



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:36 AM GMT

PDB ID : 2HRQ  
Title : Crystal structure of Human Liver Carboxylesterase 1 (hCE1) in covalent complex with the nerve agent Soman (GD)  
Authors : Fleming, C.D.; Redinbo, M.R.  
Deposited on : 2006-07-20  
Resolution : 2.70 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  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

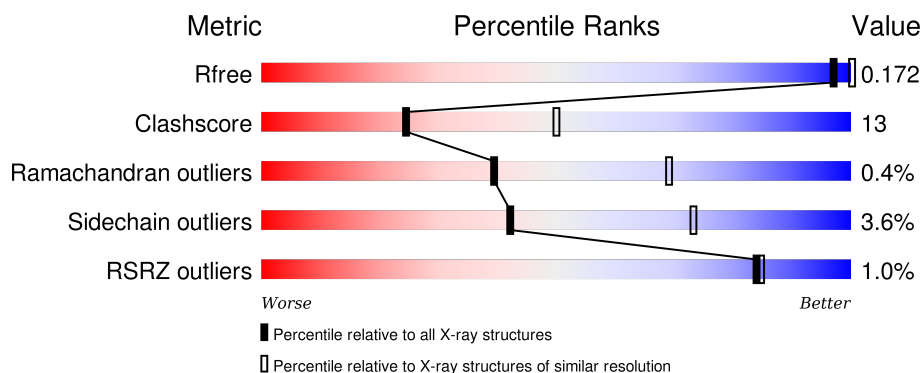
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	2103 (2.70-2.70)
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (2.70-2.70)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	532	<div> <div>75%</div> <div>22%</div> <div>.</div> </div>
1	B	532	<div> <div>74%</div> <div>24%</div> <div>.</div> </div>
1	C	532	<div> <div>71%</div> <div>27%</div> <div>.</div> </div>
1	D	532	<div> <div>73%</div> <div>25%</div> <div>.</div> </div>
1	E	532	<div> <div>75%</div> <div>23%</div> <div>.</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	532	

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
3	SIA	A	182	-	-	-	X
3	SIA	B	282	-	-	-	X
3	SIA	C	382	-	-	-	X
3	SIA	D	482	-	-	-	X
3	SIA	E	582	-	-	-	X
3	SIA	F	682	-	-	-	X
5	SO4	A	185	-	-	-	X
5	SO4	A	284	-	-	-	X
5	SO4	B	285	-	-	-	X
5	SO4	B	385	-	-	-	X
5	SO4	C	184	-	-	-	X
5	SO4	C	384	-	-	-	X
5	SO4	D	484	-	-	-	X
5	SO4	D	585	-	-	-	X
5	SO4	E	584	-	-	-	X
5	SO4	F	485	-	-	-	X
5	SO4	F	684	-	-	-	X
5	SO4	F	685	-	-	-	X
6	GD7	A	1	-	-	-	X
6	GD7	B	1	-	-	-	X
6	GD7	C	1	-	-	-	X
6	GD7	E	1	-	-	-	X
6	GD7	F	1	-	-	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 26340 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 Liver carboxylesterase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	532	Total	C	N	O	S	0	0	0
			4130	2662	685	763	20			
1	B	531	Total	C	N	O	S	0	0	0
			4124	2659	684	761	20			
1	C	531	Total	C	N	O	S	0	0	0
			4124	2659	684	761	20			
1	D	532	Total	C	N	O	S	0	0	0
			4130	2662	685	763	20			
1	E	531	Total	C	N	O	S	0	0	0
			4124	2659	684	761	20			
1	F	531	Total	C	N	O	S	0	0	0
			4125	2659	684	762	20			

There are 6 discrepancies between the modelled and reference sequences:

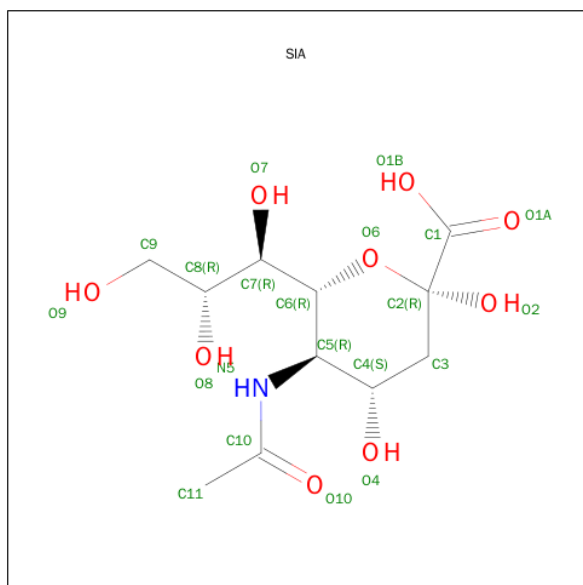
Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	GLN	DELETION	UNP Q9UK77
B	?	-	GLN	DELETION	UNP Q9UK77
C	?	-	GLN	DELETION	UNP Q9UK77
D	?	-	GLN	DELETION	UNP Q9UK77
E	?	-	GLN	DELETION	UNP Q9UK77
F	?	-	GLN	DELETION	UNP Q9UK77

