



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 12, 2017 – 04:46 AM EDT

PDB ID : 4U3C  
Title : Docking Site of Maltohexaose in the Mtb GlgE  
Authors : Ronning, D.R.; Lindenberger, J.J.  
Deposited on : unknown  
Resolution : 3.98 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20029824  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : rb-20029824

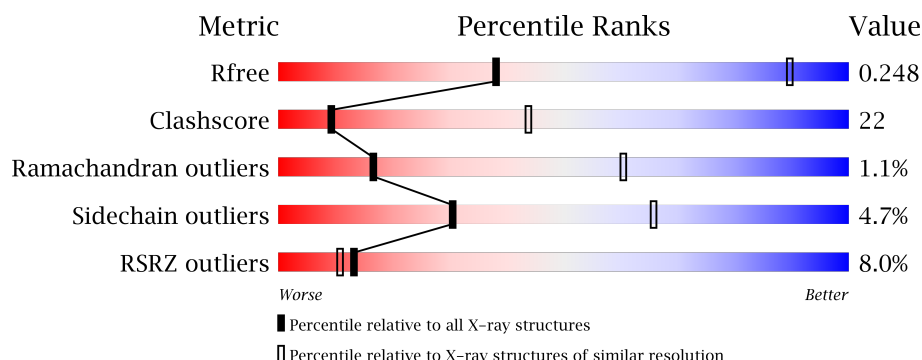
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.98 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))
$R_{free}$	100719	1066 (4.36-3.60)
Clashscore	112137	1163 (4.36-3.60)
Ramachandran outliers	110173	1119 (4.36-3.60)
Sidechain outliers	110143	1108 (4.36-3.60)
RSRZ outliers	101464	1078 (4.36-3.60)

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	723	<div> <div>7%</div> <div>55%</div> <div>33%</div> <div>9%</div> </div>
1	B	723	<div> <div>7%</div> <div>58%</div> <div>31%</div> <div>9%</div> </div>
1	C	723	<div> <div>9%</div> <div>55%</div> <div>33%</div> <div>9%</div> </div>
1	D	723	<div> <div>7%</div> <div>50%</div> <div>37%</div> <div>9%</div> </div>
1	E	723	<div> <div>7%</div> <div>53%</div> <div>34%</div> <div>9%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	723	<div><div></div><div>6%</div><div>56%</div><div>32%</div><div>•</div><div>9%</div></div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 31938 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 Alpha-1,4-glucan:maltose-1-phosphate maltosyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	660	Total	C	N	O	S	0	0	0
			5233	3359	910	950	14			
1	B	660	Total	C	N	O	S	0	0	0
			5233	3359	910	950	14			
1	C	660	Total	C	N	O	S	0	0	0
			5233	3359	910	950	14			
1	D	660	Total	C	N	O	S	0	0	0
			5233	3359	910	950	14			
1	E	660	Total	C	N	O	S	0	0	0
			5233	3359	910	950	14			
1	F	660	Total	C	N	O	S	0	0	0
			5233	3359	910	950	14			

There are 132 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-21	MET	-	expression tag	UNP P9WQ16
A	-20	GLY	-	expression tag	UNP P9WQ16
A	-19	SER	-	expression tag	UNP P9WQ16
A	-18	SER	-	expression tag	UNP P9WQ16
A	-17	HIS	-	expression tag	UNP P9WQ16
A	-16	HIS	-	expression tag	UNP P9WQ16
A	-15	HIS	-	expression tag	UNP P9WQ16
A	-14	HIS	-	expression tag	UNP P9WQ16
A	-13	HIS	-	expression tag	UNP P9WQ16
A	-12	HIS	-	expression tag	UNP P9WQ16
A	-11	SER	-	expression tag	UNP P9WQ16
A	-10	SER	-	expression tag	UNP P9WQ16
A	-9	GLY	-	expression tag	UNP P9WQ16
A	-8	LEU	-	expression tag	UNP P9WQ16
A	-7	GLU	-	expression tag	UNP P9WQ16
A	-6	VAL	-	expression tag	UNP P9WQ16
A	-5	LEU	-	expression tag	UNP P9WQ16

