



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 2, 2021 – 03:55 AM EDT

PDB ID : 3AE0  
Title : Crystal structure of the C(30) carotenoid dehydrosqualene synthase from *Staphylococcus aureus* complexed with geranylgeranyl thiopyrophosphate  
Authors : Liu, C.I.; Jeng, W.Y.; Wang, A.H.J.; Oldfield, E.  
Deposited on : 2010-01-31  
Resolution : 2.37 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

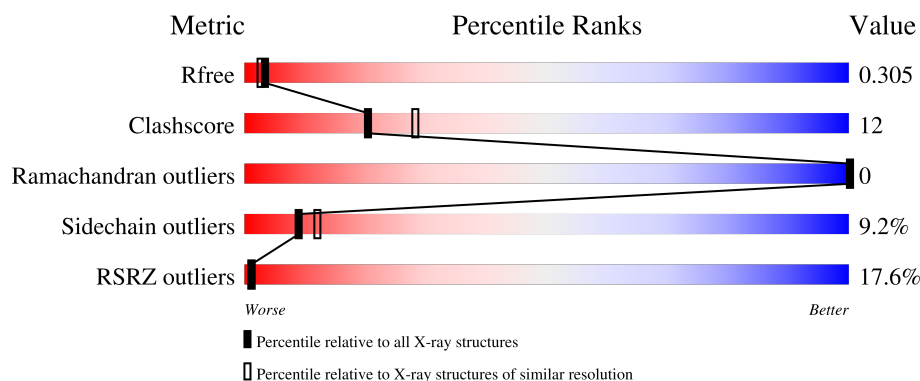
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.37 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	5509 (2.40-2.36)
Clashscore	141614	6082 (2.40-2.36)
Ramachandran outliers	138981	5973 (2.40-2.36)
Sidechain outliers	138945	5975 (2.40-2.36)
RSRZ outliers	127900	5397 (2.40-2.36)

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	293	<div> <div>16%</div> <div>72%</div> <div>23%</div> <div>• •</div> </div>
1	B	293	<div> <div>18%</div> <div>75%</div> <div>19%</div> <div>• •</div> </div>

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	GS	A	1001	-	-	X	-

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4964 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.

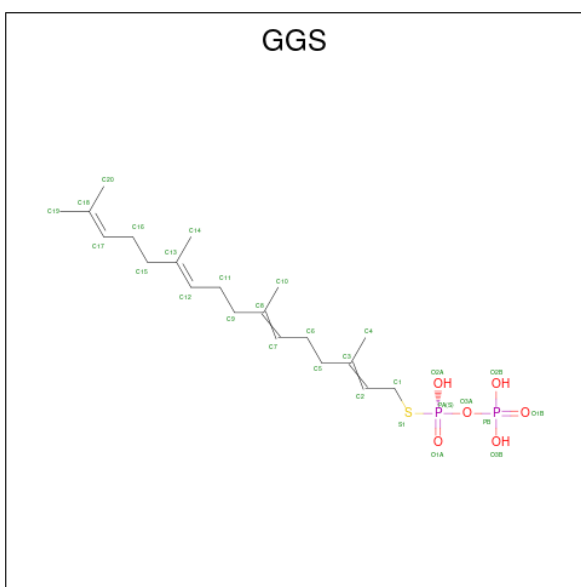
- Molecule 1 is a protein called Dehydrosqualene synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	284	Total	C	N	O	S	0	0	0
			2386	1530	400	444	12			
1	B	284	Total	C	N	O	S	0	0	0
			2386	1530	400	444	12			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	ALA	-	expression tag	UNP A9JQL9
A	-4	ALA	-	expression tag	UNP A9JQL9
A	-3	ALA	-	expression tag	UNP A9JQL9
A	-2	ALA	-	expression tag	UNP A9JQL9
A	-1	ALA	-	expression tag	UNP A9JQL9
A	0	ALA	-	expression tag	UNP A9JQL9
A	26	ALA	PHE	engineered mutation	UNP A9JQL9
B	-5	ALA	-	expression tag	UNP A9JQL9
B	-4	ALA	-	expression tag	UNP A9JQL9
B	-3	ALA	-	expression tag	UNP A9JQL9
B	-2	ALA	-	expression tag	UNP A9JQL9
B	-1	ALA	-	expression tag	UNP A9JQL9
B	0	ALA	-	expression tag	UNP A9JQL9
B	26	ALA	PHE	engineered mutation	UNP A9JQL9

- Molecule 2 is phosphonooxy-[(10E)-3,7,11,15-tetramethylhexadeca-2,6,10,14-tetraenyl]sulfonyl-phosphinic acid (three-letter code: GGS) (formula: C<sub>20</sub>H<sub>36</sub>O<sub>6</sub>P<sub>2</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total 29	C 20	O 6	P 2	S 1	0	0
2	A	1	Total 29	C 20	O 6	P 2	S 1	0	0
2	B	1	Total 29	C 20	O 6	P 2	S 1	0	0
2	B	1	Total 29	C 20	O 6	P 2	S 1	0	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	3	Total Mg 3 3	0	0
3	B	3	Total Mg 3 3	0	0

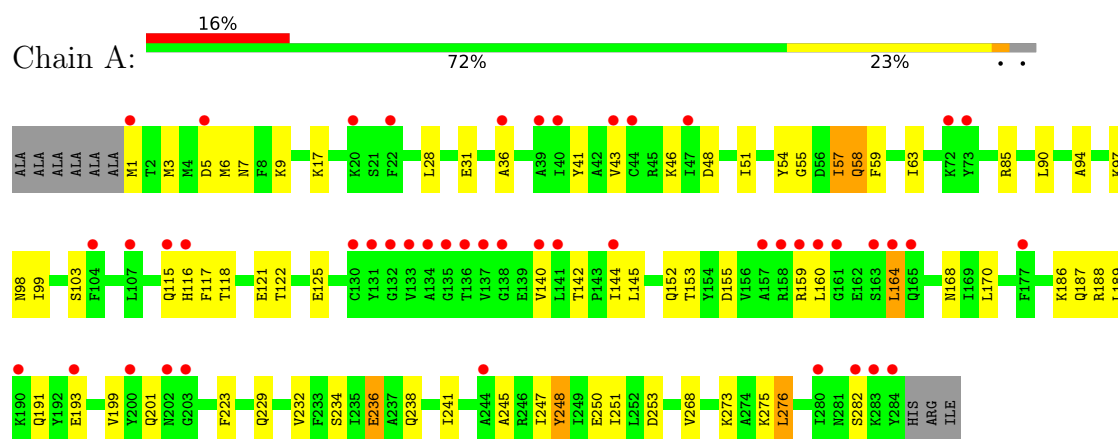
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	30	Total O 30 30	0	0
4	B	40	Total O 40 40	0	0

