



# Full wwPDB X-ray Structure Validation Report ⓘ

May 13, 2020 – 03:59 am BST

PDB ID : 5HD6  
Title : High resolution structure of 3-hydroxydecanoyl-(acyl carrier protein) dehydratase from *Yersinia pestis* at 1.35 Å  
Authors : Chang, C.; Maltseva, N.; Kim, Y.; Mulligan, R.; Kwon, K.; Anderson, W.F.; Joachimiak, A.; Center for Structural Genomics of Infectious Diseases (CS-GID)  
Deposited on : 2016-01-04  
Resolution : 1.35 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

<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

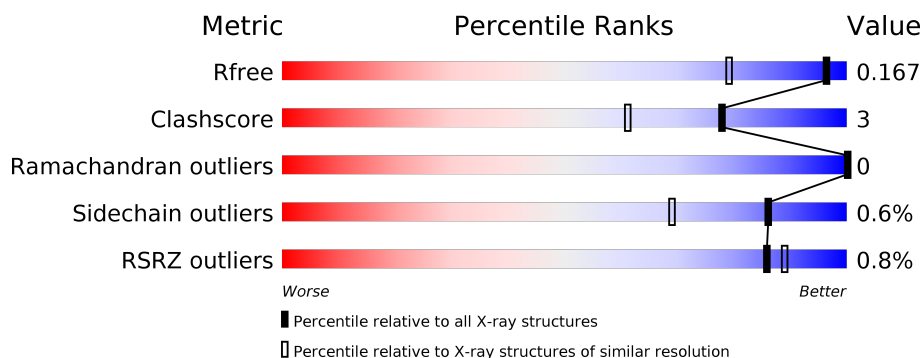
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.35 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	1509 (1.38-1.34)
Clashscore	141614	1551 (1.38-1.34)
Ramachandran outliers	138981	1530 (1.38-1.34)
Sidechain outliers	138945	1530 (1.38-1.34)
RSRZ outliers	127900	1487 (1.38-1.34)

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	175	<div> <div>%</div> <div> <div></div> <div>94%</div> <div>6%</div> </div> </div>
1	B	175	<div> <div></div> <div>93%</div> <div>7%</div> </div>
1	C	175	<div> <div>%</div> <div> <div></div> <div>95%</div> <div>5%</div> </div> </div>
1	D	175	<div> <div>%</div> <div> <div></div> <div>93%</div> <div>7%</div> </div> </div>
1	E	175	<div> <div>2%</div> <div> <div></div> <div>94%</div> <div>5%</div> <div>..</div> </div> </div>
1	F	175	<div> <div></div> <div>89%</div> <div>7%</div> <div>..</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	175	<div><div>%</div><div><div></div><div>87%</div><div>9%</div><div></div></div><div></div></div>
1	H	175	<div><div>%</div><div><div></div><div>91%</div><div>7%</div><div></div></div><div></div></div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 12776 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-hydroxydecanoyl-[acyl-carrier-protein] dehydratase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	175	Total	C	N	O	S	Se	0	6	0
			1390	891	232	255	2	10			
1	B	175	Total	C	N	O	S	Se	0	6	0
			1389	888	230	259	2	10			
1	C	175	Total	C	N	O	S	Se	0	4	0
			1372	879	228	253	2	10			
1	D	175	Total	C	N	O	S	Se	0	8	0
			1407	899	233	263	2	10			
1	E	173	Total	C	N	O	S	Se	0	6	0
			1372	884	229	247	2	10			
1	F	168	Total	C	N	O	S	Se	0	9	0
			1367	881	229	246	2	9			
1	G	168	Total	C	N	O	S	Se	0	7	0
			1347	869	224	242	2	10			
1	H	172	Total	C	N	O	S	Se	0	11	0
			1402	904	233	251	2	12			

There are 24 discrepancies between the modelled and reference sequences:

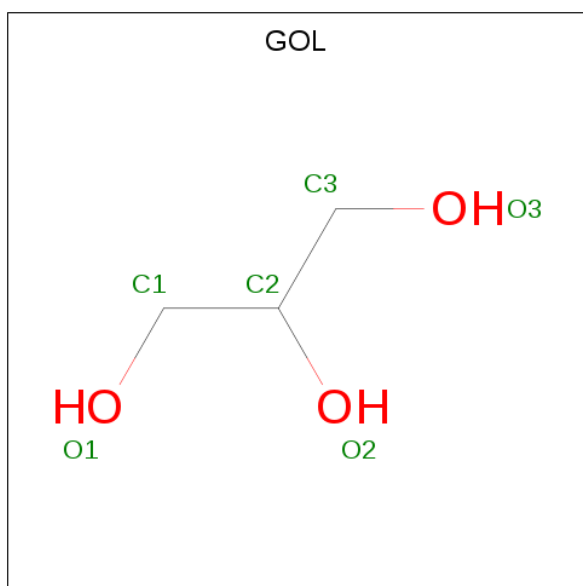
Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP A0A0B6NZG1
A	-1	ASN	-	expression tag	UNP A0A0B6NZG1
A	0	ALA	-	expression tag	UNP A0A0B6NZG1
B	-2	SER	-	expression tag	UNP A0A0B6NZG1
B	-1	ASN	-	expression tag	UNP A0A0B6NZG1
B	0	ALA	-	expression tag	UNP A0A0B6NZG1
C	-2	SER	-	expression tag	UNP A0A0B6NZG1
C	-1	ASN	-	expression tag	UNP A0A0B6NZG1
C	0	ALA	-	expression tag	UNP A0A0B6NZG1
D	-2	SER	-	expression tag	UNP A0A0B6NZG1
D	-1	ASN	-	expression tag	UNP A0A0B6NZG1
D	0	ALA	-	expression tag	UNP A0A0B6NZG1
E	-2	SER	-	expression tag	UNP A0A0B6NZG1

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	-1	ASN	-	expression tag	UNP A0A0B6NZG1
E	0	ALA	-	expression tag	UNP A0A0B6NZG1
F	-2	SER	-	expression tag	UNP A0A0B6NZG1
F	-1	ASN	-	expression tag	UNP A0A0B6NZG1
F	0	ALA	-	expression tag	UNP A0A0B6NZG1
G	-2	SER	-	expression tag	UNP A0A0B6NZG1
G	-1	ASN	-	expression tag	UNP A0A0B6NZG1
G	0	ALA	-	expression tag	UNP A0A0B6NZG1
H	-2	SER	-	expression tag	UNP A0A0B6NZG1
H	-1	ASN	-	expression tag	UNP A0A0B6NZG1
H	0	ALA	-	expression tag	UNP A0A0B6NZG1

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



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		
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		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
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	1
			12	6	6		
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 water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	204	Total	O	0	0
			204	204		
3	B	223	Total	O	0	0
			223	223		
3	C	196	Total	O	0	0
			196	196		
3	D	202	Total	O	0	0
			202	202		
3	E	197	Total	O	0	0
			197	197		
3	F	221	Total	O	0	0
			221	221		
3	G	203	Total	O	0	0
			203	203		
3	H	188	Total	O	0	0
			188	188		

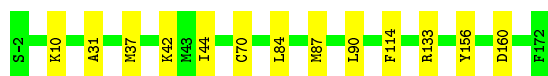
### 3 Residue-property plots [i](#)

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-hydroxydecanoyl-[acyl-carrier-protein] dehydratase



- Molecule 1: 3-hydroxydecanoyl-[acyl-carrier-protein] dehydratase



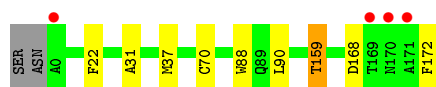
- Molecule 1: 3-hydroxydecanoyl-[acyl-carrier-protein] dehydratase




- Molecule 1: 3-hydroxydecanoyl-[acyl-carrier-protein] dehydratase



- Molecule 1: 3-hydroxydecanoyl-[acyl-carrier-protein] dehydratase




- Molecule 1: 3-hydroxydecanoyl-[acyl-carrier-protein] dehydratase

Chain F:  89% 7% . .

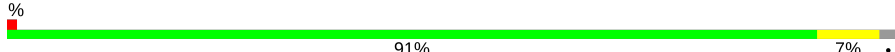


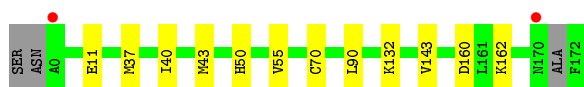
- Molecule 1: 3-hydroxydecanoyl-[acyl-carrier-protein] dehydratase

Chain G:  87% 9% .



