



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 22, 2023 – 06:23 AM EDT

PDB ID : 2OZ3  
Title : Crystal structure of L-Rhamnonate dehydratase from *Azotobacter vinelandii*  
Authors : Patskovsky, Y.; Toro, R.; Sauder, J.M.; Freeman, J.C.; Bain, K.; Gheyi, T.;  
Wu, B.; Wasserman, S.R.; Smith, D.; Gerlt, J.; Burley, S.K.; Almo, S.C.; New  
York SGX Research Center for Structural Genomics (NYSGXRC)  
Deposited on : 2007-02-23  
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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.35  
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) : 2.35

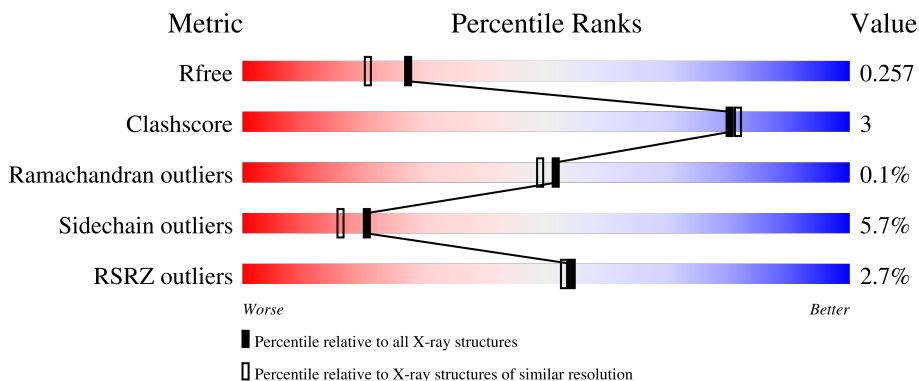
# 1 Overall quality at a glance

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))
$R_{free}$	130704	8085 (2.00-2.00)
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 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	404	
1	B	404	
1	C	404	
1	D	404	
1	E	404	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	404	<p>3% 87% 8% •</p>
1	G	404	<p>3% 85% 11% ••</p>
1	H	404	<p>3% 83% 11% • 5%</p>

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	GOL	C	3022	-	-	X	-
2	GOL	H	3012	-	-	X	-

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 26576 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 Mandelate racemase/muconate lactonizing enzyme.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	396	Total 3143	C 1996	N 554	O 572	S 5	Se 16	0	4	0
1	B	396	Total 3143	C 1997	N 553	O 573	S 5	Se 15	0	4	0
1	C	388	Total 3083	C 1959	N 544	O 560	S 5	Se 15	0	3	0
1	D	395	Total 3132	C 1991	N 550	O 571	S 5	Se 15	0	4	0
1	E	396	Total 3137	C 1991	N 556	O 570	S 5	Se 15	0	2	0
1	F	387	Total 3080	C 1959	N 541	O 560	S 5	Se 15	0	4	0
1	G	395	Total 3131	C 1992	N 549	O 570	S 5	Se 15	5	4	0
1	H	383	Total 3063	C 1954	N 539	O 550	S 5	Se 15	0	6	0

