	2.41	0.55
1:B:365:LEU:HD11	1:C:365:LEU:HG	1.90	0.54
1:C:373:ILE:HB	1:C:374:PRO:HA	1.91	0.52
1:A:179:GLN:O	1:A:274:ILE:HA	2.11	0.50
1:B:180:LEU:HD21	2:B:1001:SIO:HM9	1.95	0.49
1:A:180:LEU:HD21	2:A:1001:SIO:HM9	1.94	0.49
1:C:328:ALA:HB1	1:C:331:GLY:HA3	1.96	0.48
1:A:373:ILE:HB	1:A:374:PRO:HA	1.94	0.48
1:B:178:PRO:HD3	1:B:210:PRO:HB2	1.95	0.47
1:B:340:GLU:O	1:B:344:MET:HG3	2.15	0.47
1:C:326:TYR:CE1	1:C:338:ASP:HB2	2.49	0.47
1:B:208:LEU:HD11	1:B:256:TYR:HB3	1.97	0.46
1:A:30:HIS:HE1	1:A:343:GLN:HE21	1.64	0.46
1:A:326:TYR:CZ	1:A:338:ASP:HB2	2.51	0.45
1:B:27:ILE:HD11	1:B:380:PRO:HG2	1.98	0.45
1:B:326:TYR:CE1	1:B:338:ASP:HB2	2.52	0.45
1:A:208:LEU:HD11	1:A:256:TYR:HB3	1.97	0.45
1:B:324:LYS:HB2	1:B:325:PRO:HD2	1.99	0.45
1:A:177:ASN:ND2	1:A:278:LYS:HB2	2.32	0.44
1:C:25:LYS:HA	1:C:26:PRO:HD3	1.87	0.44
1:A:165:CYS:HB2	1:A:172:ALA:O	2.18	0.43
1:A:36:ILE:HG13	1:A:37:THR:HG23	2.01	0.43
1:B:247:ALA:HA	1:B:248:GLY:HA3	1.63	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:357:PRO:HD3	1:B:374:PRO:O	2.20	0.42
1:C:234:SER:C	1:C:236:SER:H	2.24	0.41
1:C:63:GLN:C	1:C:65:LEU:H	2.25	0.40
1:B:154:PRO:HB2	1:B:191:THR:OG1	2.20	0.40
1:A:379:CYS:HA	1:A:380:PRO:HD2	1.85	0.40
1:C:46:ALA:HB3	3:C:2001:ACY:C	2.52	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	360/377 (96%)	343 (95%)	16 (4%)	1 (0%)	41	72
1	B	360/377 (96%)	341 (95%)	19 (5%)	0	100	100
1	C	360/377 (96%)	340 (94%)	18 (5%)	2 (1%)	25	56
All	All	1080/1131 (96%)	1024 (95%)	53 (5%)	3 (0%)	41	72

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	64	GLU
1	C	235	PRO
1	A	338	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	316/331 (96%)	304 (96%)	12 (4%)	33	67
1	B	316/331 (96%)	305 (96%)	11 (4%)	36	70
1	C	316/331 (96%)	303 (96%)	13 (4%)	30	64
All	All	948/993 (96%)	912 (96%)	36 (4%)	33	67

All (36) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	81	MET
1	A	127	SER
1	A	151	LEU
1	A	188	LYS
1	A	189	LEU
1	A	190	SER
1	A	200	ASP
1	A	206	LYS
1	A	225	CYS
1	A	230	TYR
1	A	249	LEU
1	A	384	LYS
1	B	65	LEU
1	B	81	MET
1	B	105	GLU
1	B	151	LEU
1	B	188	LYS
1	B	200	ASP
1	B	208	LEU
1	B	222	SER
1	B	225	CYS
1	B	246	THR
1	B	350	VAL
1	C	65	LEU
1	C	81	MET
1	C	127	SER
1	C	151	LEU
1	C	158	THR
1	C	189	LEU
1	C	206	LYS
1	C	225	CYS
1	C	230	TYR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	246	THR
1	C	251	SER
1	C	284	THR
1	C	361	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	30	HIS
1	A	343	GLN
1	B	303	GLN
1	C	30	HIS