- Molecule 1: 3-hydroxydecanoyl-[acyl-carrier-protein] dehydratase

Chain H:  91% 7% .



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	42.66 Å 50.82 Å 138.78 Å 90.02° 90.02° 90.10°	Depositor
Resolution (Å)	26.91 – 1.35 31.83 – 1.35	Depositor EDS
% Data completeness (in resolution range)	86.3 (26.91-1.35) 94.3 (31.83-1.35)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.83 (at 1.35 Å)	Xtriage
Refinement program	PHENIX 1.10_2155	Depositor
R, $R_{free}$	0.136 , 0.170 0.133 , 0.167	Depositor DCC
$R_{free}$ test set	11804 reflections (4.88%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	9.8	Xtriage
Anisotropy	0.373	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 44.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.459 for h,-k,-l 0.458 for -h,k,-l 0.470 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	12776	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.80% 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: 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	0.35	0/1409	0.59	0/1877
1	B	0.35	0/1408	0.59	0/1878
1	C	0.35	0/1391	0.59	0/1855
1	D	0.36	0/1426	0.59	0/1901
1	E	0.37	0/1392	0.57	0/1855
1	F	0.34	0/1386	0.58	0/1844
1	G	0.35	0/1367	0.57	0/1820
1	H	0.36	0/1421	0.57	0/1889
All	All	0.35	0/11200	0.58	0/14919

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 [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	1390	0	1392	8	0
1	B	1389	0	1375	8	0
1	C	1372	0	1367	8	0
1	D	1407	0	1392	9	0
1	E	1372	0	1379	7	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	1367	0	1368	10	0
1	G	1347	0	1349	11	0
1	H	1402	0	1405	8	2
2	A	6	0	8	0	0
2	B	6	0	8	0	0
2	C	18	0	24	2	0
2	D	12	0	16	0	2
2	E	6	0	8	1	0
2	F	12	0	16	1	0
2	G	18	0	24	3	0
2	H	18	0	24	0	0
3	A	204	0	0	2	0
3	B	223	0	0	2	0
3	C	196	0	0	1	0
3	D	202	0	0	5	0
3	E	197	0	0	2	0
3	F	221	0	0	1	0
3	G	203	0	0	5	0
3	H	188	0	0	0	0
All	All	12776	0	11155	64	2

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.

All (64) 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:F:3:ASP:N	3:F:301:HOH:O	2.21	0.73
1:G:3:ASP:N	3:G:303:HOH:O	2.25	0.70
1:C:167:LYS:H	2:C:202:GOL:H11	1.59	0.67
1:G:153:LYS:NZ	3:G:304:HOH:O	2.27	0.67
1:G:22[B]:PHE:O	3:G:301:HOH:O	2.13	0.65
1:E:22[A]:PHE:O	3:E:301:HOH:O	2.14	0.65
1:E:37[A]:MSE:SE	1:E:90:LEU:HD21	2.50	0.62
1:H:37[A]:MSE:SE	1:H:90:LEU:HD21	2.50	0.61
1:A:10[A]:LYS:NZ	3:A:301:HOH:O	2.34	0.61
1:D:168[A]:ASP:OD2	3:D:301:HOH:O	2.16	0.58
1:H:143:VAL:HG22	1:H:162[B]:LYS:HG2	1.87	0.57
1:C:167:LYS:N	2:C:202:GOL:H11	2.19	0.57
1:B:42:LYS:HD3	1:B:44:ILE:HD11	1.86	0.56
1:E:88:TRP:HB3	2:E:201:GOL:H12	1.86	0.56

