



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 14, 2020 – 01:05 am BST

PDB ID : 1EC9  
Title : E. COLI GLUCARATE DEHYDRATASE BOUND TO XYLAROHYDROX-AMATE  
Authors : Gulick, A.M.; Hubbard, B.K.; Gerlt, J.A.; Rayment, I.  
Deposited on : 2000-01-25  
Resolution : 2.00 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

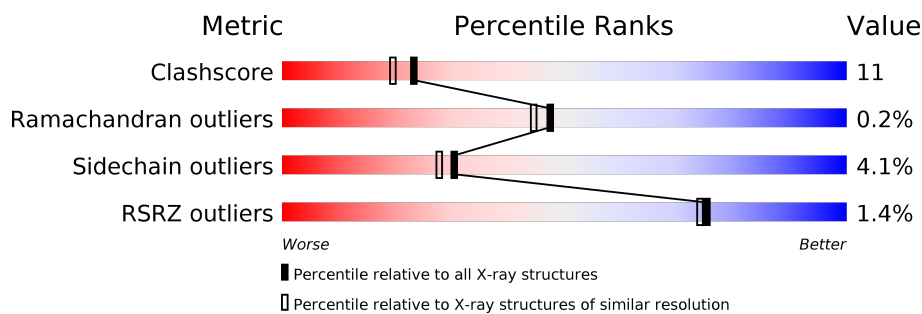
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	446	<div> <div>75%</div> <div>21%</div> <div>..</div> </div>
1	B	446	<div> <div>69%</div> <div>26%</div> <div>..</div> </div>
1	C	446	<div> <div>70%</div> <div>27%</div> <div>..</div> </div>
1	D	446	<div> <div>2%</div> <div>66%</div> <div>28%</div> <div>5%</div> <div>.</div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 14854 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.

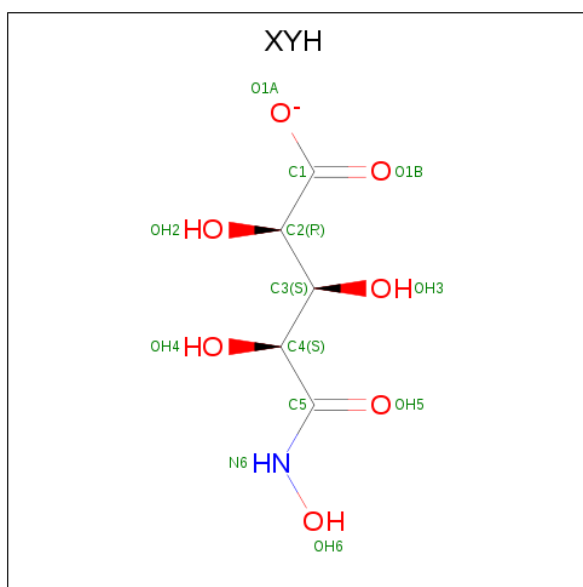
- Molecule 1 is a protein called GLUCARATE DEHYDRATASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	443	Total	C	N	O	S	0	0	0
			3426	2164	600	641	21			
1	B	443	Total	C	N	O	S	0	0	0
			3404	2152	594	637	21			
1	C	443	Total	C	N	O	S	0	1	0
			3428	2166	599	642	21			
1	D	444	Total	C	N	O	S	0	0	0
			3427	2166	602	638	21			

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

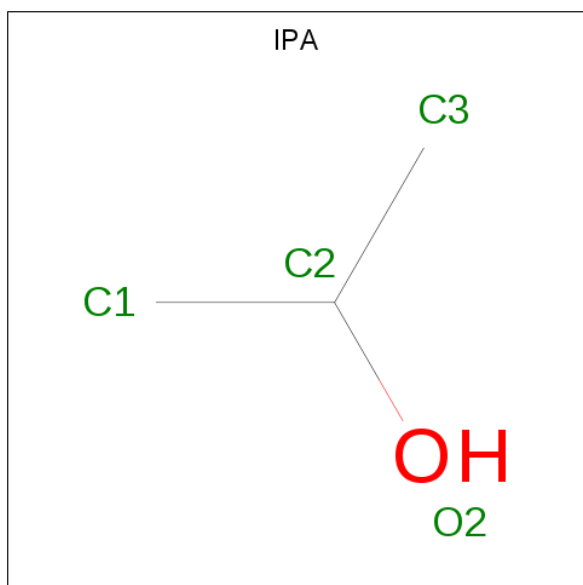
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Mg	0	0
			1	1		
2	A	1	Total	Mg	0	0
			1	1		
2	D	1	Total	Mg	0	0
			1	1		
2	C	1	Total	Mg	0	0
			1	1		

