



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 03:06 am BST

PDB ID : 6OHC  
Title : E. coli Guanine Deaminase  
Authors : Shek, R.S.; French, J.B.  
Deposited on : 2019-04-05  
Resolution : 2.30 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 : 2.11  
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) : 2.11

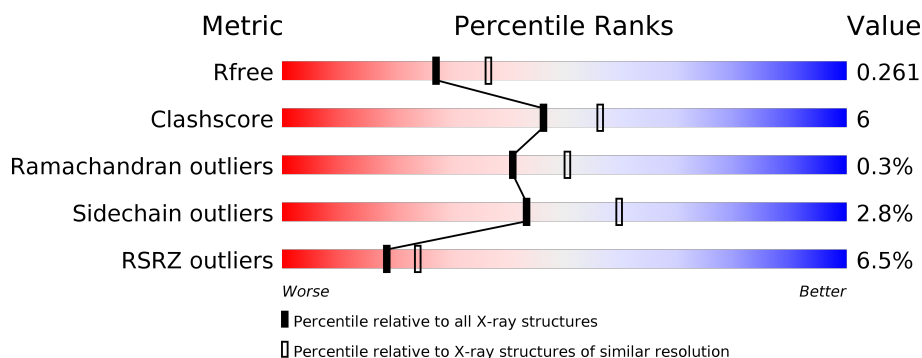
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	439	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, green 83%, yellow 14%, grey 2%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>%</span> <span>83%</span> <span>14%</span> <span>• •</span> </div> </div>
1	B	439	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, green 81%, yellow 16%, grey 2%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>%</span> <span>81%</span> <span>16%</span> <span>• •</span> </div> </div>
1	C	439	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 13%, green 86%, yellow 10%, grey 2%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>13%</span> <span>86%</span> <span>10%</span> <span>• •</span> </div> </div>
1	D	439	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 10%, green 84%, yellow 12%, grey 2%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>10%</span> <span>84%</span> <span>12%</span> <span>• •</span> </div> </div>

## 2 Entry composition [i](#)

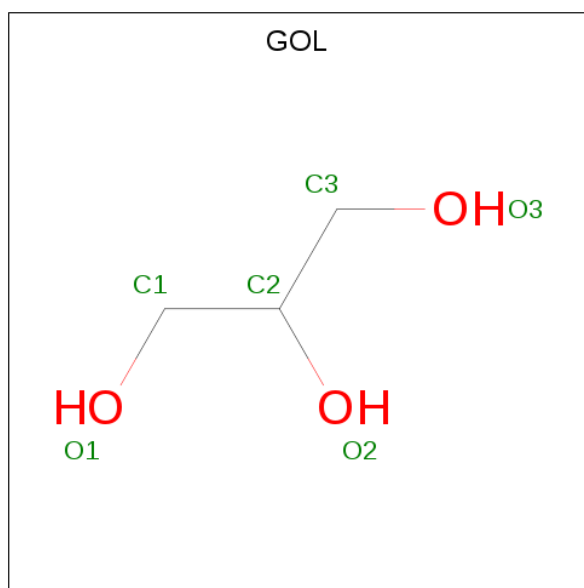
There are 4 unique types of molecules in this entry. The entry contains 13442 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Guanine deaminase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	428	Total	C	N	O	S	0	1	0
			3390	2167	573	633	17			
1	B	430	Total	C	N	O	S	0	0	0
			3428	2191	581	639	17			
1	C	425	Total	C	N	O	S	0	0	0
			3059	1948	532	565	14			
1	D	427	Total	C	N	O	S	0	0	0
			3187	2030	548	595	14			

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			6	3	3		
2	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total 1	Zn 1	0	0
3	A	1	Total 1	Zn 1	0	0
3	D	1	Total 1	Zn 1	0	0
3	C	1	Total 1	Zn 1	0	0

