



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 19, 2017 – 03:31 PM EDT

PDB ID : 5O3N  
Title : Crystal structure of E. cloacae 3,4-dihydroxybenzoic acid decarboxylase (AroY) reconstituted with prFMN  
Authors : Marshall, S.A.; Leys, D.  
Deposited on : unknown  
Resolution : 2.05 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

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030345

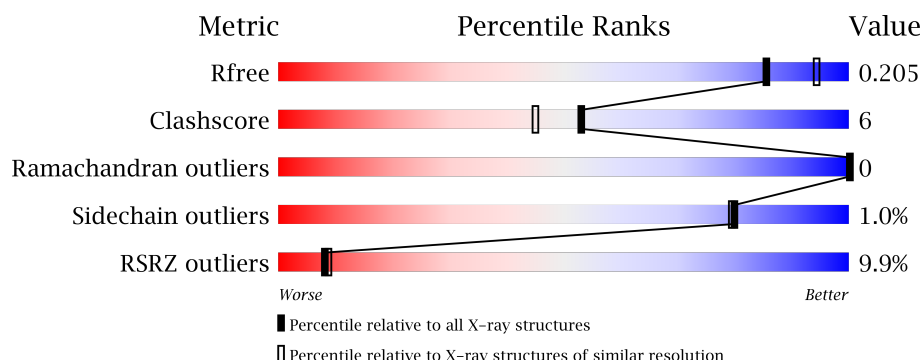
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.05 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))
$R_{free}$	100719	1316 (2.04-2.04)
Clashscore	112137	1394 (2.04-2.04)
Ramachandran outliers	110173	1383 (2.04-2.04)
Sidechain outliers	110143	1383 (2.04-2.04)
RSRZ outliers	101464	1319 (2.04-2.04)

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. 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	515	<div> <div>7%</div> <div>81%</div> <div>13%</div> <div>5%</div> </div>
1	B	515	<div> <div>12%</div> <div>80%</div> <div>14%</div> <div>5%</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
4	GOL	A	503	-	-	-	X
4	GOL	A	505	-	-	-	X
4	GOL	A	506	-	-	-	X
4	GOL	A	507	-	-	-	X
4	GOL	B	503	-	-	-	X
5	4LU	B	504	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 7934 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 3,4-dihydroxybenzoate decarboxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	488	Total	C	N	O	S	0	5	0
			3746	2371	658	698	19			
1	B	488	Total	C	N	O	S	0	2	0
			3723	2356	654	693	20			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	initiating methionine	UNP B2DCZ6
A	-18	GLY	-	expression tag	UNP B2DCZ6
A	-17	SER	-	expression tag	UNP B2DCZ6
A	-16	SER	-	expression tag	UNP B2DCZ6
A	-15	HIS	-	expression tag	UNP B2DCZ6
A	-14	HIS	-	expression tag	UNP B2DCZ6
A	-13	HIS	-	expression tag	UNP B2DCZ6
A	-12	HIS	-	expression tag	UNP B2DCZ6
A	-11	HIS	-	expression tag	UNP B2DCZ6
A	-10	HIS	-	expression tag	UNP B2DCZ6
A	-9	SER	-	expression tag	UNP B2DCZ6
A	-8	SER	-	expression tag	UNP B2DCZ6
A	-7	GLY	-	expression tag	UNP B2DCZ6
A	-6	LEU	-	expression tag	UNP B2DCZ6
A	-5	VAL	-	expression tag	UNP B2DCZ6
A	-4	PRO	-	expression tag	UNP B2DCZ6
A	-3	ARG	-	expression tag	UNP B2DCZ6
A	-2	GLY	-	expression tag	UNP B2DCZ6
A	-1	SER	-	expression tag	UNP B2DCZ6
A	0	HIS	-	expression tag	UNP B2DCZ6
B	-19	MET	-	initiating methionine	UNP B2DCZ6
B	-18	GLY	-	expression tag	UNP B2DCZ6
B	-17	SER	-	expression tag	UNP B2DCZ6
B	-16	SER	-	expression tag	UNP B2DCZ6
B	-15	HIS	-	expression tag	UNP B2DCZ6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-14	HIS	-	expression tag	UNP B2DCZ6
B	-13	HIS	-	expression tag	UNP B2DCZ6
B	-12	HIS	-	expression tag	UNP B2DCZ6
B	-11	HIS	-	expression tag	UNP B2DCZ6
B	-10	HIS	-	expression tag	UNP B2DCZ6
B	-9	SER	-	expression tag	UNP B2DCZ6
B	-8	SER	-	expression tag	UNP B2DCZ6
B	-7	GLY	-	expression tag	UNP B2DCZ6
B	-6	LEU	-	expression tag	UNP B2DCZ6
B	-5	VAL	-	expression tag	UNP B2DCZ6
B	-4	PRO	-	expression tag	UNP B2DCZ6
B	-3	ARG	-	expression tag	UNP B2DCZ6
B	-2	GLY	-	expression tag	UNP B2DCZ6
B	-1	SER	-	expression tag	UNP B2DCZ6
B	0	HIS	-	expression tag	UNP B2DCZ6