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:37[B]:MSE:HE3	1:G:90:LEU:HD21	1.88	0.55
1:H:43[B]:MSE:SE	1:H:55:VAL:HG13	2.56	0.55
1:G:43[B]:MSE:SE	1:G:55:VAL:HG13	2.58	0.54
1:F:103:LYS:HE2	1:F:169:THR:OG1	2.07	0.54
1:A:167:LYS:NZ	3:A:303:HOH:O	2.41	0.53
1:B:10:LYS:NZ	3:B:305:HOH:O	2.40	0.53
1:C:20:GLU:OE2	3:C:302:HOH:O	2.19	0.52
1:C:101:GLU:HB2	1:C:167:LYS:HE3	1.92	0.51
1:F:37[A]:MSE:SE	1:F:90:LEU:HD21	2.61	0.51
2:G:202[B]:GOL:H12	3:G:311:HOH:O	2.12	0.50
1:F:37[B]:MSE:HE1	1:F:89:GLN:HB3	1.95	0.48
1:A:37[A]:MSE:SE	1:A:90:LEU:HD21	2.64	0.48
1:A:72:PHE:HZ	2:G:202[B]:GOL:H11	1.77	0.48
1:C:114:PHE:HD1	1:C:156:TYR:CD1	2.31	0.48
1:F:37[B]:MSE:HE1	1:F:89:GLN:CB	2.44	0.48
1:B:37[A]:MSE:SE	1:B:90:LEU:HD21	2.64	0.47
1:A:72:PHE:CZ	2:G:202[B]:GOL:H11	2.49	0.47
1:C:31:ALA:O	1:E:70:CYS:HB2	2.16	0.46
1:A:114:PHE:HD1	1:A:156:TYR:CD1	2.34	0.45
1:A:31:ALA:O	1:G:70:CYS:HB2	2.16	0.45
1:B:31:ALA:O	1:F:70:CYS:HB2	2.17	0.45
1:C:37[A]:MSE:SE	1:C:90:LEU:HD21	2.66	0.45
1:D:153:LYS:NZ	3:D:306:HOH:O	2.38	0.45
1:D:37[A]:MSE:SE	1:D:90:LEU:HD21	2.66	0.44
1:G:159[A]:THR:HG21	3:G:381:HOH:O	2.17	0.44
1:D:31:ALA:O	1:H:70:CYS:HB2	2.18	0.44
1:D:101[B]:GLU:HG2	3:D:320:HOH:O	2.17	0.44
1:G:133:ARG:HG3	1:G:135:ILE:HG13	2.00	0.44
1:H:50:HIS:CD2	1:H:132:LYS:HG2	2.53	0.44
1:F:128:ARG:NE	1:F:147[B]:GLU:OE2	2.49	0.43
1:A:-1:ASN:HB2	1:A:101:GLU:CG	2.49	0.43
1:F:88:TRP:HB3	2:F:202:GOL:C3	2.48	0.43
1:D:-1:ASN:HB2	1:D:101[A]:GLU:CG	2.48	0.43
1:G:123[A]:LYS:HD2	1:G:123[A]:LYS:HA	1.75	0.43
1:B:114:PHE:HD1	1:B:156:TYR:CD1	2.37	0.43
1:E:37[B]:MSE:HB2	1:E:37[B]:MSE:HE2	1.94	0.42
1:F:133[B]:ARG:HG3	1:F:135[B]:ILE:HG13	2.02	0.42
1:B:133:ARG:NH1	3:B:310:HOH:O	2.52	0.42
1:E:159[A]:THR:HG23	3:E:390:HOH:O	2.18	0.42
1:H:37[B]:MSE:HE2	1:H:37[B]:MSE:HB2	1.99	0.42
1:H:40:ILE:HG21	1:H:43[B]:MSE:SE	2.68	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:114:PHE:HD1	1:D:156:TYR:CD1	2.37	0.42
1:D:10:LYS:NZ	3:D:302:HOH:O	2.32	0.41
1:G:50:HIS:CD2	1:G:132:LYS:HG2	2.54	0.41
1:G:40:ILE:HG21	1:G:43[B]:MSE:SE	2.71	0.41
1:C:70:CYS:HB2	1:E:31:ALA:O	2.21	0.41
1:D:113[B]:LYS:NZ	3:D:310:HOH:O	2.54	0.41
1:B:84[B]:LEU:HD23	1:B:87:MSE:CE	2.52	0.40
1:B:70:CYS:HB2	1:F:31:ALA:O	2.21	0.40
1:H:160:ASP:O	1:H:162[A]:LYS:HG3	2.21	0.40

