W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Oct 31, 2023 – 06:36 PM JST

PDB ID	:	5GTG
Title	:	Crystal structure of onion lachrymatory factor synthase (LFS) containing 1,2-propanediol
Authors	:	Arakawa, T.; Sato, Y.; Takabe, J.; Masamura, N.; Tsuge, N.; Imai, S.; Fushinobu, S.
Deposited on	:	2016-08-20
Resolution	:	1.70 Å(reported)

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

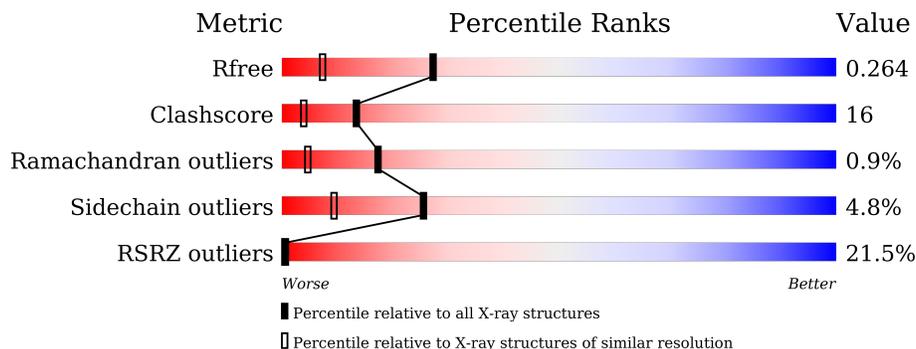|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.70 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		0.264
Clashscore		16
Ramachandran outliers		0.9%
Sidechain outliers		4.8%
RSRZ outliers		21.5%

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	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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 >=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	175	5% / 50% · 30% · 6% · 15%
1	B	175	52% · 29% · · 15%
1	C	175	47% / 53% · 29% · · · 15%
1	D	175	23% / 53% · 32% · · 15%
1	E	175	14% / 55% · 23% · 6% · 15%
1	F	175	9% / 43% · 32% · 9% · 15%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	175	17% / 58% / 23% / · / 15%
1	H	175	33% / 53% / 28% / 5% / 15%
2	I	2	50% / 50%

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
5	MES	B	204	-	X	-	-

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 10289 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 Lachrymatory-factor synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	149	Total	C	N	O	S	0	0	0
			1205	776	192	227	10			
1	B	149	Total	C	N	O	S	0	0	0
			1205	776	192	227	10			
1	C	149	Total	C	N	O	S	0	0	0
			1205	776	192	227	10			
1	D	149	Total	C	N	O	S	0	0	0
			1205	776	192	227	10			
1	E	149	Total	C	N	O	S	0	0	0
			1205	776	192	227	10			
1	F	149	Total	C	N	O	S	0	0	0
			1205	776	192	227	10			
1	G	149	Total	C	N	O	S	0	0	0
			1205	776	192	227	10			
1	H	149	Total	C	N	O	S	0	0	0
			1205	776	192	227	10			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	170	HIS	-	expression tag	UNP P59082
A	171	HIS	-	expression tag	UNP P59082
A	172	HIS	-	expression tag	UNP P59082
A	173	HIS	-	expression tag	UNP P59082
A	174	HIS	-	expression tag	UNP P59082
A	175	HIS	-	expression tag	UNP P59082
B	170	HIS	-	expression tag	UNP P59082
B	171	HIS	-	expression tag	UNP P59082
B	172	HIS	-	expression tag	UNP P59082
B	173	HIS	-	expression tag	UNP P59082
B	174	HIS	-	expression tag	UNP P59082
B	175	HIS	-	expression tag	UNP P59082
C	170	HIS	-	expression tag	UNP P59082

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	171	HIS	-	expression tag	UNP P59082
C	172	HIS	-	expression tag	UNP P59082
C	173	HIS	-	expression tag	UNP P59082
C	174	HIS	-	expression tag	UNP P59082
C	175	HIS	-	expression tag	UNP P59082
D	170	HIS	-	expression tag	UNP P59082
D	171	HIS	-	expression tag	UNP P59082
D	172	HIS	-	expression tag	UNP P59082
D	173	HIS	-	expression tag	UNP P59082
D	174	HIS	-	expression tag	UNP P59082
D	175	HIS	-	expression tag	UNP P59082
E	170	HIS	-	expression tag	UNP P59082
E	171	HIS	-	expression tag	UNP P59082
E	172	HIS	-	expression tag	UNP P59082
E	173	HIS	-	expression tag	UNP P59082
E	174	HIS	-	expression tag	UNP P59082
E	175	HIS	-	expression tag	UNP P59082
F	170	HIS	-	expression tag	UNP P59082
F	171	HIS	-	expression tag	UNP P59082
F	172	HIS	-	expression tag	UNP P59082
F	173	HIS	-	expression tag	UNP P59082
F	174	HIS	-	expression tag	UNP P59082
F	175	HIS	-	expression tag	UNP P59082
G	170	HIS	-	expression tag	UNP P59082
G	171	HIS	-	expression tag	UNP P59082
G	172	HIS	-	expression tag	UNP P59082
G	173	HIS	-	expression tag	UNP P59082
G	174	HIS	-	expression tag	UNP P59082
G	175	HIS	-	expression tag	UNP P59082
H	170	HIS	-	expression tag	UNP P59082
H	171	HIS	-	expression tag	UNP P59082
H	172	HIS	-	expression tag	UNP P59082
H	173	HIS	-	expression tag	UNP P59082
H	174	HIS	-	expression tag	UNP P59082
H	175	HIS	-	expression tag	UNP P59082