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total Mn 1 1	0	0
2	A	1	Total Mn 1 1	0	0

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

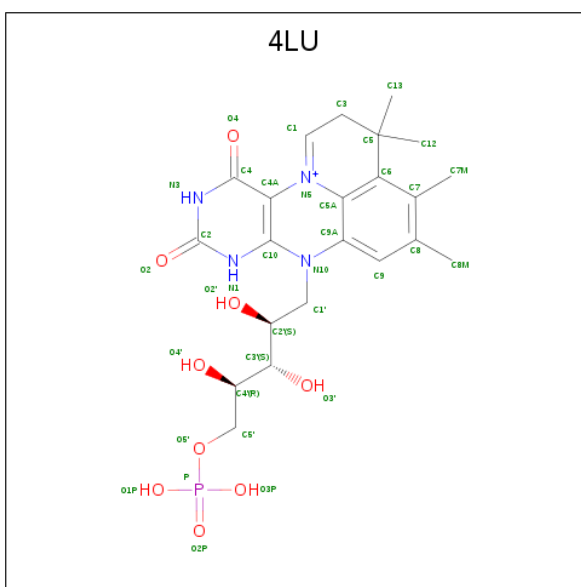
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Na 1 1	0	0
3	A	1	Total Na 1 1	0	0

- Molecule 4 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
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is 1-deoxy-5-O-phosphono-1-(3,3,4,5-tetramethyl-9,11-dioxo-2,3,8,9,10,11-hexahydro-7H-quinolino[1,8-fg]pteridin-12-ium-7-yl)-D-ribitol (three-letter code: 4LU) (formula: C<sub>22</sub>H<sub>30</sub>N<sub>4</sub>O<sub>9</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total 36	C 22	N 4	O 9	P 1	0	0
5	B	1	Total 36	C 22	N 4	O 9	P 1	0	0

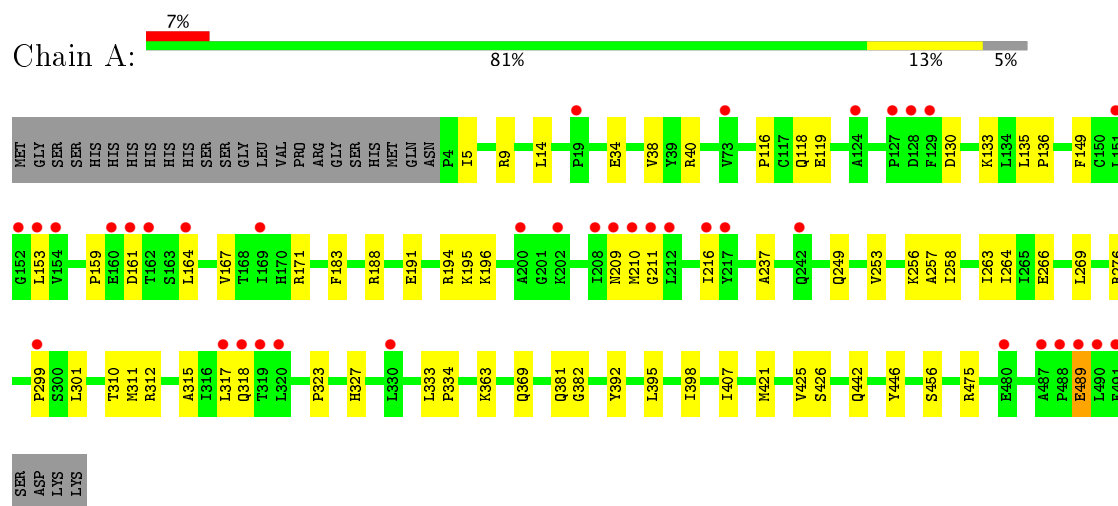
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	196	Total O 196 196	0	0
6	B	157	Total O 157 157	0	0