There are 208 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MSE	-	cloning artifact	UNP Q4J3N3
A	0	SER	-	cloning artifact	UNP Q4J3N3
A	1	LEU	-	cloning artifact	UNP Q4J3N3
A	38	MSE	MET	modified residue	UNP Q4J3N3
A	108	MSE	MET	modified residue	UNP Q4J3N3
A	171	MSE	MET	modified residue	UNP Q4J3N3
A	178	MSE	MET	modified residue	UNP Q4J3N3
A	201	MSE	MET	modified residue	UNP Q4J3N3
A	212	MSE	MET	modified residue	UNP Q4J3N3
A	217	MSE	MET	modified residue	UNP Q4J3N3
A	262	MSE	MET	modified residue	UNP Q4J3N3
A	263	MSE	MET	modified residue	UNP Q4J3N3
A	278	MSE	MET	modified residue	UNP Q4J3N3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	282	MSE	MET	modified residue	UNP Q4J3N3
A	341	MSE	MET	modified residue	UNP Q4J3N3
A	342	MSE	MET	modified residue	UNP Q4J3N3
A	352	MSE	MET	modified residue	UNP Q4J3N3
A	368	MSE	MET	modified residue	UNP Q4J3N3
A	395	GLU	-	cloning artifact	UNP Q4J3N3
A	396	GLY	-	cloning artifact	UNP Q4J3N3
A	397	HIS	-	cloning artifact	UNP Q4J3N3
A	398	HIS	-	cloning artifact	UNP Q4J3N3
A	399	HIS	-	cloning artifact	UNP Q4J3N3
A	400	HIS	-	cloning artifact	UNP Q4J3N3
A	401	HIS	-	cloning artifact	UNP Q4J3N3
A	402	HIS	-	cloning artifact	UNP Q4J3N3
B	-1	MSE	-	cloning artifact	UNP Q4J3N3
B	0	SER	-	cloning artifact	UNP Q4J3N3
B	1	LEU	-	cloning artifact	UNP Q4J3N3
B	38	MSE	MET	modified residue	UNP Q4J3N3
B	108	MSE	MET	modified residue	UNP Q4J3N3
B	171	MSE	MET	modified residue	UNP Q4J3N3
B	178	MSE	MET	modified residue	UNP Q4J3N3
B	201	MSE	MET	modified residue	UNP Q4J3N3
B	212	MSE	MET	modified residue	UNP Q4J3N3
B	217	MSE	MET	modified residue	UNP Q4J3N3
B	262	MSE	MET	modified residue	UNP Q4J3N3
B	263	MSE	MET	modified residue	UNP Q4J3N3
B	278	MSE	MET	modified residue	UNP Q4J3N3
B	282	MSE	MET	modified residue	UNP Q4J3N3
B	341	MSE	MET	modified residue	UNP Q4J3N3
B	342	MSE	MET	modified residue	UNP Q4J3N3
B	352	MSE	MET	modified residue	UNP Q4J3N3
B	368	MSE	MET	modified residue	UNP Q4J3N3
B	395	GLU	-	cloning artifact	UNP Q4J3N3
B	396	GLY	-	cloning artifact	UNP Q4J3N3
B	397	HIS	-	cloning artifact	UNP Q4J3N3
B	398	HIS	-	cloning artifact	UNP Q4J3N3
B	399	HIS	-	cloning artifact	UNP Q4J3N3
B	400	HIS	-	cloning artifact	UNP Q4J3N3
B	401	HIS	-	cloning artifact	UNP Q4J3N3
B	402	HIS	-	cloning artifact	UNP Q4J3N3
C	-1	MSE	-	cloning artifact	UNP Q4J3N3
C	0	SER	-	cloning artifact	UNP Q4J3N3
C	1	LEU	-	cloning artifact	UNP Q4J3N3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	38	MSE	MET	modified residue	UNP Q4J3N3
C	108	MSE	MET	modified residue	UNP Q4J3N3
C	171	MSE	MET	modified residue	UNP Q4J3N3
C	178	MSE	MET	modified residue	UNP Q4J3N3
C	201	MSE	MET	modified residue	UNP Q4J3N3
C	212	MSE	MET	modified residue	UNP Q4J3N3
C	217	MSE	MET	modified residue	UNP Q4J3N3
C	262	MSE	MET	modified residue	UNP Q4J3N3
C	263	MSE	MET	modified residue	UNP Q4J3N3
C	278	MSE	MET	modified residue	UNP Q4J3N3
C	282	MSE	MET	modified residue	UNP Q4J3N3
C	341	MSE	MET	modified residue	UNP Q4J3N3
C	342	MSE	MET	modified residue	UNP Q4J3N3
C	352	MSE	MET	modified residue	UNP Q4J3N3
C	368	MSE	MET	modified residue	UNP Q4J3N3
C	395	GLU	-	cloning artifact	UNP Q4J3N3
C	396	GLY	-	cloning artifact	UNP Q4J3N3
C	397	HIS	-	cloning artifact	UNP Q4J3N3
C	398	HIS	-	cloning artifact	UNP Q4J3N3
C	399	HIS	-	cloning artifact	UNP Q4J3N3
C	400	HIS	-	cloning artifact	UNP Q4J3N3
C	401	HIS	-	cloning artifact	UNP Q4J3N3
C	402	HIS	-	cloning artifact	UNP Q4J3N3
D	-1	MSE	-	cloning artifact	UNP Q4J3N3
D	0	SER	-	cloning artifact	UNP Q4J3N3
D	1	LEU	-	cloning artifact	UNP Q4J3N3
D	38	MSE	MET	modified residue	UNP Q4J3N3
D	108	MSE	MET	modified residue	UNP Q4J3N3
D	171	MSE	MET	modified residue	UNP Q4J3N3
D	178	MSE	MET	modified residue	UNP Q4J3N3
D	201	MSE	MET	modified residue	UNP Q4J3N3
D	212	MSE	MET	modified residue	UNP Q4J3N3
D	217	MSE	MET	modified residue	UNP Q4J3N3
D	262	MSE	MET	modified residue	UNP Q4J3N3
D	263	MSE	MET	modified residue	UNP Q4J3N3
D	278	MSE	MET	modified residue	UNP Q4J3N3
D	282	MSE	MET	modified residue	UNP Q4J3N3
D	341	MSE	MET	modified residue	UNP Q4J3N3
D	342	MSE	MET	modified residue	UNP Q4J3N3
D	352	MSE	MET	modified residue	UNP Q4J3N3
D	368	MSE	MET	modified residue	UNP Q4J3N3
D	395	GLU	-	cloning artifact	UNP Q4J3N3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	396	GLY	-	cloning artifact	UNP Q4J3N3
D	397	HIS	-	cloning artifact	UNP Q4J3N3
D	398	HIS	-	cloning artifact	UNP Q4J3N3
D	399	HIS	-	cloning artifact	UNP Q4J3N3
D	400	HIS	-	cloning artifact	UNP Q4J3N3
D	401	HIS	-	cloning artifact	UNP Q4J3N3
D	402	HIS	-	cloning artifact	UNP Q4J3N3
E	-1	MSE	-	cloning artifact	UNP Q4J3N3
E	0	SER	-	cloning artifact	UNP Q4J3N3
E	1	LEU	-	cloning artifact	UNP Q4J3N3
E	38	MSE	MET	modified residue	UNP Q4J3N3
E	108	MSE	MET	modified residue	UNP Q4J3N3
E	171	MSE	MET	modified residue	UNP Q4J3N3
E	178	MSE	MET	modified residue	UNP Q4J3N3
E	201	MSE	MET	modified residue	UNP Q4J3N3
E	212	MSE	MET	modified residue	UNP Q4J3N3
E	217	MSE	MET	modified residue	UNP Q4J3N3
E	262	MSE	MET	modified residue	UNP Q4J3N3
E	263	MSE	MET	modified residue	UNP Q4J3N3
E	278	MSE	MET	modified residue	UNP Q4J3N3
E	282	MSE	MET	modified residue	UNP Q4J3N3
E	341	MSE	MET	modified residue	UNP Q4J3N3
E	342	MSE	MET	modified residue	UNP Q4J3N3
E	352	MSE	MET	modified residue	UNP Q4J3N3
E	368	MSE	MET	modified residue	UNP Q4J3N3
E	395	GLU	-	cloning artifact	UNP Q4J3N3
E	396	GLY	-	cloning artifact	UNP Q4J3N3
E	397	HIS	-	cloning artifact	UNP Q4J3N3
E	398	HIS	-	cloning artifact	UNP Q4J3N3
E	399	HIS	-	cloning artifact	UNP Q4J3N3
E	400	HIS	-	cloning artifact	UNP Q4J3N3
E	401	HIS	-	cloning artifact	UNP Q4J3N3
E	402	HIS	-	cloning artifact	UNP Q4J3N3
F	-1	MSE	-	cloning artifact	UNP Q4J3N3
F	0	SER	-	cloning artifact	UNP Q4J3N3
F	1	LEU	-	cloning artifact	UNP Q4J3N3
F	38	MSE	MET	modified residue	UNP Q4J3N3
F	108	MSE	MET	modified residue	UNP Q4J3N3
F	171	MSE	MET	modified residue	UNP Q4J3N3
F	178	MSE	MET	modified residue	UNP Q4J3N3
F	201	MSE	MET	modified residue	UNP Q4J3N3
F	212	MSE	MET	modified residue	UNP Q4J3N3