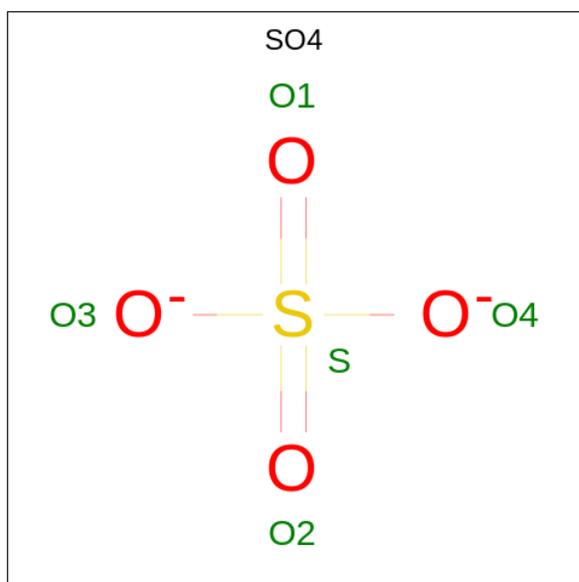- Molecule 2 is an oligosaccharide called alpha-D-glucopyranose-(1-1)-alpha-D-glucopyranose.
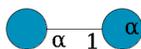
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	I	2	Total 23	C 12	O 11	0	0	0

- Molecule 3 is S-1,2-PROPANEDIOL (three-letter code: PGO) (formula: $C_3H_8O_2$).
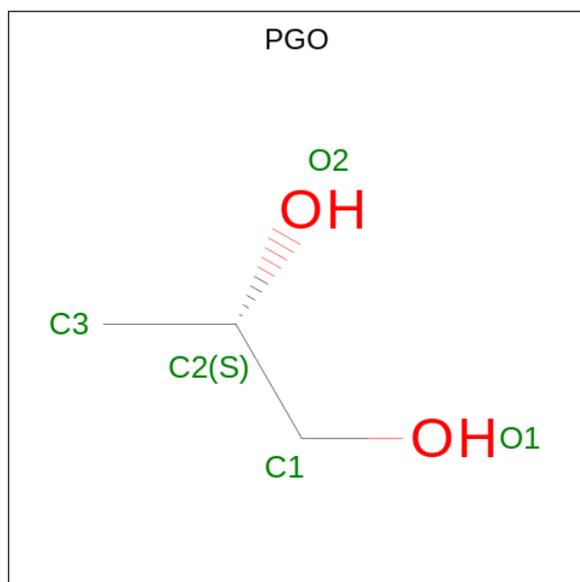
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 5	C 3	O 2	0	0
3	B	1	Total 5	C 3	O 2	0	0
3	D	1	Total 5	C 3	O 2	0	0
3	G	1	Total 5	C 3	O 2	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

SO4

O1

O3 O⁻ — S — O⁻ O4

S

O2

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total 5	O 4	S 1	0	0
4	B	1	Total 5	O 4	S 1	0	0

- Molecule 5 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: $C_6H_{13}NO_4S$).

MES

O1

C2   C6

C3   C5

H
N⁺
N4

C7

C8

O2S O = S = O O1S

O⁻

O3S
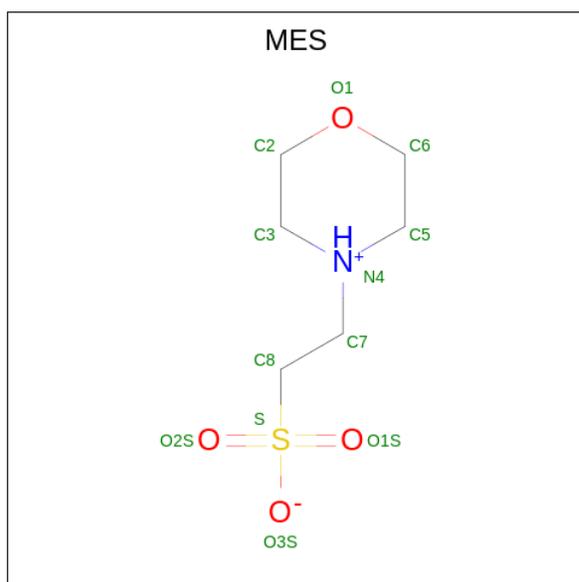
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	B	1	Total 12	C 6	N 1	O 4	S 1	0	0

- Molecule 6 is water.