### 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](#)

22 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	SIO	A	1001	-	25,28,28	1.31	2 (8%)	32,40,40	1.70	9 (28%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	SIO	C	1001	-	25,28,28	1.29	2 (8%)	32,40,40	1.67	4 (12%)
3	ACY	A	2001	-	1,3,3	1.03	0	0,3,3	-	-
3	ACY	C	2001	-	1,3,3	1.22	0	0,3,3	-	-
3	ACY	B	2001	-	1,3,3	1.21	0	0,3,3	-	-
4	NAG	B	5154	1	14,14,15	0.48	0	17,19,21	0.95	1 (5%)
4	NAG	B	5157	1	14,14,15	0.54	0	17,19,21	1.05	1 (5%)
4	NAG	C	5158	1	14,14,15	0.56	0	17,19,21	1.42	3 (17%)
4	NAG	C	5155	1	14,14,15	0.78	1 (7%)	17,19,21	1.69	5 (29%)
4	NAG	A	5156	1	14,14,15	0.41	0	17,19,21	1.24	1 (5%)
2	SIO	B	1001	-	25,28,28	1.23	2 (8%)	32,40,40	1.96	9 (28%)
4	NAG	A	5158	1	14,14,15	0.81	0	17,19,21	1.10	1 (5%)
4	NAG	A	5154	1	14,14,15	0.39	0	17,19,21	1.16	2 (11%)
4	NAG	B	5156	1	14,14,15	0.61	0	17,19,21	0.98	1 (5%)
4	NAG	C	5154	1	14,14,15	0.58	0	17,19,21	0.97	0
4	NAG	A	5157	1	14,14,15	0.47	0	17,19,21	1.65	2 (11%)
4	NAG	C	5156	1	14,14,15	0.48	0	17,19,21	2.02	5 (29%)
4	NAG	C	5157	1	14,14,15	0.63	0	17,19,21	1.30	1 (5%)
4	NAG	A	5155	1	14,14,15	0.55	0	17,19,21	0.99	0
4	NAG	B	5158	1	14,14,15	0.61	0	17,19,21	1.19	2 (11%)
4	NAG	B	5159	1	14,14,15	0.60	0	17,19,21	1.05	1 (5%)
4	NAG	B	5155	1	14,14,15	0.92	1 (7%)	17,19,21	1.35	2 (11%)

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	SIO	A	1001	-	-	2/24/48/48	0/1/1/1
2	SIO	C	1001	-	-	10/24/48/48	0/1/1/1
4	NAG	B	5154	1	-	1/6/23/26	0/1/1/1
4	NAG	B	5157	1	-	0/6/23/26	0/1/1/1
4	NAG	C	5158	1	-	2/6/23/26	0/1/1/1
4	NAG	C	5155	1	-	1/6/23/26	0/1/1/1
4	NAG	A	5156	1	-	2/6/23/26	0/1/1/1
2	SIO	B	1001	-	-	6/24/48/48	0/1/1/1
4	NAG	A	5158	1	-	0/6/23/26	0/1/1/1
4	NAG	A	5154	1	-	2/6/23/26	0/1/1/1

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	B	5156	1	-	0/6/23/26	0/1/1/1
4	NAG	C	5154	1	-	2/6/23/26	0/1/1/1
4	NAG	A	5157	1	-	0/6/23/26	0/1/1/1
4	NAG	C	5156	1	-	2/6/23/26	0/1/1/1
4	NAG	C	5157	1	-	2/6/23/26	0/1/1/1
4	NAG	A	5155	1	-	2/6/23/26	0/1/1/1
4	NAG	B	5158	1	-	2/6/23/26	0/1/1/1
4	NAG	B	5159	1	-	2/6/23/26	0/1/1/1
4	NAG	B	5155	1	-	2/6/23/26	0/1/1/1

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1001	SIO	O4-CA4	5.06	1.46	1.35
2	B	1001	SIO	O4-CA4	4.83	1.46	1.35
2	A	1001	SIO	O4-CA4	4.82	1.46	1.35
4	B	5155	NAG	C1-C2	2.70	1.56	1.52
2	A	1001	SIO	O9-CA9	2.46	1.45	1.33
2	C	1001	SIO	O9-CA9	2.42	1.45	1.33
2	B	1001	SIO	O9-CA9	2.37	1.45	1.33
4	C	5155	NAG	C1-C2	2.32	1.55	1.52