*Continued on next page...*



*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	217	MSE	MET	modified residue	UNP Q4J3N3
F	262	MSE	MET	modified residue	UNP Q4J3N3
F	263	MSE	MET	modified residue	UNP Q4J3N3
F	278	MSE	MET	modified residue	UNP Q4J3N3
F	282	MSE	MET	modified residue	UNP Q4J3N3
F	341	MSE	MET	modified residue	UNP Q4J3N3
F	342	MSE	MET	modified residue	UNP Q4J3N3
F	352	MSE	MET	modified residue	UNP Q4J3N3
F	368	MSE	MET	modified residue	UNP Q4J3N3
F	395	GLU	-	cloning artifact	UNP Q4J3N3
F	396	GLY	-	cloning artifact	UNP Q4J3N3
F	397	HIS	-	cloning artifact	UNP Q4J3N3
F	398	HIS	-	cloning artifact	UNP Q4J3N3
F	399	HIS	-	cloning artifact	UNP Q4J3N3
F	400	HIS	-	cloning artifact	UNP Q4J3N3
F	401	HIS	-	cloning artifact	UNP Q4J3N3
F	402	HIS	-	cloning artifact	UNP Q4J3N3
G	-1	MSE	-	cloning artifact	UNP Q4J3N3
G	0	SER	-	cloning artifact	UNP Q4J3N3
G	1	LEU	-	cloning artifact	UNP Q4J3N3
G	38	MSE	MET	modified residue	UNP Q4J3N3
G	108	MSE	MET	modified residue	UNP Q4J3N3
G	171	MSE	MET	modified residue	UNP Q4J3N3
G	178	MSE	MET	modified residue	UNP Q4J3N3
G	201	MSE	MET	modified residue	UNP Q4J3N3
G	212	MSE	MET	modified residue	UNP Q4J3N3
G	217	MSE	MET	modified residue	UNP Q4J3N3
G	262	MSE	MET	modified residue	UNP Q4J3N3
G	263	MSE	MET	modified residue	UNP Q4J3N3
G	278	MSE	MET	modified residue	UNP Q4J3N3
G	282	MSE	MET	modified residue	UNP Q4J3N3
G	341	MSE	MET	modified residue	UNP Q4J3N3
G	342	MSE	MET	modified residue	UNP Q4J3N3
G	352	MSE	MET	modified residue	UNP Q4J3N3
G	368	MSE	MET	modified residue	UNP Q4J3N3
G	395	GLU	-	cloning artifact	UNP Q4J3N3
G	396	GLY	-	cloning artifact	UNP Q4J3N3
G	397	HIS	-	cloning artifact	UNP Q4J3N3
G	398	HIS	-	cloning artifact	UNP Q4J3N3
G	399	HIS	-	cloning artifact	UNP Q4J3N3
G	400	HIS	-	cloning artifact	UNP Q4J3N3
G	401	HIS	-	cloning artifact	UNP Q4J3N3