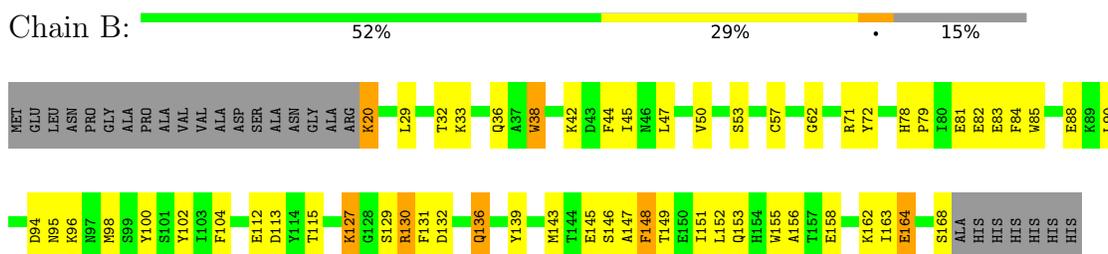| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	147	Total 147	O 147	0	0
6	B	111	Total 111	O 111	0	0
6	C	24	Total 24	O 24	0	0
6	D	42	Total 42	O 42	0	0
6	E	54	Total 54	O 54	0	0
6	F	86	Total 86	O 86	0	0
6	G	66	Total 66	O 66	0	0
6	H	54	Total 54	O 54	0	0

3 Residue-property plots ⓘ

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.
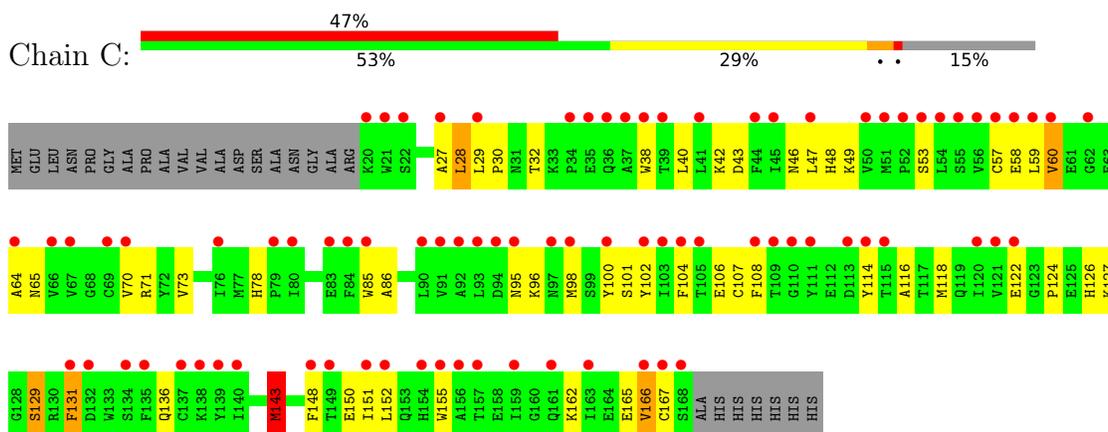
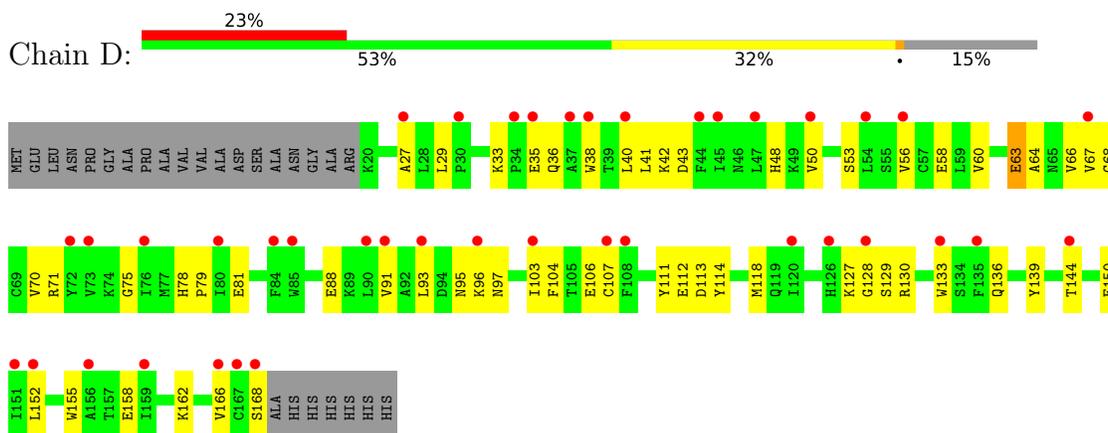- Molecule 1: Lachrymatory-factor synthase
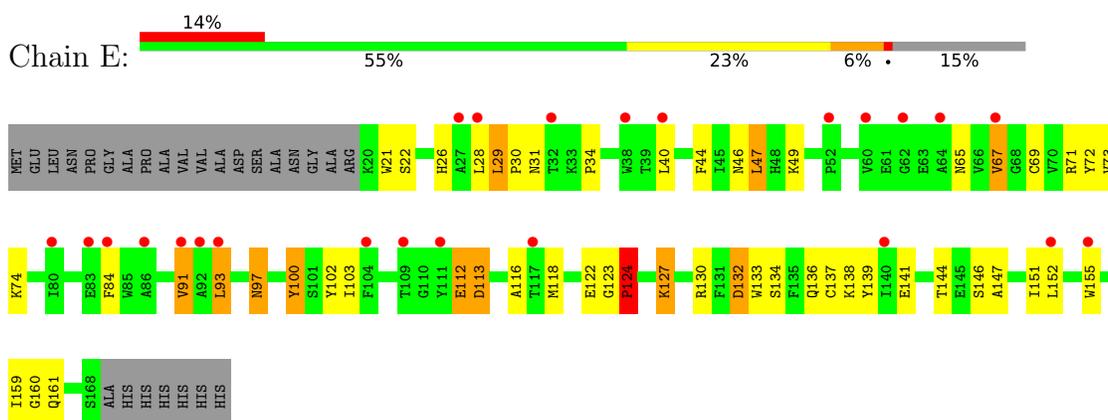
Chain A:

- Molecule 1: Lachrymatory-factor synthase
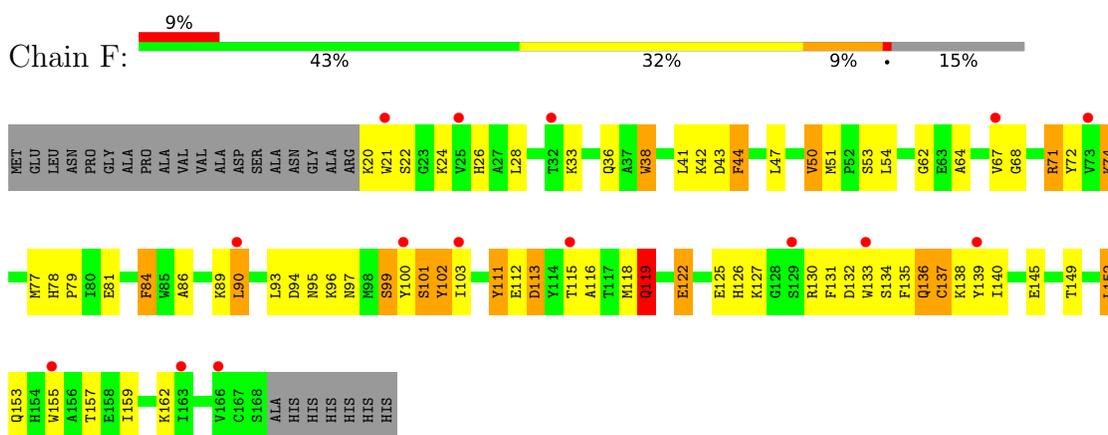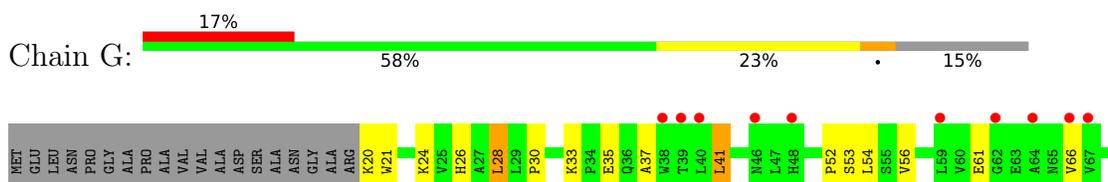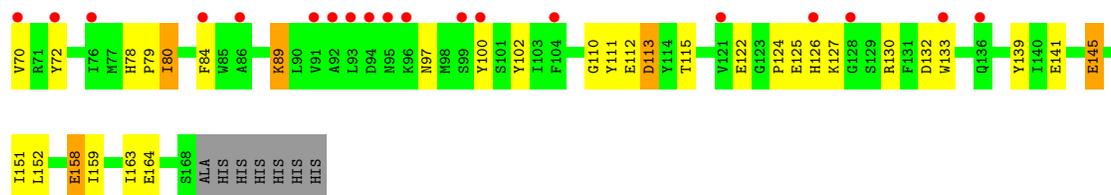
Chain B:

- Molecule 1: Lachrymatory-factor synthase

Chain C:

• Molecule 1: Lachrymatory-factor synthase

Chain D:

23%

53% 32% 15%

MET GLU LEU ASN PRO GLY ALA PRO ALA VAL VAL ALA ASP SER ALA ASN GLY ALA ARG K20 A27 L28 L29 P30 K33 P34 E35 Q36 A37 W38 T39 L40 L41 K42 D43 F44 I45 N46 L47 H48 K49 V50 S53 L54 S55 V56 C57 E58 L59 V60 E63 A64 N65 V66 V67 G68

C69 V70 R71 Y72 V73 K74 G75 I76 M77 H78 P79 I80 E81 F84 W85 E88 K89 L90 V91 A92 L93 D94 N95 K96 N97 I103 F104 T105 E106 C107 F108 Y111 E112 D113 Y114 M118 Q119 I120 H126 K127 G128 S129 R130 W133 S134 F135 Q136 Y139 T144 E150

I151 L152 W155 A156 T157 E158 I159 K162 V166 C167 S168 ALA HIS HIS HIS HIS HIS

• Molecule 1: Lachrymatory-factor synthase

Chain E:

14%

55% 23% 6% 15%

MET GLU LEU ASN PRO GLY ALA PRO ALA VAL VAL ALA ASP SER ALA ASN GLY ALA ARG K20 W21 S22 H26 A27 L28 L29 P30 N31 T32 K33 P34 W38 T39 L40 F44 I45 N46 L47 H48 K49 P52 V60 E61 G62 E63 A64 N65 V66 V67 G68 C69 V70 R71 Y72 V73

K74 I80 E83 F84 W85 A86 V91 A92 L93 N97 Y100 S101 Y102 I103 F104 T109 G110 Y111 E112 D113 A116 T117 M118 E122 G123 P124 K127 R130 F131 D132 W133 S134 F135 Q136 C137 K138 Y139 I140 E141 T144 E145 S146 A147 I151 L152 W155