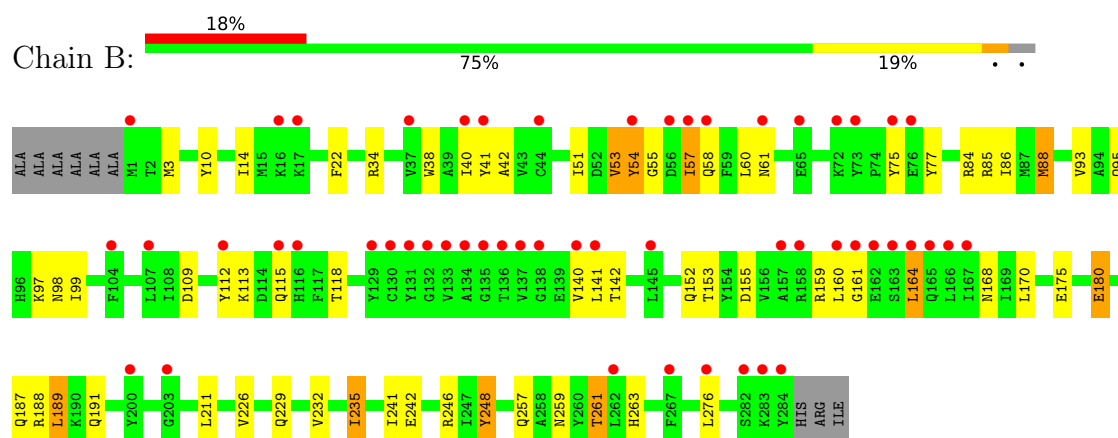
### 3 Residue-property plots [i](#)

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: Dehydrosqualene synthase



- Molecule 1: Dehydrosqualene synthase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	80.12Å 80.12Å 183.62Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 2.37 28.00 – 2.37	Depositor EDS
% Data completeness (in resolution range)	99.0 (30.00-2.37) 99.0 (28.00-2.37)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.12 (at 2.36Å)	Xtriage
Refinement program	CNS, REFMAC 5.5.0072	Depositor
R, $R_{free}$	0.219 , 0.281 0.260 , 0.305	Depositor DCC
$R_{free}$ test set	1428 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	50.4	Xtriage
Anisotropy	0.347	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 36.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.037 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	4964	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 90.11 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.6219e-08. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG, GGS

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	0.37	0/2441	0.71	0/3292
1	B	0.37	0/2441	0.68	0/3292
All	All	0.37	0/4882	0.70	0/6584

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2386	0	2317	44	0
1	B	2386	0	2317	46	0
2	A	58	0	66	29	0
2	B	58	0	66	28	0
3	A	3	0	0	0	0
3	B	3	0	0	0	0
4	A	30	0	0	2	0
4	B	40	0	0	2	0
All	All	4964	0	4766	120	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 12.

All (120) 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:B:1003:GGS:H19B	2:B:1004:GGS:C19	1.85	1.06
2:B:1003:GGS:C19	2:B:1004:GGS:H19B	1.87	1.04
2:B:1003:GGS:H19B	2:B:1004:GGS:H19B	1.41	1.01
2:A:1001:GGS:H19	2:A:1002:GGS:C18	1.94	0.97
1:B:152:GLN:HE21	1:B:229:GLN:HE21	1.22	0.87
1:A:6:MET:HG2	4:B:544:HOH:O	1.74	0.85
1:A:188:ARG:HH11	1:A:191:GLN:HE22	1.24	0.83
2:B:1003:GGS:C19	2:B:1004:GGS:C19	2.49	0.83
1:A:152:GLN:HE21	1:A:229:GLN:HE21	1.28	0.77
1:B:164:LEU:HB3	2:B:1003:GGS:C6	2.17	0.73
1:B:235:ILE:O	1:B:235:ILE:HD13	1.89	0.73
1:A:268:VAL:HG23	1:A:273:LYS:HE3	1.72	0.71
2:A:1001:GGS:H19A	2:A:1002:GGS:H19B	1.73	0.71
1:A:41:TYR:CD1	2:A:1002:GGS:H11	2.27	0.70
1:A:247:ILE:HG23	1:A:268:VAL:HG21	1.74	0.69
2:A:1001:GGS:C19	2:A:1002:GGS:H19B	2.22	0.69
1:B:175:GLU:HG3	4:B:531:HOH:O	1.93	0.69
1:B:164:LEU:HB3	2:B:1003:GGS:H6A	1.75	0.67
1:B:41:TYR:CE1	2:B:1004:GGS:H10	2.30	0.67
1:A:9:LYS:NZ	1:B:180:GLU:OE2	2.26	0.67
2:B:1004:GGS:H1	2:B:1004:GGS:O3B	1.94	0.67
1:B:141:LEU:HD11	2:B:1004:GGS:H14B	1.75	0.67
1:A:168:ASN:HB2	2:A:1001:GGS:H4B	1.76	0.67
2:B:1003:GGS:H19B	2:B:1004:GGS:H19	1.76	0.65
2:A:1001:GGS:H19	2:A:1002:GGS:C19	2.27	0.65
2:A:1001:GGS:C19	2:A:1002:GGS:C19	2.76	0.64
2:A:1001:GGS:C2	2:A:1002:GGS:H4B	2.28	0.64
1:B:41:TYR:CD1	2:B:1004:GGS:H11	2.34	0.63
1:B:34:ARG:HD2	1:B:38:TRP:CZ2	2.34	0.62
1:B:53:VAL:HG22	1:B:54:TYR:HD1	1.65	0.62
2:B:1003:GGS:H19	2:B:1004:GGS:C18	2.29	0.62
1:A:155:ASP:O	1:A:159:ARG:HG3	2.00	0.61
2:B:1004:GGS:H4	2:B:1004:GGS:C7	2.29	0.61
1:A:103:SER:HB2	1:A:140:VAL:HG23	1.81	0.61
1:B:142:THR:HG23	1:B:153:THR:HG22	1.82	0.61
2:B:1003:GGS:H1	2:B:1004:GGS:S1	2.42	0.60
1:A:188:ARG:NH1	1:A:191:GLN:HE22	1.96	0.60
1:A:48:ASP:CG	2:A:1002:GGS:H2	2.22	0.60
1:B:77:TYR:OH	1:B:84:ARG:NH1	2.36	0.59
2:B:1003:GGS:C19	2:B:1004:GGS:C18	2.83	0.57