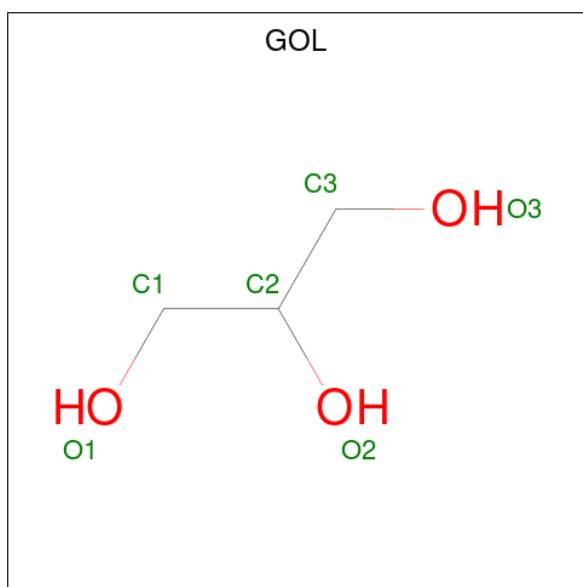
*Continued on next page...*



*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	402	HIS	-	cloning artifact	UNP Q4J3N3
H	-1	MSE	-	cloning artifact	UNP Q4J3N3
H	0	SER	-	cloning artifact	UNP Q4J3N3
H	1	LEU	-	cloning artifact	UNP Q4J3N3
H	38	MSE	MET	modified residue	UNP Q4J3N3
H	108	MSE	MET	modified residue	UNP Q4J3N3
H	171	MSE	MET	modified residue	UNP Q4J3N3
H	178	MSE	MET	modified residue	UNP Q4J3N3
H	201	MSE	MET	modified residue	UNP Q4J3N3
H	212	MSE	MET	modified residue	UNP Q4J3N3
H	217	MSE	MET	modified residue	UNP Q4J3N3
H	262	MSE	MET	modified residue	UNP Q4J3N3
H	263	MSE	MET	modified residue	UNP Q4J3N3
H	278	MSE	MET	modified residue	UNP Q4J3N3
H	282	MSE	MET	modified residue	UNP Q4J3N3
H	341	MSE	MET	modified residue	UNP Q4J3N3
H	342	MSE	MET	modified residue	UNP Q4J3N3
H	352	MSE	MET	modified residue	UNP Q4J3N3
H	368	MSE	MET	modified residue	UNP Q4J3N3
H	395	GLU	-	cloning artifact	UNP Q4J3N3
H	396	GLY	-	cloning artifact	UNP Q4J3N3
H	397	HIS	-	cloning artifact	UNP Q4J3N3
H	398	HIS	-	cloning artifact	UNP Q4J3N3
H	399	HIS	-	cloning artifact	UNP Q4J3N3
H	400	HIS	-	cloning artifact	UNP Q4J3N3
H	401	HIS	-	cloning artifact	UNP Q4J3N3
H	402	HIS	-	cloning artifact	UNP Q4J3N3

- 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 6 3 3	0	0
2	A	1	Total C O 6 3 3	0	0
2	A	1	Total C O 6 3 3	0	0
2	A	1	Total C O 6 3 3	0	0
2	A	1	Total C O 6 3 3	0	0
2	A	1	Total C O 6 3 3	0	0
2	A	1	Total C O 6 3 3	0	0
2	A	1	Total C O 6 3 3	0	0
2	A	1	Total C O 6 3 3	0	0
2	A	1	Total C O 6 3 3	0	0
2	B	1	Total C O 6 3 3	0	0
2	B	1	Total C O 6 3 3	0	0
2	B	1	Total C O 6 3 3	0	0
2	B	1	Total C O 6 3 3	0	0
2	C	1	Total C O 6 3 3	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	C	1	Total	C	O	0	0
			6	3	3		
2	C	1	Total	C	O	0	0
			6	3	3		
2	C	1	Total	C	O	0	0
			6	3	3		
2	C	1	Total	C	O	0	0
			6	3	3		
2	D	1	Total	C	O	0	0
			6	3	3		
2	D	1	Total	C	O	0	0
			6	3	3		
2	D	1	Total	C	O	0	0
			6	3	3		
2	D	1	Total	C	O	0	0
			6	3	3		
2	E	1	Total	C	O	0	0
			6	3	3		
2	F	1	Total	C	O	0	0
			6	3	3		
2	F	1	Total	C	O	0	0
			6	3	3		
2	G	1	Total	C	O	0	0
			6	3	3		
2	G	1	Total	C	O	0	0
			6	3	3		
2	H	1	Total	C	O	0	0
			6	3	3		
2	H	1	Total	C	O	0	0
			6	3	3		
2	H	1	Total	C	O	0	0
			6	3	3		

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

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

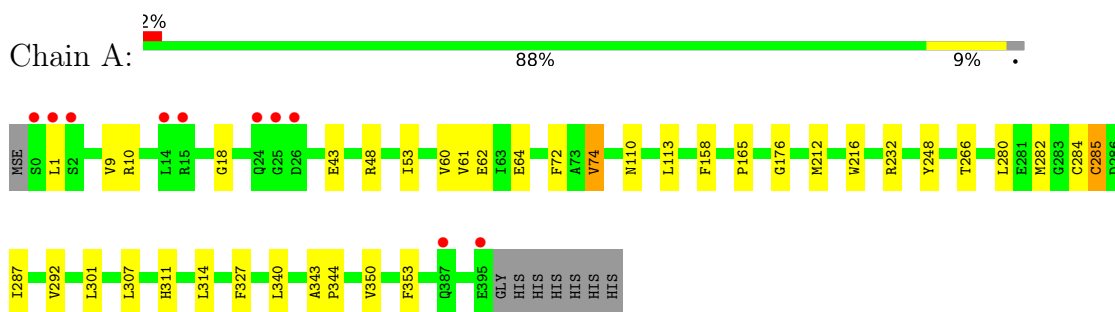
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	242	Total O 242 242	0	0
4	B	228	Total O 228 228	0	0
4	C	228	Total O 228 228	0	0
4	D	149	Total O 149 149	0	0
4	E	158	Total O 158 158	0	0
4	F	166	Total O 166 166	0	0
4	G	149	Total O 149 149	0	0
4	H	156	Total O 156 156	0	0