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	PHE	-	expression tag	UNP P9WQ16
A	-3	GLN	-	expression tag	UNP P9WQ16
A	-2	GLY	-	expression tag	UNP P9WQ16
A	-1	PRO	-	expression tag	UNP P9WQ16
A	0	HIS	-	expression tag	UNP P9WQ16
B	-21	MET	-	expression tag	UNP P9WQ16
B	-20	GLY	-	expression tag	UNP P9WQ16
B	-19	SER	-	expression tag	UNP P9WQ16
B	-18	SER	-	expression tag	UNP P9WQ16
B	-17	HIS	-	expression tag	UNP P9WQ16
B	-16	HIS	-	expression tag	UNP P9WQ16
B	-15	HIS	-	expression tag	UNP P9WQ16
B	-14	HIS	-	expression tag	UNP P9WQ16
B	-13	HIS	-	expression tag	UNP P9WQ16
B	-12	HIS	-	expression tag	UNP P9WQ16
B	-11	SER	-	expression tag	UNP P9WQ16
B	-10	SER	-	expression tag	UNP P9WQ16
B	-9	GLY	-	expression tag	UNP P9WQ16
B	-8	LEU	-	expression tag	UNP P9WQ16
B	-7	GLU	-	expression tag	UNP P9WQ16
B	-6	VAL	-	expression tag	UNP P9WQ16
B	-5	LEU	-	expression tag	UNP P9WQ16
B	-4	PHE	-	expression tag	UNP P9WQ16
B	-3	GLN	-	expression tag	UNP P9WQ16
B	-2	GLY	-	expression tag	UNP P9WQ16
B	-1	PRO	-	expression tag	UNP P9WQ16
B	0	HIS	-	expression tag	UNP P9WQ16
C	-21	MET	-	expression tag	UNP P9WQ16
C	-20	GLY	-	expression tag	UNP P9WQ16
C	-19	SER	-	expression tag	UNP P9WQ16
C	-18	SER	-	expression tag	UNP P9WQ16
C	-17	HIS	-	expression tag	UNP P9WQ16
C	-16	HIS	-	expression tag	UNP P9WQ16
C	-15	HIS	-	expression tag	UNP P9WQ16
C	-14	HIS	-	expression tag	UNP P9WQ16
C	-13	HIS	-	expression tag	UNP P9WQ16
C	-12	HIS	-	expression tag	UNP P9WQ16
C	-11	SER	-	expression tag	UNP P9WQ16
C	-10	SER	-	expression tag	UNP P9WQ16
C	-9	GLY	-	expression tag	UNP P9WQ16
C	-8	LEU	-	expression tag	UNP P9WQ16
C	-7	GLU	-	expression tag	UNP P9WQ16

*Continued on next page...*

*Continued from previous page...*

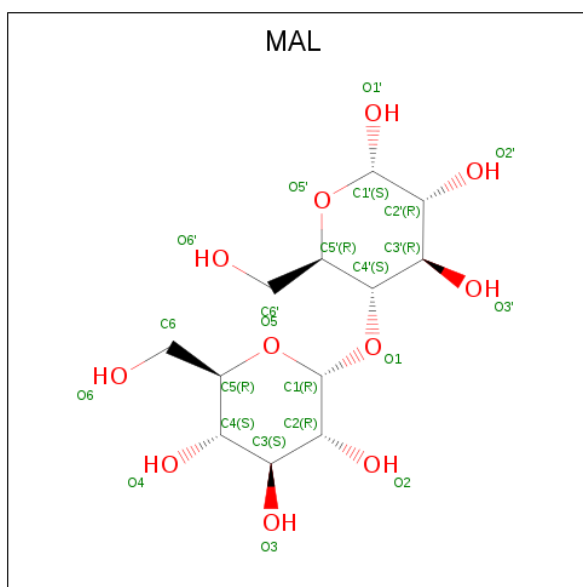
Chain	Residue	Modelled	Actual	Comment	Reference
C	-6	VAL	-	expression tag	UNP P9WQ16
C	-5	LEU	-	expression tag	UNP P9WQ16
C	-4	PHE	-	expression tag	UNP P9WQ16
C	-3	GLN	-	expression tag	UNP P9WQ16
C	-2	GLY	-	expression tag	UNP P9WQ16
C	-1	PRO	-	expression tag	UNP P9WQ16
C	0	HIS	-	expression tag	UNP P9WQ16
D	-21	MET	-	expression tag	UNP P9WQ16
D	-20	GLY	-	expression tag	UNP P9WQ16
D	-19	SER	-	expression tag	UNP P9WQ16
D	-18	SER	-	expression tag	UNP P9WQ16
D	-17	HIS	-	expression tag	UNP P9WQ16
D	-16	HIS	-	expression tag	UNP P9WQ16
D	-15	HIS	-	expression tag	UNP P9WQ16
D	-14	HIS	-	expression tag	UNP P9WQ16
D	-13	HIS	-	expression tag	UNP P9WQ16
D	-12	HIS	-	expression tag	UNP P9WQ16
D	-11	SER	-	expression tag	UNP P9WQ16
D	-10	SER	-	expression tag	UNP P9WQ16
D	-9	GLY	-	expression tag	UNP P9WQ16
D	-8	LEU	-	expression tag	UNP P9WQ16
D	-7	GLU	-	expression tag	UNP P9WQ16
D	-6	VAL	-	expression tag	UNP P9WQ16
D	-5	LEU	-	expression tag	UNP P9WQ16
D	-4	PHE	-	expression tag	UNP P9WQ16
D	-3	GLN	-	expression tag	UNP P9WQ16
D	-2	GLY	-	expression tag	UNP P9WQ16
D	-1	PRO	-	expression tag	UNP P9WQ16
D	0	HIS	-	expression tag	UNP P9WQ16
E	-21	MET	-	expression tag	UNP P9WQ16
E	-20	GLY	-	expression tag	UNP P9WQ16
E	-19	SER	-	expression tag	UNP P9WQ16
E	-18	SER	-	expression tag	UNP P9WQ16
E	-17	HIS	-	expression tag	UNP P9WQ16
E	-16	HIS	-	expression tag	UNP P9WQ16
E	-15	HIS	-	expression tag	UNP P9WQ16
E	-14	HIS	-	expression tag	UNP P9WQ16
E	-13	HIS	-	expression tag	UNP P9WQ16
E	-12	HIS	-	expression tag	UNP P9WQ16
E	-11	SER	-	expression tag	UNP P9WQ16
E	-10	SER	-	expression tag	UNP P9WQ16
E	-9	GLY	-	expression tag	UNP P9WQ16