- Molecule 3 is XYLAROHYDROXAMATE (three-letter code: XYH) (formula: C<sub>5</sub>H<sub>8</sub>NO<sub>7</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			13	5	1	7		
3	B	1	Total	C	N	O	0	0
			13	5	1	7		
3	C	1	Total	C	N	O	0	0
			13	5	1	7		
3	D	1	Total	C	N	O	0	0
			13	5	1	7		

- Molecule 4 is ISOPROPYL ALCOHOL (three-letter code: IPA) (formula: C<sub>3</sub>H<sub>8</sub>O).



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

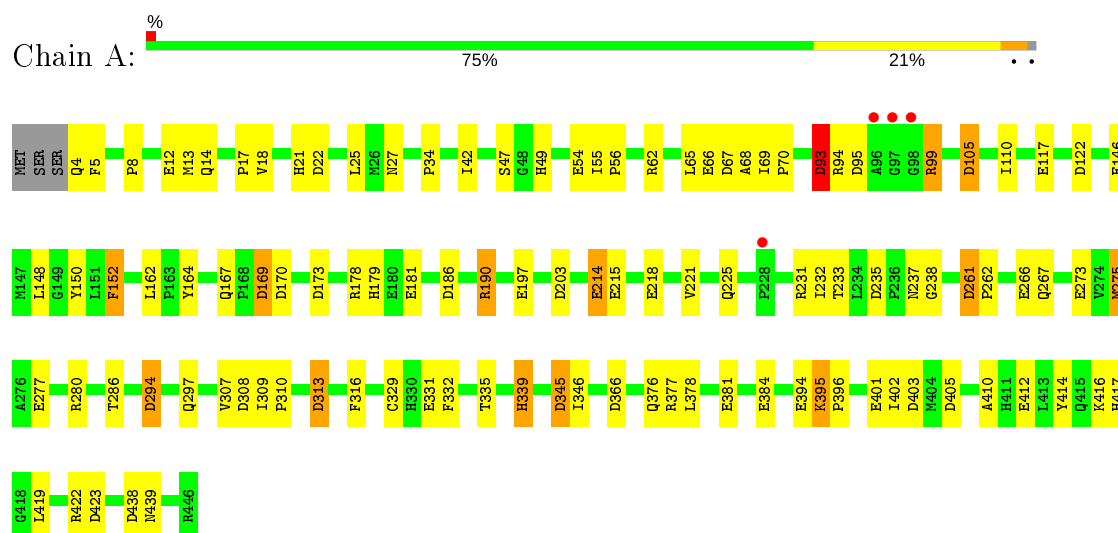
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	316	Total O 316 316	0	0
5	B	231	Total O 231 231	0	0
5	C	302	Total O 302 302	0	0
5	D	248	Total O 248 248	0	0