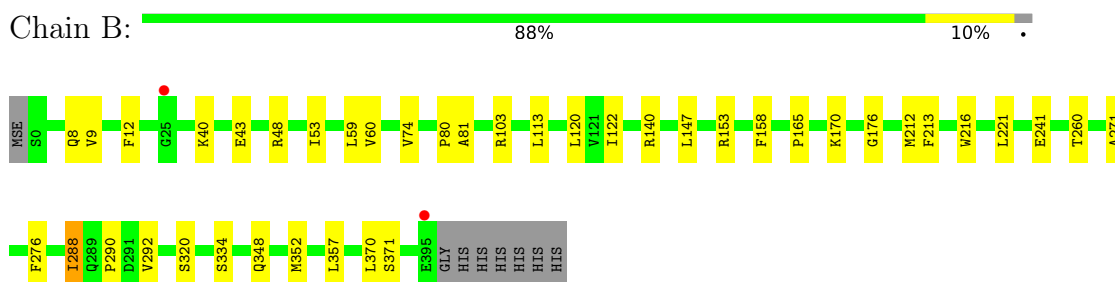
### 3 Residue-property plots [i](#)

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.

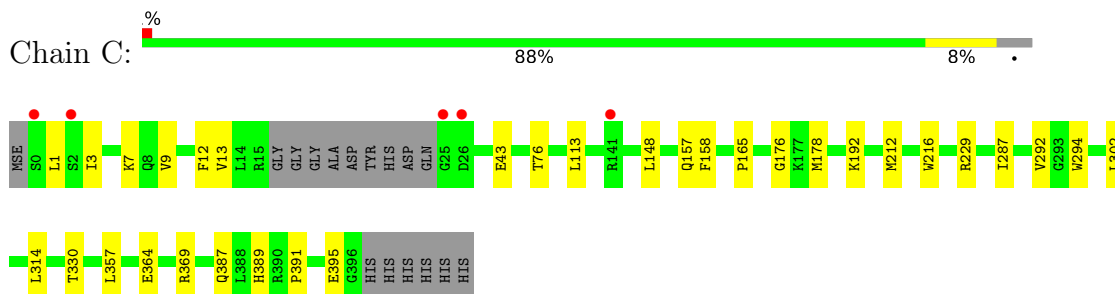
- Molecule 1: Mandelate racemase/muconate lactonizing enzyme



- Molecule 1: Mandelate racemase/muconate lactonizing enzyme

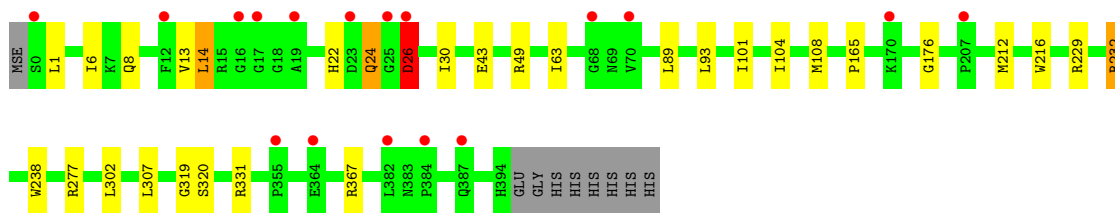


- Molecule 1: Mandelate racemase/muconate lactonizing enzyme

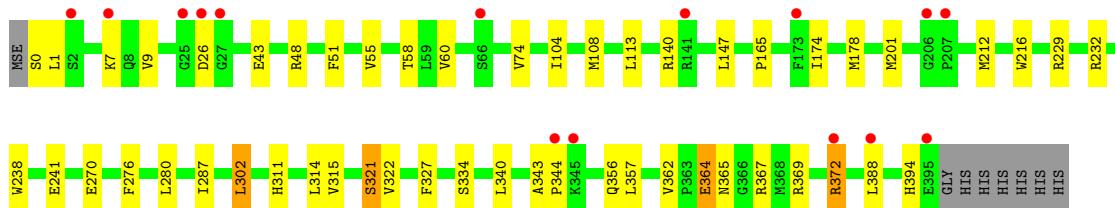
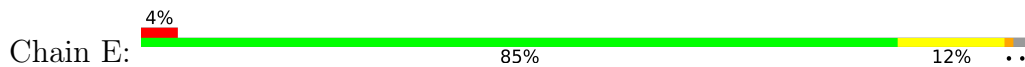


- Molecule 1: Mandelate racemase/muconate lactonizing enzyme

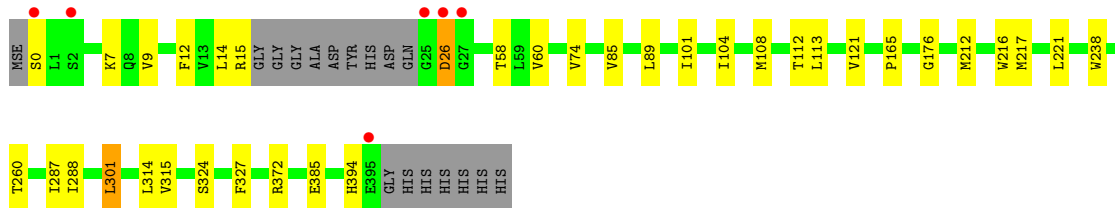
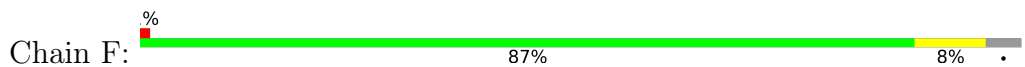