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	-8	LEU	-	expression tag	UNP P9WQ16
E	-7	GLU	-	expression tag	UNP P9WQ16
E	-6	VAL	-	expression tag	UNP P9WQ16
E	-5	LEU	-	expression tag	UNP P9WQ16
E	-4	PHE	-	expression tag	UNP P9WQ16
E	-3	GLN	-	expression tag	UNP P9WQ16
E	-2	GLY	-	expression tag	UNP P9WQ16
E	-1	PRO	-	expression tag	UNP P9WQ16
E	0	HIS	-	expression tag	UNP P9WQ16
F	-21	MET	-	expression tag	UNP P9WQ16
F	-20	GLY	-	expression tag	UNP P9WQ16
F	-19	SER	-	expression tag	UNP P9WQ16
F	-18	SER	-	expression tag	UNP P9WQ16
F	-17	HIS	-	expression tag	UNP P9WQ16
F	-16	HIS	-	expression tag	UNP P9WQ16
F	-15	HIS	-	expression tag	UNP P9WQ16
F	-14	HIS	-	expression tag	UNP P9WQ16
F	-13	HIS	-	expression tag	UNP P9WQ16
F	-12	HIS	-	expression tag	UNP P9WQ16
F	-11	SER	-	expression tag	UNP P9WQ16
F	-10	SER	-	expression tag	UNP P9WQ16
F	-9	GLY	-	expression tag	UNP P9WQ16
F	-8	LEU	-	expression tag	UNP P9WQ16
F	-7	GLU	-	expression tag	UNP P9WQ16
F	-6	VAL	-	expression tag	UNP P9WQ16
F	-5	LEU	-	expression tag	UNP P9WQ16
F	-4	PHE	-	expression tag	UNP P9WQ16
F	-3	GLN	-	expression tag	UNP P9WQ16
F	-2	GLY	-	expression tag	UNP P9WQ16
F	-1	PRO	-	expression tag	UNP P9WQ16
F	0	HIS	-	expression tag	UNP P9WQ16

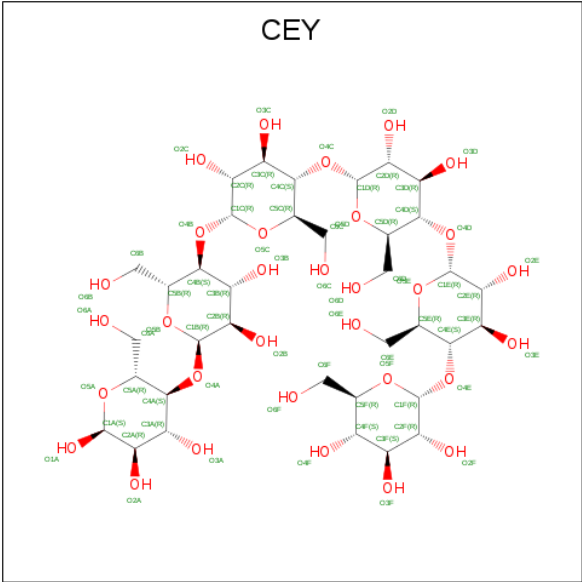
- Molecule 2 is MALTOSE (three-letter code: MAL) (formula:  $C_{12}H_{22}O_{11}$ ).