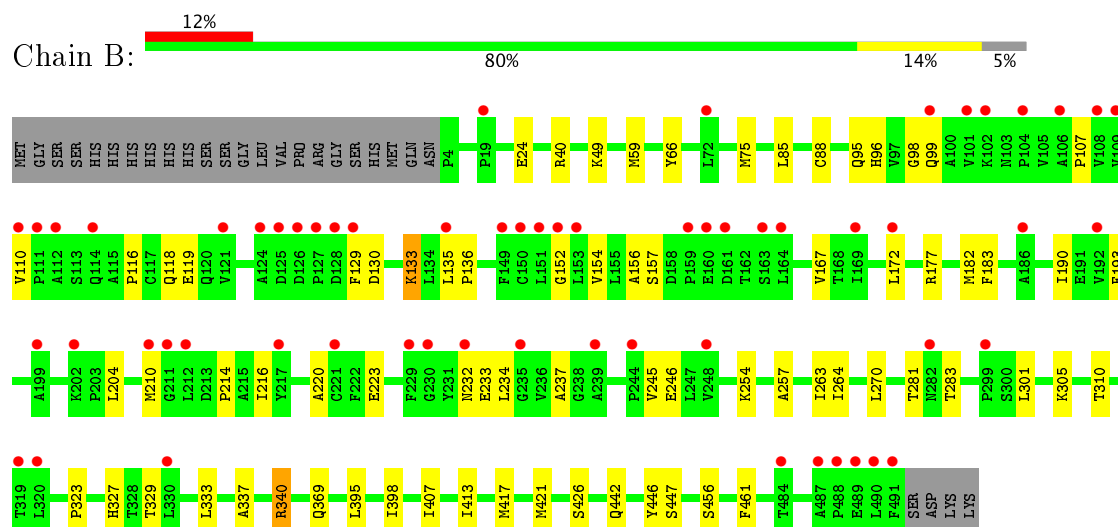
### 3 Residue-property plots [i](#)

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 ( $RSRZ > 2$ ). 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: 3,4-dihydroxybenzoate decarboxylase



- Molecule 1: 3,4-dihydroxybenzoate decarboxylase





## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	208.31Å 208.31Å 157.91Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.07 – 2.05 30.07 – 2.05	Depositor EDS
% Data completeness (in resolution range)	99.4 (30.07-2.05) 99.4 (30.07-2.05)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.51 (at 2.05Å)	Xtriage
Refinement program	PHENIX (dev_2689: ???)	Depositor
R, $R_{free}$	0.184 , 0.205 0.182 , 0.205	Depositor DCC
$R_{free}$ test set	4126 reflections (5.07%)	DCC
Wilson B-factor (Å <sup>2</sup> )	41.9	Xtriage
Anisotropy	0.246	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 51.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	7934	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.11% 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, MN, 4LU, NA

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	0.27	0/3840	0.48	0/5230
1	B	0.30	0/3813	0.50	0/5195
All	All	0.28	0/7653	0.49	0/10425

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	3746	0	3745	42	0
1	B	3723	0	3722	47	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	30	0	40	0	0
4	B	6	0	8	0	0
5	A	36	0	28	6	0
5	B	36	0	28	4	0
6	A	196	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	B	157	0	0	0	0
All	All	7934	0	7571	94	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.