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:H:11[A]:GLU:OE2	2:D:201:GOL:C3[1_655]	1.37	0.83
1:H:11[A]:GLU:OE2	2:D:201:GOL:O3[1_655]	1.89	0.31

## 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	179/175 (102%)	176 (98%)	3 (2%)	0	100	100
1	B	179/175 (102%)	176 (98%)	3 (2%)	0	100	100
1	C	177/175 (101%)	174 (98%)	3 (2%)	0	100	100
1	D	181/175 (103%)	178 (98%)	3 (2%)	0	100	100
1	E	177/175 (101%)	170 (96%)	7 (4%)	0	100	100
1	F	174/175 (99%)	172 (99%)	2 (1%)	0	100	100
1	G	172/175 (98%)	169 (98%)	3 (2%)	0	100	100
1	H	180/175 (103%)	177 (98%)	3 (2%)	0	100	100

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1419/1400 (101%)	1392 (98%)	27 (2%)	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	144/129 (112%)	144 (100%)	0	100	100
1	B	144/129 (112%)	142 (99%)	2 (1%)	67	36
1	C	142/129 (110%)	142 (100%)	0	100	100
1	D	146/129 (113%)	144 (99%)	2 (1%)	67	36
1	E	141/129 (109%)	137 (97%)	4 (3%)	43	10
1	F	141/129 (109%)	139 (99%)	2 (1%)	67	36
1	G	139/129 (108%)	139 (100%)	0	100	100
1	H	143/129 (111%)	143 (100%)	0	100	100
All	All	1140/1032 (110%)	1130 (99%)	10 (1%)	86	53

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	160[A]	ASP
1	B	160[B]	ASP
1	D	160[A]	ASP
1	D	160[B]	ASP
1	E	159[A]	THR
1	E	159[B]	THR
1	E	168	ASP
1	E	172	PHE
1	F	37[A]	MSE
1	F	37[B]	MSE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no

such sidechains identified.

### 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 ⓘ

16 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	GOL	G	201	-	5,5,5	0.48	0	5,5,5	0.45	0
2	GOL	D	201	-	5,5,5	0.35	0	5,5,5	0.34	0
2	GOL	A	201	-	5,5,5	0.48	0	5,5,5	0.52	0
2	GOL	F	201	-	5,5,5	0.49	0	5,5,5	0.50	0
2	GOL	C	201	-	5,5,5	0.44	0	5,5,5	0.56	0
2	GOL	F	202	-	5,5,5	0.39	0	5,5,5	0.55	0
2	GOL	B	201	-	5,5,5	0.47	0	5,5,5	0.53	0
2	GOL	D	202	-	5,5,5	0.45	0	5,5,5	0.46	0
2	GOL	G	202[B]	-	5,5,5	0.35	0	5,5,5	0.50	0
2	GOL	H	203	-	5,5,5	0.38	0	5,5,5	0.28	0
2	GOL	G	202[A]	-	5,5,5	0.36	0	5,5,5	0.29	0
2	GOL	C	203	-	5,5,5	0.47	0	5,5,5	0.39	0
2	GOL	H	202	-	5,5,5	0.34	0	5,5,5	0.31	0
2	GOL	H	201	-	5,5,5	0.52	0	5,5,5	0.44	0
2	GOL	C	202	-	5,5,5	0.35	0	5,5,5	0.26	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	GOL	E	201	-	5,5,5	0.40	0	5,5,5	0.26	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	201	-	-	1/4/4/4	-
2	GOL	D	201	-	-	4/4/4/4	-
2	GOL	A	201	-	-	2/4/4/4	-
2	GOL	F	201	-	-	1/4/4/4	-
2	GOL	C	201	-	-	1/4/4/4	-
2	GOL	F	202	-	-	0/4/4/4	-
2	GOL	B	201	-	-	0/4/4/4	-
2	GOL	D	202	-	-	2/4/4/4	-
2	GOL	G	202[B]	-	-	2/4/4/4	-
2	GOL	H	203	-	-	2/4/4/4	-
2	GOL	G	202[A]	-	-	2/4/4/4	-
2	GOL	C	203	-	-	1/4/4/4	-
2	GOL	H	202	-	-	0/4/4/4	-
2	GOL	H	201	-	-	0/4/4/4	-
2	GOL	C	202	-	-	2/4/4/4	-
2	GOL	E	201	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (22) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	201	GOL	C1-C2-C3-O3
2	G	202[B]	GOL	O1-C1-C2-C3
2	G	202[A]	GOL	O1-C1-C2-C3
2	H	203	GOL	O1-C1-C2-O2
2	H	203	GOL	O1-C1-C2-C3
2	C	202	GOL	C1-C2-C3-O3
2	E	201	GOL	O1-C1-C2-C3
2	D	201	GOL	O1-C1-C2-C3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
2	A	201	GOL	O1-C1-C2-C3
2	C	201	GOL	O1-C1-C2-C3
2	D	202	GOL	O1-C1-C2-C3
2	G	202[B]	GOL	O1-C1-C2-O2
2	C	202	GOL	O2-C2-C3-O3
2	E	201	GOL	O1-C1-C2-O2
2	D	201	GOL	O1-C1-C2-O2
2	D	201	GOL	O2-C2-C3-O3
2	G	202[A]	GOL	O1-C1-C2-O2
2	F	201	GOL	O1-C1-C2-C3
2	D	202	GOL	O1-C1-C2-O2
2	C	203	GOL	O1-C1-C2-C3
2	A	201	GOL	O1-C1-C2-O2
2	G	201	GOL	O1-C1-C2-C3