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

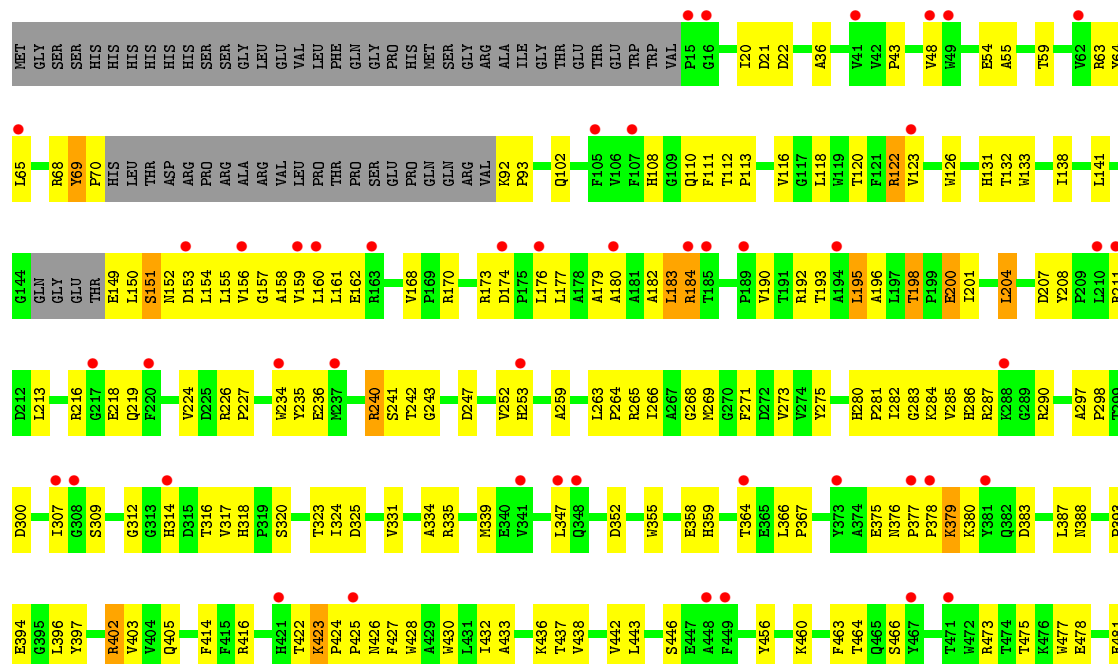
- Molecule 3 is alpha-D-glucopyranosyl-(1->4)-alpha-D-glucopyranosyl-(1->4)-alpha-D-glucopyranosyl-(1->4)-alpha-D-glucopyranosyl-(1->4)-alpha-D-glucopyranose (three-letter code: CEY) (formula: C<sub>36</sub>H<sub>62</sub>O<sub>31</sub>).