- Molecule 2 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



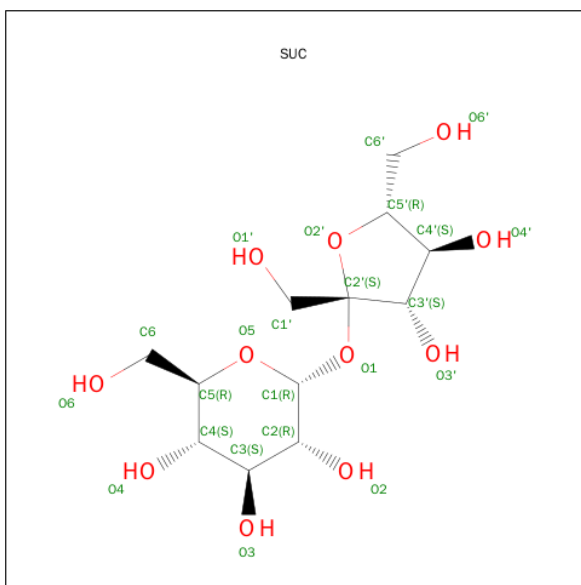
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	C	1	Total	C	N	O	0	0
			14	8	1	5		
2	D	1	Total	C	N	O	0	0
			14	8	1	5		
2	E	1	Total	C	N	O	0	0
			14	8	1	5		
2	F	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 3 is SUGAR (O-SIALIC ACID) (three-letter code: SIA) (formula: C<sub>11</sub>H<sub>19</sub>NO<sub>9</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			21	11	1	9		
3	B	1	Total	C	N	O	0	0
			21	11	1	9		
3	C	1	Total	C	N	O	0	0
			21	11	1	9		
3	D	1	Total	C	N	O	0	0
			21	11	1	9		
3	E	1	Total	C	N	O	0	0
			21	11	1	9		
3	F	1	Total	C	N	O	0	0
			21	11	1	9		

- Molecule 4 is SUGAR (SUCROSE) (three-letter code: SUC) (formula: C<sub>12</sub>H<sub>22</sub>O<sub>11</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			23	12	11		
4	B	1	Total	C	O	0	0
			23	12	11		
4	C	1	Total	C	O	0	0
			23	12	11		
4	D	1	Total	C	O	0	0
			23	12	11		
4	E	1	Total	C	O	0	0
			23	12	11		
4	F	1	Total	C	O	0	0
			23	12	11		

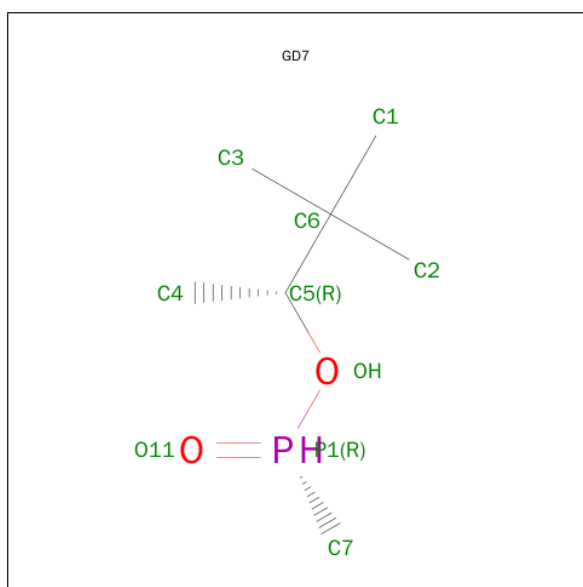
- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total	O	S	0	0
			5	4	1		
5	A	1	Total	O	S	0	0
			5	4	1		
5	A	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	C	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	D	1	Total	O	S	0	0
			5	4	1		
5	F	1	Total	O	S	0	0
			5	4	1		
5	E	1	Total	O	S	0	0
			5	4	1		
5	D	1	Total	O	S	0	0
			5	4	1		
5	F	1	Total	O	S	0	0
			5	4	1		
5	F	1	Total	O	S	0	0
			5	4	1		

- Molecule 6 is (1R)-1,2,2-TRIMETHYLPROPYL (R)-METHYLPHOSPHINATE (three-letter code: GD7) (formula: C<sub>7</sub>H<sub>17</sub>O<sub>2</sub>P).