All (50) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1001	SIO	O4-CA4-CM4	6.11	122.34	111.09
2	C	1001	SIO	O4-CA4-CM4	5.81	121.78	111.09
4	A	5157	NAG	C1-O5-C5	5.61	119.79	112.19
4	C	5156	NAG	C1-O5-C5	5.39	119.50	112.19
2	A	1001	SIO	O4-CA4-CM4	4.53	119.43	111.09
4	A	5156	NAG	C1-O5-C5	4.04	117.67	112.19
2	B	1001	SIO	O6-C6-C5	-3.81	106.06	109.78
2	C	1001	SIO	C4-C5-N5	-3.54	102.94	110.62
2	B	1001	SIO	O4-CA4-OA4	-3.50	116.00	122.96
4	C	5157	NAG	C1-O5-C5	3.31	116.67	112.19
4	C	5156	NAG	C3-C4-C5	-3.28	104.38	110.24
4	C	5158	NAG	C1-O5-C5	3.27	116.62	112.19
4	C	5155	NAG	C3-C4-C5	-3.09	104.73	110.24
2	B	1001	SIO	C4-C5-N5	-3.04	104.03	110.62
4	B	5155	NAG	C4-C3-C2	3.02	115.44	111.02

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	5158	NAG	C4-C3-C2	3.00	115.41	111.02
4	B	5159	NAG	O5-C5-C6	2.98	111.88	107.20
4	C	5156	NAG	C4-C3-C2	-2.98	106.65	111.02
2	C	1001	SIO	O4-CA4-OA4	-2.97	117.07	122.96
2	A	1001	SIO	C3-C4-C5	2.95	115.08	110.48
4	B	5158	NAG	O5-C5-C6	2.94	111.81	107.20
2	A	1001	SIO	C3-C2-C1	-2.91	109.08	114.12
2	A	1001	SIO	C6-C5-N5	-2.89	106.12	110.91
4	C	5158	NAG	C1-C2-N2	2.86	115.38	110.49
4	B	5154	NAG	C1-O5-C5	2.86	116.07	112.19
2	A	1001	SIO	C2-C3-C4	2.78	116.41	110.58
2	B	1001	SIO	C4-O4-CA4	-2.76	113.14	117.46
4	B	5157	NAG	C1-O5-C5	2.58	115.68	112.19
4	C	5155	NAG	O5-C5-C6	2.57	111.23	107.20
4	C	5155	NAG	C2-N2-C7	2.54	126.52	122.90
2	A	1001	SIO	CM2-O2-C2	-2.53	112.91	115.35
4	B	5158	NAG	C2-N2-C7	2.45	126.39	122.90
4	C	5155	NAG	C1-O5-C5	2.43	115.48	112.19
2	B	1001	SIO	C3-C4-C5	2.37	114.17	110.48
2	B	1001	SIO	O6-C2-C3	-2.34	107.78	111.18
2	A	1001	SIO	O4-CA4-OA4	-2.33	118.33	122.96
4	B	5156	NAG	C1-O5-C5	2.32	115.33	112.19
2	B	1001	SIO	CM2-O2-C2	-2.29	113.15	115.35
2	C	1001	SIO	O6-C2-C3	-2.28	107.87	111.18
4	A	5154	NAG	C2-N2-C7	2.26	126.12	122.90
4	C	5158	NAG	O5-C5-C6	2.20	110.66	107.20
4	A	5154	NAG	C1-O5-C5	2.19	115.16	112.19
4	C	5156	NAG	O4-C4-C5	2.16	114.66	109.30
4	C	5155	NAG	O7-C7-N2	2.11	125.82	121.95
4	A	5157	NAG	C1-C2-N2	2.09	114.06	110.49
2	A	1001	SIO	C2-O6-C6	2.08	118.87	114.19
2	B	1001	SIO	C8-C7-C6	-2.04	109.16	113.03
2	A	1001	SIO	C4-C5-N5	-2.03	106.22	110.62
4	B	5155	NAG	O5-C1-C2	2.01	114.47	111.29
4	C	5156	NAG	O5-C5-C6	2.00	110.35	107.20

There are no chirality outliers.

All (40) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1001	SIO	OA4-CA4-O4-C4
2	B	1001	SIO	C3-C2-O2-CM2