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:164:LEU:HB3	2:A:1001:GGS:H6A	1.87	0.56
2:B:1004:GGS:S1	2:B:1004:GGS:H4B	2.45	0.56
2:A:1001:GGS:H9A	2:A:1002:GGS:H12	1.87	0.56
1:B:112:TYR:C	1:B:112:TYR:CD2	2.80	0.55
1:B:57:ILE:HG13	1:B:57:ILE:O	2.07	0.55
1:B:42:ALA:HB1	1:B:86:ILE:HD12	1.88	0.54
2:A:1001:GGS:C13	2:A:1001:GGS:H20B	2.37	0.54
1:B:113:LYS:C	1:B:115:GLN:H	2.11	0.54
2:B:1004:GGS:H4	2:B:1004:GGS:H7	1.89	0.53
2:A:1001:GGS:H20B	2:A:1001:GGS:C12	2.38	0.53
2:A:1001:GGS:H19	2:A:1002:GGS:C20	2.39	0.53
1:B:168:ASN:ND2	2:B:1003:GGS:H2	2.24	0.53
2:B:1003:GGS:H10	2:B:1004:GGS:H9A	1.91	0.52
1:B:10:TYR:CE2	1:B:14:ILE:HD11	2.44	0.52
2:A:1001:GGS:H1	2:A:1002:GGS:H4B	1.90	0.52
1:B:164:LEU:HB3	2:B:1003:GGS:H6	1.90	0.52
1:A:223:PHE:HE1	1:A:245:ALA:HB1	1.75	0.51
1:B:10:TYR:CZ	1:B:14:ILE:HD11	2.45	0.51
1:A:46:LYS:HE2	4:A:541:HOH:O	2.09	0.51
1:B:160:LEU:HD23	2:B:1003:GGS:H16A	1.92	0.51
1:A:57:ILE:HG13	1:A:58:GLN:N	2.22	0.51
1:A:223:PHE:CE1	1:A:245:ALA:HB1	2.46	0.51
1:A:51:ILE:O	1:A:55:GLY:HA2	2.11	0.50
1:B:58:GLN:HA	1:B:61:ASN:HD22	1.77	0.50
1:B:160:LEU:HD13	1:B:226:VAL:HG11	1.93	0.50
2:B:1003:GGS:H9A	2:B:1004:GGS:H12	1.95	0.49
1:A:187:GLN:NE2	4:A:557:HOH:O	2.45	0.49
1:A:250:GLU:HG3	1:A:276:LEU:HD21	1.96	0.48
1:A:43:VAL:HG21	1:A:90:LEU:HD22	1.96	0.48
1:A:41:TYR:HB2	2:A:1002:GGS:H14A	1.95	0.48
1:A:115:GLN:C	1:A:117:PHE:H	2.17	0.48
2:A:1001:GGS:H2	2:A:1002:GGS:H4B	1.95	0.48
1:A:241:ILE:HG12	2:A:1001:GGS:H17	1.96	0.47
1:A:59:PHE:O	1:A:63:ILE:HG13	2.13	0.47
1:A:236:GLU:H	1:A:236:GLU:CD	2.17	0.47
1:A:122:THR:HG23	1:A:125:GLU:H	1.79	0.47
2:A:1002:GGS:C4	2:A:1002:GGS:H7	2.44	0.47
1:B:42:ALA:CB	1:B:86:ILE:HD12	2.45	0.47
1:B:152:GLN:HE21	1:B:229:GLN:NE2	2.02	0.47
1:A:41:TYR:CE1	2:A:1002:GGS:H11	2.50	0.47
1:B:188:ARG:HH11	1:B:191:GLN:HE22	1.63	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:160:LEU:HG	2:A:1001:GGS:C14	2.46	0.46
1:A:51:ILE:O	1:A:55:GLY:N	2.48	0.46
1:B:160:LEU:HG	2:B:1003:GGS:H14B	1.97	0.45
1:B:242:GLU:OE2	1:B:246:ARG:NH1	2.49	0.45
1:B:188:ARG:HD3	1:B:191:GLN:NE2	2.32	0.45
1:B:75:TYR:CE1	1:B:95:GLN:HG2	2.52	0.45
1:B:40:ILE:HD12	1:B:140:VAL:CG1	2.48	0.44
1:B:53:VAL:HG22	1:B:54:TYR:CD1	2.49	0.44
1:A:7:ASN:ND2	1:A:85:ARG:HG2	2.33	0.43
1:B:14:ILE:HD12	1:B:86:ILE:HD11	2.00	0.43
1:A:142:THR:HG23	1:A:153:THR:HG22	2.00	0.43
1:A:94:ALA:HB2	1:A:99:ILE:HD12	2.01	0.43
2:A:1002:GGS:H16A	2:A:1002:GGS:H14	1.71	0.43
1:B:241:ILE:HG23	2:B:1003:GGS:H17	2.00	0.43
1:A:1:MET:HG3	1:A:5:ASP:CB	2.49	0.42
2:A:1001:GGS:C1	2:A:1002:GGS:H4B	2.49	0.42
1:B:51:ILE:O	1:B:55:GLY:HA2	2.19	0.42
1:B:84:ARG:O	1:B:88:MET:HB2	2.19	0.42
1:A:160:LEU:HD23	2:A:1001:GGS:H16A	2.01	0.42
1:A:234:SER:O	1:A:238:GLN:HG3	2.19	0.42
1:B:93:VAL:HG12	1:B:99:ILE:HD11	2.01	0.42
1:B:188:ARG:HD3	1:B:191:GLN:HE21	1.84	0.42
1:B:161:GLY:N	2:B:1003:GGS:H14A	2.34	0.42
1:A:36:ALA:HB1	1:A:144:ILE:HG12	2.01	0.42
1:A:121:GLU:HA	1:A:186:LYS:HB2	2.02	0.42
1:B:189:LEU:HD13	1:B:211:LEU:HD21	2.01	0.41
2:B:1003:GGS:C5	2:B:1003:GGS:S1	3.08	0.41
1:A:41:TYR:CE1	2:A:1002:GGS:H10	2.56	0.41
2:A:1001:GGS:S1	2:A:1001:GGS:C5	3.08	0.41
1:B:155:ASP:O	1:B:159:ARG:HG3	2.20	0.41
1:A:145:LEU:HD11	2:A:1001:GGS:H19A	2.03	0.41
1:A:168:ASN:HD21	2:A:1002:GGS:PA	2.44	0.41
1:A:188:ARG:HA	1:A:188:ARG:HD3	1.93	0.41
1:B:41:TYR:CE1	2:B:1004:GGS:H11	2.55	0.41
1:B:241:ILE:HG12	2:B:1003:GGS:H17	2.03	0.41
1:A:160:LEU:HG	2:A:1001:GGS:H14	2.03	0.40
1:B:261:THR:HG22	1:B:263:HIS:H	1.86	0.40
1:B:22:PHE:CZ	1:B:248:TYR:CE1	3.09	0.40
1:A:248:TYR:O	1:A:251:ILE:HB	2.22	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	282/293 (96%)	273 (97%)	9 (3%)	0	100	100
1	B	282/293 (96%)	270 (96%)	12 (4%)	0	100	100
All	All	564/586 (96%)	543 (96%)	21 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains