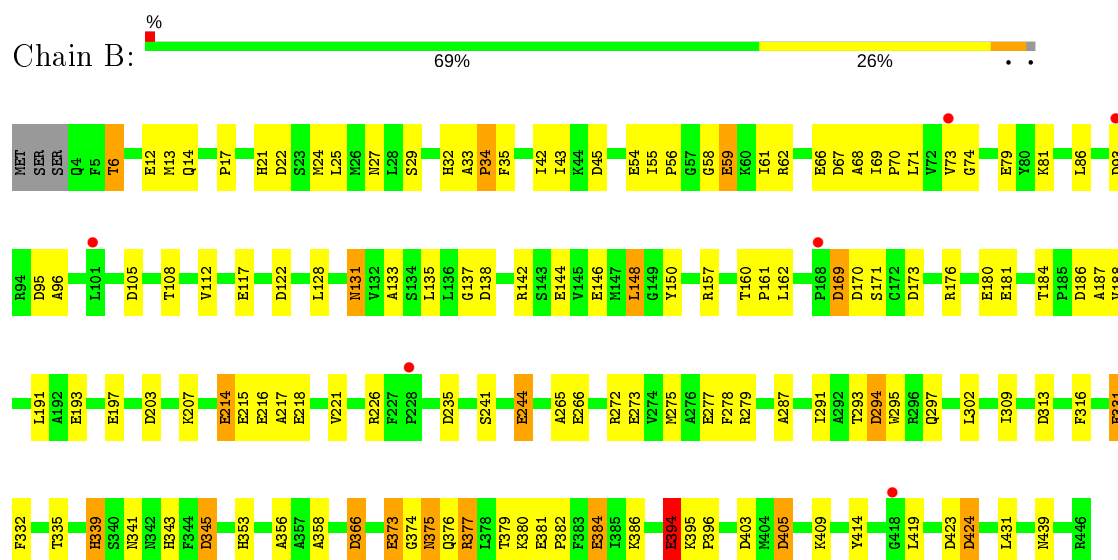
### 3 Residue-property plots

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 ( $RSRZ > 2$ ). 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: GLUCARATE DEHYDRATASE