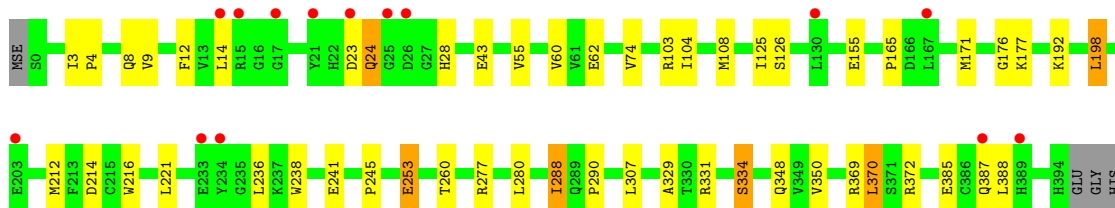
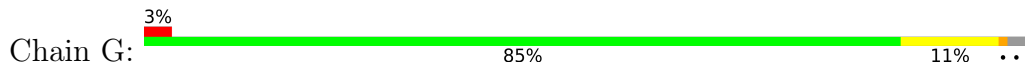
● Molecule 1: Mandelate racemase/muconate lactonizing enzyme



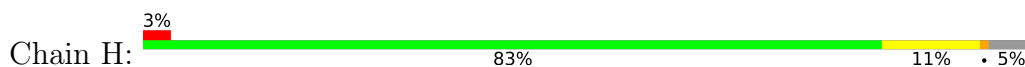
● Molecule 1: Mandelate racemase/muconate lactonizing enzyme

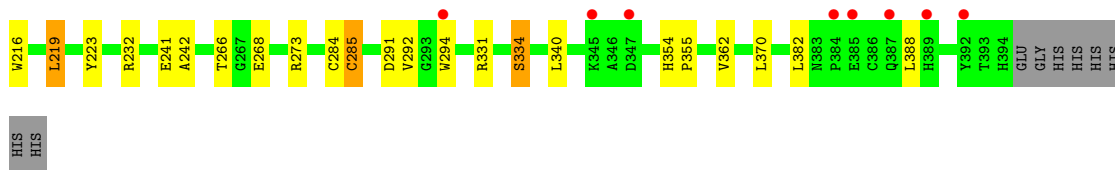


● Molecule 1: Mandelate racemase/muconate lactonizing enzyme



● Molecule 1: Mandelate racemase/muconate lactonizing enzyme





HIS  
HIS



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	170.77Å 115.40Å 164.11Å 90.00° 96.06° 90.00°	Depositor
Resolution (Å)	20.00 – 2.00 31.46 – 2.00	Depositor EDS
% Data completeness (in resolution range)	96.9 (20.00-2.00) 96.9 (31.46-2.00)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.06 (at 2.00Å)	Xtrriage
Refinement program	REFMAC 5.3.0028	Depositor
R, $R_{free}$	0.196 , 0.257 0.197 , 0.257	Depositor DCC
$R_{free}$ test set	6227 reflections (3.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	33.3	Xtrriage
Anisotropy	0.333	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 60.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	26576	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	38.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.00% 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: NA, GOL

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	0.43	0/3223	0.58	0/4346
1	B	0.42	0/3223	0.59	0/4346
1	C	0.42	0/3154	0.61	0/4251
1	D	0.40	0/3212	0.56	0/4333
1	E	0.40	0/3211	0.56	0/4329
1	F	0.39	0/3157	0.55	0/4255
1	G	0.39	0/3211	0.55	0/4331
1	H	0.40	0/3146	0.57	0/4240
All	All	0.40	0/25537	0.57	0/34431

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	2
1	B	0	1
1	C	0	3
1	D	1	1
1	E	0	1
1	F	0	2
1	G	0	1
1	H	0	3
All	All	1	14

There are no bond length outliers.

There are no bond angle outliers.

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	D	26	ASP	CA

5 of 14 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	158	PHE	Peptide
1	A	284	CYS	Peptide
1	B	158	PHE	Peptide
1	C	12	PHE	Peptide
1	C	158	PHE	Peptide

## 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	3143	0	3074	15	0
1	B	3143	0	3074	18	0
1	C	3083	0	3019	14	0
1	D	3132	0	3063	15	0
1	E	3137	0	3065	18	0
1	F	3080	0	3022	12	0
1	G	3131	0	3066	22	0
1	H	3063	0	3028	19	0
2	A	54	0	72	1	0
2	B	24	0	32	2	0
2	C	36	0	48	7	0
2	D	24	0	32	1	0
2	E	6	0	8	0	0
2	F	12	0	16	0	0
2	G	12	0	16	0	0
2	H	18	0	24	4	0
3	B	1	0	0	0	0
3	D	1	0	0	0	0
4	A	242	0	0	2	0
4	B	228	0	0	1	0
4	C	228	0	0	1	0
4	D	149	0	0	1	0
4	E	158	0	0	2	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	F	166	0	0	0	0
4	G	149	0	0	0	0
4	H	156	0	0	0	0
All	All	26576	0	24659	132	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 3.