*Continued on next page...*



*Continued from previous page...*

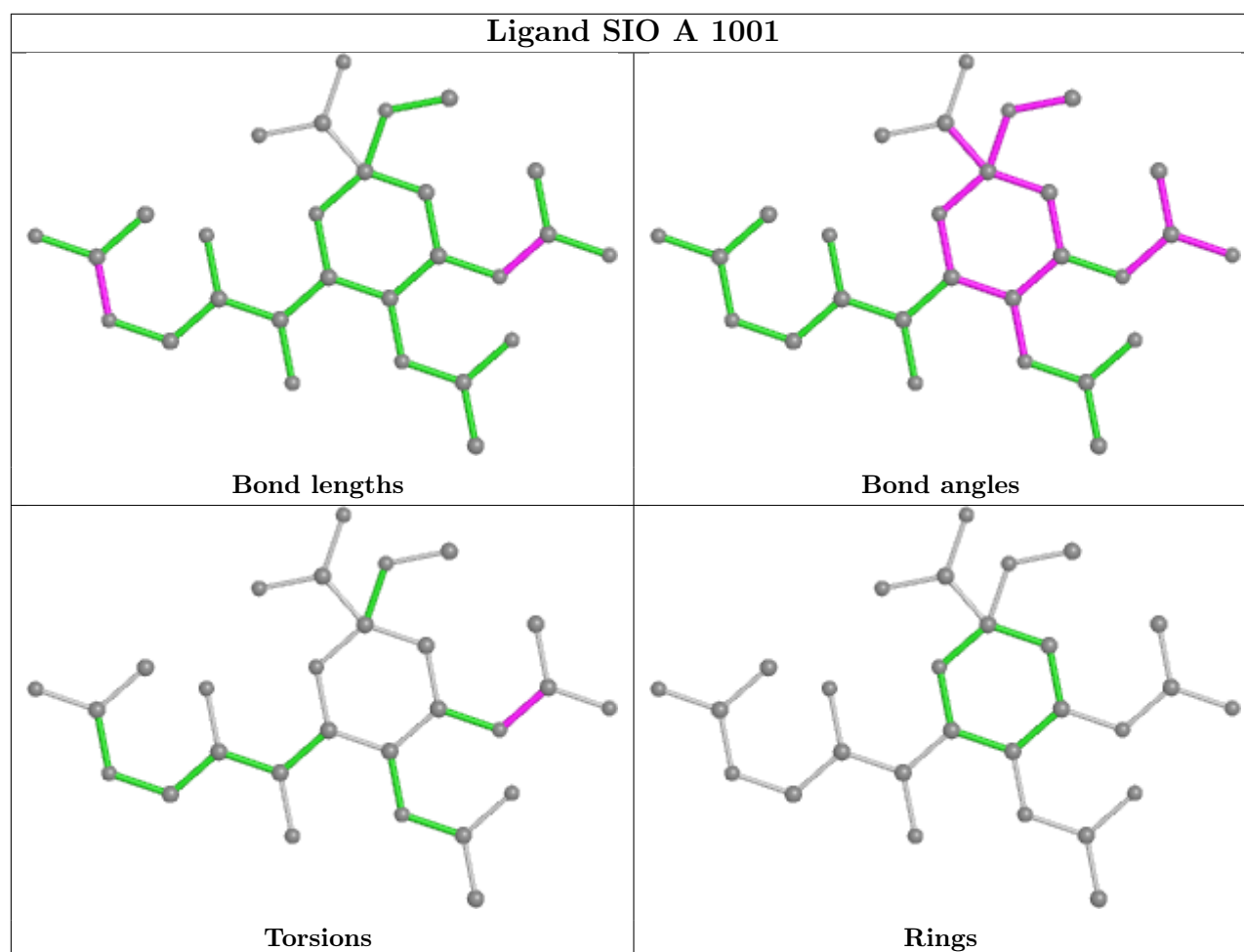
Mol	Chain	Res	Type	Atoms
2	B	1001	SIO	C1-C2-O2-CM2
2	B	1001	SIO	O6-C2-O2-CM2
2	C	1001	SIO	OA4-CA4-O4-C4
2	C	1001	SIO	CM4-CA4-O4-C4
2	C	1001	SIO	C3-C2-O2-CM2
2	C	1001	SIO	C1-C2-O2-CM2
2	C	1001	SIO	O6-C2-O2-CM2
2	A	1001	SIO	CM4-CA4-O4-C4
2	B	1001	SIO	CM4-CA4-O4-C4
2	A	1001	SIO	OA4-CA4-O4-C4
4	C	5154	NAG	C4-C5-C6-O6
4	C	5154	NAG	O5-C5-C6-O6
4	C	5156	NAG	O5-C5-C6-O6
4	A	5154	NAG	O5-C5-C6-O6
4	C	5158	NAG	O5-C5-C6-O6
4	C	5158	NAG	C4-C5-C6-O6
4	B	5158	NAG	O5-C5-C6-O6
4	C	5156	NAG	C4-C5-C6-O6
4	A	5154	NAG	C4-C5-C6-O6
4	C	5157	NAG	O5-C5-C6-O6
4	C	5157	NAG	C4-C5-C6-O6
4	B	5158	NAG	C4-C5-C6-O6
4	B	5159	NAG	C4-C5-C6-O6
4	B	5154	NAG	O5-C5-C6-O6
4	B	5155	NAG	C4-C5-C6-O6
2	C	1001	SIO	C7-C8-C9-O9
4	B	5159	NAG	O5-C5-C6-O6
4	A	5155	NAG	C4-C5-C6-O6
2	C	1001	SIO	CM9-CA9-O9-C9
4	A	5155	NAG	O5-C5-C6-O6
4	B	5155	NAG	O5-C5-C6-O6
2	C	1001	SIO	O8-C8-C9-O9
4	A	5156	NAG	C4-C5-C6-O6
4	A	5156	NAG	O5-C5-C6-O6
2	B	1001	SIO	C5-C4-O4-CA4
2	C	1001	SIO	C5-C4-O4-CA4
4	C	5155	NAG	C4-C5-C6-O6
2	C	1001	SIO	OA9-CA9-O9-C9

