



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 25, 2020 – 03:42 PM BST

PDB ID : 3KAK  
Title : Structure of homogluthathione synthetase from Glycine max in open conformation with gamma-glutamyl-cysteine bound.  
Authors : Galant, A.; Arkus, K.A.J.; Zubieta, C.; Cahoon, R.E.; Jez, J.M.  
Deposited on : 2009-10-19  
Resolution : 2.11 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](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 : 2.13  
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) : 2.13

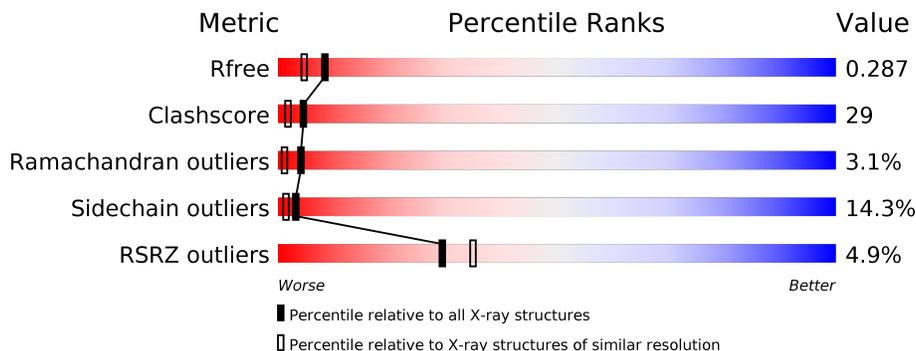
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.11 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	6241 (2.14-2.10)
Clashscore	141614	6778 (2.14-2.10)
Ramachandran outliers	138981	6705 (2.14-2.10)
Sidechain outliers	138945	6706 (2.14-2.10)
RSRZ outliers	127900	6112 (2.14-2.10)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	499	 3% 50% 28% 8% • 11%
1	B	499	 5% 50% 26% 8% • 12%

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	3GC	A	501	X	-	X	-
2	3GC	B	501	X	-	-	-
2	3GC	B	502	X	-	X	-

## 2 Entry composition [i](#)

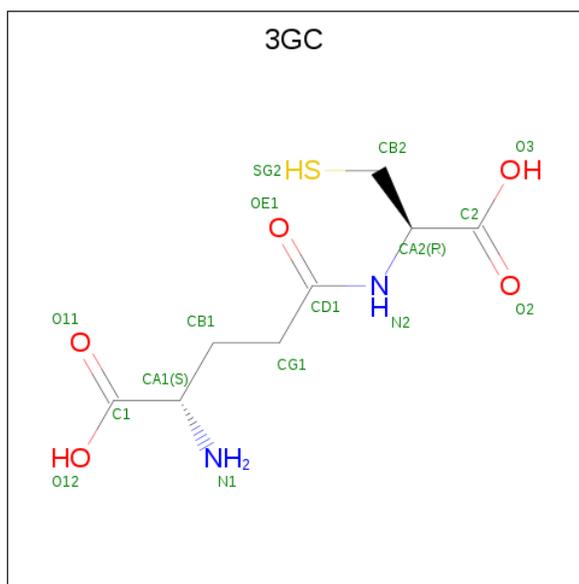
There are 3 unique types of molecules in this entry. The entry contains 7414 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Homoglutathione synthetase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	443	Total 3574	C 2275	N 619	O 665	S 15	0	6	0
1	B	437	Total 3511	C 2230	N 608	O 658	S 15	0	6	0

- Molecule 2 is GAMMA-GLUTAMYLCYSTEINE (three-letter code: 3GC) (formula:  $C_8H_{14}N_2O_5S$ ).

