WORLDWIDE PDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

May 26, 2020 – 12:50 pm BST

PDB ID	:	4L7Q
Title	:	Crystal structure of gamma glutamyl hydrolase (wild-type) from zebrafish
Authors	:	Chuankhayan, P.; Kao, T.-T.; Chen, C.-J.; Fu, T.-F.
Deposited on	:	2013-06-14
Resolution	:	2.10 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

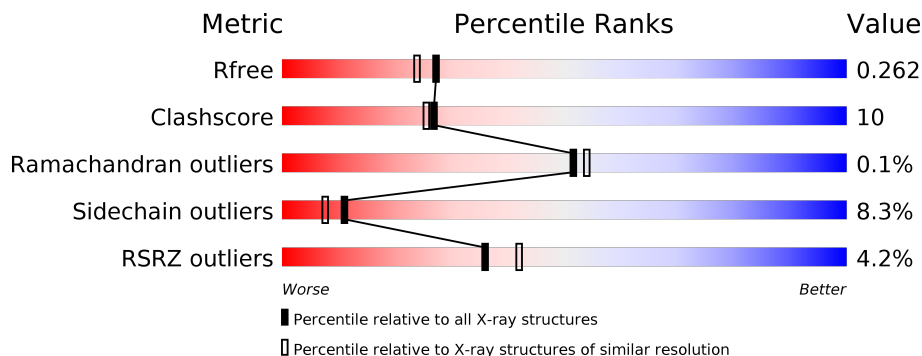|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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	Percentile Ranks	Value
Rfree		0.262
Clashscore		10
Ramachandran outliers		0.1%
Sidechain outliers		8.3%
RSRZ outliers		4.2%

Worse   Better

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 >=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 <=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.

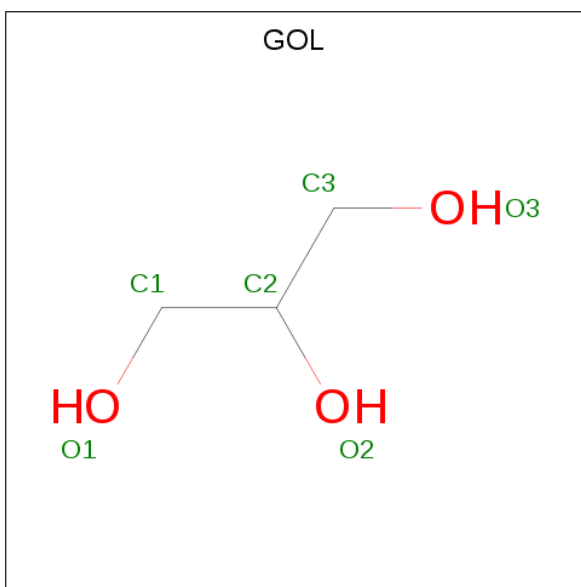| Mol	Chain	Length	Quality of chain
1	A	312	5% / 69% 19% · 8%
1	B	312	3% / 73% 14% ·· 8%
1	C	312	2% / 73% 13% 5% 8%
1	D	312	3% / 72% 16% ·· 8%
1	E	312	5% / 69% 21% · 8%
1	F	312	5% / 68% 21% ·· 8%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 14641 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 Gamma-glutamyl hydrolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	287	Total	C	N	O	S	0	0	0
			2309	1492	370	442	5			
1	D	287	Total	C	N	O	S	0	0	0
			2309	1492	370	442	5			
1	A	287	Total	C	N	O	S	0	0	0
			2309	1492	370	442	5			
1	B	287	Total	C	N	O	S	0	0	0
			2309	1492	370	442	5			
1	E	287	Total	C	N	O	S	0	0	0
			2309	1492	370	442	5			
1	F	287	Total	C	N	O	S	0	0	0
			2309	1492	370	442	5			

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

GOL

C3

OH O3

C1   C2

HO   OH

O1   O2

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

- Molecule 3 is water.