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	256/259 (99%)	232 (91%)	24 (9%)	8	11
1	B	256/259 (99%)	233 (91%)	23 (9%)	9	12
All	All	512/518 (99%)	465 (91%)	47 (9%)	9	12

All (47) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	3	MET
1	A	17	LYS
1	A	28	LEU
1	A	31	GLU
1	A	54	TYR
1	A	57	ILE
1	A	58	GLN
1	A	97	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	98	ASN
1	A	116	HIS
1	A	118	THR
1	A	164	LEU
1	A	170	LEU
1	A	189	LEU
1	A	193	GLU
1	A	199	VAL
1	A	201	GLN
1	A	232	VAL
1	A	236	GLU
1	A	248	TYR
1	A	253	ASP
1	A	275	LYS
1	A	276	LEU
1	A	282	SER
1	B	3	MET
1	B	53	VAL
1	B	54	TYR
1	B	57	ILE
1	B	60	LEU
1	B	85	ARG
1	B	88	MET
1	B	97	LYS
1	B	98	ASN
1	B	109	ASP
1	B	118	THR
1	B	164	LEU
1	B	170	LEU
1	B	180	GLU
1	B	187	GLN
1	B	189	LEU
1	B	232	VAL
1	B	235	ILE
1	B	248	TYR
1	B	257	GLN
1	B	259	ASN
1	B	261	THR
1	B	276	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (19) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	7	ASN
1	A	58	GLN
1	A	92	HIS
1	A	96	HIS
1	A	152	GLN
1	A	165	GLN
1	A	187	GLN
1	A	191	GLN
1	A	201	GLN
1	B	7	ASN
1	B	58	GLN
1	B	61	ASN
1	B	98	ASN
1	B	102	GLN
1	B	115	GLN
1	B	152	GLN
1	B	191	GLN
1	B	238	GLN
1	B	259	ASN

### 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](#)

Of 10 ligands modelled in this entry, 6 are monoatomic - leaving 4 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the

expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	GGs	A	1001	3	24,28,28	3.17	2 (8%)	29,37,37	2.18	12 (41%)
2	GGs	A	1002	3	24,28,28	3.09	2 (8%)	29,37,37	1.78	8 (27%)
2	GGs	B	1004	3	24,28,28	3.08	2 (8%)	29,37,37	1.92	6 (20%)
2	GGs	B	1003	3	24,28,28	3.15	1 (4%)	29,37,37	1.93	11 (37%)

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	GGs	A	1001	3	-	18/25/31/31	-
2	GGs	A	1002	3	-	17/25/31/31	-
2	GGs	B	1004	3	-	13/25/31/31	-
2	GGs	B	1003	3	-	17/25/31/31	-

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1001	GGs	C1-S1	-15.07	1.66	1.84
2	B	1003	GGs	C1-S1	-14.95	1.66	1.84
2	A	1002	GGs	C1-S1	-14.68	1.67	1.84
2	B	1004	GGs	C1-S1	-14.58	1.67	1.84
2	B	1004	GGs	PA-O2A	-2.16	1.51	1.56
2	A	1002	GGs	PA-O2A	-2.16	1.51	1.56
2	A	1001	GGs	PA-O2A	-2.08	1.51	1.56

All (37) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1002	GGs	C10-C8-C9	5.39	124.34	115.27
2	B	1004	GGs	C10-C8-C9	4.51	122.85	115.27
2	B	1004	GGs	C15-C16-C17	4.38	126.26	111.88
2	A	1001	GGs	C11-C12-C13	-4.23	117.46	127.66
2	A	1001	GGs	C10-C8-C9	4.03	122.05	115.27
2	B	1003	GGs	C15-C16-C17	3.65	123.87	111.88
2	A	1001	GGs	C14-C13-C15	3.61	121.34	115.27