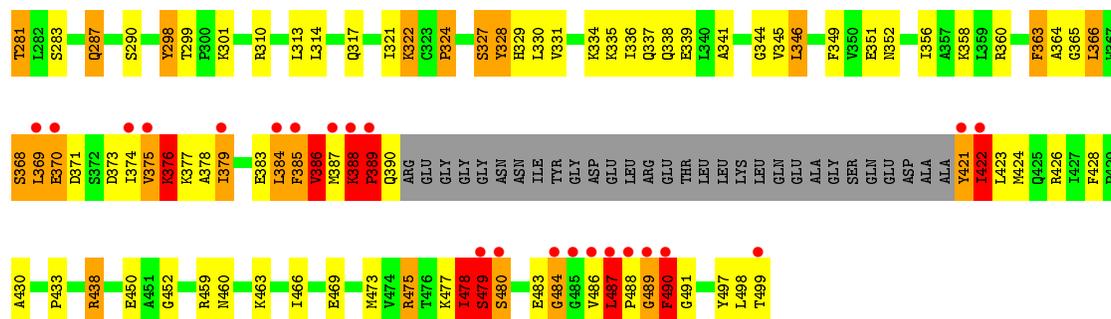


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
2	A	1	Total 16	C 8	N 2	O 5	S 1	0	0
2	B	1	Total 16	C 8	N 2	O 5	S 1	0	0
2	B	1	Total 16	C 8	N 2	O 5	S 1	0	0

- Molecule 3 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
3	A	130	Total 130	O 130	0	0
3	B	151	Total 151	O 151	0	0





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	64.88Å 80.95Å 89.12Å 90.00° 95.96° 90.00°	Depositor
Resolution (Å)	19.80 – 2.11 19.80 – 2.11	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.80-2.11) 99.7 (19.80-2.11)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.92 (at 2.11Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.207 , 0.288 0.210 , 0.287	Depositor DCC
$R_{free}$ test set	2635 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.8	Xtrriage
Anisotropy	0.188	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 56.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7414	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 20.16 % 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 9.2346e-03. 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: 3GC

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	1.30	20/3639 (0.5%)	1.25	37/4913 (0.8%)
1	B	1.30	10/3578 (0.3%)	1.27	26/4835 (0.5%)
All	All	1.30	30/7217 (0.4%)	1.26	63/9748 (0.6%)

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	7
1	B	0	10
All	All	0	17

The worst 5 of 30 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	383	GLU	CB-CG	10.16	1.71	1.52
1	B	328	TYR	CD1-CE1	8.04	1.51	1.39
1	A	54	CYS	CB-SG	7.89	1.95	1.82
1	A	383	GLU	CG-CD	7.45	1.63	1.51
1	A	402	GLU	CB-CG	7.02	1.65	1.52

The worst 5 of 63 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	148	ARG	NE-CZ-NH2	-13.07	113.77	120.30
1	B	148	ARG	NE-CZ-NH1	9.48	125.04	120.30
1	A	264	ARG	NE-CZ-NH1	-8.97	115.82	120.30
1	A	255	ARG	NE-CZ-NH2	8.16	124.38	120.30

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	148	ARG	NE-CZ-NH1	8.16	124.38	120.30

There are no chirality outliers.

5 of 17 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	28	LEU	Peptide
1	A	373	ASP	Peptide
1	A	374	ILE	Peptide
1	A	375	VAL	Peptide
1	A	380	GLU	Mainchain

## 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	3574	0	3616	173	0
1	B	3511	0	3541	236	0
2	A	16	0	11	7	0
2	B	32	0	22	17	0
3	A	130	0	0	15	0
3	B	151	0	0	29	0
All	All	7414	0	7190	409	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 29.

The worst 5 of 409 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:384:LEU:O	1:B:426:ARG:HB2	1.26	1.29
1:A:398:ILE:HG21	1:A:405:GLU:OE1	1.28	1.28
1:A:39:LEU:HD12	1:A:442:TRP:CZ3	1.69	1.25
1:B:374:ILE:O	1:B:376:LYS:N	1.72	1.19
1:B:479:SER:HB3	3:B:640:HOH:O	1.51	1.11

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	437/499 (88%)	406 (93%)	23 (5%)	8 (2%)	8	3
1	B	437/499 (88%)	387 (89%)	31 (7%)	19 (4%)	2	0
All	All	874/998 (88%)	793 (91%)	54 (6%)	27 (3%)	4	1

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	404	ARG
1	A	461	LYS
1	B	197	LYS
1	B	370	GLU
1	B	375	VAL

### 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	397/431 (92%)	337 (85%)	60 (15%)	3	1
1	B	390/431 (90%)	336 (86%)	54 (14%)	3	1
All	All	787/862 (91%)	673 (86%)	114 (14%)	3	1