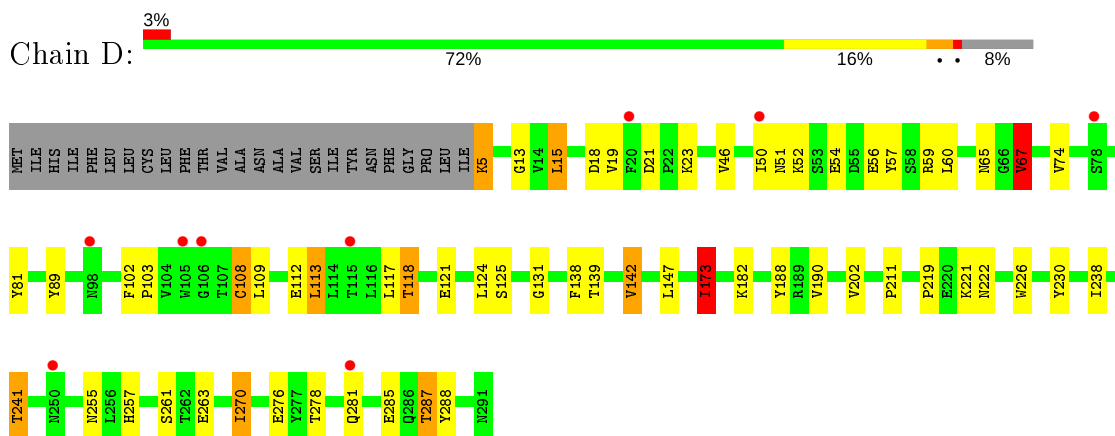| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	174	Total 174	O 174	0	0
3	D	129	Total 129	O 129	0	0
3	A	125	Total 125	O 125	0	0
3	B	125	Total 125	O 125	0	0
3	E	93	Total 93	O 93	0	0
3	F	105	Total 105	O 105	0	0

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.
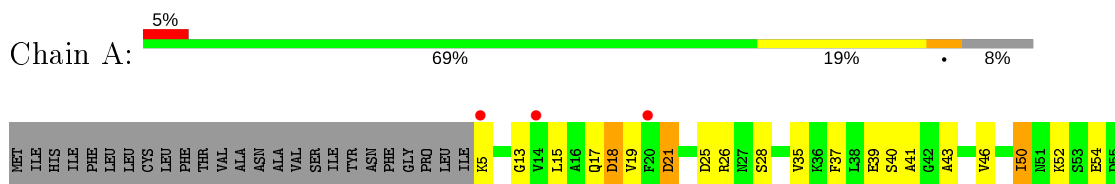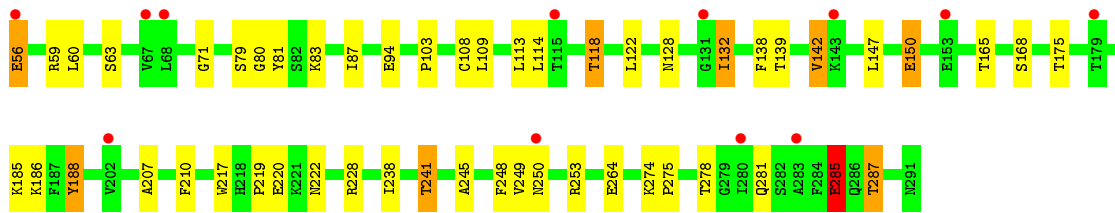
- Molecule 1: Gamma-glutamyl hydrolase
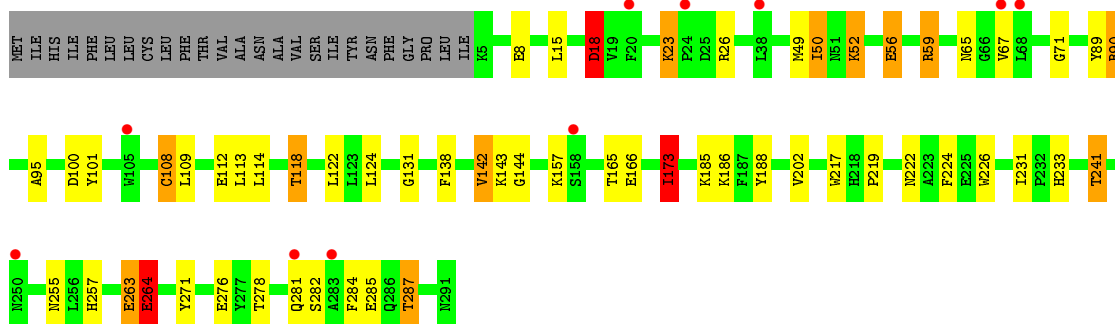