I159 G160 Q161 S168 ALA HIS HIS HIS HIS HIS

• Molecule 1: Lachrymatory-factor synthase

Chain F:

9%

43% 32% 9% 15%

MET GLU LEU ASN PRO GLY ALA PRO ALA VAL VAL ALA ASP SER ALA ASN GLY ALA ARG K20 W21 S22 G23 K24 V25 H26 A27 L28 T32 K33 Q36 A37 W38 L41 K42 D43 F44 L47 V50 M51 P52 S53 L54 G62 E63 A64 V67 G68 R71 Y72 V73 K74

M77 H78 P79 I80 E81 F84 W85 A86 K89 L90 L93 D94 N95 K96 N97 M98 S99 Y100 S101 Y102 I103 Y111 E112 D113 Y114 T115 A116 M118 Q119 E122 E125 H126 K127 G128 S129 R130 F131 D132 W133 S134 F135 Q136 C137 K138 Y139 I140 E145 T149 L152

Q153 H154 W155 A156 T157 E158 I159 K162 I163 V166 C167 S168 ALA HIS HIS HIS HIS HIS

• Molecule 1: Lachrymatory-factor synthase

Chain G:

17%

58% 23% 15%

MET GLU LEU ASN PRO GLY ALA PRO ALA VAL VAL ALA ASP SER ALA ASN GLY ALA ARG K20 W21 K24 V25 H26 A27 L28 L29 P30 K33 P34 E35 Q36 A37 W38 T39 L40 L41 N46 L47 H48 P52 S53 L54 S55 V56 L59 V60 E61 G62 E63 A64 N65 V66 V67

• Molecule 1: Lachrymatory-factor synthase

Chain H:

• Molecule 2: alpha-D-glucopyranose-(1-1)-alpha-D-glucopyranose

Chain I:

4   Data and refinement statistics   ⓘ

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, α, β, γ	172.86Å 172.86Å 64.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.94 – 1.70 47.94 – 1.70	Depositor EDS
% Data completeness (in resolution range)	97.3 (47.94-1.70) 97.5 (47.94-1.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.73 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.8.0155	Depositor
R, R_{free}	0.230 , 0.260 0.239 , 0.264	Depositor DCC
R_{free} test set	10273 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	17.4	Xtriage
Anisotropy	0.083	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 46.5	EDS
L-test for twinning[2]	$< |L| > = 0.48$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.087 for h,-k,-l	Xtriage
Reported twinning fraction	0.501 for H, K, L 0.499 for -H, K, -L	Depositor
Outliers	1 of 205386 reflections (0.000%)	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	10289	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.17% 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: SO4, PGO, GLC, MES

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	2.08	36/1237 (2.9%)	1.64	18/1674 (1.1%)
1	B	1.95	26/1237 (2.1%)	1.69	23/1674 (1.4%)
1	C	1.17	1/1237 (0.1%)	1.18	4/1674 (0.2%)
1	D	1.36	5/1237 (0.4%)	1.26	4/1674 (0.2%)
1	E	1.45	8/1237 (0.6%)	1.39	12/1674 (0.7%)
1	F	1.64	13/1237 (1.1%)	1.52	17/1674 (1.0%)
1	G	1.41	3/1237 (0.2%)	1.33	4/1674 (0.2%)
1	H	1.44	4/1237 (0.3%)	1.30	3/1674 (0.2%)
All	All	1.59	96/9896 (1.0%)	1.42	85/13392 (0.6%)

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	C	0	1
1	F	0	2
All	All	0	5

The worst 5 of 96 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	20	LYS	N-CA	11.98	1.70	1.46
1	B	112	GLU	CD-OE2	-10.85	1.13	1.25
1	A	83	GLU	CD-OE1	10.11	1.36	1.25
1	A	133	TRP	CG-CD1	9.77	1.50	1.36
1	A	62	GLY	C-O	8.37	1.37	1.23

The worst 5 of 85 bond angle outliers are listed below:

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	132	ASP	CB-CG-OD1	10.98	128.18	118.30
1	B	127	LYS	CD-CE-NZ	9.58	133.72	111.70
1	F	152	LEU	CB-CG-CD1	9.11	126.48	111.00
1	F	71	ARG	NE-CZ-NH2	-9.11	115.75	120.30
1	B	113	ASP	CB-CG-OD1	8.21	125.69	118.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	139	TYR	Mainchain
1	A	97	ASN	Mainchain
1	C	58	GLU	Peptide
1	F	135	PHE	Peptide
1	F	86	ALA	Peptide

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	1205	0	1165	26	0
1	B	1205	0	1165	26	0
1	C	1205	0	1165	33	0
1	D	1205	0	1165	46	0
1	E	1205	0	1165	49	0
1	F	1205	0	1165	56	0
1	G	1205	0	1165	31	0
1	H	1205	0	1165	54	0
2	I	23	0	20	3	0
3	A	5	0	8	2	0
3	B	5	0	7	0	0
3	D	5	0	8	1	0
3	G	5	0	8	1	0
4	B	10	0	0	1	0
5	B	12	0	13	4	0
6	A	147	0	0	11	0
6	B	111	0	0	7	0
6	C	24	0	0	5	0

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	D	42	0	0	16	0
6	E	54	0	0	17	0
6	F	86	0	0	18	0
6	G	66	0	0	8	0
6	H	54	0	0	19	0
All	All	10289	0	9384	308	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 16.