5 of 114 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	421	TYR
1	B	60	LYS
1	B	422	ILE
1	A	431	THR
1	A	477	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	40	GLN
1	B	142	ASN
1	B	287	GLN
1	A	468	ASN
1	B	337	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](#)

3 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	3GC	B	501	-	8,15,15	2.20	2 (25%)	8,19,19	3.39	8 (100%)
2	3GC	A	501	-	8,15,15	2.26	2 (25%)	8,19,19	3.51	2 (25%)
2	3GC	B	502	-	8,15,15	3.18	2 (25%)	8,19,19	5.82	8 (100%)

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	3GC	B	501	-	1/1/5/6	2/11/19/19	-
2	3GC	A	501	-	1/1/5/6	5/11/19/19	-
2	3GC	B	502	-	1/1/5/6	8/11/19/19	-

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	502	3GC	OE1-CD1	7.67	1.38	1.23
2	A	501	3GC	OE1-CD1	4.95	1.33	1.23
2	B	501	3GC	OE1-CD1	4.81	1.33	1.23
2	B	502	3GC	CD1-N2	-4.04	1.25	1.34
2	A	501	3GC	CD1-N2	-3.51	1.26	1.34

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	501	3GC	CB2-CA2-N2	-9.63	97.55	111.28
2	B	502	3GC	CA2-CB2-SG2	-8.45	104.69	114.19
2	B	502	3GC	CG1-CD1-N2	-8.37	101.31	115.83
2	B	502	3GC	CB1-CG1-CD1	-6.10	99.41	113.04
2	B	501	3GC	CA2-N2-CD1	-6.06	113.68	123.33

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	B	501	3GC	CA1
2	A	501	3GC	CA1
2	B	502	3GC	CA1

5 of 15 torsion outliers are listed below:

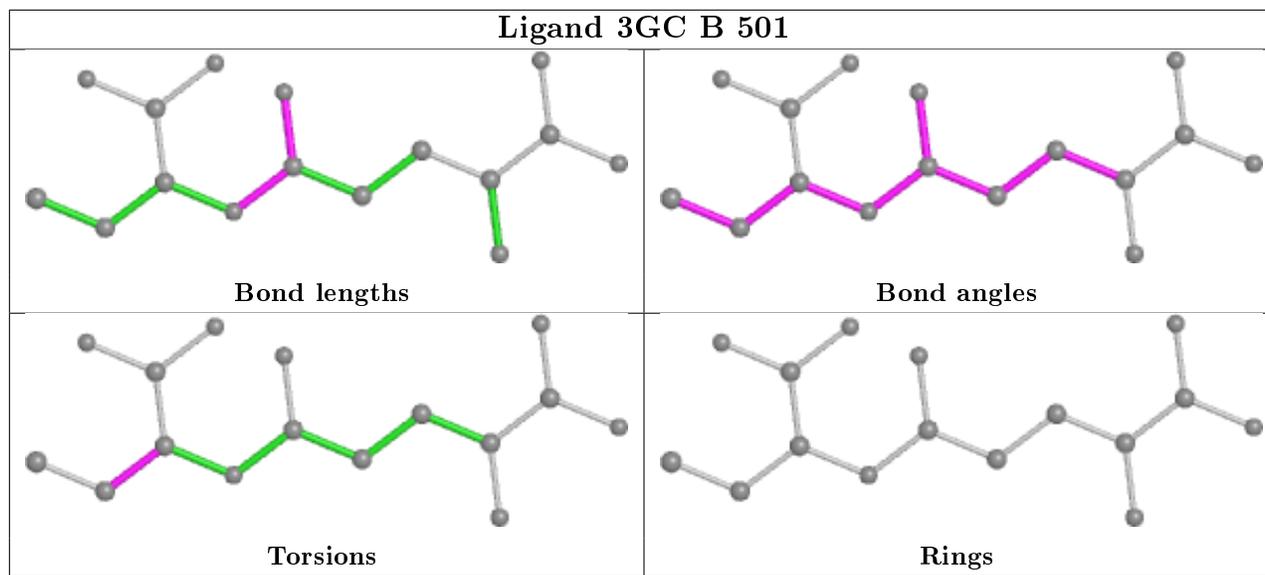
Mol	Chain	Res	Type	Atoms
2	B	501	3GC	N2-CA2-CB2-SG2
2	B	501	3GC	C2-CA2-CB2-SG2
2	A	501	3GC	C1-CA1-CB1-CG1
2	A	501	3GC	N2-CA2-CB2-SG2
2	A	501	3GC	C2-CA2-CB2-SG2