The worst 5 of 132 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:C:229[B]:ARG:HH21	1:C:229[B]:ARG:HG2	1.05	1.17
1:C:229[B]:ARG:HG2	1:C:229[B]:ARG:NH2	1.78	0.89
1:D:229:ARG:HG3	1:D:232:ARG:HH21	1.43	0.83
2:C:3022:GOL:H31	1:D:302:LEU:HD22	1.61	0.82
1:B:292:VAL:HG23	4:B:3233:HOH:O	1.86	0.74

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	398/404 (98%)	384 (96%)	14 (4%)	0	100	100
1	B	398/404 (98%)	391 (98%)	7 (2%)	0	100	100
1	C	387/404 (96%)	376 (97%)	10 (3%)	1 (0%)	41	37
1	D	397/404 (98%)	382 (96%)	13 (3%)	2 (0%)	29	23
1	E	396/404 (98%)	386 (98%)	9 (2%)	1 (0%)	41	37
1	F	387/404 (96%)	376 (97%)	11 (3%)	0	100	100
1	G	397/404 (98%)	385 (97%)	12 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	H	385/404 (95%)	375 (97%)	10 (3%)	0	100	100
All	All	3145/3232 (97%)	3055 (97%)	86 (3%)	4 (0%)	51	49

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	13	VAL
1	D	26	ASP
1	E	321	SER
1	D	319	GLY

### 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	329/316 (104%)	317 (96%)	12 (4%)	35	34
1	B	329/316 (104%)	311 (94%)	18 (6%)	21	17
1	C	323/316 (102%)	308 (95%)	15 (5%)	27	23
1	D	328/316 (104%)	312 (95%)	16 (5%)	25	21
1	E	327/316 (104%)	305 (93%)	22 (7%)	16	11
1	F	324/316 (102%)	307 (95%)	17 (5%)	23	19
1	G	328/316 (104%)	303 (92%)	25 (8%)	13	8
1	H	323/316 (102%)	299 (93%)	24 (7%)	13	9
All	All	2611/2528 (103%)	2462 (94%)	149 (6%)	20	16

5 of 149 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	280	LEU
1	H	334	SER
1	G	350	VAL
1	H	14	LEU
1	D	14	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	257	ASN
1	G	257	ASN
1	H	257	ASN
1	G	311	HIS
1	C	157	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](#)

Of 33 ligands modelled in this entry, 2 are monoatomic - leaving 31 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	G	3013	-	5,5,5	0.37	0	5,5,5	0.19	0
2	GOL	C	3001	-	5,5,5	0.39	0	5,5,5	0.32	0
2	GOL	E	3010	-	5,5,5	0.36	0	5,5,5	0.32	0
2	GOL	A	3028	-	5,5,5	0.38	0	5,5,5	0.30	0
2	GOL	C	3020	-	5,5,5	0.45	0	5,5,5	0.45	0
2	GOL	F	3011	-	5,5,5	0.37	0	5,5,5	0.38	0
2	GOL	B	3029	-	5,5,5	0.37	0	5,5,5	0.16	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	GOL	A	3014	-	5,5,5	0.37	0	5,5,5	0.29	0
2	GOL	C	3026	-	5,5,5	0.35	0	5,5,5	0.25	0
2	GOL	C	3019	-	5,5,5	0.37	0	5,5,5	0.34	0
2	GOL	D	3005	-	5,5,5	0.34	0	5,5,5	0.47	0
2	GOL	G	3009	-	5,5,5	0.35	0	5,5,5	0.24	0
2	GOL	C	3022	-	5,5,5	0.36	0	5,5,5	0.71	0
2	GOL	A	3024	-	5,5,5	0.37	0	5,5,5	0.16	0
2	GOL	H	3018	-	5,5,5	0.37	0	5,5,5	0.28	0
2	GOL	A	3016	-	5,5,5	0.37	0	5,5,5	0.39	0
2	GOL	H	3025	-	5,5,5	0.34	0	5,5,5	0.36	0
2	GOL	B	3002	-	5,5,5	0.37	0	5,5,5	0.32	0
2	GOL	D	3017	-	5,5,5	0.37	0	5,5,5	0.22	0
2	GOL	C	3027	-	5,5,5	0.39	0	5,5,5	0.33	0
2	GOL	H	3012	-	5,5,5	0.58	0	5,5,5	0.31	0
2	GOL	A	3032	-	5,5,5	0.39	0	5,5,5	0.32	0
2	GOL	A	3030	-	5,5,5	0.38	0	5,5,5	0.26	0
2	GOL	A	3031	-	5,5,5	0.36	0	5,5,5	0.41	0
2	GOL	B	3004	-	5,5,5	0.33	0	5,5,5	0.42	0
2	GOL	D	3007	-	5,5,5	0.41	0	5,5,5	0.25	0
2	GOL	F	3008	-	5,5,5	0.37	0	5,5,5	0.29	0
2	GOL	A	3023	-	5,5,5	0.40	0	5,5,5	0.32	0
2	GOL	D	3006	-	5,5,5	0.33	0	5,5,5	0.30	0
2	GOL	A	3003	-	5,5,5	0.34	0	5,5,5	0.20	0
2	GOL	B	3021	-	5,5,5	0.40	0	5,5,5	0.27	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	G	3013	-	-	2/4/4/4	-
2	GOL	C	3001	-	-	4/4/4/4	-
2	GOL	E	3010	-	-	2/4/4/4	-
2	GOL	A	3028	-	-	2/4/4/4	-
2	GOL	C	3020	-	-	2/4/4/4	-
2	GOL	F	3011	-	-	2/4/4/4	-
2	GOL	B	3029	-	-	3/4/4/4	-
2	GOL	A	3014	-	-	2/4/4/4	-
2	GOL	C	3026	-	-	4/4/4/4	-