The worst 5 of 308 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:A:20:LYS:N	1:A:20:LYS:CA	1.70	1.52
1:C:155:TRP:HD1	6:C:201:HOH:O	1.27	1.16
1:B:127:LYS:HE2	6:B:346:HOH:O	1.49	1.11
1:H:151:ILE:HG12	6:H:205:HOH:O	1.51	1.08
1:A:46:ASN:HD21	2:I:1:GLC:H61	1.19	1.07

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	147/175 (84%)	144 (98%)	3 (2%)	0	100	100
1	B	147/175 (84%)	146 (99%)	1 (1%)	0	100	100
1	C	147/175 (84%)	135 (92%)	8 (5%)	4 (3%)	5	0
1	D	147/175 (84%)	137 (93%)	9 (6%)	1 (1%)	22	8
1	E	147/175 (84%)	137 (93%)	8 (5%)	2 (1%)	11	2

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	147/175 (84%)	142 (97%)	4 (3%)	1 (1%)	22	8
1	G	147/175 (84%)	143 (97%)	4 (3%)	0	100	100
1	H	147/175 (84%)	136 (92%)	8 (5%)	3 (2%)	7	1
All	All	1176/1400 (84%)	1120 (95%)	45 (4%)	11 (1%)	17	5

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	139	TYR
1	E	124	PRO
1	F	42	LYS
1	C	59	LEU
1	C	64	ALA

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	132/150 (88%)	130 (98%)	2 (2%)	65	51
1	B	132/150 (88%)	129 (98%)	3 (2%)	50	33
1	C	132/150 (88%)	122 (92%)	10 (8%)	13	3
1	D	132/150 (88%)	129 (98%)	3 (2%)	50	33
1	E	132/150 (88%)	123 (93%)	9 (7%)	16	4
1	F	132/150 (88%)	122 (92%)	10 (8%)	13	3
1	G	132/150 (88%)	124 (94%)	8 (6%)	18	5
1	H	132/150 (88%)	126 (96%)	6 (4%)	27	10
All	All	1056/1200 (88%)	1005 (95%)	51 (5%)	25	9

5 of 51 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	53	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	F	136	GLN
1	H	152	LEU
1	F	84	PHE
1	F	125	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	154	HIS
1	H	31	ASN
1	H	136	GLN
1	D	78	HIS
1	C	95	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

2 monosaccharides 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	GLC	I	1	2	11,11,12	1.55	2 (18%)	15,15,17	2.27	7 (46%)
2	GLC	I	2	2	12,12,12	1.45	1 (8%)	17,17,17	2.29	7 (41%)

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	GLC	I	1	2	-	2/2/19/22	0/1/1/1
2	GLC	I	2	2	-	1/2/22/22	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	I	1	GLC	O5-C1	3.68	1.49	1.43
2	I	2	GLC	O5-C5	-2.93	1.37	1.44
2	I	1	GLC	C6-C5	2.08	1.58	1.51

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	I	2	GLC	C3-C4-C5	4.61	118.47	110.24
2	I	1	GLC	O5-C5-C4	4.30	121.30	110.83
2	I	1	GLC	C6-C5-C4	-4.25	103.05	113.00
2	I	2	GLC	C4-C3-C2	4.17	118.11	110.82
2	I	2	GLC	O4-C4-C5	-4.08	99.16	109.30

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	I	1	GLC	C4-C5-C6-O6
2	I	1	GLC	O5-C5-C6-O6
2	I	2	GLC	O5-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	I	1	GLC	3	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