All (94) 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)
5:A:508:4LU:H13	5:A:508:4LU:H14	1.30	1.12
5:B:504:4LU:H13	5:B:504:4LU:H14	1.30	1.12
1:B:234:LEU:HD12	1:B:245:VAL:HG21	1.38	1.04
1:B:130:ASP:HB3	1:B:133:LYS:HG3	1.50	0.94
5:B:504:4LU:C12	5:B:504:4LU:H14	2.08	0.81
1:A:9:ARG:HH21	1:A:312:ARG:HD2	1.48	0.78
1:B:234:LEU:CD1	1:B:245:VAL:HG21	2.16	0.73
5:A:508:4LU:C7M	5:A:508:4LU:H13	2.15	0.73
1:B:281:THR:HB	1:B:283:THR:HG23	1.72	0.71
1:A:257:ALA:HB1	1:A:264:ILE:HD13	1.75	0.68
1:A:211:GLY:HA2	1:A:318:GLN:HE21	1.60	0.67
1:A:159:PRO:HD2	1:A:196:LYS:HE3	1.79	0.64
1:B:234:LEU:HD12	1:B:245:VAL:CG2	2.23	0.64
1:A:209:ASN:ND2	1:A:258:ILE:H	1.96	0.63
1:A:171:ARG:HB3	5:A:508:4LU:H10	1.79	0.63
1:B:118:GLN:HA	1:B:310:THR:HG23	1.81	0.62
1:B:167:VAL:HG21	1:B:234:LEU:CD2	2.30	0.62
1:B:129:PHE:O	1:B:177:ARG:NH1	2.32	0.62
1:B:154:VAL:HG21	1:B:182:MET:HE2	1.83	0.59
1:B:135:LEU:HD22	1:B:263:ILE:HD13	1.84	0.59
1:B:167:VAL:HG21	1:B:234:LEU:HD23	1.85	0.59
1:A:196:LYS:HZ2	1:A:196:LYS:HB3	1.68	0.59
1:A:475:ARG:HD3	1:B:49:LYS:HD3	1.86	0.57
1:A:118:GLN:HG2	1:A:310:THR:HG21	1.87	0.57
5:A:508:4LU:C12	5:A:508:4LU:H14	2.13	0.56
1:B:323:PRO:HB2	1:B:327:HIS:HB2	1.86	0.55
1:B:110:VAL:HG21	1:B:246:GLU:HB3	1.88	0.55
1:A:118:GLN:HA	1:A:310:THR:HG23	1.89	0.55
1:A:323:PRO:HB2	1:A:327:HIS:HB2	1.89	0.54
1:A:149:PHE:CD1	1:A:210:MET:HE1	2.43	0.54
1:B:257:ALA:HB1	1:B:264:ILE:HD12	1.89	0.54
1:A:311:MET:HE3	1:A:315:ALA:HB3	1.89	0.54