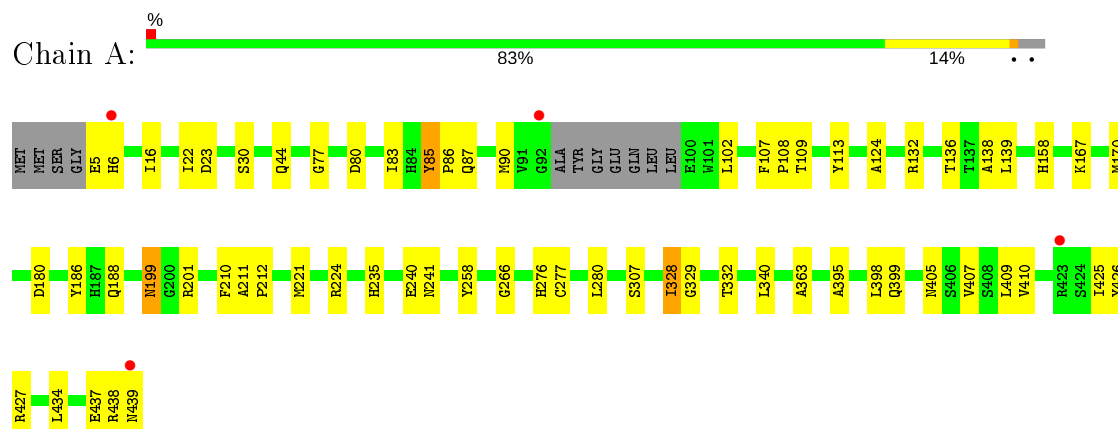
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	134	Total 134	O 134	0	0
4	B	148	Total 148	O 148	0	0
4	C	19	Total 19	O 19	0	0
4	D	61	Total 61	O 61	0	0

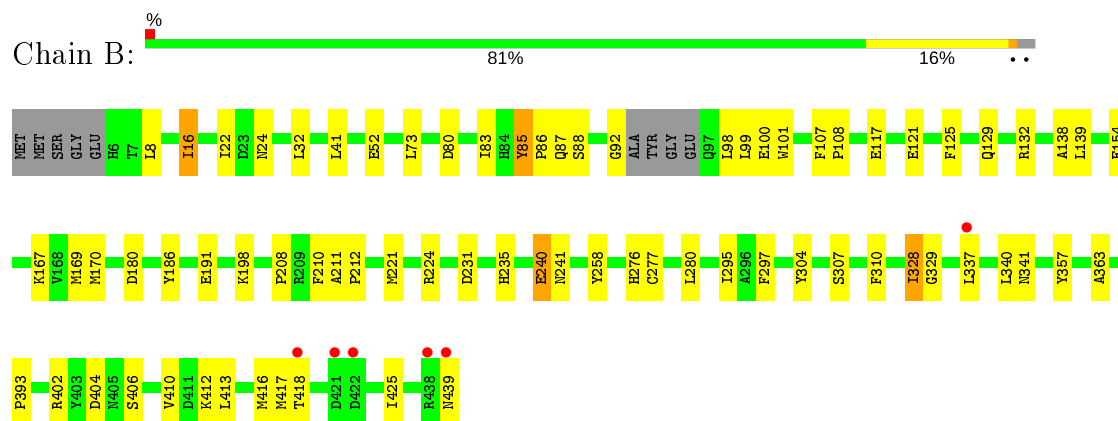
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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 ( $\text{RSRZ} > 2$ ). 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.