Oligosaccharide Chain I

Bond lengths

Bond angles

Torsions

Rings
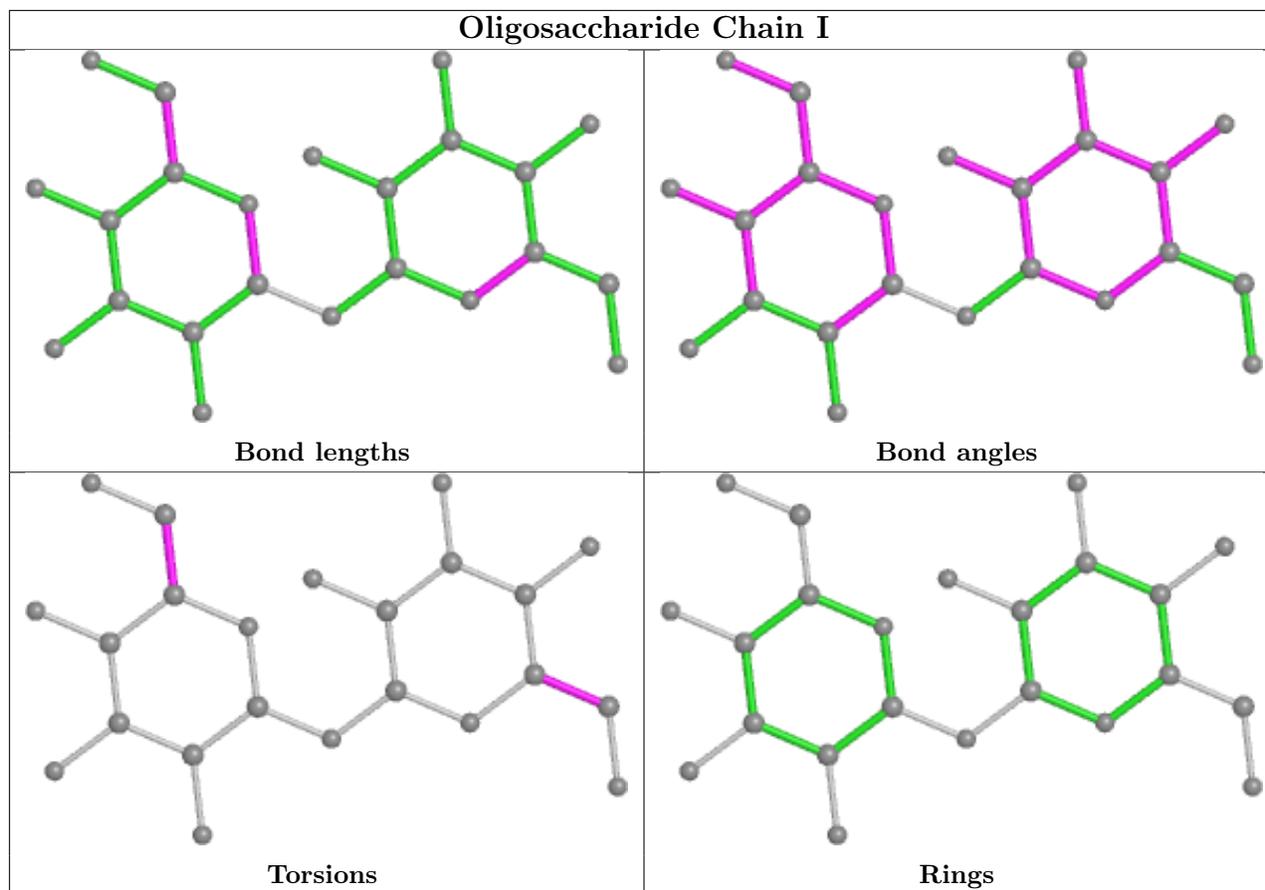
5.6 Ligand geometry ⓘ

7 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	#$|Z| > 2$
3	PGO	A	201	-	3,4,4	0.45	0	1,4,4	1.52	0
5	MES	B	204	-	12,12,12	2.36	5 (41%)	14,16,16	4.11	9 (64%)
4	SO4	B	201	-	4,4,4	0.90	0	6,6,6	0.60	0
3	PGO	G	201	-	3,4,4	0.43	0	1,4,4	0.93	0
4	SO4	B	202	-	4,4,4	0.44	0	6,6,6	1.12	0
3	PGO	D	201	-	3,4,4	0.20	0	1,4,4	0.27	0
3	PGO	B	203	-	3,4,4	0.54	0	1,4,4	1.46	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	PGO	A	201	-	-	2/2/2/2	-
5	MES	B	204	-	-	5/6/14/14	0/1/1/1
3	PGO	G	201	-	-	2/2/2/2	-
3	PGO	D	201	-	-	2/2/2/2	-
3	PGO	B	203	-	-	2/2/2/2	-

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	204	MES	C8-S	-6.40	1.68	1.77
5	B	204	MES	C7-N4	2.48	1.53	1.47
5	B	204	MES	O2S-S	2.41	1.52	1.45
5	B	204	MES	O1S-S	2.18	1.51	1.45
5	B	204	MES	C5-C6	2.09	1.58	1.50

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	204	MES	O1S-S-C8	12.39	121.83	106.92
5	B	204	MES	C5-N4-C3	-3.80	100.28	108.83
5	B	204	MES	O3S-S-O1S	-3.70	102.23	111.27
5	B	204	MES	O1-C2-C3	3.69	119.91	111.80
5	B	204	MES	C2-C3-N4	3.43	115.30	110.10

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	203	PGO	O1-C1-C2-O2
3	G	201	PGO	O1-C1-C2-O2
5	B	204	MES	C8-C7-N4-C3
5	B	204	MES	C7-C8-S-O2S
5	B	204	MES	C7-C8-S-O3S

There are no ring outliers.

5 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	201	PGO	2	0
5	B	204	MES	4	0
3	G	201	PGO	1	0
4	B	202	SO4	1	0
3	D	201	PGO	1	0

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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(Å^2)	Q<0.9
1	A	149/175 (85%)	0.84	8 (5%)	25	28	4, 10, 19, 28	0
1	B	149/175 (85%)	0.59	0	100	100	5, 11, 21, 25	0
1	C	149/175 (85%)	2.53	83 (55%)	0	0	19, 29, 38, 42	149 (100%)
1	D	149/175 (85%)	1.52	40 (26%)	0	0	15, 23, 33, 43	149 (100%)
1	E	149/175 (85%)	1.21	24 (16%)	1	1	12, 20, 28, 36	149 (100%)
1	F	149/175 (85%)	1.19	15 (10%)	7	8	10, 18, 27, 33	149 (100%)
1	G	149/175 (85%)	1.28	29 (19%)	1	1	13, 22, 31, 40	149 (100%)
1	H	149/175 (85%)	1.86	57 (38%)	0	0	13, 23, 33, 43	149 (100%)
All	All	1192/1400 (85%)	1.38	256 (21%)	0	0	4, 20, 33, 43	894 (75%)