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	O	P	0	0
			10	7	2	1		
6	B	1	Total	C	O	P	0	0
			10	7	2	1		
6	C	1	Total	C	O	P	0	0
			10	7	2	1		
6	D	1	Total	C	O	P	0	0
			10	7	2	1		
6	E	1	Total	C	O	P	0	0
			10	7	2	1		
6	F	1	Total	C	O	P	0	0
			10	7	2	1		

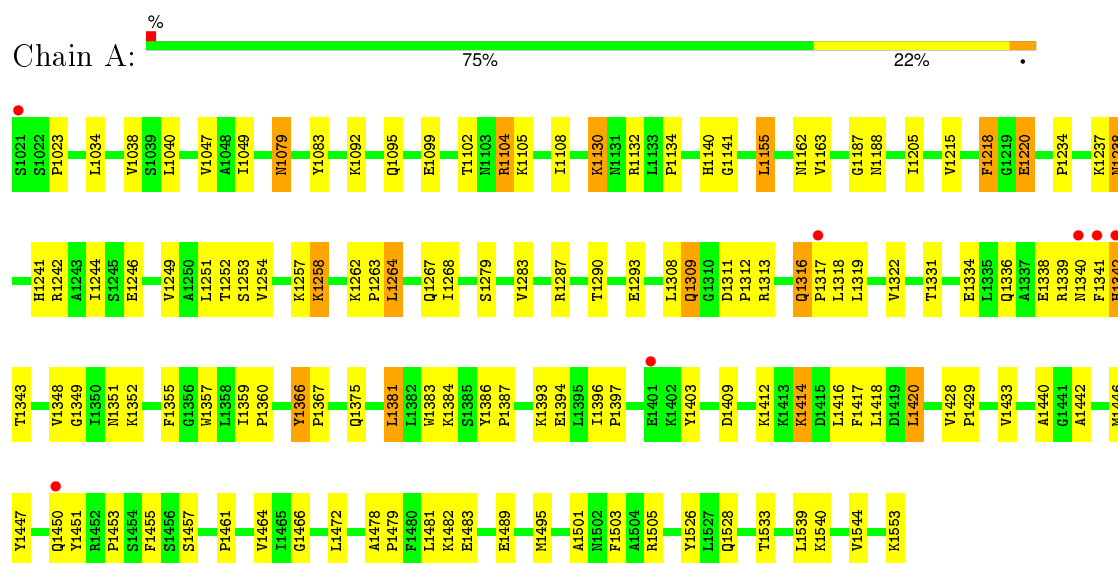
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	193	Total	O	0	0
			193	193		
7	B	166	Total	O	0	0
			166	166		
7	C	205	Total	O	0	0
			205	205		
7	D	186	Total	O	0	0
			186	186		
7	E	153	Total	O	0	0
			153	153		
7	F	212	Total	O	0	0
			212	212		

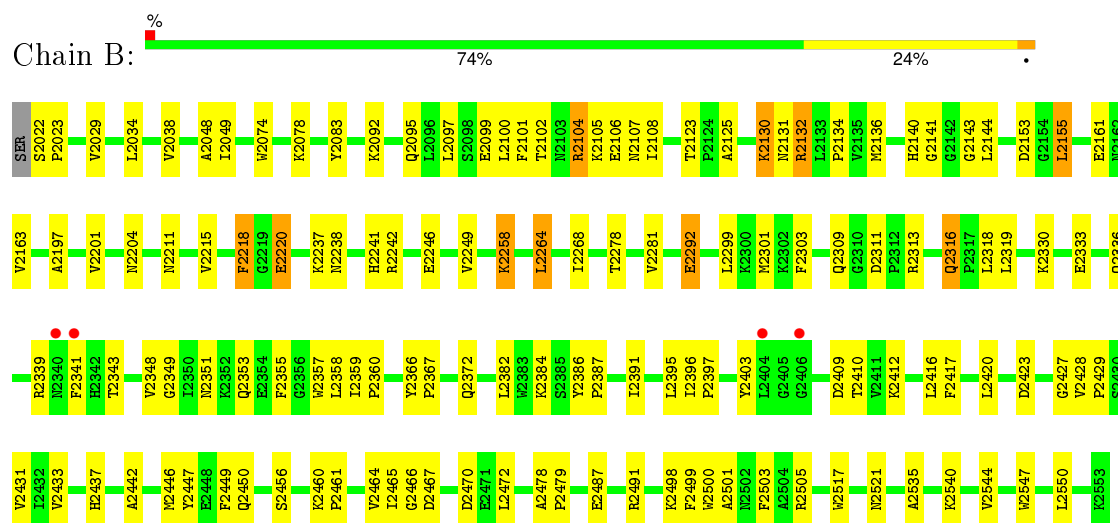
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: Liver carboxylesterase 1