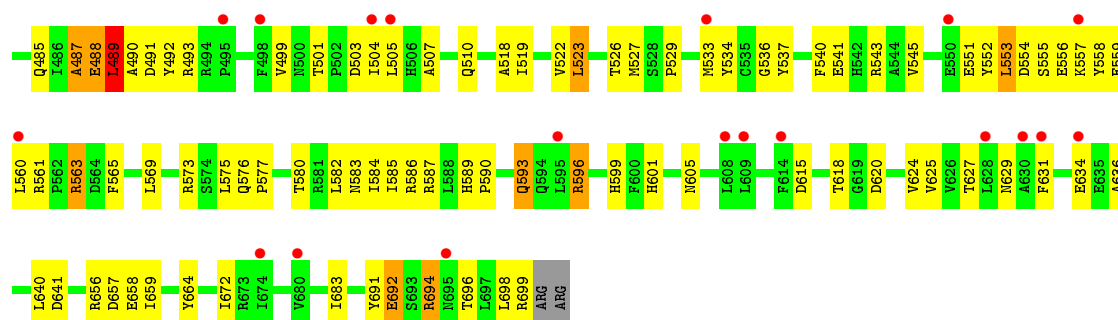




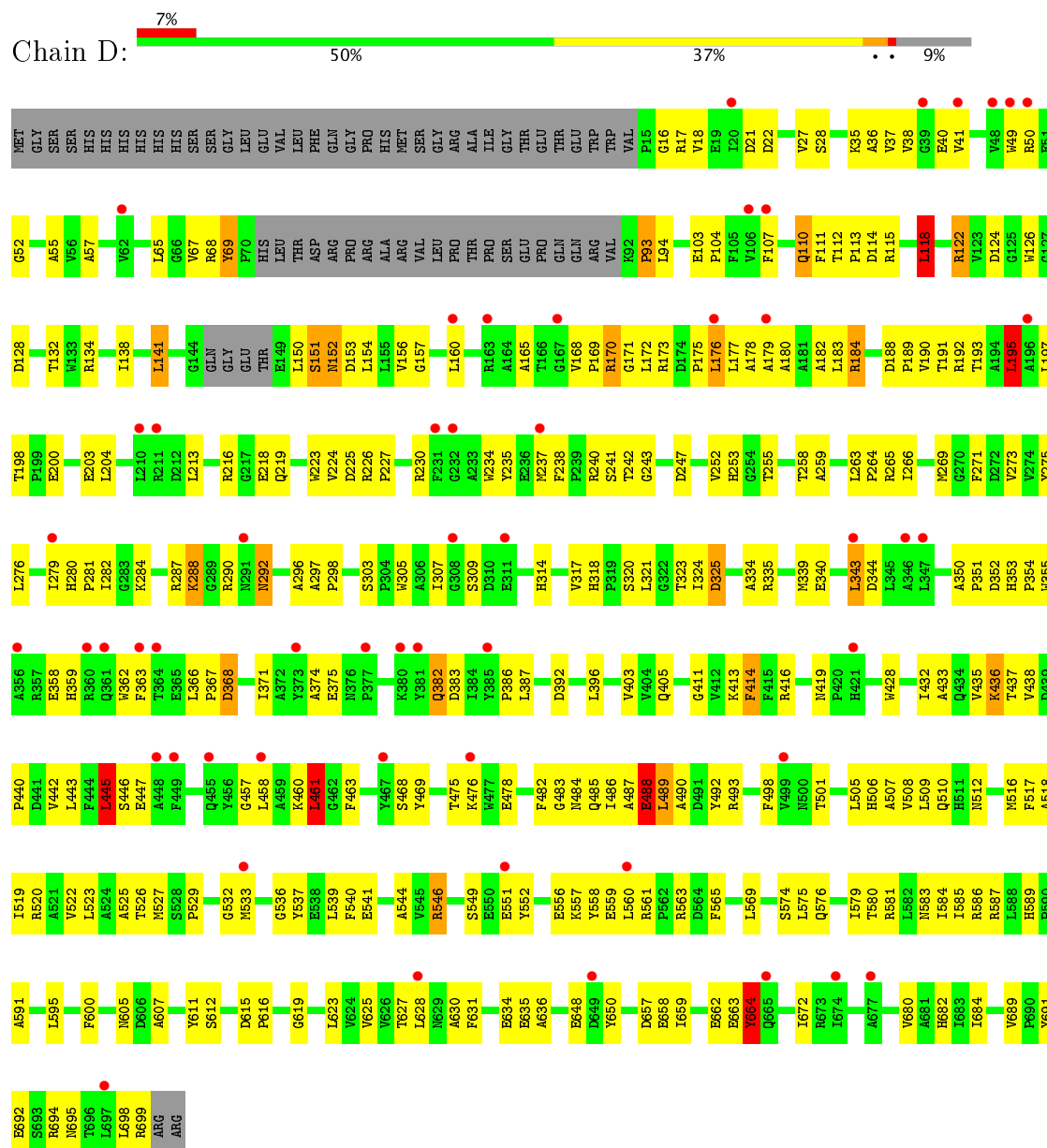
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			67	36	31		
3	B	1	Total	C	O	0	0
			67	36	31		
3	C	1	Total	C	O	0	0
			67	36	31		
3	D	1	Total	C	O	0	0
			67	36	31		
3	E	1	Total	C	O	0	0
			67	36	31		
3	F	1	Total	C	O	0	0
			67	36	31		