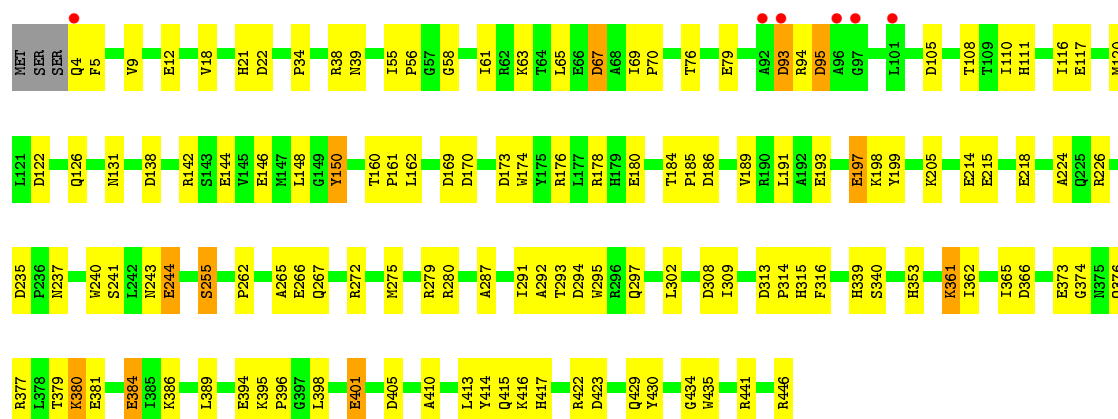


#### • Molecule 1: GLUCARATE DEHYDRATASE

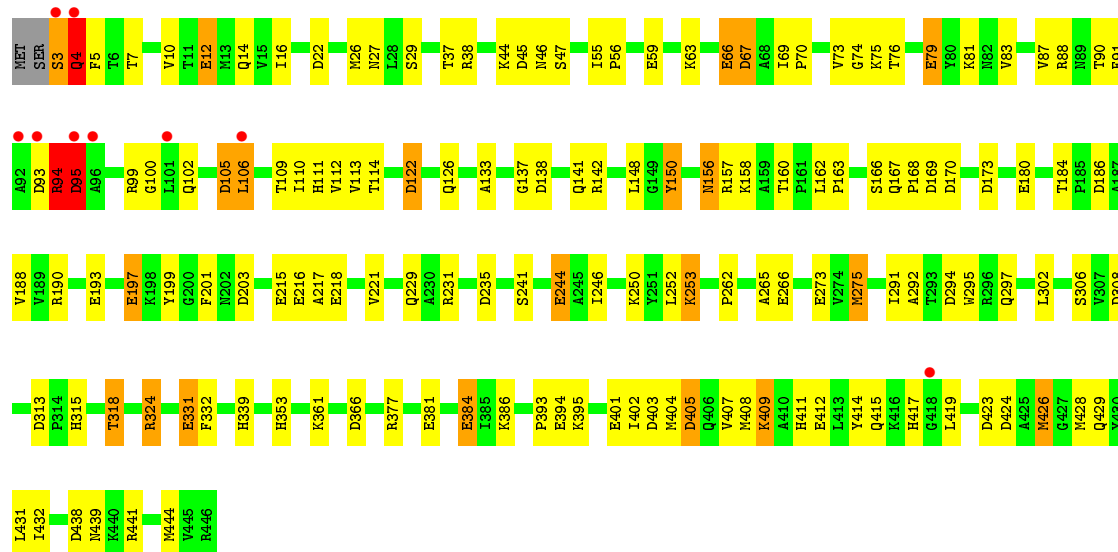


#### • Molecule 1: GLUCARATE DEHYDRATASE





### • Molecule 1: GLUCARATE DEHYDRATASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	71.29Å 84.84Å 98.99Å 103.14° 94.31° 113.20°	Depositor
Resolution (Å)	30.00 – 2.00 19.93 – 2.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) (30.00-2.00) 94.9 (19.93-2.00)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.20 (at 2.01Å)	Xtriage
Refinement program	TNT 5E	Depositor
R, $R_{free}$	0.179 , 0.281 0.174 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	21.0	Xtriage
Anisotropy	0.135	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 82.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	14854	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.44% 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: XYH, MG, IPA

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	1.01	20/3505 (0.6%)	1.23	42/4752 (0.9%)
1	B	0.97	24/3483 (0.7%)	1.21	42/4726 (0.9%)
1	C	0.98	13/3507 (0.4%)	1.19	38/4755 (0.8%)
1	D	0.97	18/3506 (0.5%)	1.19	42/4752 (0.9%)
All	All	0.98	75/14001 (0.5%)	1.21	164/18985 (0.9%)

The worst 5 of 75 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	214	GLU	CD-OE2	8.84	1.35	1.25
1	A	215	GLU	CD-OE2	7.69	1.34	1.25
1	B	54	GLU	CD-OE2	7.57	1.33	1.25
1	A	214	GLU	CD-OE2	6.76	1.33	1.25
1	C	384	GLU	CD-OE2	6.74	1.33	1.25

The worst 5 of 164 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	294	ASP	CB-CG-OD2	-9.46	109.78	118.30
1	A	377	ARG	NE-CZ-NH1	9.03	124.81	120.30
1	A	294	ASP	CB-CG-OD1	8.93	126.34	118.30
1	C	366	ASP	CB-CG-OD2	-8.73	110.44	118.30
1	B	366	ASP	CB-CG-OD2	-8.72	110.45	118.30

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	3426	0	3347	60	0
1	B	3404	0	3309	84	0
1	C	3428	0	3346	67	0
1	D	3427	0	3352	110	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	13	0	7	3	0
3	B	13	0	7	1	0
3	C	13	0	6	1	0
3	D	13	0	6	0	0
4	A	8	0	16	3	0
4	B	4	0	8	2	0
4	D	4	0	8	3	0
5	A	316	0	0	8	0
5	B	231	0	0	3	0
5	C	302	0	0	6	0
5	D	248	0	0	4	0
All	All	14854	0	13412	307	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 11.

The worst 5 of 307 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:D:100:GLY:HA3	1:D:106:LEU:HD12	1.43	0.95
1:B:293:THR:H	1:B:297:GLN:HE21	1.09	0.95
1:B:128:LEU:HD23	1:D:7:THR:HG21	1.50	0.93
1:B:302:LEU:HD22	4:B:604:IPA:H33	1.53	0.90
1:D:384:GLU:HB2	1:D:386:LYS:HE3	1.56	0.88

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	441/446 (99%)	432 (98%)	8 (2%)	1 (0%)	47	44
1	B	441/446 (99%)	423 (96%)	18 (4%)	0	100	100
1	C	442/446 (99%)	426 (96%)	16 (4%)	0	100	100
1	D	442/446 (99%)	428 (97%)	11 (2%)	3 (1%)	22	16
All	All	1766/1784 (99%)	1709 (97%)	53 (3%)	4 (0%)	47	44

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	4	GLN
1	D	94	ARG
1	A	93	ASP
1	D	46	ASN

### 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	357/362 (99%)	347 (97%)	10 (3%)	43	44
1	B	352/362 (97%)	334 (95%)	18 (5%)	24	19
1	C	357/362 (99%)	345 (97%)	12 (3%)	37	36
1	D	356/362 (98%)	338 (95%)	18 (5%)	24	19
All	All	1422/1448 (98%)	1364 (96%)	58 (4%)	30	28