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:98:GLY:HA2	1:B:337:ALA:HB1	1.91	0.53
1:B:369:GLN:HB2	1:B:407:ILE:HG21	1.90	0.53
1:B:223:GLU:OE2	5:B:504:4LU:O3'	2.27	0.53
1:A:161:ASP:HB3	1:A:164:LEU:HD12	1.90	0.53
1:B:118:GLN:HG2	1:B:310:THR:HG21	1.91	0.52
1:B:172:LEU:HD21	1:B:182:MET:HE3	1.92	0.51
1:A:116:PRO:O	1:A:312:ARG:HG2	2.12	0.50
1:B:216:ILE:HA	1:B:237:ALA:HB2	1.93	0.50
1:B:183:PHE:HA	1:B:301:LEU:HD22	1.93	0.50
1:A:5:ILE:HD11	1:A:14:LEU:HD13	1.94	0.50
1:B:24:GLU:HG2	1:B:59:MET:HB3	1.93	0.50
1:A:116:PRO:HB2	1:A:312:ARG:NE	2.27	0.49
1:B:156:ALA:HB1	1:B:204:LEU:HD21	1.94	0.49
1:A:249:GLN:HG2	1:A:256:LYS:HZ3	1.77	0.49
1:A:363[B]:LYS:NZ	1:A:392:TYR:OH	2.39	0.49
1:A:135:LEU:HD22	1:A:263:ILE:HD13	1.95	0.48
1:A:489:GLU:CD	1:A:489:GLU:H	2.17	0.48
1:A:253:VAL:HG22	1:A:266:GLU:HB3	1.96	0.48
1:A:183:PHE:HA	1:A:301:LEU:HD22	1.95	0.48
1:B:421:MET:HG3	1:B:426:SER:HB2	1.96	0.48
1:B:442:GLN:CD	1:B:456:SER:HB2	2.34	0.48
1:B:190:ILE:HA	1:B:193:PHE:CD1	2.48	0.47
1:B:116:PRO:O	1:B:119:GLU:HG3	2.15	0.47
1:A:421:MET:HG2	1:A:426:SER:HB2	1.95	0.47
1:B:220:ALA:HA	1:B:233:GLU:HB3	1.97	0.47
1:A:381:GLN:HG2	1:A:382:GLY:N	2.30	0.47
1:A:188:ARG:HE	5:A:508:4LU:H7	1.64	0.46
1:A:194:ARG:HG3	1:A:269:LEU:HD13	1.98	0.46
1:A:257:ALA:HB1	1:A:264:ILE:CD1	2.42	0.46
1:B:329:THR:O	1:B:333:LEU:HG	2.15	0.45
1:B:395:LEU:HD13	1:B:398:ILE:HD11	1.98	0.45
1:B:270:LEU:HD11	1:B:305:LYS:HB2	1.99	0.45
1:B:85:LEU:HD13	1:B:214:PRO:HG2	1.98	0.45
5:B:504:4LU:C12	5:B:504:4LU:C7M	2.85	0.44
1:A:196:LYS:NZ	1:A:196:LYS:HB3	2.28	0.44
1:B:75:MET:HB2	1:B:75:MET:HE2	1.83	0.44
1:A:395:LEU:HD13	1:A:398:ILE:HD11	1.98	0.44
1:A:153:LEU:HG	1:A:167:VAL:HG22	1.98	0.44
1:A:369:GLN:HB2	1:A:407:ILE:HG21	1.99	0.44
1:A:34:GLU:O	1:A:38:VAL:HG23	2.18	0.44
1:B:107:PRO:HB3	1:B:245:VAL:HG22	2.01	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:66:TYR:CD1	1:B:136:PRO:HB3	2.52	0.43
1:A:333:LEU:HB2	1:A:334:PRO:HD3	2.01	0.43
1:A:442:GLN:CD	1:A:456:SER:HB2	2.39	0.43
1:B:152:GLY:HA2	1:B:210:MET:HG2	2.01	0.43
1:B:88[C]:CYS:HG	1:B:96:HIS:CG	2.33	0.43
1:B:270:LEU:HD21	1:B:305:LYS:HB2	1.99	0.43
5:A:508:4LU:C7M	5:A:508:4LU:C12	2.85	0.42
1:B:254:LYS:HA	1:B:254:LYS:HD3	1.84	0.42
1:B:95:GLN:HA	1:B:340[A]:ARG:HD3	2.01	0.42
1:A:276:ARG:HG3	1:A:299:PRO:HA	2.01	0.42
1:A:136:PRO:O	1:A:317:LEU:HD22	2.20	0.42
1:B:421:MET:HE1	1:B:461:PHE:HD1	1.85	0.42
1:B:99:GLN:HE22	1:B:340[A]:ARG:NH1	2.18	0.42
1:A:216:ILE:HA	1:A:237:ALA:HB2	2.02	0.42
1:B:154:VAL:HG11	1:B:182:MET:HE1	2.01	0.41
1:A:116:PRO:O	1:A:119:GLU:HG3	2.21	0.41
1:A:130:ASP:HB3	1:A:133:LYS:HG3	2.02	0.41
1:B:413:ILE:O	1:B:417:MET:HG3	2.21	0.41
1:A:191:GLU:O	1:A:195:LYS:HG3	2.21	0.41
1:B:232:ASN:OD1	1:B:234:LEU:HB2	2.21	0.40
1:A:425:VAL:HG21	1:B:447:SER:HB2	2.04	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	491/515 (95%)	478 (97%)	13 (3%)	0	100	100
1	B	489/515 (95%)	475 (97%)	14 (3%)	0	100	100
All	All	980/1030 (95%)	953 (97%)	27 (3%)	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	401/420 (96%)	398 (99%)	3 (1%)	87	87
1	B	398/420 (95%)	392 (98%)	6 (2%)	70	68
All	All	799/840 (95%)	790 (99%)	9 (1%)	80	77

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	40	ARG
1	A	446	TYR
1	A	489	GLU
1	B	40	ARG
1	B	133	LYS
1	B	157	SER
1	B	340[A]	ARG
1	B	340[B]	ARG
1	B	446	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	95	GLN
1	A	209	ASN
1	A	249	GLN
1	A	318	GLN
1	A	453	ASN
1	B	16	GLN
1	B	383	GLN
1	B	453	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