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1001	GGs	C15-C16-C17	3.48	123.33	111.88
2	B	1003	GGs	C10-C8-C9	3.47	121.11	115.27
2	B	1003	GGs	C5-C6-C7	3.44	123.20	111.88
2	B	1003	GGs	C4-C3-C5	3.19	120.63	115.27
2	A	1001	GGs	C5-C6-C7	3.15	122.24	111.88
2	A	1002	GGs	C2-C1-S1	2.96	126.82	112.00
2	A	1001	GGs	C9-C11-C12	2.92	121.47	111.88
2	A	1002	GGs	C10-C8-C7	-2.84	116.39	123.68
2	B	1004	GGs	C5-C6-C7	2.71	120.80	111.88
2	B	1004	GGs	C6-C7-C8	-2.70	121.17	127.66
2	B	1003	GGs	C11-C12-C13	-2.69	121.18	127.66
2	A	1002	GGs	C14-C13-C15	2.61	119.66	115.27
2	A	1001	GGs	C6-C5-C3	-2.55	104.59	112.98
2	B	1003	GGs	C2-C1-S1	2.46	124.32	112.00
2	A	1001	GGs	C6-C7-C8	-2.45	121.75	127.66
2	A	1002	GGs	C14-C13-C12	-2.38	117.58	123.68
2	B	1004	GGs	C16-C17-C18	-2.37	119.66	127.75
2	A	1002	GGs	O3B-PB-O2B	2.35	116.62	107.64
2	A	1001	GGs	C14-C13-C12	-2.30	117.78	123.68
2	A	1001	GGs	O2A-PA-O1A	2.29	115.36	109.82
2	A	1001	GGs	O3A-PB-O1B	-2.26	98.65	111.19
2	B	1003	GGs	C4-C3-C2	-2.23	117.96	123.68
2	A	1002	GGs	C9-C11-C12	2.20	119.11	111.88
2	A	1001	GGs	O3B-PB-O2B	2.19	116.02	107.64
2	B	1003	GGs	C14-C13-C15	2.14	118.87	115.27
2	B	1003	GGs	C6-C5-C3	-2.13	105.96	112.98
2	B	1003	GGs	O2A-PA-O1A	2.13	114.97	109.82
2	A	1002	GGs	O2A-PA-O1A	2.05	114.77	109.82
2	B	1003	GGs	C20-C18-C19	2.02	119.06	114.60
2	B	1004	GGs	C4-C3-C5	2.02	118.66	115.27

There are no chirality outliers.

All (65) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1001	GGs	C1-C2-C3-C4
2	A	1001	GGs	C1-C2-C3-C5
2	A	1001	GGs	C5-C6-C7-C8
2	A	1001	GGs	C6-C7-C8-C9
2	A	1001	GGs	C6-C7-C8-C10
2	A	1001	GGs	PA-O3A-PB-O2B
2	A	1001	GGs	C15-C16-C17-C18

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
2	A	1001	GGs	C16-C17-C18-C19
2	A	1002	GGs	C6-C7-C8-C9
2	A	1002	GGs	C6-C7-C8-C10
2	A	1002	GGs	C10-C8-C9-C11
2	A	1002	GGs	C12-C13-C15-C16
2	A	1002	GGs	C14-C13-C15-C16
2	A	1002	GGs	C16-C17-C18-C19
2	B	1003	GGs	C1-C2-C3-C4
2	B	1003	GGs	C1-C2-C3-C5
2	B	1003	GGs	C2-C3-C5-C6
2	B	1003	GGs	C4-C3-C5-C6
2	B	1003	GGs	C3-C5-C6-C7
2	B	1003	GGs	C5-C6-C7-C8
2	B	1003	GGs	PA-O3A-PB-O2B
2	B	1003	GGs	C15-C16-C17-C18
2	B	1003	GGs	C16-C17-C18-C20
2	B	1004	GGs	C1-C2-C3-C4
2	B	1004	GGs	C1-C2-C3-C5
2	B	1004	GGs	C11-C12-C13-C14
2	A	1001	GGs	C16-C17-C18-C20
2	A	1002	GGs	C16-C17-C18-C20
2	B	1003	GGs	C16-C17-C18-C19
2	B	1004	GGs	C16-C17-C18-C19
2	B	1004	GGs	C16-C17-C18-C20
2	A	1002	GGs	C7-C8-C9-C11
2	A	1002	GGs	C11-C12-C13-C14
2	B	1003	GGs	C6-C7-C8-C10
2	A	1002	GGs	C11-C12-C13-C15
2	B	1003	GGs	C6-C7-C8-C9
2	B	1004	GGs	C11-C12-C13-C15
2	B	1004	GGs	C10-C8-C9-C11
2	B	1004	GGs	C7-C8-C9-C11
2	A	1001	GGs	C3-C5-C6-C7
2	A	1001	GGs	C12-C11-C9-C8
2	A	1002	GGs	C12-C11-C9-C8
2	A	1002	GGs	C13-C15-C16-C17
2	B	1004	GGs	C12-C11-C9-C8
2	A	1001	GGs	C4-C3-C5-C6
2	A	1002	GGs	C3-C5-C6-C7
2	B	1003	GGs	PA-O3A-PB-O1B
2	A	1001	GGs	C2-C3-C5-C6
2	B	1003	GGs	C10-C8-C9-C11