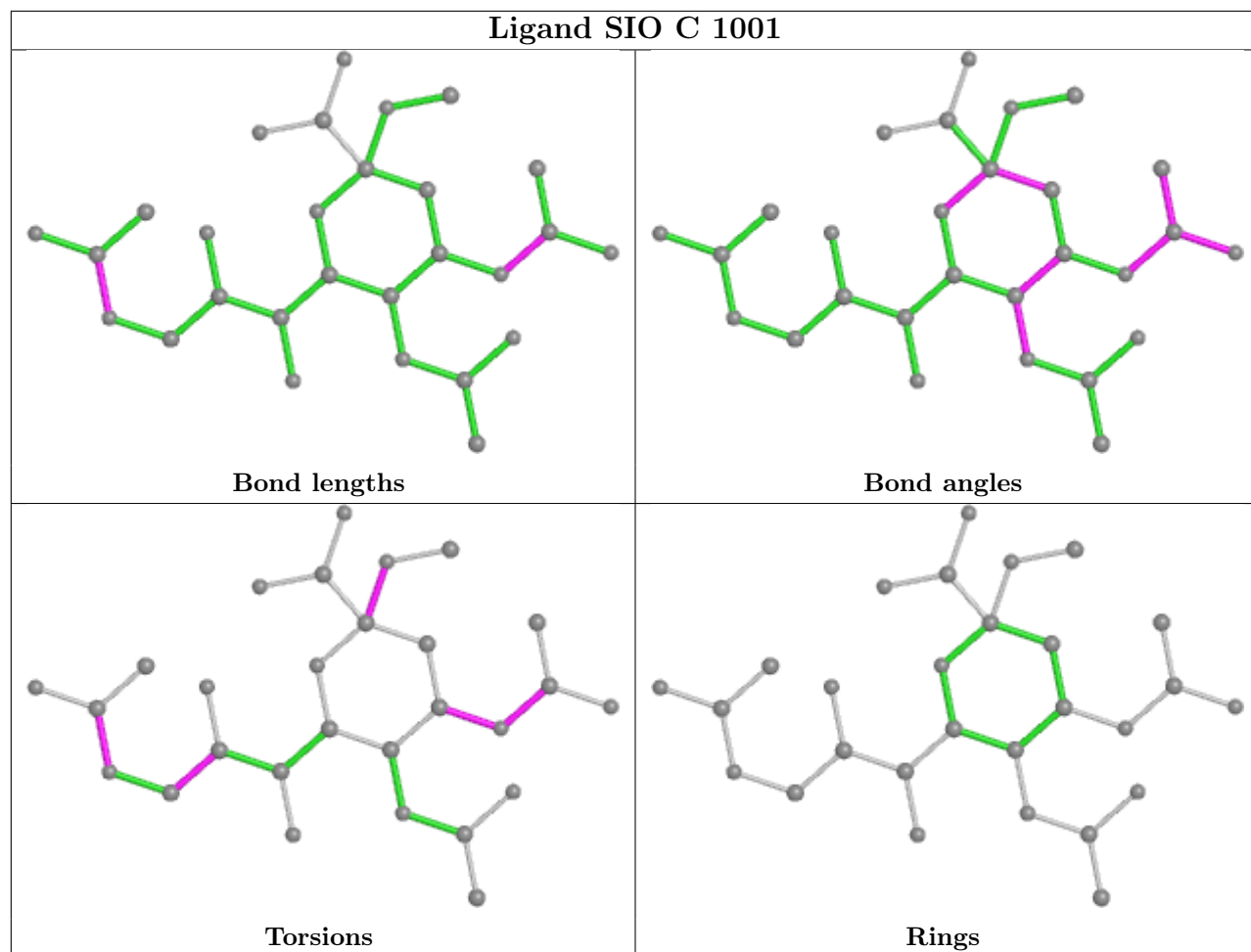
There are no ring outliers.