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	GOL	A	503	-	5,5,5	0.90	0	5,5,5	0.96	0
4	GOL	A	504	-	5,5,5	0.92	0	5,5,5	0.88	0
4	GOL	A	505	-	5,5,5	0.88	0	5,5,5	0.98	0
4	GOL	A	506	-	5,5,5	0.95	0	5,5,5	0.98	0
4	GOL	A	507	-	5,5,5	0.86	0	5,5,5	1.07	0
5	4LU	A	508	3,2	33,39,39	1.85	9 (27%)	36,62,62	1.76	8 (22%)
4	GOL	B	503	-	5,5,5	0.81	0	5,5,5	0.93	0
5	4LU	B	504	3,2	33,39,39	1.64	7 (21%)	36,62,62	1.79	7 (19%)

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
4	GOL	A	503	-	-	0/4/4/4	0/0/0/0
4	GOL	A	504	-	-	0/4/4/4	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	A	505	-	-	0/4/4/4	0/0/0/0
4	GOL	A	506	-	-	0/4/4/4	0/0/0/0
4	GOL	A	507	-	-	0/4/4/4	0/0/0/0
5	4LU	A	508	3,2	-	0/16/30/30	0/3/4/4
4	GOL	B	503	-	-	0/4/4/4	0/0/0/0
5	4LU	B	504	3,2	-	0/16/30/30	0/3/4/4

All (16) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	504	4LU	C5-C6	-4.69	1.50	1.54
5	A	508	4LU	C5-C6	-4.25	1.50	1.54
5	A	508	4LU	C1'-N10	-4.06	1.44	1.48
5	A	508	4LU	C2-N3	-3.93	1.30	1.38
5	B	504	4LU	C2-N3	-3.41	1.31	1.38
5	B	504	4LU	C1'-N10	-3.31	1.45	1.48
5	B	504	4LU	C2-N1	-3.27	1.31	1.38
5	A	508	4LU	C2-N1	-3.00	1.32	1.38
5	B	504	4LU	C6-C5A	-2.94	1.38	1.43
5	A	508	4LU	C6-C5A	-2.91	1.38	1.43
5	A	508	4LU	O4-C4	-2.76	1.17	1.24
5	B	504	4LU	O4-C4	-2.40	1.18	1.24
5	A	508	4LU	P-O2P	-2.25	1.43	1.50
5	B	504	4LU	C9-C9A	-2.15	1.36	1.40
5	A	508	4LU	C5A-C9A	-2.09	1.37	1.40
5	A	508	4LU	C9-C9A	-2.01	1.36	1.40

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	508	4LU	C12-C5-C3	-3.63	103.12	109.15
5	A	508	4LU	C4-C4A-C10	-3.18	117.39	119.96
5	B	504	4LU	C12-C5-C3	-3.14	103.94	109.15
5	A	508	4LU	C13-C5-C3	-2.84	104.44	109.15
5	B	504	4LU	C4-C4A-C10	-2.74	117.75	119.96
5	B	504	4LU	C4A-C4-N3	-2.60	119.78	123.48
5	A	508	4LU	C4A-C10-N10	-2.26	118.95	120.52
5	A	508	4LU	O3P-P-O5'	2.02	112.11	106.73
5	B	504	4LU	O5'-P-O2P	2.13	112.45	106.47
5	B	504	4LU	P-O5'-C5'	2.56	125.36	118.30
5	A	508	4LU	C1'-N10-C9A	2.74	120.86	118.35
5	A	508	4LU	P-O5'-C5'	2.80	126.00	118.30

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	504	4LU	C1'-N10-C9A	3.22	121.30	118.35
5	A	508	4LU	C4-N3-C2	5.41	119.89	115.16
5	B	504	4LU	C4-N3-C2	6.75	121.06	115.16

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	508	4LU	6	0
5	B	504	4LU	4	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

### 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	488/515 (94%)	0.24	37 (7%) 15 16	27, 54, 79, 108	0
1	B	488/515 (94%)	0.53	60 (12%) 5 4	28, 60, 98, 132	0
All	All	976/1030 (94%)	0.38	97 (9%) 8 9	27, 56, 94, 132	0