There are no ring outliers.

5 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	201	GOL	0	2
2	F	202	GOL	1	0
2	G	202[B]	GOL	3	0
2	C	202	GOL	2	0
2	E	201	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 ⓘ

### 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	166/175 (94%)	-0.60	2 (1%) 79 83	7, 10, 21, 28	0
1	B	166/175 (94%)	-0.64	0 100 100	7, 10, 21, 26	0
1	C	166/175 (94%)	-0.61	1 (0%) 89 91	7, 10, 21, 28	0
1	D	166/175 (94%)	-0.62	1 (0%) 89 91	7, 10, 21, 28	0
1	E	164/175 (93%)	-0.51	4 (2%) 59 65	7, 11, 27, 51	0
1	F	160/175 (91%)	-0.61	0 100 100	7, 11, 21, 37	0
1	G	160/175 (91%)	-0.60	1 (0%) 89 91	7, 10, 22, 37	0
1	H	163/175 (93%)	-0.55	2 (1%) 79 83	7, 11, 24, 50	0
All	All	1311/1400 (93%)	-0.59	11 (0%) 86 89	7, 10, 22, 51	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	171	ALA	8.5
1	H	0	ALA	5.9
1	E	0	ALA	2.9
1	C	139	LEU	2.8
1	G	169	THR	2.8
1	A	139	LEU	2.7
1	D	139	LEU	2.5
1	H	170	ASN	2.4
1	E	170	ASN	2.1
1	E	169	THR	2.1
1	A	138[A]	LYS	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. 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	E	201	6/6	0.69	0.19	49,50,51,52	0
2	GOL	C	202	6/6	0.70	0.17	47,50,50,52	0
2	GOL	H	202	6/6	0.73	0.17	29,31,34,36	0
2	GOL	D	201	6/6	0.75	0.28	63,65,66,66	0
2	GOL	F	202	6/6	0.77	0.20	44,44,45,45	0
2	GOL	G	202[A]	6/6	0.79	0.19	23,28,29,30	6
2	GOL	G	202[B]	6/6	0.79	0.19	19,24,26,27	6
2	GOL	H	203	6/6	0.86	0.11	26,33,33,36	0
2	GOL	A	201	6/6	0.87	0.11	17,18,19,19	0
2	GOL	C	201	6/6	0.88	0.11	18,18,19,19	0
2	GOL	F	201	6/6	0.89	0.09	16,17,19,19	0
2	GOL	H	201	6/6	0.90	0.11	15,17,19,20	0
2	GOL	B	201	6/6	0.92	0.11	16,18,19,20	0
2	GOL	C	203	6/6	0.92	0.11	16,17,19,19	0
2	GOL	D	202	6/6	0.93	0.09	19,20,20,20	0
2	GOL	G	201	6/6	0.94	0.10	15,16,18,19	0

### 6.5 Other polymers ⓘ

There are no such residues in this entry.