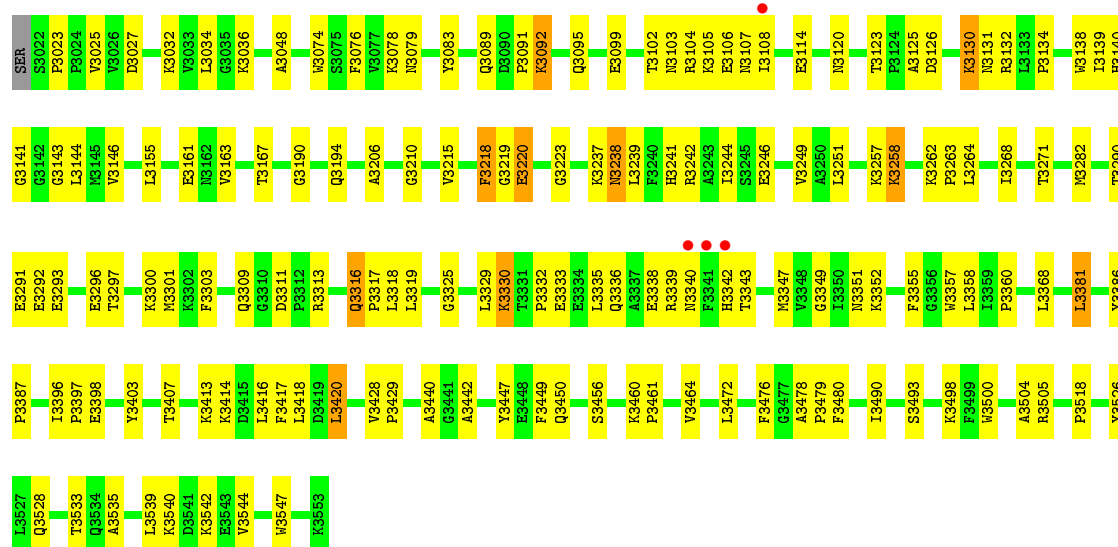


#### • Molecule 1: Liver carboxylesterase 1

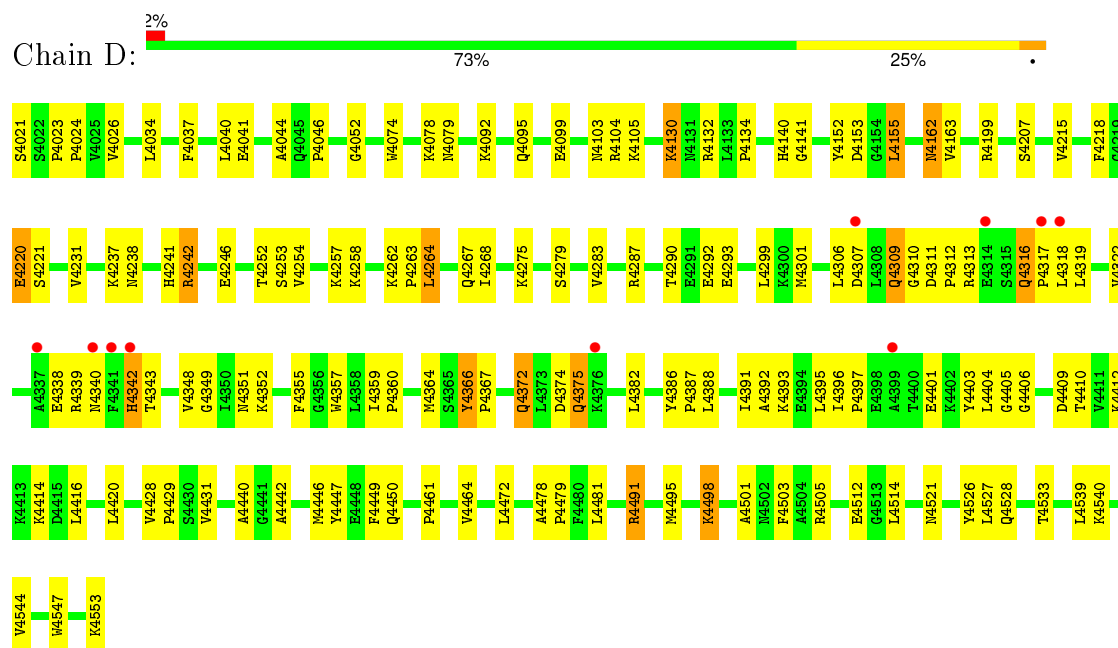


#### • Molecule 1: Liver carboxylesterase 1



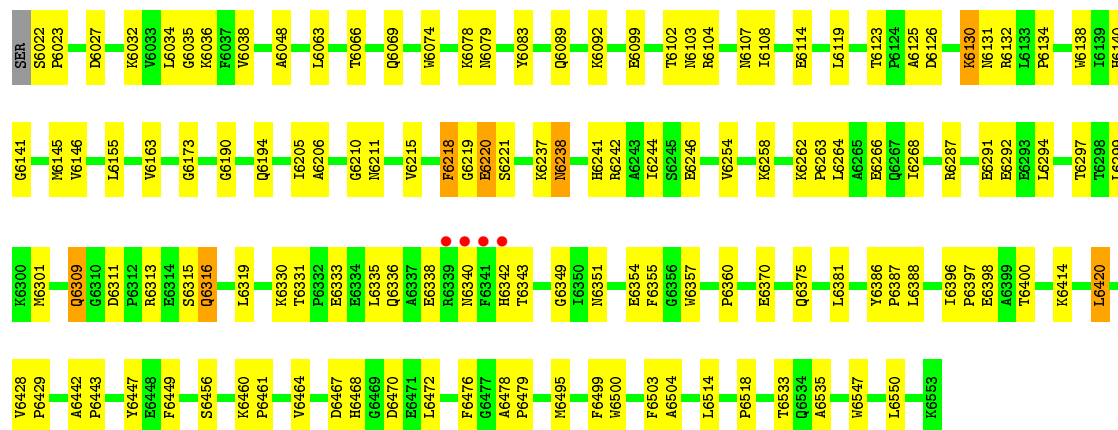
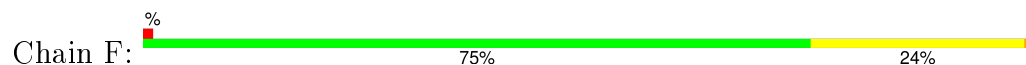


• Molecule 1: Liver carboxylesterase 1





• Molecule 1: Liver carboxylesterase 1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	55.46Å 181.19Å 203.05Å 90.00° 89.99° 90.00°	Depositor
Resolution (Å)	42.91 – 2.70 42.90 – 2.70	Depositor EDS
% Data completeness (in resolution range)	98.5 (42.91-2.70) 99.8 (42.90-2.70)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	7.79 (at 2.69Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.170 , 0.225 0.175 , 0.172	Depositor DCC
$R_{free}$ test set	5468 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	29.9	Xtriage
Anisotropy	0.287	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 23.9	EDS
Estimated twinning fraction	0.457 for h,-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	0 of 109259 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	26340	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.50% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: SO4, SIA, SUC, NAG, GD7

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	0.36	1/4236 (0.0%)	0.59	1/5754 (0.0%)
1	B	0.35	1/4230 (0.0%)	0.58	1/5746 (0.0%)
1	C	0.38	1/4230 (0.0%)	0.61	1/5746 (0.0%)
1	D	0.36	1/4236 (0.0%)	0.59	1/5754 (0.0%)
1	E	0.35	1/4230 (0.0%)	0.58	1/5746 (0.0%)