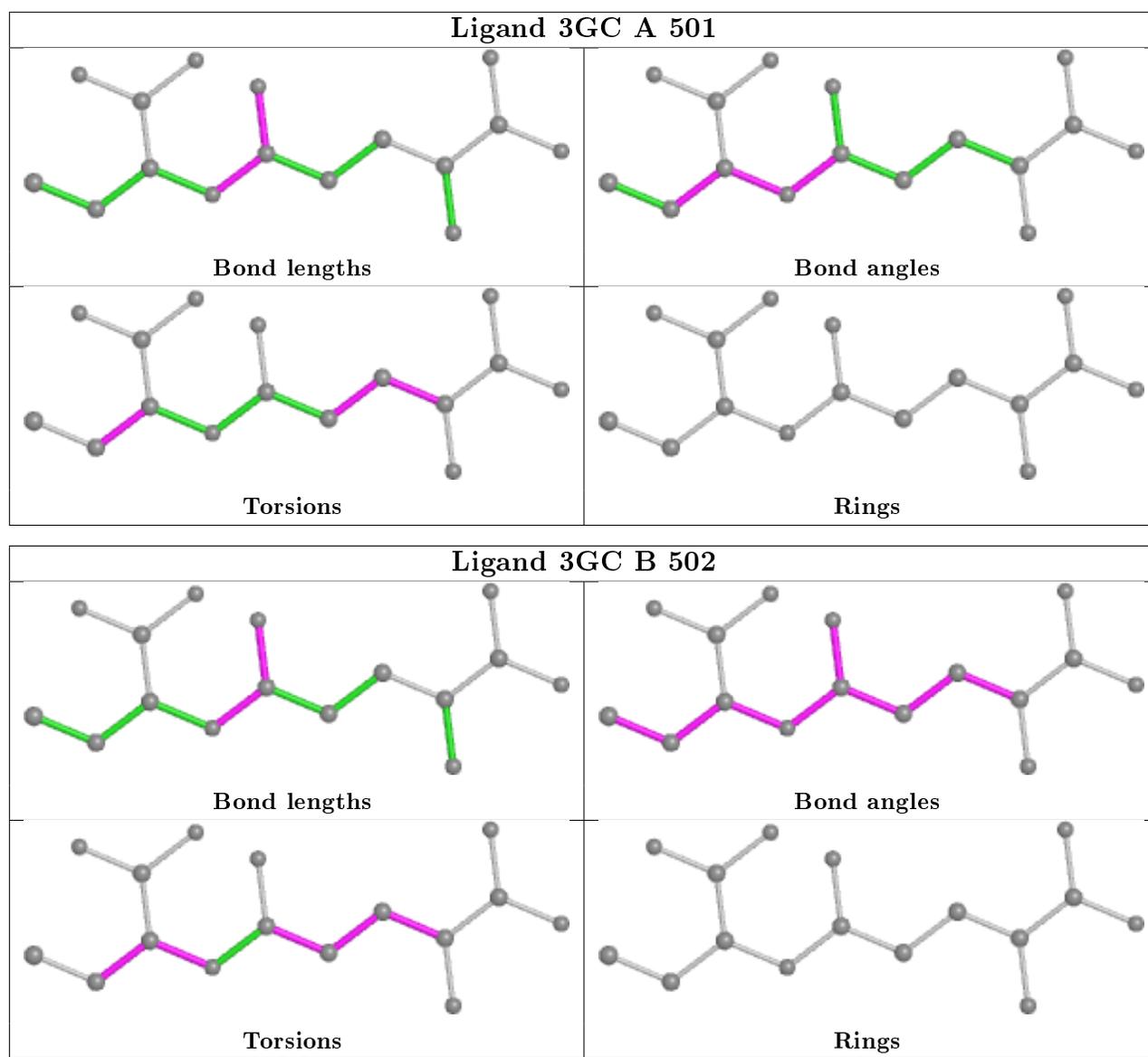
There are no ring outliers.