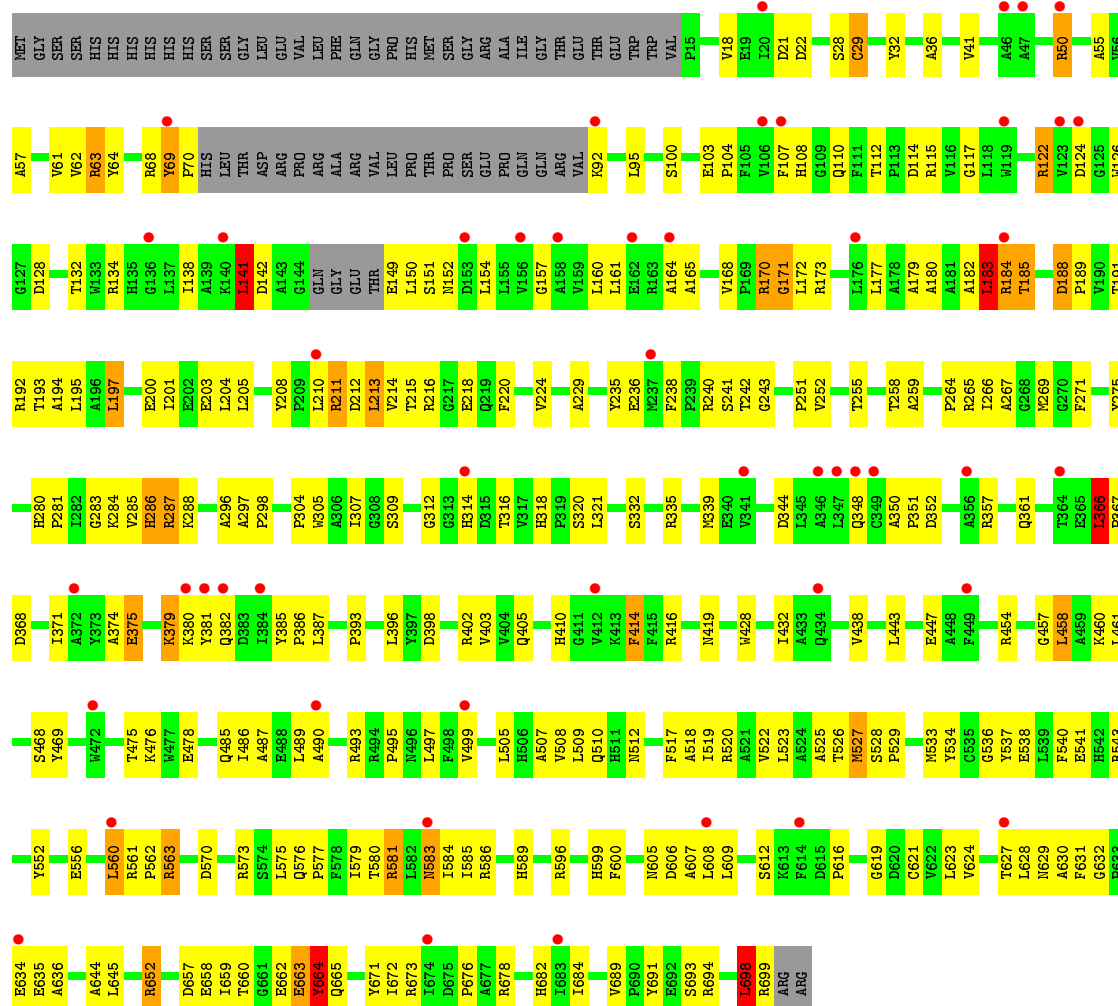


- Molecule 1: Alpha-1,4-glucan:maltose-1-phosphate maltosyltransferase

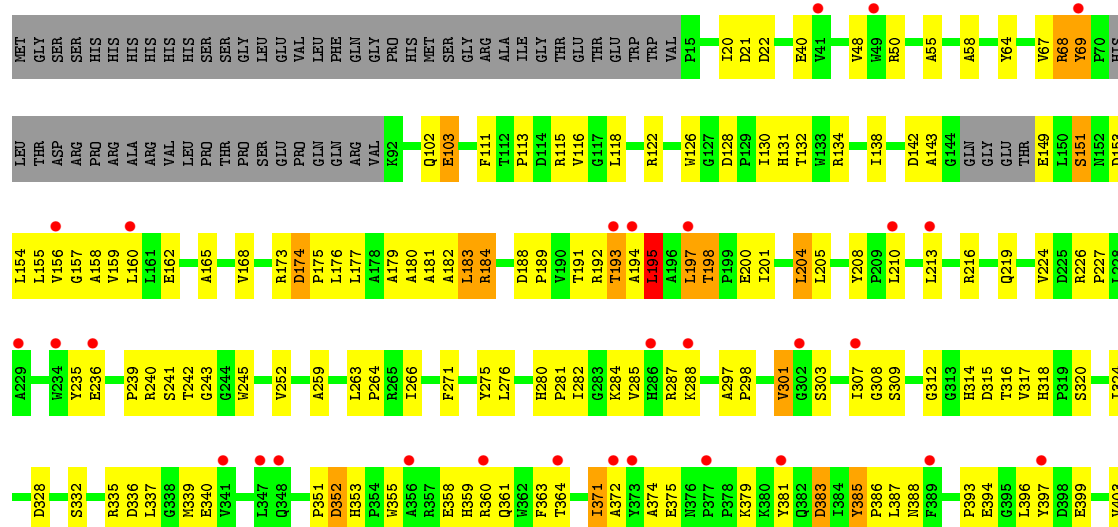


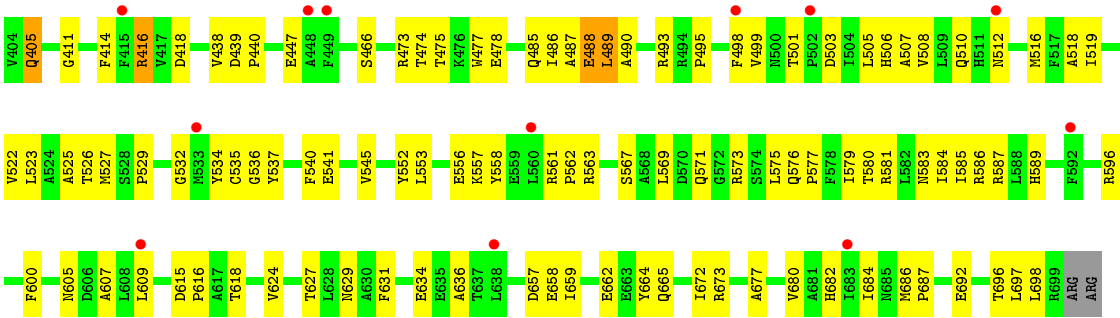
- Molecule 1: Alpha-1,4-glucan:maltose-1-phosphate maltosyltransferase





● Molecule 1: Alpha-1,4-glucan:maltose-1-phosphate maltosyltransferase





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	338.42Å 239.42Å 239.45Å 90.00° 134.90° 90.00°	Depositor
Resolution (Å)	43.73 – 3.98 43.74 – 3.98	Depositor EDS
% Data completeness (in resolution range)	97.9 (43.73-3.98) 88.0 (43.74-3.98)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.05 (at 4.00Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.224 , 0.256 0.221 , 0.248	Depositor DCC
$R_{free}$ test set	5103 reflections (5.03%)	DCC
Wilson B-factor (Å <sup>2</sup> )	110.8	Xtriage
Anisotropy	0.115	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 44.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.000 for h+2*k,-h-l 0.000 for k+l,h+l,-l 0.000 for -k+l,-h-l,-l 0.000 for -h+k-l,-l,-k 0.000 for -h-k-l,l,k 0.407 for h-k+l,l,-h-l 0.407 for -k-l,-h-l,k 0.408 for h+k+l,-l,-h-l 0.409 for k-l,h+l,-k 0.000 for h,-k,-h-l 0.409 for -h-2*k,-k,l	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	31938	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	121.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.02% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MAL, CEY

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	0.38	0/5396	0.74	7/7380 (0.1%)
1	B	0.39	0/5396	0.76	6/7380 (0.1%)