The worst 5 of 256 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	91	VAL	8.3
1	C	155	TRP	7.9
1	C	47	LEU	6.8
1	C	92	ALA	6.3
1	C	45	ILE	6.3

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

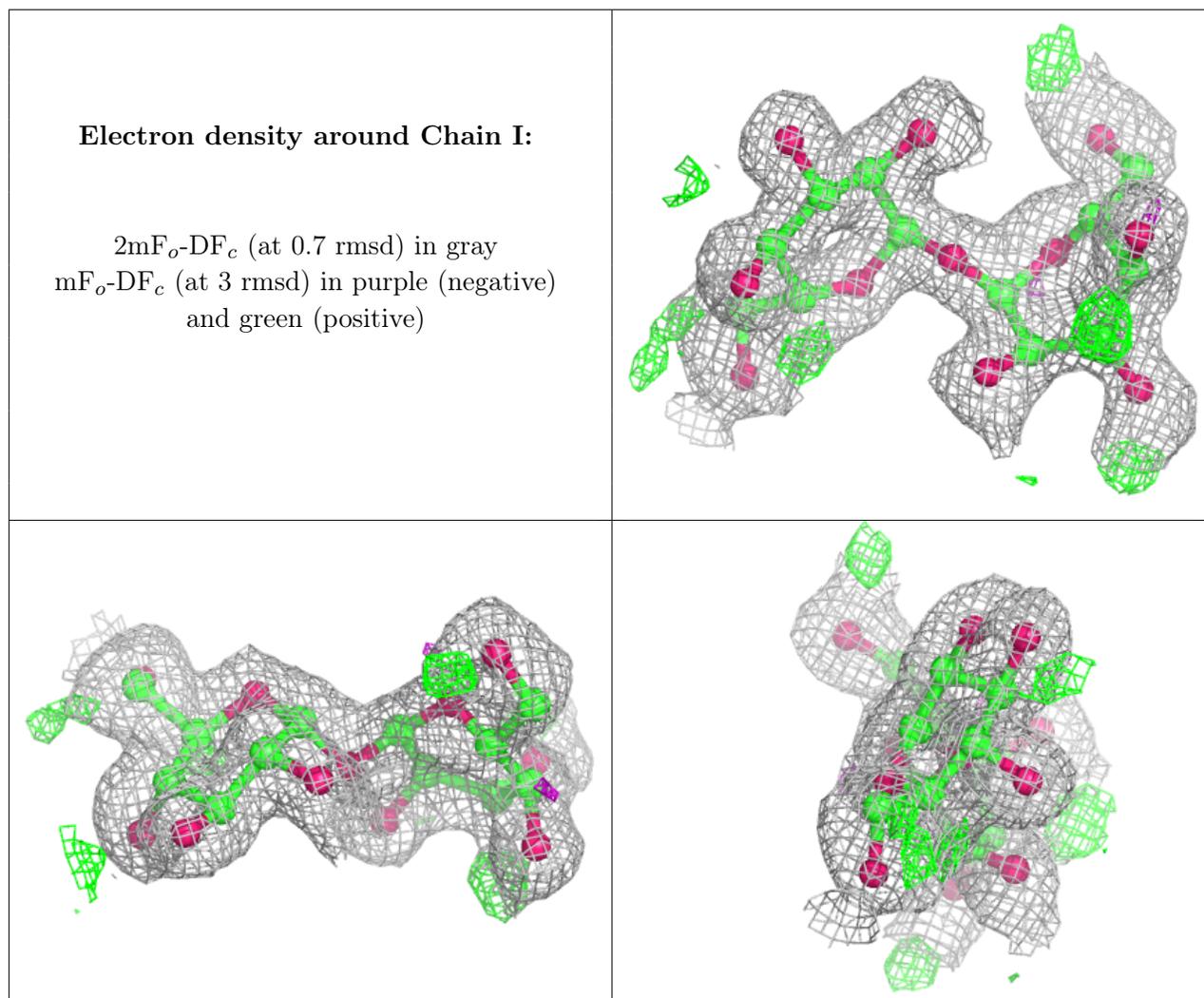There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

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	GLC	I	2	12/12	0.86	0.17	20,22,25,26	12
2	GLC	I	1	11/12	0.90	0.13	17,18,20,20	11

The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around Chain I:

$2\text{mF}_o\text{-DF}_c$ (at 0.7 rmsd) in gray
$\text{mF}_o\text{-DF}_c$ (at 3 rmsd) in purple (negative) and green (positive)

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, 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
5	MES	B	204	12/12	0.76	0.22	21,24,27,27	12
3	PGO	D	201	5/5	0.89	0.11	23,26,27,28	5
3	PGO	G	201	5/5	0.92	0.11	23,25,25,31	5
3	PGO	B	203	5/5	0.92	0.21	15,15,18,23	5
3	PGO	A	201	5/5	0.97	0.25	14,14,15,17	5
4	SO4	B	202	5/5	0.98	0.12	23,24,28,28	5
4	SO4	B	201	5/5	0.98	0.15	18,18,20,20	5

6.5　Other polymers ⓘ

There are no such residues in this entry.