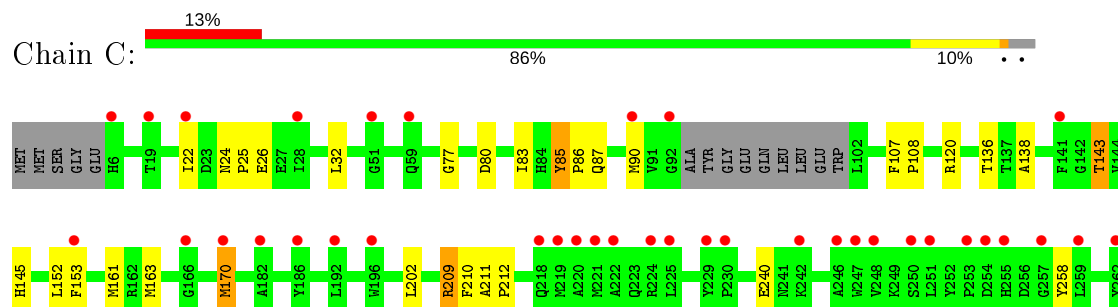
#### • Molecule 1: Guanine deaminase

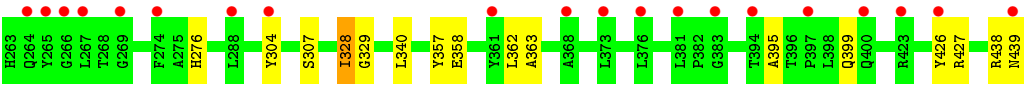


#### • Molecule 1: Guanine deaminase

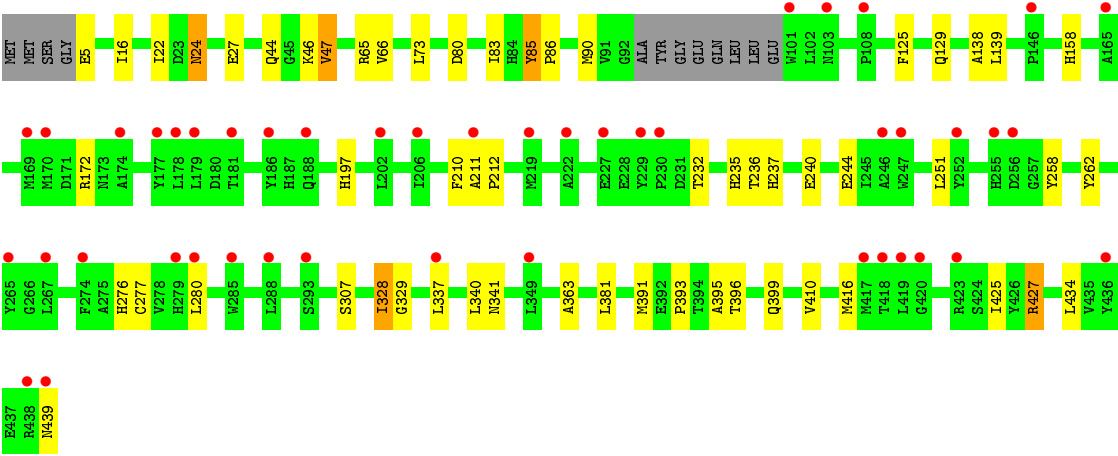
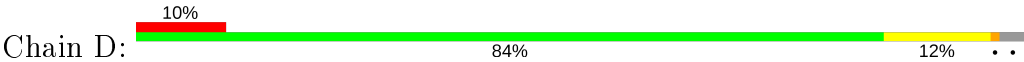


#### • Molecule 1: Guanine deaminase





● Molecule 1: Guanine deaminase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.63 Å   80.59 Å   101.43 Å 104.80°   105.72°   105.83°	Depositor
Resolution (Å)	48.91 – 2.30 48.91 – 2.30	Depositor EDS
% Data completeness (in resolution range)	97.8 (48.91-2.30) 97.8 (48.91-2.30)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.52 (at 2.29 Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, $R_{free}$	0.209   ,   0.258 0.213   ,   0.261	Depositor DCC
$R_{free}$ test set	3958 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	46.6	Xtriage
Anisotropy	0.172	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 48.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	13442	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.32% 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

### 5.1 Standard geometry

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

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	0.78	0/3477	0.93	1/4722 (0.0%)
1	B	0.82	6/3512 (0.2%)	0.92	3/4761 (0.1%)
1	C	0.72	0/3136	0.83	0/4285
1	D	0.71	0/3267	0.85	1/4459 (0.0%)
All	All	0.76	6/13392 (0.0%)	0.89	5/18227 (0.0%)

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	1
1	B	0	1
1	C	0	1
1	D	0	1
All	All	0	4

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	52	GLU	CD-OE1	6.66	1.32	1.25
1	B	191	GLU	CD-OE1	6.15	1.32	1.25
1	B	117	GLU	CD-OE1	5.82	1.32	1.25
1	B	154	GLU	CD-OE2	5.74	1.31	1.25
1	B	121	GLU	CD-OE2	5.48	1.31	1.25