- Molecule 1: Gamma-glutamyl hydrolase

- Molecule 1: Gamma-glutamyl hydrolase

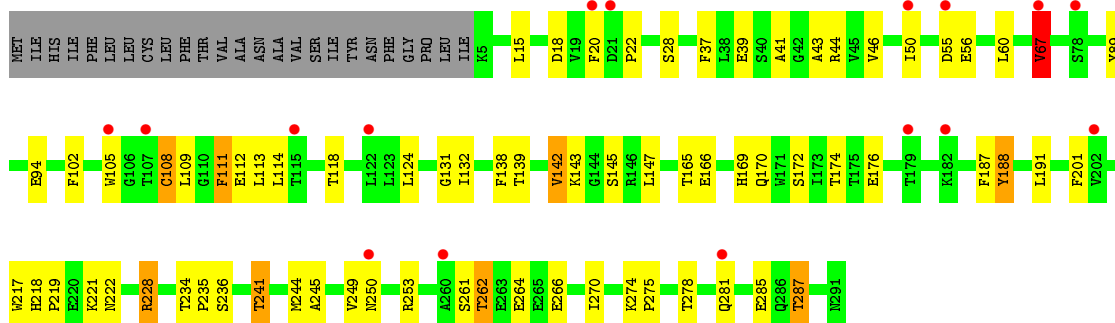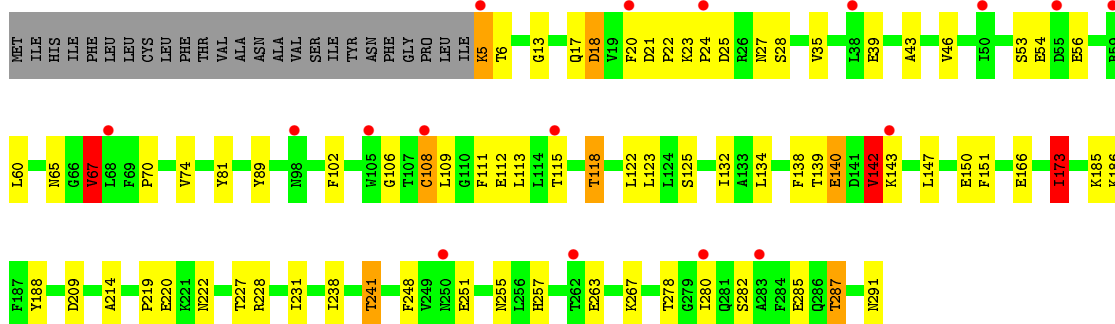E56 R59 L60 S63 V67 L68 G71 S79 G80 Y81 S82 K83 I87 E94 P103 C108 L109 L113 L114 T115 T118 L122 N128 G131 I132 F138 T139 V142 K143 L147 E150 E153 T165 S168 T175 T179

K185 K186 F187 Y188 V202 A207 F210 W217 H218 P219 E220 K221 N222 R228 I238 T241 A245 F248 V249 N250 R253 E264 K274 P275 T278 G279 I280 Q281 S282 A283 F284 E285 Q286 T287 N291

• Molecule 1: Gamma-glutamyl hydrolase

Chain B:

3%
73%   14%   • •   8%

MET HIS ILE PHE LEU LEU CYS LEU PHE THR VAL ALA ASN ALA VAL SER ILE TYR ASN PHE GLY PRO LEU ILE K5 E8 L15 D18 V19 F20 K23 P24 D25 R26 L38 M49 I50 N51 K52 E56 R59 N65 G66 V67 L68 G71 Y89 R90

A95 D100 Y101 W105 C108 L109 E112 L113 L114 T118 L122 L123 L124 G131 F138 V142 K143 G144 K157 S158 T165 E166 I173 K185 K186 F187 Y188 V202 W217 H218 P219 N222 A223 F224 E225 W226 I231 P232 H233 T241

N250 N255 L256 H257 E263 E264 Y271 E276 Y277 T278 Q281 S282 A283 F284 E285 Q286 T287 N291

• Molecule 1: Gamma-glutamyl hydrolase

Chain E:

5%
69%   21%   •   8%

MET HIS ILE PHE LEU LEU CYS LEU PHE THR VAL ALA ASN ALA VAL SER ILE TYR ASN PHE GLY PRO LEU ILE K5 L15 D18 V19 F20 D21 P22 S28 F37 L38 E39 S40 A41 G42 A43 R44 V45 V46 I50 D55 E56 L60 V67 S78 Y89

E94 F102 W105 G106 T107 C108 L109 G110 F111 E112 L113 L114 T115 T118 L122 L123 L124 G131 I132 F138 T139 V142 K143 G144 S145 R146 L147 T165 E166 H169 Q170 W171 S172 I173 T174 T175 E176 T179 K182 F187 Y188 L191 F201 V202