1	C	0.39	0/5396	0.76	5/7380 (0.1%)
1	D	0.41	0/5396	0.81	10/7380 (0.1%)
1	E	0.40	0/5396	0.78	6/7380 (0.1%)
1	F	0.38	0/5396	0.78	8/7380 (0.1%)
All	All	0.39	0/32376	0.77	42/44280 (0.1%)

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	D	0	3
1	E	0	2
All	All	0	7

There are no bond length outliers.

The worst 5 of 42 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	366	LEU	CA-CB-CG	8.78	135.49	115.30
1	E	698	LEU	CA-CB-CG	8.69	135.29	115.30
1	B	505	LEU	CA-CB-CG	8.41	134.65	115.30
1	A	560	LEU	CA-CB-CG	8.29	134.37	115.30
1	F	195	LEU	CA-CB-CG	7.42	132.37	115.30



There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	202	GLU	Peptide
1	B	292	ASN	Sidechain
1	D	292	ASN	Sidechain
1	D	484	ASN	Sidechain
1	D	488	GLU	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	5233	0	5044	235	0
1	B	5233	0	5044	223	0
1	C	5233	0	5044	229	0
1	D	5233	0	5044	269	0
1	E	5233	0	5045	232	1
1	F	5233	0	5045	209	0
2	A	23	0	22	0	0
2	B	23	0	22	0	0
2	C	23	0	22	2	0
2	D	23	0	22	4	0
2	E	23	0	22	2	0
2	F	23	0	22	1	0
3	A	67	0	62	3	0
3	B	67	0	62	3	0
3	C	67	0	62	3	0
3	D	67	0	62	4	0
3	E	67	0	62	4	0
3	F	67	0	62	5	0
All	All	31938	0	30770	1380	1

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 22.

The worst 5 of 1380 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:D:487:ALA:HA	1:D:488:GLU:HB2	1.13	1.11
1:C:487:ALA:HB1	1:C:488:GLU:HB3	1.30	1.08
1:D:154:LEU:HD22	1:D:183:LEU:HD22	1.38	1.03
1:D:487:ALA:HA	1:D:488:GLU:CB	1.90	1.02
1:B:165:ALA:HB1	1:B:173:ARG:HD2	1.43	0.97

All (1) 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:E:29:CYS:CB	1:E:29:CYS:SG[2_556]	1.75	0.45