1	F	0.37	1/4231 (0.0%)	0.61	1/5746 (0.0%)
All	All	0.36	6/25393 (0.0%)	0.59	6/34492 (0.0%)

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	3092	LYS	CE-NZ	-6.21	1.33	1.49
1	F	6092	LYS	CE-NZ	-6.18	1.33	1.49
1	B	2092	LYS	CE-NZ	-6.13	1.33	1.49
1	E	5092	LYS	CE-NZ	-6.09	1.33	1.49
1	A	1092	LYS	CE-NZ	-6.08	1.33	1.49

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1092	LYS	CD-CE-NZ	5.58	124.52	111.70
1	D	4092	LYS	CD-CE-NZ	5.56	124.48	111.70
1	C	3092	LYS	CD-CE-NZ	5.55	124.46	111.70
1	F	6092	LYS	CD-CE-NZ	5.51	124.38	111.70
1	E	5092	LYS	CD-CE-NZ	5.35	124.00	111.70

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4130	0	4129	106	0
1	B	4124	0	4125	102	0
1	C	4124	0	4125	114	0
1	D	4130	0	4129	109	0
1	E	4124	0	4125	98	0
1	F	4125	0	4124	101	0
2	A	14	0	13	5	0
2	B	14	0	13	0	0
2	C	14	0	13	0	0
2	D	14	0	13	0	0
2	E	14	0	13	1	0
2	F	14	0	13	2	0
3	A	21	0	18	6	0
3	B	21	0	18	5	0
3	C	21	0	18	5	0
3	D	21	0	18	2	0
3	E	21	0	18	6	0
3	F	21	0	18	4	0
4	A	23	0	22	0	0
4	B	23	0	22	0	0
4	C	23	0	22	1	0
4	D	23	0	22	0	0
4	E	23	0	22	0	0
4	F	23	0	22	0	0
5	A	10	0	0	0	0
5	B	10	0	0	0	0
5	C	10	0	0	1	0
5	D	10	0	0	0	0
5	E	5	0	0	0	0
5	F	15	0	0	0	0
6	A	10	0	16	0	0
6	B	10	0	16	0	0
6	C	10	0	16	0	0
6	D	10	0	16	0	0
6	E	10	0	16	2	0
6	F	10	0	16	0	0
7	A	193	0	0	15	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	B	166	0	0	11	0
7	C	205	0	0	8	0
7	D	186	0	0	11	0
7	E	153	0	0	6	0
7	F	212	0	0	9	0
All	All	26340	0	25171	649	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 13.