5 of 58 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	394	GLU
1	C	191	LEU
1	D	306	SER
1	B	439	ASN
1	C	12	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 46 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	353	HIS
1	C	39	ASN
1	D	156	ASN
1	B	375	ASN
1	B	439	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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 12 ligands modelled in this entry, 4 are monoatomic - leaving 8 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
4	IPA	D	602	-	3,3,3	0.56	0	3,3,3	0.60	0
3	XYH	D	502	2	9,12,12	3.31	1 (11%)	10,16,16	3.95	6 (60%)
3	XYH	C	501	2	9,12,12	2.73	2 (22%)	10,16,16	5.90	5 (50%)
4	IPA	A	601	-	3,3,3	0.44	0	3,3,3	0.40	0
3	XYH	A	499	2	9,12,12	2.93	2 (22%)	10,16,16	5.53	8 (80%)
3	XYH	B	500	2	9,12,12	3.10	2 (22%)	10,16,16	3.84	6 (60%)
4	IPA	B	604	-	3,3,3	0.64	0	3,3,3	0.28	0
4	IPA	A	603	-	3,3,3	0.51	0	3,3,3	0.43	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
3	XYH	C	501	2	-	7/14/18/18	-
3	XYH	A	499	2	-	4/14/18/18	-
3	XYH	D	502	2	-	7/14/18/18	-
3	XYH	B	500	2	-	7/14/18/18	-

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	502	XYH	OH5-C5	9.69	1.42	1.23
3	B	500	XYH	OH5-C5	8.70	1.40	1.23
3	A	499	XYH	OH5-C5	8.08	1.39	1.23
3	C	501	XYH	OH5-C5	7.32	1.37	1.23
3	B	500	XYH	C5-N6	-2.55	1.30	1.33

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	501	XYH	OH4-C4-C5	-15.44	77.36	110.63
3	A	499	XYH	OH4-C4-C5	-14.87	78.60	110.63
3	B	500	XYH	OH4-C4-C5	-10.07	88.93	110.63
3	D	502	XYH	OH4-C4-C5	-9.06	91.10	110.63
3	C	501	XYH	OH4-C4-C3	8.81	129.13	110.45

There are no chirality outliers.

5 of 25 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	502	XYH	C1-C2-C3-C4
3	D	502	XYH	OH2-C2-C3-C4
3	D	502	XYH	C3-C4-C5-OH5
3	D	502	XYH	C3-C4-C5-N6
3	C	501	XYH	C1-C2-C3-C4

There are no ring outliers.

7 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	602	IPA	3	0
3	C	501	XYH	1	0
4	A	601	IPA	2	0
3	A	499	XYH	3	0
3	B	500	XYH	1	0
4	B	604	IPA	2	0
4	A	603	IPA	1	0

## 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	443/446 (99%)	-0.55	4 (0%) 84 83	15, 24, 51, 87	0
1	B	443/446 (99%)	-0.40	6 (1%) 75 74	15, 30, 60, 88	0
1	C	443/446 (99%)	-0.53	6 (1%) 75 74	14, 24, 50, 83	0
1	D	444/446 (99%)	-0.34	9 (2%) 65 63	14, 28, 61, 100	0
All	All	1773/1784 (99%)	-0.45	25 (1%) 75 74	14, 27, 57, 100	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	3	SER	7.7
1	D	93	ASP	5.7
1	D	4	GLN	4.7
1	D	92	ALA	4.4
1	C	96	ALA	3.7

### 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
3	XYH	A	499	13/13	0.94	0.12	19,30,57,64	0
4	IPA	B	604	4/4	0.94	0.15	16,42,45,69	0
3	XYH	C	501	13/13	0.95	0.10	20,29,75,78	0
4	IPA	D	602	4/4	0.95	0.12	21,33,41,60	0
3	XYH	D	502	13/13	0.95	0.09	19,24,47,57	0
4	IPA	A	603	4/4	0.95	0.15	30,31,47,48	0
2	MG	B	498	1/1	0.97	0.03	24,24,24,24	0
3	XYH	B	500	13/13	0.97	0.10	21,26,55,61	0
2	MG	D	498	1/1	0.98	0.07	23,23,23,23	0
4	IPA	A	601	4/4	0.98	0.13	12,31,43,100	0
2	MG	A	498	1/1	0.99	0.06	21,21,21,21	0
2	MG	C	498	1/1	0.99	0.04	23,23,23,23	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.