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GOL	C	3019	-	-	4/4/4/4	-
2	GOL	D	3005	-	-	2/4/4/4	-
2	GOL	G	3009	-	-	1/4/4/4	-
2	GOL	C	3022	-	-	4/4/4/4	-
2	GOL	A	3024	-	-	0/4/4/4	-
2	GOL	H	3018	-	-	4/4/4/4	-
2	GOL	A	3016	-	-	2/4/4/4	-
2	GOL	H	3025	-	-	2/4/4/4	-
2	GOL	B	3002	-	-	4/4/4/4	-
2	GOL	D	3017	-	-	2/4/4/4	-
2	GOL	C	3027	-	-	2/4/4/4	-
2	GOL	H	3012	-	-	4/4/4/4	-
2	GOL	A	3032	-	-	4/4/4/4	-
2	GOL	A	3030	-	-	2/4/4/4	-
2	GOL	A	3031	-	-	0/4/4/4	-
2	GOL	B	3004	-	-	4/4/4/4	-
2	GOL	D	3007	-	-	3/4/4/4	-
2	GOL	F	3008	-	-	2/4/4/4	-
2	GOL	A	3023	-	-	4/4/4/4	-
2	GOL	D	3006	-	-	2/4/4/4	-
2	GOL	A	3003	-	-	0/4/4/4	-
2	GOL	B	3021	-	-	4/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 79 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	3016	GOL	C1-C2-C3-O3
2	A	3028	GOL	O1-C1-C2-C3
2	A	3030	GOL	O1-C1-C2-C3
2	A	3032	GOL	C1-C2-C3-O3
2	A	3032	GOL	O2-C2-C3-O3

There are no ring outliers.

7 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	3020	GOL	2	0
2	C	3022	GOL	5	0
2	H	3012	GOL	4	0
2	B	3004	GOL	1	0
2	D	3007	GOL	1	0
2	A	3023	GOL	1	0
2	B	3021	GOL	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	381/404 (94%)	-0.27	10 (2%) 56 54	13, 26, 58, 88	0
1	B	381/404 (94%)	-0.24	2 (0%) 91 90	15, 30, 52, 77	0
1	C	373/404 (92%)	-0.33	5 (1%) 77 76	15, 30, 51, 94	0
1	D	380/404 (94%)	0.14	17 (4%) 33 32	19, 39, 71, 104	0
1	E	381/404 (94%)	0.23	15 (3%) 39 38	25, 42, 66, 88	0
1	F	372/404 (92%)	-0.16	6 (1%) 72 70	18, 34, 59, 106	0
1	G	380/404 (94%)	0.15	14 (3%) 41 41	22, 41, 69, 94	0
1	H	368/404 (91%)	0.13	13 (3%) 44 43	20, 40, 62, 104	0
All	All	3016/3232 (93%)	-0.04	82 (2%) 54 53	13, 36, 63, 106	0

The worst 5 of 82 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	25	GLY	8.2
1	A	25	GLY	6.1
1	G	25	GLY	5.2
1	F	26	ASP	4.9
1	A	26	ASP	4.8

### 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(Å <sup>2</sup> )	Q<0.9
2	GOL	D	3007	6/6	0.72	0.22	48,62,68,69	0
3	NA	B	2001	1/1	0.72	0.20	62,62,62,62	0
2	GOL	A	3032	6/6	0.80	0.17	49,58,67,68	0
2	GOL	C	3020	6/6	0.81	0.31	31,61,64,66	0
2	GOL	G	3013	6/6	0.83	0.23	38,40,45,50	0
2	GOL	F	3011	6/6	0.83	0.24	61,73,74,80	0
3	NA	D	2002	1/1	0.83	0.52	104,104,104,104	0
2	GOL	C	3022	6/6	0.84	0.22	35,40,52,54	0
2	GOL	A	3023	6/6	0.84	0.19	38,62,66,69	0
2	GOL	H	3018	6/6	0.85	0.25	53,55,63,68	0
2	GOL	H	3012	6/6	0.86	0.19	53,65,72,81	0
2	GOL	E	3010	6/6	0.86	0.26	61,66,73,74	0
2	GOL	H	3025	6/6	0.86	0.20	46,59,65,65	0
2	GOL	A	3014	6/6	0.86	0.17	54,60,62,70	0
2	GOL	A	3024	6/6	0.86	0.17	54,57,64,64	0
2	GOL	D	3005	6/6	0.87	0.21	56,56,57,60	0
2	GOL	F	3008	6/6	0.88	0.21	41,54,58,74	0
2	GOL	B	3021	6/6	0.89	0.21	38,54,58,60	0
2	GOL	A	3016	6/6	0.90	0.17	57,63,68,74	0
2	GOL	D	3006	6/6	0.90	0.15	46,57,62,64	0
2	GOL	D	3017	6/6	0.91	0.27	46,51,58,63	0
2	GOL	C	3027	6/6	0.91	0.25	45,62,67,70	0
2	GOL	G	3009	6/6	0.91	0.15	35,39,56,58	0
2	GOL	C	3026	6/6	0.92	0.22	56,69,75,81	0
2	GOL	C	3001	6/6	0.92	0.15	42,55,60,65	0
2	GOL	A	3028	6/6	0.92	0.26	48,60,69,75	0
2	GOL	A	3030	6/6	0.92	0.34	57,60,71,72	0
2	GOL	B	3002	6/6	0.93	0.15	35,50,66,82	0
2	GOL	B	3004	6/6	0.93	0.13	37,41,47,55	0
2	GOL	C	3019	6/6	0.93	0.10	47,52,57,62	0
2	GOL	A	3031	6/6	0.94	0.23	34,53,62,69	0
2	GOL	B	3029	6/6	0.94	0.13	44,62,70,71	0
2	GOL	A	3003	6/6	0.95	0.09	31,54,58,60	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.