The worst 5 of 649 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)
2:A:179:NAG:H82	7:A:8924:HOH:O	1.48	1.12
1:C:3343:THR:HB	1:C:3442:ALA:HB2	1.23	1.11
1:F:6343:THR:HB	1:F:6442:ALA:HB2	1.14	1.10
1:D:4491:ARG:HH11	1:D:4491:ARG:HB2	1.17	1.08
1:A:1079:ASN:H	3:A:182:SIA:H112	1.23	1.00

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	530/532 (100%)	497 (94%)	29 (6%)	4 (1%)	24	51
1	B	529/532 (99%)	495 (94%)	34 (6%)	0	100	100
1	C	529/532 (99%)	502 (95%)	25 (5%)	2 (0%)	39	69
1	D	530/532 (100%)	496 (94%)	32 (6%)	2 (0%)	39	69
1	E	529/532 (99%)	498 (94%)	27 (5%)	4 (1%)	24	51

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	529/532 (99%)	500 (94%)	28 (5%)	1 (0%)	52	80
All	All	3176/3192 (100%)	2988 (94%)	175 (6%)	13 (0%)	39	69

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	5238	ASN
1	F	6238	ASN
1	C	3076	PHE
1	C	3238	ASN
1	D	4253	SER

### 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	448/448 (100%)	430 (96%)	18 (4%)	38	69
1	B	447/448 (100%)	432 (97%)	15 (3%)	44	75
1	C	447/448 (100%)	432 (97%)	15 (3%)	44	75
1	D	448/448 (100%)	426 (95%)	22 (5%)	31	61
1	E	447/448 (100%)	433 (97%)	14 (3%)	47	78
1	F	447/448 (100%)	435 (97%)	12 (3%)	52	82
All	All	2684/2688 (100%)	2588 (96%)	96 (4%)	42	73

5 of 96 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	3358	LEU
1	D	4162	ASN
1	F	6264	LEU
1	C	3381	LEU
1	D	4021	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 73 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	3241	HIS
1	D	4140	HIS
1	F	6340	ASN
1	C	3450	GLN
1	D	4241	HIS