*Continued on next page...*

*Continued from previous page...*

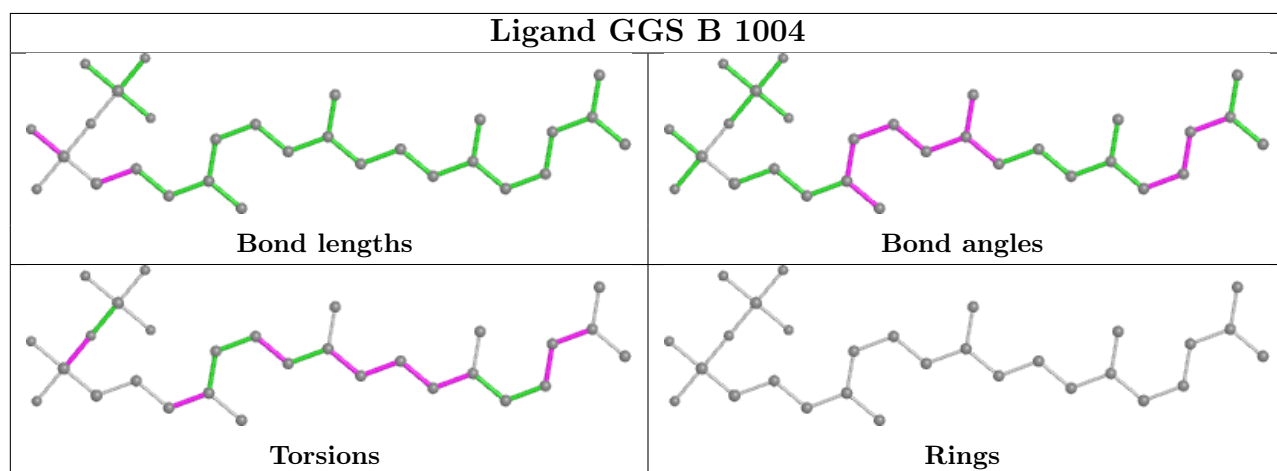
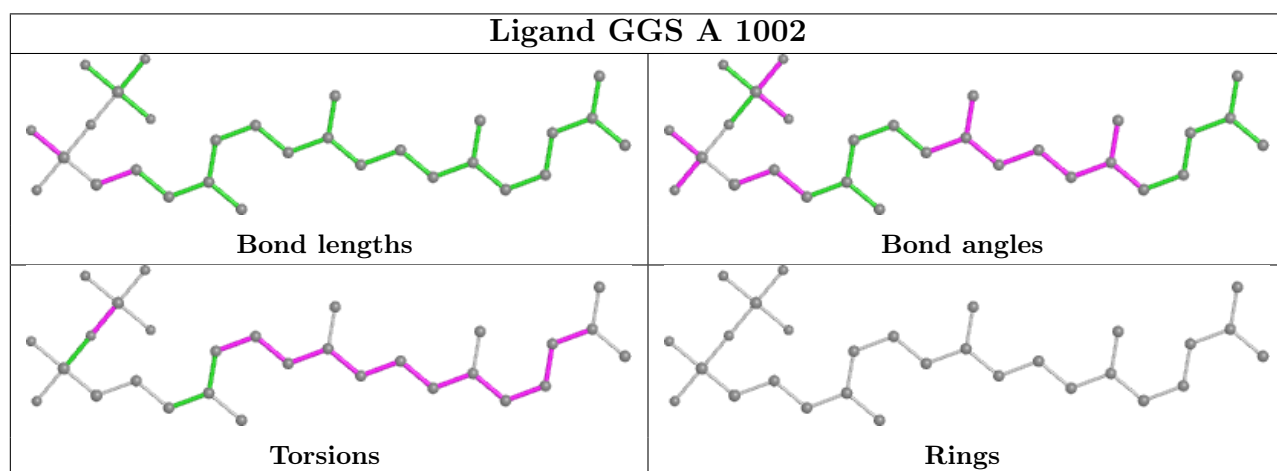
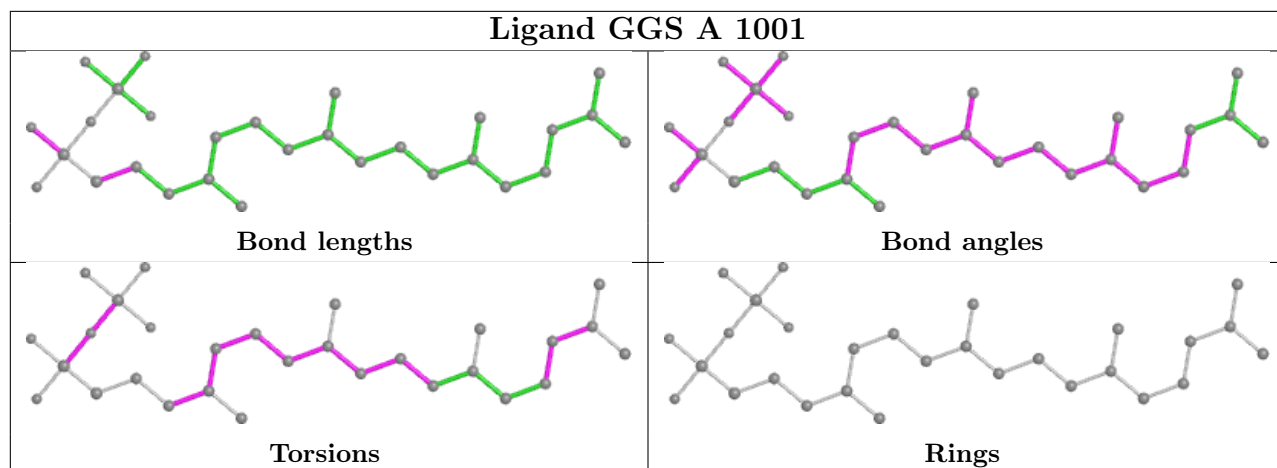
Mol	Chain	Res	Type	Atoms
2	B	1003	GGG	C7-C8-C9-C11
2	A	1002	GGG	PA-O3A-PB-O2B
2	A	1001	GGG	C10-C8-C9-C11
2	A	1001	GGG	C7-C8-C9-C11
2	A	1001	GGG	PB-O3A-PA-O2A
2	B	1003	GGG	PB-O3A-PA-O2A
2	B	1004	GGG	PB-O3A-PA-O1A
2	A	1001	GGG	PA-O3A-PB-O1B
2	A	1001	GGG	C9-C11-C12-C13
2	A	1002	GGG	C5-C6-C7-C8
2	A	1002	GGG	C15-C16-C17-C18
2	B	1003	GGG	C9-C11-C12-C13
2	A	1002	GGG	C9-C11-C12-C13
2	B	1004	GGG	C5-C6-C7-C8
2	B	1004	GGG	C9-C11-C12-C13
2	B	1004	GGG	C15-C16-C17-C18

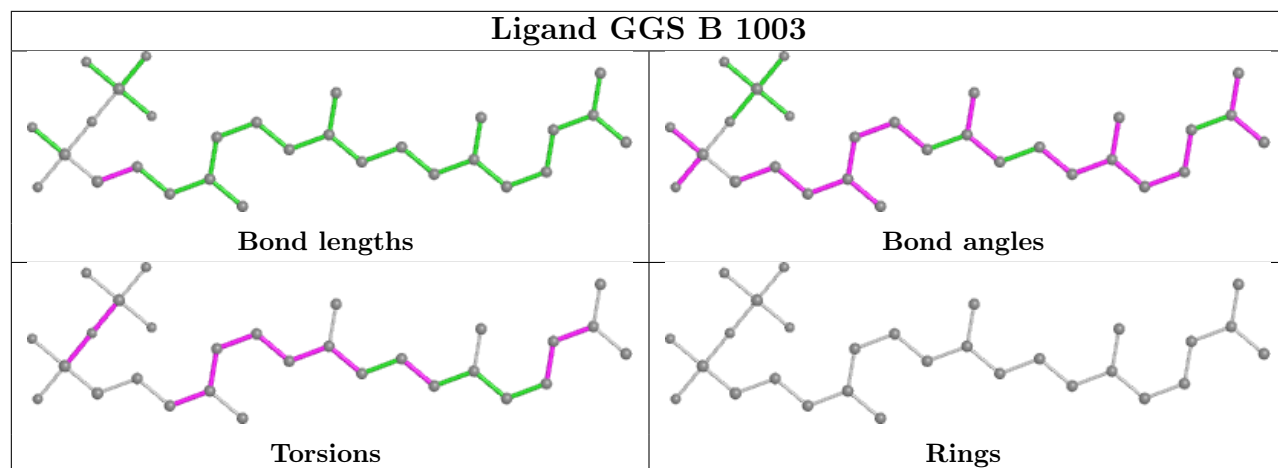
There are no ring outliers.