W217 H218 P219 E220 K221 N222 R228 T234 P235 S236 T241 M244 A245 V249 N250 R253 A260 S261 T262 E263 E264 E265 E266 I270 K274 P275 T278 Q281 E285 Q286 T287 N291

• Molecule 1: Gamma-glutamyl hydrolase

Chain F:

5%
68%   21%   • •   8%

MET HIS ILE PHE LEU LEU CYS LEU PHE THR VAL ALA ASN ALA VAL SER ILE TYR ASN PHE GLY PRO LEU ILE K5 T6 G13 Q17 D18 V19 F20 D21 P22 K23 P24 D25 R26 N27 S28 V35 L38 E39 A43 V46 I50 S53 E54 D55 E56 R59

L60 N65 G66 V67 L68 F69 P70 V74 Y81 Y89 N98 F102 W105 G106 T107 C108 L109 G110 F111 E112 L113 L114 T115 T118 L122 L123 L124 S125 I132 A133 L134 F138 T139 E140 D141 V142 K143 L147 E150 F151 E166 I173 K185 K186

F187 Y188 D209 A214 P219 E220 K221 N222 T227 R228 I231 I238 T241 F248 V249 N250 E251 N255 L256 H257 T262 E263 K267 T278 G279 I280 Q281 S282 A283 F284 E285 Q286 T287 N291

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	271.61Å 62.05Å 123.94Å 90.00° 95.41° 90.00°	Depositor
Resolution (Å)	30.00 – 2.10 28.20 – 2.10	Depositor EDS
% Data completeness (in resolution range)	95.7 (30.00-2.10) 97.1 (28.20-2.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.22 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.198 , 0.263 0.201 , 0.262	Depositor DCC
R_{free} test set	5869 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	29.2	Xtriage
Anisotropy	0.044	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 47.2	EDS
L-test for twinning[2]	$< |L| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	14641	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.61% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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

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	1.14	6/2374 (0.3%)	0.92	1/3221 (0.0%)
1	B	1.18	8/2374 (0.3%)	1.02	8/3221 (0.2%)
1	C	1.27	9/2374 (0.4%)	1.05	14/3221 (0.4%)
1	D	1.06	2/2374 (0.1%)	0.95	2/3221 (0.1%)
1	E	0.98	1/2374 (0.0%)	0.92	3/3221 (0.1%)
1	F	1.04	5/2374 (0.2%)	0.92	4/3221 (0.1%)
All	All	1.12	31/14244 (0.2%)	0.97	32/19326 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	1
1	C	0	1
1	E	0	1
All	All	0	4

The worst 5 of 31 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	228	ARG	C-N	10.43	1.54	1.34
1	C	150	GLU	CD-OE1	9.04	1.35	1.25
1	B	263	GLU	CG-CD	-9.00	1.38	1.51
1	B	56	GLU	CB-CG	8.52	1.68	1.52
1	A	285	GLU	CD-OE1	-7.95	1.17	1.25

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	228	ARG	O-C-N	-10.25	101.62	121.10
1	B	59	ARG	NE-CZ-NH2	-9.70	115.45	120.30
1	C	131	GLY	N-CA-C	7.57	132.03	113.10
1	B	50	ILE	C-N-CA	-7.53	102.87	121.70
1	B	173	ILE	CB-CA-C	-7.06	97.49	111.60

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	132	ILE	Mainchain
1	B	144	GLY	Peptide
1	C	144	GLY	Peptide
1	E	228	ARG	Mainchain

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	2309	0	2229	46	0
1	B	2309	0	2229	53	0
1	C	2309	0	2227	40	0
1	D	2309	0	2229	44	0
1	E	2309	0	2229	46	0
1	F	2309	0	2229	57	0
2	A	6	0	8	1	0
2	B	6	0	8	0	0
2	C	6	0	8	0	0
2	D	6	0	8	2	0
2	E	6	0	8	0	0
2	F	6	0	8	0	0
3	A	125	0	0	3	0
3	B	125	0	0	5	0
3	C	174	0	0	5	0
3	D	129	0	0	1	0
3	E	93	0	0	1	0
3	F	105	0	0	4	0
All	All	14641	0	13420	278	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 10.