### 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](#)

36 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	$\# Z  > 2$
6	GD7	A	1	1	6,9,9	1.01	0	6,13,13	0.38	0
4	SUC	A	11	-	24,24,24	0.98	1 (4%)	36,36,36	0.84	1 (2%)
2	NAG	A	179	1	14,14,15	0.79	1 (7%)	15,19,21	0.63	0
3	SIA	A	182	-	17,21,21	0.95	1 (5%)	19,31,31	0.81	1 (5%)
5	SO4	A	185	-	4,4,4	0.22	0	6,6,6	0.09	0
5	SO4	A	284	-	4,4,4	0.22	0	6,6,6	0.07	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	GD7	B	1	1	6,9,9	1.04	0	6,13,13	0.39	0
4	SUC	B	22	-	24,24,24	0.99	1 (4%)	36,36,36	0.71	0
2	NAG	B	279	1	14,14,15	0.73	0	15,19,21	0.81	1 (6%)
3	SIA	B	282	-	17,21,21	1.01	2 (11%)	19,31,31	0.96	1 (5%)
5	SO4	B	285	-	4,4,4	0.24	0	6,6,6	0.15	0
5	SO4	B	385	-	4,4,4	0.22	0	6,6,6	0.10	0
6	GD7	C	1	1	6,9,9	1.03	0	6,13,13	0.39	0
5	SO4	C	184	-	4,4,4	0.21	0	6,6,6	0.12	0
4	SUC	C	33	-	24,24,24	0.90	1 (4%)	36,36,36	0.72	0
2	NAG	C	379	1	14,14,15	0.65	0	15,19,21	0.78	1 (6%)
3	SIA	C	382	-	17,21,21	0.80	0	19,31,31	0.75	1 (5%)
5	SO4	C	384	-	4,4,4	0.22	0	6,6,6	0.07	0
6	GD7	D	1	1	6,9,9	1.02	0	6,13,13	0.30	0
4	SUC	D	44	-	24,24,24	0.97	1 (4%)	36,36,36	0.69	0
2	NAG	D	479	1	14,14,15	0.82	1 (7%)	15,19,21	0.70	1 (6%)
3	SIA	D	482	-	17,21,21	0.94	1 (5%)	19,31,31	1.13	1 (5%)
5	SO4	D	484	-	4,4,4	0.22	0	6,6,6	0.08	0
5	SO4	D	585	-	4,4,4	0.22	0	6,6,6	0.06	0
6	GD7	E	1	1	6,9,9	0.97	0	6,13,13	0.41	0
4	SUC	E	55	-	24,24,24	0.98	1 (4%)	36,36,36	0.67	0
2	NAG	E	579	1	14,14,15	0.57	0	15,19,21	0.70	0
3	SIA	E	582	-	17,21,21	1.03	1 (5%)	19,31,31	1.05	1 (5%)
5	SO4	E	584	-	4,4,4	0.24	0	6,6,6	0.07	0
6	GD7	F	1	1	6,9,9	1.10	0	6,13,13	0.41	0
5	SO4	F	485	-	4,4,4	0.24	0	6,6,6	0.08	0
4	SUC	F	66	-	24,24,24	0.91	1 (4%)	36,36,36	0.77	0
2	NAG	F	679	1	14,14,15	0.63	0	15,19,21	0.81	1 (6%)
3	SIA	F	682	-	17,21,21	0.88	0	19,31,31	0.64	0
5	SO4	F	684	-	4,4,4	0.20	0	6,6,6	0.07	0
5	SO4	F	685	-	4,4,4	0.25	0	6,6,6	0.10	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
6	GD7	A	1	1	-	0/6/10/10	0/0/0/0
4	SUC	A	11	-	-	0/12/51/51	0/2/2/2
2	NAG	A	179	1	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	SIA	A	182	-	-	0/14/38/38	0/1/1/1
5	SO4	A	185	-	-	0/0/0/0	0/0/0/0
5	SO4	A	284	-	-	0/0/0/0	0/0/0/0
6	GD7	B	1	1	-	0/6/10/10	0/0/0/0
4	SUC	B	22	-	-	0/12/51/51	0/2/2/2
2	NAG	B	279	1	-	0/6/23/26	0/1/1/1
3	SIA	B	282	-	-	1/14/38/38	0/1/1/1
5	SO4	B	285	-	-	0/0/0/0	0/0/0/0
5	SO4	B	385	-	-	0/0/0/0	0/0/0/0
6	GD7	C	1	1	-	0/6/10/10	0/0/0/0
5	SO4	C	184	-	-	0/0/0/0	0/0/0/0
4	SUC	C	33	-	-	0/12/51/51	0/2/2/2
2	NAG	C	379	1	-	1/6/23/26	0/1/1/1
3	SIA	C	382	-	-	0/14/38/38	0/1/1/1
5	SO4	C	384	-	-	0/0/0/0	0/0/0/0
6	GD7	D	1	1	-	0/6/10/10	0/0/0/0
4	SUC	D	44	-	-	0/12/51/51	0/2/2/2
2	NAG	D	479	1	-	0/6/23/26	0/1/1/1
3	SIA	D	482	-	-	0/14/38/38	0/1/1/1
5	SO4	D	484	-	-	0/0/0/0	0/0/0/0
5	SO4	D	585	-	-	0/0/0/0	0/0/0/0
6	GD7	E	1	1	-	0/6/10/10	0/0/0/0
4	SUC	E	55	-	-	0/12/51/51	0/2/2/2
2	NAG	E	579	1	-	0/6/23/26	0/1/1/1
3	SIA	E	582	-	-	0/14/38/38	0/1/1/1
5	SO4	E	584	-	-	0/0/0/0	0/0/0/0
6	GD7	F	1	1	-	0/6/10/10	0/0/0/0
5	SO4	F	485	-	-	0/0/0/0	0/0/0/0
4	SUC	F	66	-	-	0/12/51/51	0/2/2/2
2	NAG	F	679	1	-	2/6/23/26	0/1/1/1
3	SIA	F	682	-	-	0/14/38/38	0/1/1/1
5	SO4	F	684	-	-	0/0/0/0	0/0/0/0
5	SO4	F	685	-	-	0/0/0/0	0/0/0/0

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	282	SIA	C7-C6	2.10	1.55	1.52
3	E	582	SIA	O6-C2	2.26	1.44	1.42
3	B	282	SIA	O6-C2	2.26	1.44	1.42
2	A	179	NAG	C1-C2	2.33	1.55	1.52
3	A	182	SIA	O6-C2	2.33	1.45	1.42

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	482	SIA	C7-C6-C5	-3.55	108.95	114.32
2	D	479	NAG	C2-N2-C7	-2.15	120.28	123.04
2	B	279	NAG	C2-N2-C7	-2.11	120.33	123.04
3	A	182	SIA	C7-C6-C5	-2.11	111.13	114.32
2	F	679	NAG	C2-N2-C7	-2.10	120.34	123.04

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	379	NAG	O7-C7-N2-C2
2	F	679	NAG	C8-C7-N2-C2
2	F	679	NAG	O7-C7-N2-C2
3	B	282	SIA	C4-C5-N5-C10

There are no ring outliers.