3 monomers are involved in 24 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	501	3GC	5	0
2	A	501	3GC	7	0
2	B	502	3GC	12	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 [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, 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	443/499 (88%)	0.09	16 (3%) 42 49	14, 27, 46, 71	0
1	B	437/499 (87%)	0.14	27 (6%) 20 25	12, 25, 59, 78	0
All	All	880/998 (88%)	0.11	43 (4%) 29 35	12, 26, 52, 78	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	484	GLY	9.2
1	B	487	LEU	7.1
1	B	369	LEU	6.7
1	B	375	VAL	6.3
1	A	374	ILE	5.6

### 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](#)

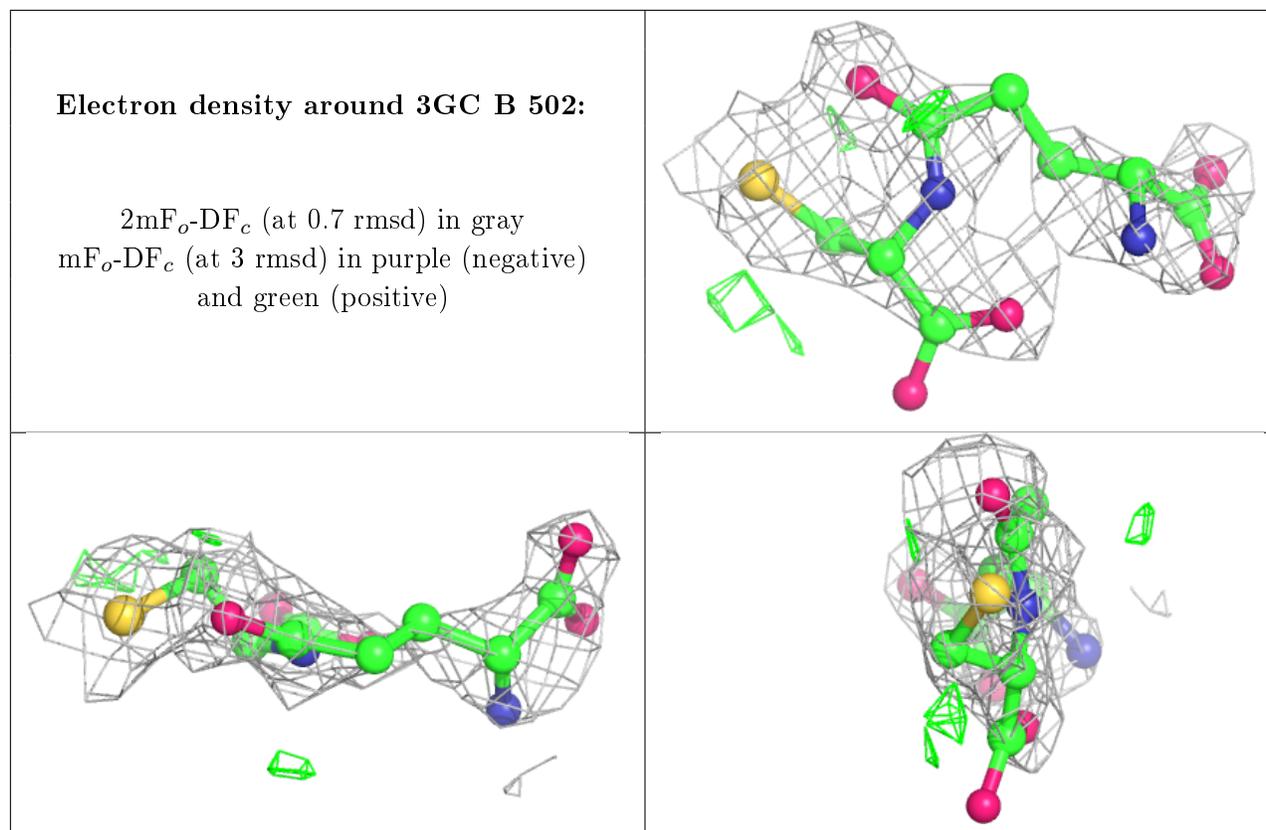
There are no monosaccharides in this entry.