4 monomers are involved in 57 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1001	GGG	21	0
2	A	1002	GGG	19	0
2	B	1004	GGG	18	0
2	B	1003	GGG	20	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	284/293 (96%)	0.92	47 (16%)  	15, 29, 42, 56	0
1	B	284/293 (96%)	1.04	53 (18%)  	19, 29, 47, 61	0
All	All	568/586 (96%)	0.98	100 (17%)  	15, 29, 44, 61	0

All (100) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	134	ALA	6.3
1	B	137	VAL	6.0
1	A	137	VAL	5.8
1	B	73	TYR	5.0
1	A	136	THR	5.0
1	B	138	GLY	5.0
1	A	134	ALA	4.9
1	B	130	CYS	4.8
1	B	136	THR	4.7
1	B	141	LEU	4.6
1	A	164	LEU	4.5
1	A	40	ILE	4.4
1	B	200	TYR	4.4
1	A	280	ILE	4.4
1	A	141	LEU	4.3
1	A	43	VAL	4.3
1	A	140	VAL	4.3
1	B	164	LEU	4.3
1	A	161	GLY	4.2
1	B	160	LEU	4.1
1	A	282	SER	4.1
1	B	283	LYS	4.0
1	B	282	SER	3.9
1	A	138	GLY	3.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	116	HIS	3.8
1	B	161	GLY	3.8
1	B	135	GLY	3.8
1	A	284	TYR	3.8
1	A	131	TYR	3.7
1	A	1	MET	3.7
1	B	133	VAL	3.7
1	B	140	VAL	3.7
1	B	56	ASP	3.7
1	A	160	LEU	3.5
1	B	107	LEU	3.5
1	A	73	TYR	3.5
1	B	58	GLN	3.5
1	A	39	ALA	3.4
1	A	130	CYS	3.4
1	B	131	TYR	3.4
1	B	40	ILE	3.3
1	B	44	CYS	3.3
1	B	1	MET	3.3
1	B	112	TYR	3.2
1	A	283	LYS	3.2
1	A	133	VAL	3.2
1	B	104	PHE	3.1
1	B	116	HIS	3.1
1	A	165	GLN	3.1
1	A	135	GLY	3.0
1	A	244	ALA	3.0
1	B	41	TYR	3.0
1	B	267	PHE	3.0
1	A	44	CYS	3.0
1	A	157	ALA	2.9
1	B	76	GLU	2.9
1	B	72	LYS	2.9
1	B	157	ALA	2.9
1	B	284	TYR	2.9
1	B	37	VAL	2.8
1	A	107	LEU	2.8
1	A	158	ARG	2.8
1	A	132	GLY	2.7
1	B	61	ASN	2.7
1	A	104	PHE	2.7
1	B	276	LEU	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	145	LEU	2.7
1	B	65	GLU	2.7
1	B	163	SER	2.6
1	B	17	LYS	2.6
1	B	54	TYR	2.6
1	A	200	TYR	2.6
1	B	132	GLY	2.6
1	A	144	ILE	2.6
1	B	167	ILE	2.6
1	A	163	SER	2.5
1	B	262	LEU	2.5
1	B	115	GLN	2.5
1	A	202	ASN	2.5
1	A	193	GLU	2.4
1	B	203	GLY	2.4
1	B	158	ARG	2.4
1	A	22	PHE	2.4
1	B	166	LEU	2.4
1	A	115	GLN	2.4
1	A	203	GLY	2.3
1	A	47	ILE	2.3
1	B	162	GLU	2.3
1	B	165	GLN	2.3
1	A	20	LYS	2.2
1	B	129	TYR	2.1
1	A	177	PHE	2.1
1	B	57	ILE	2.1
1	A	5	ASP	2.1
1	B	75	TYR	2.1
1	A	36	ALA	2.1
1	A	190	LYS	2.1
1	B	16	LYS	2.1
1	A	159	ARG	2.0
1	A	72	LYS	2.0

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

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](#)

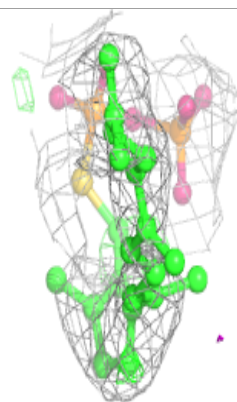
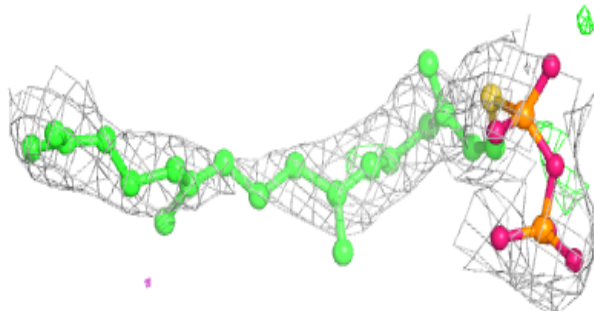
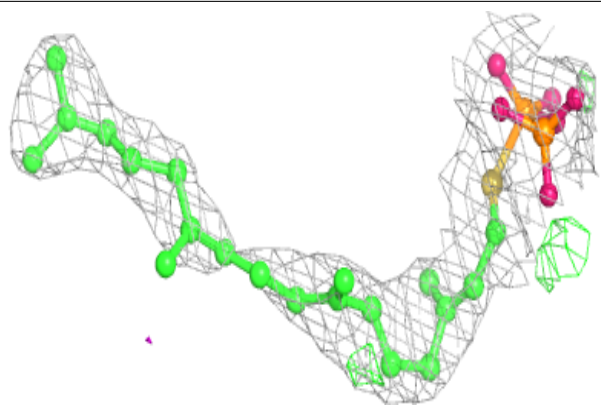
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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
3	MG	A	1006	1/1	0.80	0.11	55,55,55,55	0
3	MG	B	1009	1/1	0.80	0.22	59,59,59,59	0
3	MG	A	1007	1/1	0.81	0.26	65,65,65,65	0
2	GGS	A	1002	29/29	0.90	0.34	53,60,68,70	0
2	GGS	A	1001	29/29	0.90	0.38	48,70,75,76	0
3	MG	B	1010	1/1	0.90	0.27	67,67,67,67	0
2	GGS	B	1004	29/29	0.91	0.35	60,69,80,82	0
2	GGS	B	1003	29/29	0.92	0.35	52,70,77,77	0
3	MG	B	1008	1/1	0.93	0.27	44,44,44,44	0
3	MG	A	1005	1/1	0.95	0.21	42,42,42,42	0