12 monomers are involved in 38 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	179	NAG	5	0
3	A	182	SIA	6	0
3	B	282	SIA	5	0
5	C	184	SO4	1	0
4	C	33	SUC	1	0
3	C	382	SIA	5	0
3	D	482	SIA	2	0
6	E	1	GD7	2	0
2	E	579	NAG	1	0
3	E	582	SIA	6	0
2	F	679	NAG	2	0
3	F	682	SIA	4	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [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	532/532 (100%)	-0.18	7 (1%) 79 79	10, 27, 69, 89	0
1	B	531/532 (99%)	-0.18	4 (0%) 87 88	12, 32, 68, 90	0
1	C	531/532 (99%)	-0.32	4 (0%) 87 88	6, 24, 46, 83	0
1	D	532/532 (100%)	-0.18	10 (1%) 70 70	9, 27, 69, 89	0
1	E	531/532 (99%)	-0.16	2 (0%) 93 94	11, 31, 69, 88	0
1	F	531/532 (99%)	-0.31	4 (0%) 87 88	8, 25, 46, 92	0
All	All	3188/3192 (99%)	-0.22	31 (0%) 84 85	6, 27, 64, 92	0

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1342	HIS	6.4
1	B	2340	ASN	5.3
1	D	4341	PHE	5.1
1	D	4340	ASN	4.7
1	A	1340	ASN	4.7

### 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 carbohydrates in this entry.

## 6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
5	SO4	A	185	5/5	0.85	0.49	18.37	133,133,134,134	0
5	SO4	B	385	5/5	0.79	0.35	17.65	109,110,110,111	0
5	SO4	F	685	5/5	0.66	0.54	14.82	130,131,131,131	0
5	SO4	E	584	5/5	0.88	0.24	12.12	96,96,96,97	0
5	SO4	F	684	5/5	0.76	0.37	10.72	105,106,106,106	0
5	SO4	C	384	5/5	0.86	0.24	10.48	99,100,100,100	0
3	SIA	D	482	21/21	0.61	0.41	10.23	90,97,100,100	0
5	SO4	D	484	5/5	0.86	0.38	10.09	128,128,128,129	0
5	SO4	D	585	5/5	0.82	0.43	9.21	120,121,121,121	0
5	SO4	B	285	5/5	0.90	0.27	6.63	91,92,92,93	0
3	SIA	A	182	21/21	0.71	0.38	5.52	80,90,91,92	0
5	SO4	C	184	5/5	0.81	0.42	5.23	122,122,122,123	0
3	SIA	E	582	21/21	0.79	0.30	5.17	36,54,60,61	0
5	SO4	A	284	5/5	0.85	0.38	5.00	125,125,125,125	0
3	SIA	C	382	21/21	0.79	0.29	4.75	67,78,82,82	0
6	GD7	A	1	10/10	0.92	0.30	4.74	45,50,53,55	0
3	SIA	B	282	21/21	0.80	0.32	4.55	46,59,64,65	0
3	SIA	F	682	21/21	0.84	0.26	4.23	58,65,68,68	0
6	GD7	F	1	10/10	0.95	0.27	4.03	34,41,41,42	0
5	SO4	F	485	5/5	0.89	0.33	3.94	115,115,115,115	0
6	GD7	C	1	10/10	0.96	0.28	3.53	39,44,45,45	0
6	GD7	B	1	10/10	0.94	0.25	3.30	39,43,44,45	0
6	GD7	E	1	10/10	0.96	0.23	2.42	33,40,43,44	0
6	GD7	D	1	10/10	0.91	0.27	1.56	44,49,51,53	0
4	SUC	F	66	23/23	0.94	0.15	0.65	27,31,35,39	0
4	SUC	D	44	23/23	0.87	0.21	0.38	61,65,72,77	0
4	SUC	C	33	23/23	0.96	0.15	-0.18	27,29,34,38	0
4	SUC	E	55	23/23	0.89	0.18	-0.18	66,68,71,73	0
4	SUC	A	11	23/23	0.91	0.19	-0.19	57,62,71,71	0
4	SUC	B	22	23/23	0.92	0.17	-0.50	63,65,66,67	0
2	NAG	F	679	14/15	0.78	0.33	-	64,69,70,71	0
2	NAG	A	179	14/15	0.61	0.36	-	59,62,64,64	0
2	NAG	D	479	14/15	0.76	0.35	-	59,62,65,65	0
2	NAG	C	379	14/15	0.65	0.40	-	68,73,75,75	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	NAG	B	279	14/15	0.83	0.47	-	69,73,77,77	0
2	NAG	E	579	14/15	0.83	0.34	-	63,68,71,71	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.