All (97) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	164	LEU	6.8
1	B	230	GLY	5.2
1	A	151	LEU	4.9
1	B	229	PHE	4.6
1	B	490	LEU	4.6
1	B	491	PHE	4.6
1	B	127	PRO	4.5
1	A	153	LEU	4.5
1	A	491	PHE	4.4
1	B	320	LEU	4.3
1	A	212	LEU	4.3
1	B	101	VAL	4.2
1	B	111	PRO	4.1
1	B	109	VAL	4.0
1	A	490	LEU	4.0
1	B	169	ILE	3.9
1	B	163	SER	3.8
1	A	127	PRO	3.8
1	B	151	LEU	3.8
1	B	239	ALA	3.8
1	A	169	ILE	3.7
1	A	154	VAL	3.7
1	B	160	GLU	3.6
1	B	202	LYS	3.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	153	LEU	3.3
1	B	235	GLY	3.3
1	B	161	ASP	3.3
1	B	488	PRO	3.3
1	B	199	ALA	3.3
1	A	320	LEU	3.2
1	A	217	TYR	3.0
1	B	19	PRO	3.0
1	B	102	LYS	3.0
1	A	208	ILE	2.9
1	B	150	CYS	2.9
1	B	124	ALA	2.9
1	A	330	LEU	2.9
1	B	212	LEU	2.9
1	A	19	PRO	2.9
1	B	112	ALA	2.8
1	A	319	THR	2.8
1	B	110	VAL	2.8
1	B	192	VAL	2.7
1	B	121	VAL	2.7
1	B	282	ASN	2.7
1	B	319	THR	2.7
1	B	106	ALA	2.7
1	B	487	ALA	2.7
1	A	162	THR	2.7
1	A	216	ILE	2.6
1	B	149	PHE	2.6
1	A	318	GLN	2.6
1	B	104	PRO	2.6
1	A	152	GLY	2.6
1	A	488	PRO	2.6
1	B	114	GLN	2.6
1	A	73	VAL	2.5
1	A	489	GLU	2.5
1	B	489	GLU	2.5
1	B	330	LEU	2.4
1	A	211	GLY	2.4
1	A	128	ASP	2.4
1	B	210	MET	2.4
1	A	164	LEU	2.4
1	A	129	PHE	2.4
1	A	160	GLU	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	299	PRO	2.4
1	B	186	ALA	2.4
1	A	317	LEU	2.3
1	B	108	VAL	2.3
1	A	124	ALA	2.3
1	B	217	TYR	2.3
1	B	126	ASP	2.3
1	B	248	VAL	2.3
1	A	200	ALA	2.3
1	B	99	GLN	2.3
1	A	202	LYS	2.3
1	A	480	GLU	2.3
1	B	211	GLY	2.2
1	B	232	ASN	2.2
1	A	161	ASP	2.2
1	B	299	PRO	2.2
1	B	125	ASP	2.2
1	A	487	ALA	2.2
1	B	159	PRO	2.2
1	B	244	PRO	2.2
1	B	135	LEU	2.1
1	B	221	CYS	2.1
1	B	128	ASP	2.1
1	A	210	MET	2.1
1	B	129	PHE	2.1
1	A	242	GLN	2.1
1	B	484	THR	2.1
1	B	152	GLY	2.1
1	A	209	ASN	2.0
1	B	72	LEU	2.0
1	B	172	LEU	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 ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	GOL	A	507	6/6	0.74	0.24	9.21	67,70,72,73	0
4	GOL	A	506	6/6	0.89	0.25	3.92	54,61,65,69	0
5	4LU	B	504	36/36	0.94	0.37	3.22	20,20,20,20	0
4	GOL	A	505	6/6	0.90	0.18	2.52	62,71,74,77	0
4	GOL	A	503	6/6	0.86	0.22	2.48	58,61,62,62	0
4	GOL	B	503	6/6	0.87	0.20	2.31	51,61,65,76	0
4	GOL	A	504	6/6	0.95	0.13	0.87	55,57,58,60	0
3	NA	A	502	1/1	0.96	0.14	-0.40	46,46,46,46	0
5	4LU	A	508	36/36	0.94	0.12	-0.77	30,30,30,30	0
2	MN	A	501	1/1	0.97	0.04	-2.02	61,61,61,61	0
2	MN	B	501	1/1	0.91	0.11	-2.44	81,81,81,81	0
3	NA	B	502	1/1	0.51	0.10	-	66,66,66,66	0

## 6.5 Other polymers

There are no such residues in this entry.