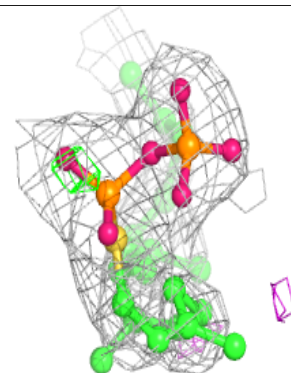
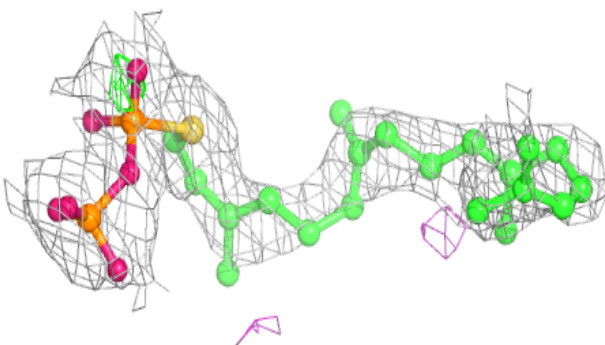
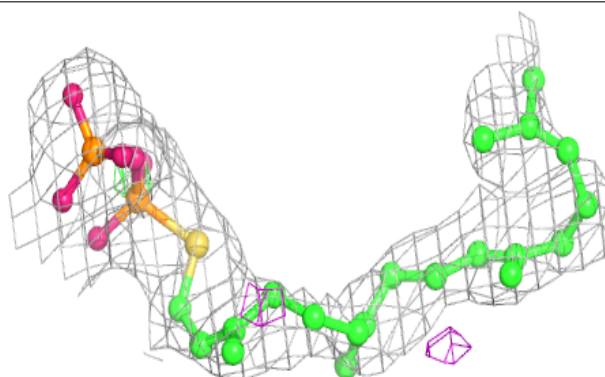
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around GGS 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)

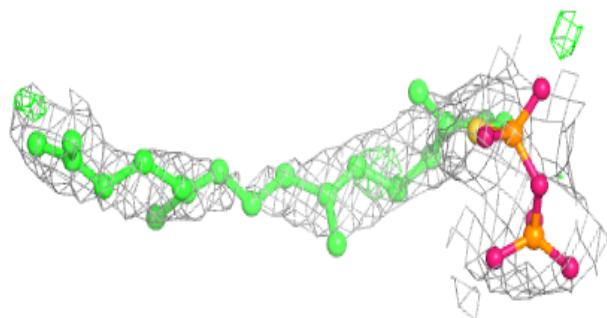
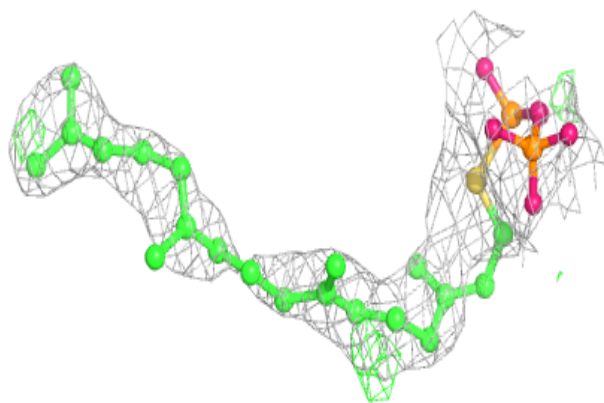
**Electron density around GGS A 1001:**

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

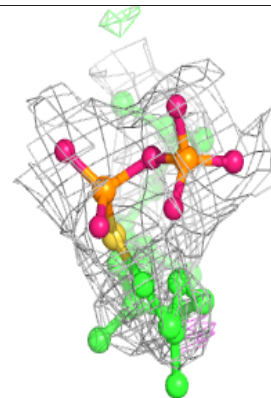
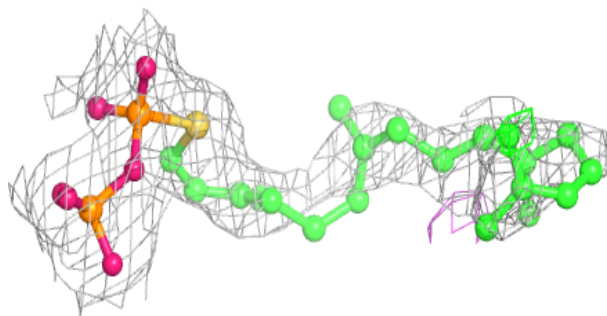
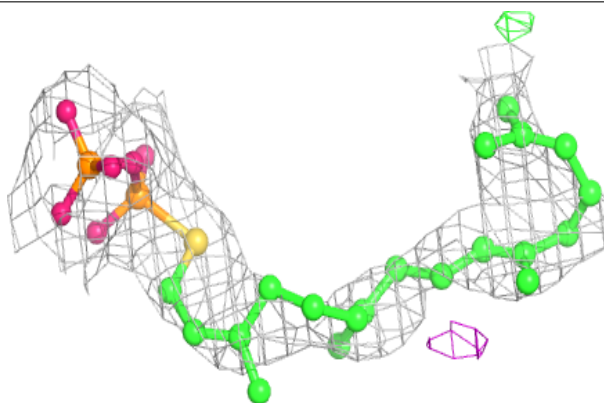


**Electron density around GGS B 1004:**

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

**Electron density around GGS B 1003:**

$2mF_o-DF_c$  (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.