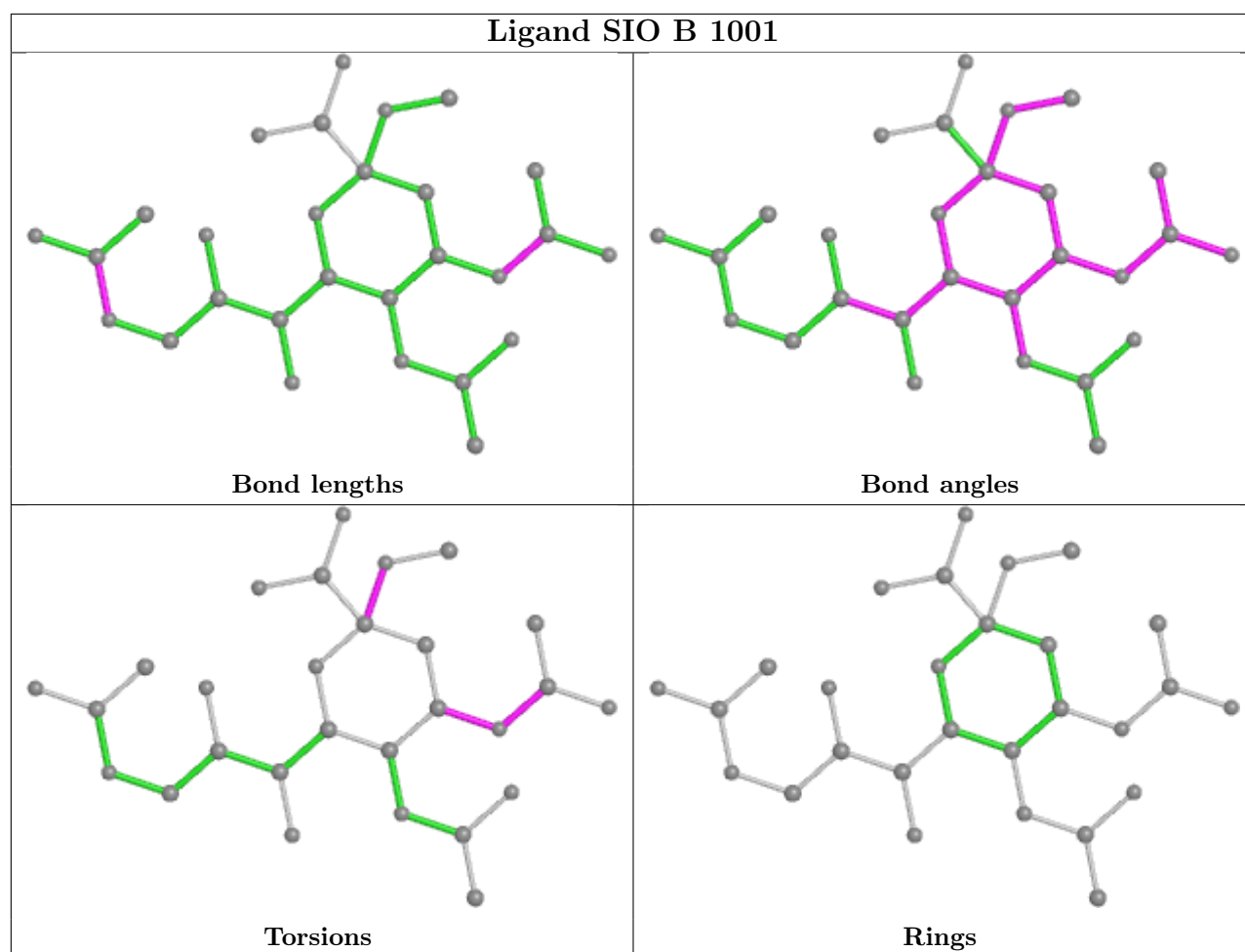
3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1001	SIO	1	0
3	C	2001	ACY	1	0
2	B	1001	SIO	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 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.