### 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.

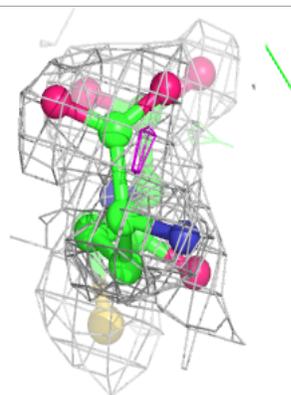
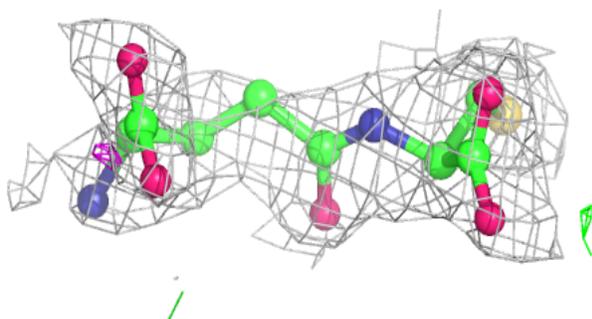
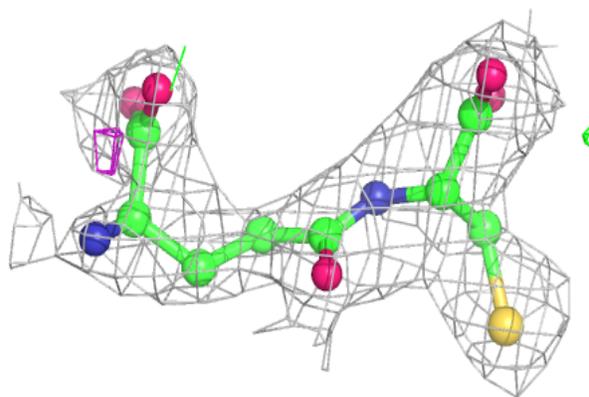
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	3GC	B	502	16/16	0.88	0.22	13,20,24,24	16
2	3GC	A	501	16/16	0.89	0.18	46,51,57,61	0
2	3GC	B	501	16/16	0.89	0.18	35,43,47,50	0

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.

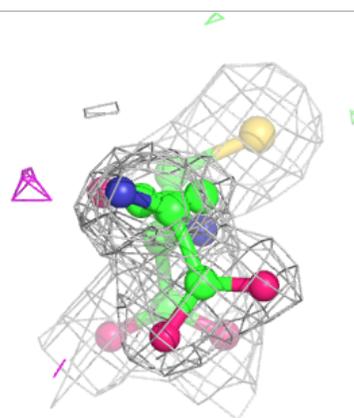
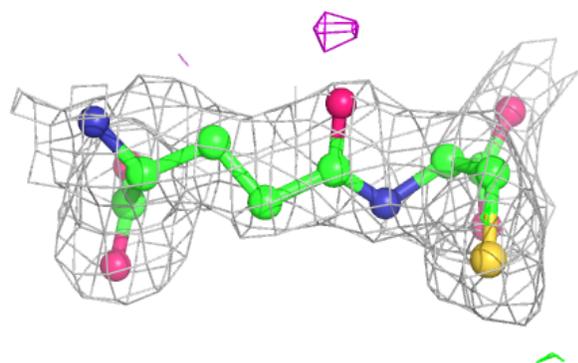
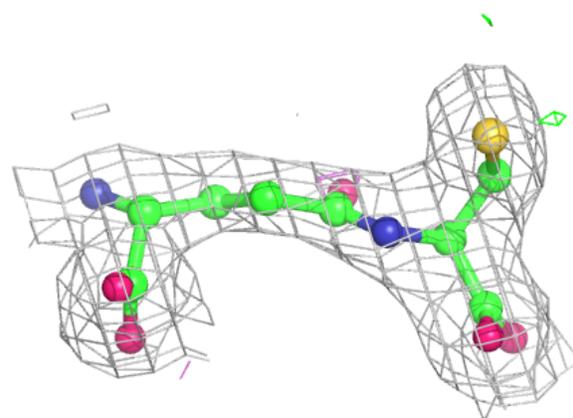


**Electron density around 3GC A 501:**

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

**Electron density around 3GC 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)



## 6.5 Other polymers

There are no such residues in this entry.