All (5) bond angle outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	132	ARG	CG-CD-NE	-5.76	99.71	111.80
1	A	201	ARG	NE-CZ-NH1	5.44	123.02	120.30
1	B	304	TYR	CB-CG-CD1	5.25	124.15	121.00
1	B	304	TYR	CB-CG-CD2	-5.24	117.86	121.00
1	D	427	ARG	NE-CZ-NH2	-5.05	117.77	120.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	328	ILE	Peptide
1	B	328	ILE	Peptide
1	C	328	ILE	Peptide
1	D	328	ILE	Peptide

## 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	3390	0	3234	47	0
1	B	3428	0	3312	48	0
1	C	3059	0	2604	31	0
1	D	3187	0	2846	42	0
2	A	6	0	8	0	0
2	B	6	0	8	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	134	0	0	4	0
4	B	148	0	0	0	0
4	C	19	0	0	0	0
4	D	61	0	0	1	0
All	All	13442	0	12012	157	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 6.

The worst 5 of 157 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:16:ILE:HD11	1:B:32:LEU:HD11	1.33	1.07
1:A:186:TYR:HB2	1:A:221:MET:HE2	1.49	0.92
1:B:8:LEU:HD12	1:B:41:LEU:HD22	1.60	0.83
1:B:16:ILE:CD1	1:B:32:LEU:HD11	2.11	0.80
1:C:358:GLU:O	1:C:362:LEU:HD23	1.81	0.80

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	425/439 (97%)	411 (97%)	13 (3%)	1 (0%)	47	58
1	B	426/439 (97%)	411 (96%)	13 (3%)	2 (0%)	29	35
1	C	421/439 (96%)	408 (97%)	12 (3%)	1 (0%)	47	58
1	D	423/439 (96%)	409 (97%)	13 (3%)	1 (0%)	47	58
All	All	1695/1756 (96%)	1639 (97%)	51 (3%)	5 (0%)	41	50

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	99	LEU
1	A	276	HIS
1	B	276	HIS
1	C	276	HIS
1	D	276	HIS

### 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	351/379 (93%)	339 (97%)	12 (3%)	37	51
1	B	361/379 (95%)	353 (98%)	8 (2%)	52	69
1	C	256/379 (68%)	249 (97%)	7 (3%)	44	61
1	D	294/379 (78%)	286 (97%)	8 (3%)	44	61
All	All	1262/1516 (83%)	1227 (97%)	35 (3%)	43	60

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	170	MET
1	B	410	VAL
1	D	244	GLU
1	B	210	PHE
1	B	241	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	87	GLN
1	C	218	GLN
1	D	188	GLN
1	B	348	GLN
1	B	439	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

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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	GOL	A	501	-	5,5,5	0.08	0	5,5,5	0.36	0
2	GOL	B	501	-	5,5,5	0.29	0	5,5,5	0.47	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GOL	A	501	-	-	1/4/4/4	-
2	GOL	B	501	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	501	GOL	O1-C1-C2-O2

There are no ring outliers.

No monomer is involved in short contacts.

## 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	428/439 (97%)	-0.10	4 (0%) 84 88	36, 47, 65, 104	0
1	B	430/439 (97%)	-0.05	6 (1%) 75 80	34, 46, 62, 88	0
1	C	425/439 (96%)	0.75	57 (13%) 3 4	54, 86, 116, 131	0
1	D	427/439 (97%)	0.55	45 (10%) 6 8	39, 76, 122, 136	0
All	All	1710/1756 (97%)	0.29	112 (6%) 18 24	34, 57, 113, 136	0

The worst 5 of 112 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	222	ALA	9.5
1	C	257	GLY	8.1
1	C	253	PRO	5.8
1	C	247	TRP	5.7
1	A	439	ASN	5.4

### 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 carbohydrates 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	GOL	B	501	6/6	0.90	0.20	59,63,64,65	0
2	GOL	A	501	6/6	0.92	0.19	59,59,60,60	0
3	ZN	C	501	1/1	0.95	0.07	73,73,73,73	0
3	ZN	D	501	1/1	0.98	0.05	67,67,67,67	0
3	ZN	A	502	1/1	0.99	0.13	37,37,37,37	0
3	ZN	B	502	1/1	1.00	0.10	40,40,40,40	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.