## 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

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	362/377 (96%)	0.22	21 (5%)	23 15	17, 23, 29, 35	9 (2%)
1	B	362/377 (96%)	0.14	9 (2%)	57 47	17, 26, 33, 42	6 (1%)
1	C	362/377 (96%)	0.20	10 (2%)	53 43	15, 23, 29, 36	8 (2%)
All	All	1086/1131 (96%)	0.19	40 (3%)	41 31	15, 24, 31, 42	23 (2%)

All (40) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	331	GLY	5.4
1	A	55	THR	5.0
1	A	53	LYS	3.8
1	C	337	GLY	3.4
1	A	328	ALA	3.3
1	C	189	LEU	3.3
1	C	332	ALA	3.3
1	A	58	SER	3.2
1	B	366	TYR	3.1
1	A	157	PHE	3.1
1	A	386	PHE	3.1
1	B	365	LEU	2.9
1	B	246	THR	2.8
1	A	59	LEU	2.7
1	B	337	GLY	2.7
1	A	188	LYS	2.7
1	A	383	TYR	2.7
1	C	188	LYS	2.6
1	A	189	LEU	2.6
1	A	332	ALA	2.6
1	A	57	LYS	2.6
1	C	349	LEU	2.5
1	C	383	TYR	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	330	ASN	2.5
1	A	158	THR	2.4
1	B	56	PRO	2.3
1	C	183	PHE	2.3
1	C	155	THR	2.3
1	A	327	LEU	2.3
1	B	383	TYR	2.3
1	B	53	LYS	2.2
1	B	157	PHE	2.2
1	A	365	LEU	2.2
1	C	331	GLY	2.1
1	A	185	ALA	2.1
1	A	156	ASN	2.1
1	A	301	ASN	2.1
1	B	58	SER	2.1
1	C	272	PHE	2.1
1	A	56	PRO	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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
4	NAG	B	5158	14/15	0.84	0.21	60,66,68,68	0
4	NAG	B	5155	14/15	0.86	0.21	56,62,63,63	0
4	NAG	B	5159	14/15	0.86	0.31	58,64,65,66	0
4	NAG	A	5158	14/15	0.88	0.32	54,60,64,65	0
4	NAG	A	5155	14/15	0.89	0.18	48,54,57,57	0
4	NAG	C	5157	14/15	0.89	0.26	44,51,53,54	0

*Continued on next page...*

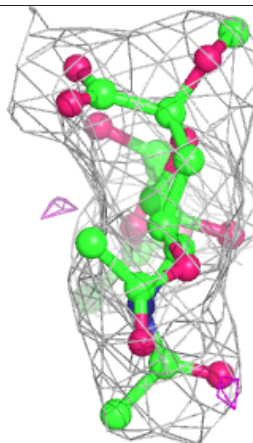
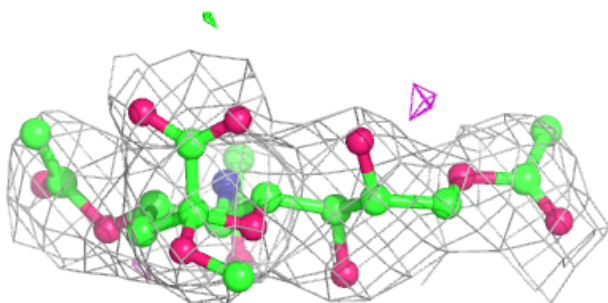
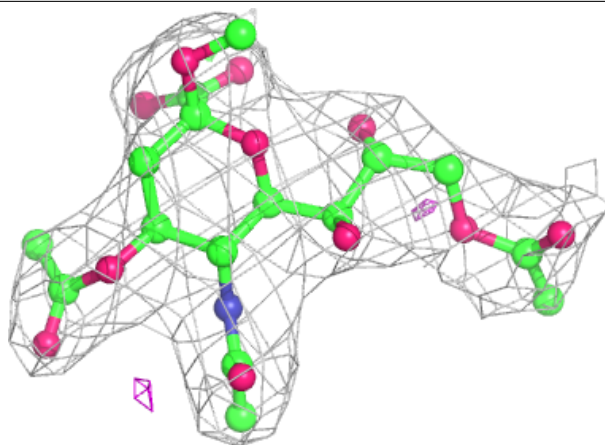
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NAG	A	5157	14/15	0.90	0.19	48,53,55,56	0
4	NAG	B	5156	14/15	0.90	0.26	59,65,67,67	0
4	NAG	B	5157	14/15	0.90	0.25	52,59,61,61	0
4	NAG	C	5155	14/15	0.91	0.16	50,55,57,57	0
4	NAG	C	5156	14/15	0.91	0.33	56,63,64,64	0
4	NAG	B	5154	14/15	0.91	0.24	49,55,58,58	0
4	NAG	A	5156	14/15	0.92	0.25	50,55,58,59	0
4	NAG	C	5154	14/15	0.93	0.21	44,48,53,53	0
4	NAG	A	5154	14/15	0.94	0.14	47,53,55,56	0
3	ACY	B	2001	4/4	0.95	0.37	70,70,70,70	0
4	NAG	C	5158	14/15	0.95	0.23	47,52,57,58	0
2	SIO	C	1001	28/28	0.96	0.20	59,61,63,64	0
2	SIO	A	1001	28/28	0.96	0.15	51,54,56,57	0
2	SIO	B	1001	28/28	0.96	0.18	50,57,58,59	0
3	ACY	A	2001	4/4	0.98	0.33	56,56,56,56	0
3	ACY	C	2001	4/4	0.98	0.25	49,49,49,50	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 SIO C 1001:**