The worst 5 of 278 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:56:GLU:HG3	1:B:59:ARG:NH2	1.41	1.32
1:F:65:ASN:HD22	1:F:255:ASN:ND2	1.50	1.07
1:F:132:ILE:HD13	1:F:134:LEU:HD21	1.42	1.02
1:B:56:GLU:CG	1:B:59:ARG:NH2	2.24	1.00
1:B:56:GLU:HG3	1:B:59:ARG:HH22	1.14	0.97

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	285/312 (91%)	271 (95%)	14 (5%)	0	100	100
1	B	285/312 (91%)	273 (96%)	11 (4%)	1 (0%)	34	32
1	C	285/312 (91%)	276 (97%)	9 (3%)	0	100	100
1	D	285/312 (91%)	269 (94%)	16 (6%)	0	100	100
1	E	285/312 (91%)	264 (93%)	20 (7%)	1 (0%)	34	32
1	F	285/312 (91%)	274 (96%)	11 (4%)	0	100	100
All	All	1710/1872 (91%)	1627 (95%)	81 (5%)	2 (0%)	51	54

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	131	GLY
1	E	131	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	250/272 (92%)	227 (91%)	23 (9%)	9	6
1	B	250/272 (92%)	233 (93%)	17 (7%)	16	13
1	C	250/272 (92%)	235 (94%)	15 (6%)	19	16
1	D	250/272 (92%)	226 (90%)	24 (10%)	8	5
1	E	250/272 (92%)	230 (92%)	20 (8%)	12	8
1	F	250/272 (92%)	225 (90%)	25 (10%)	7	5
All	All	1500/1632 (92%)	1376 (92%)	124 (8%)	11	7

5 of 124 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	186	LYS
1	B	143	LYS
1	F	142	VAL
1	A	238	ILE
1	B	15	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	250	ASN
1	B	272	ASN
1	F	216	GLN
1	A	167	ASN
1	F	255	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)

6 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	A	301	-	5,5,5	1.07	0	5,5,5	1.03	0
2	GOL	C	301	-	5,5,5	0.79	0	5,5,5	0.96	0
2	GOL	E	301	-	5,5,5	0.62	0	5,5,5	1.13	0
2	GOL	B	301	-	5,5,5	0.37	0	5,5,5	0.28	0
2	GOL	D	301	-	5,5,5	0.91	0	5,5,5	1.05	0
2	GOL	F	301	-	5,5,5	0.43	0	5,5,5	1.10	1 (20%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GOL	A	301	-	-	0/4/4/4	-
2	GOL	C	301	-	-	2/4/4/4	-
2	GOL	E	301	-	-	2/4/4/4	-
2	GOL	B	301	-	-	2/4/4/4	-
2	GOL	D	301	-	-	2/4/4/4	-
2	GOL	F	301	-	-	3/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	301	GOL	O2-C2-C3	2.15	118.61	109.12

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	301	GOL	C1-C2-C3-O3
2	E	301	GOL	C1-C2-C3-O3
2	B	301	GOL	O1-C1-C2-C3
2	D	301	GOL	C1-C2-C3-O3
2	F	301	GOL	C1-C2-C3-O3

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	301	GOL	1	0
2	D	301	GOL	2	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, 95th 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(Å²)	Q<0.9
1	A	287/312 (91%)	0.12	15 (5%)	27	32	17, 29, 46, 59	9 (3%)
1	B	287/312 (91%)	0.01	10 (3%)	44	50	13, 28, 43, 49	5 (1%)
1	C	287/312 (91%)	-0.16	6 (2%)	63	68	7, 22, 35, 46	9 (3%)
1	D	287/312 (91%)	0.13	9 (3%)	49	55	13, 29, 49, 59	5 (1%)
1	E	287/312 (91%)	0.26	16 (5%)	24	29	20, 36, 50, 58	9 (3%)
1	F	287/312 (91%)	0.26	17 (5%)	22	27	19, 32, 54, 61	5 (1%)
All	All	1722/1872 (91%)	0.10	73 (4%)	36	42	7, 29, 48, 61	42 (2%)

The worst 5 of 73 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	202	VAL	5.5
1	B	20	PHE	5.4
1	A	179	THR	4.5
1	E	20	PHE	4.1
1	E	50	ILE	4.1

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 (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, 95th 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^2)	Q<0.9
2	GOL	B	301	6/6	0.70	0.21	42,45,46,46	0
2	GOL	E	301	6/6	0.80	0.22	41,46,48,48	0
2	GOL	D	301	6/6	0.82	0.14	31,33,35,36	0
2	GOL	A	301	6/6	0.85	0.17	25,30,33,35	0
2	GOL	C	301	6/6	0.87	0.16	21,28,31,32	0
2	GOL	F	301	6/6	0.89	0.14	35,44,47,47	0

6.5 Other polymers (i)

There are no such residues in this entry.