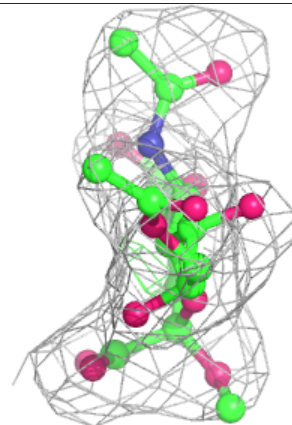
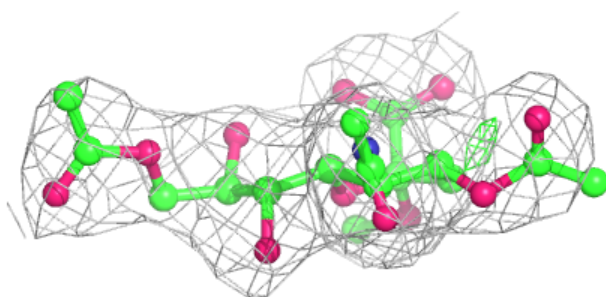
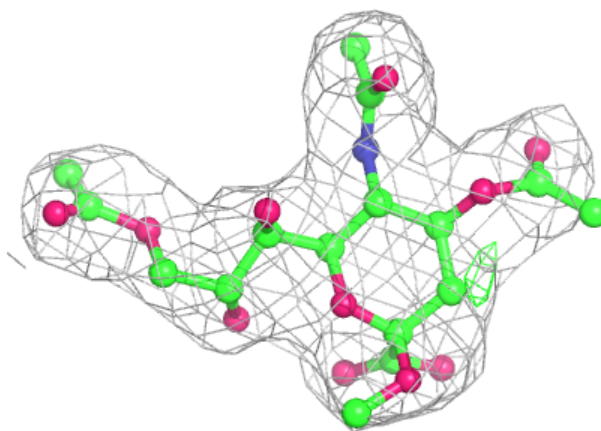
$2mF_o - DF_c$  (at 0.7 rmsd) in gray  
 $mF_o - DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



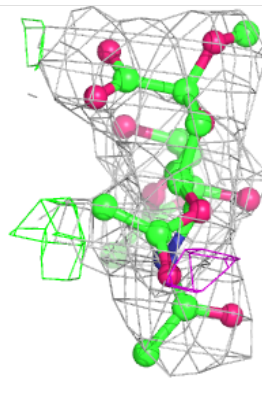
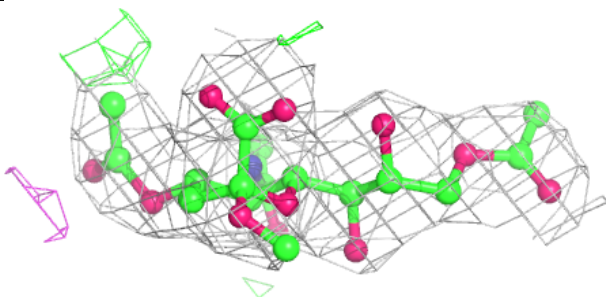
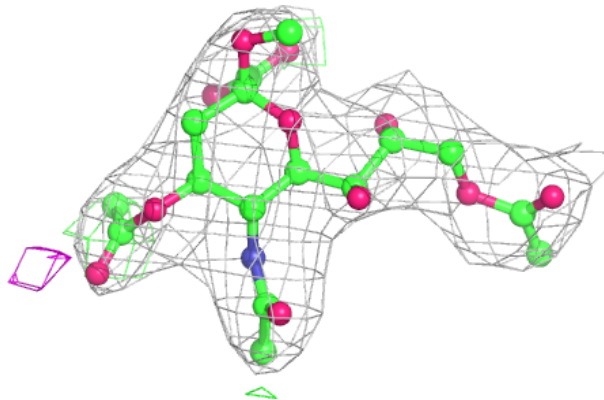


**Electron density around SIO A 1001:**

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

**Electron density around SIO 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)



## 6.5 Other polymers [i](#)

There are no such residues in this entry.