## 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	654/723 (90%)	629 (96%)	20 (3%)	5 (1%)	22	66
1	B	654/723 (90%)	624 (95%)	21 (3%)	9 (1%)	13	55
1	C	654/723 (90%)	628 (96%)	20 (3%)	6 (1%)	20	63
1	D	654/723 (90%)	621 (95%)	24 (4%)	9 (1%)	13	55
1	E	654/723 (90%)	624 (95%)	22 (3%)	8 (1%)	15	59
1	F	654/723 (90%)	629 (96%)	19 (3%)	6 (1%)	20	63
All	All	3924/4338 (90%)	3755 (96%)	126 (3%)	43 (1%)	17	60

5 of 43 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	69	TYR
1	A	151	SER
1	B	68	ARG
1	B	69	TYR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	150	LEU

### 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	542/597 (91%)	514 (95%)	28 (5%)	27	63
1	B	542/597 (91%)	525 (97%)	17 (3%)	45	75
1	C	542/597 (91%)	515 (95%)	27 (5%)	28	64
1	D	542/597 (91%)	520 (96%)	22 (4%)	35	69
1	E	542/597 (91%)	508 (94%)	34 (6%)	21	57
1	F	542/597 (91%)	516 (95%)	26 (5%)	30	65
All	All	3252/3582 (91%)	3098 (95%)	154 (5%)	30	66

5 of 154 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	692	GLU
1	D	436	LYS
1	F	383	ASP
1	C	696	THR
1	D	195	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	280	HIS
1	D	405	GLN
1	E	455	GLN
1	D	314	HIS
1	D	419	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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

12 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	MAL	A	801	-	24,24,24	0.58	0	35,35,35	1.42	5 (14%)
3	CEY	A	802	-	72,72,72	2.19	31 (43%)	107,107,107	1.32	15 (14%)
2	MAL	B	801	-	24,24,24	0.55	0	35,35,35	0.91	2 (5%)
3	CEY	B	802	-	72,72,72	2.20	32 (44%)	107,107,107	1.28	19 (17%)
2	MAL	C	801	-	24,24,24	0.55	0	35,35,35	1.04	2 (5%)
3	CEY	C	802	-	72,72,72	2.17	29 (40%)	107,107,107	1.23	9 (8%)
2	MAL	D	801	-	24,24,24	0.56	0	35,35,35	1.15	2 (5%)
3	CEY	D	802	-	72,72,72	2.21	34 (47%)	107,107,107	1.34	15 (14%)
2	MAL	E	801	-	24,24,24	0.52	0	35,35,35	1.07	3 (8%)
3	CEY	E	802	-	72,72,72	2.17	30 (41%)	107,107,107	1.20	7 (6%)
2	MAL	F	801	-	24,24,24	0.47	0	35,35,35	0.70	1 (2%)
3	CEY	F	802	-	72,72,72	2.20	32 (44%)	107,107,107	1.31	15 (14%)

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	MAL	A	801	-	-	0/8/48/48	0/2/2/2
3	CEY	A	802	-	-	0/32/152/152	0/6/6/6
2	MAL	B	801	-	-	0/8/48/48	0/2/2/2
3	CEY	B	802	-	-	0/32/152/152	0/6/6/6
2	MAL	C	801	-	-	0/8/48/48	0/2/2/2
3	CEY	C	802	-	-	0/32/152/152	0/6/6/6
2	MAL	D	801	-	-	0/8/48/48	0/2/2/2
3	CEY	D	802	-	-	0/32/152/152	0/6/6/6
2	MAL	E	801	-	-	0/8/48/48	0/2/2/2
3	CEY	E	802	-	-	0/32/152/152	0/6/6/6
2	MAL	F	801	-	-	0/8/48/48	0/2/2/2
3	CEY	F	802	-	-	0/32/152/152	0/6/6/6

The worst 5 of 188 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	802	CEY	C6C-C5C	-3.65	1.39	1.51
3	A	802	CEY	C6C-C5C	-3.63	1.39	1.51
3	F	802	CEY	C6C-C5C	-3.62	1.39	1.51
3	C	802	CEY	C6C-C5C	-3.57	1.39	1.51
3	E	802	CEY	C6C-C5C	-3.54	1.39	1.51

The worst 5 of 95 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	802	CEY	C1E-O4D-C4D	-3.98	108.29	118.00
3	F	802	CEY	C1D-O4C-C4C	-3.54	109.38	118.00
3	A	802	CEY	C1D-O4C-C4C	-3.38	109.76	118.00
3	B	802	CEY	O6B-C6B-C5B	-3.37	100.00	111.34
3	E	802	CEY	C1E-O4D-C4D	-3.22	110.15	118.00

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

10 monomers are involved in 31 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	802	CEY	3	0
3	B	802	CEY	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	801	MAL	2	0
3	C	802	CEY	3	0
2	D	801	MAL	4	0
3	D	802	CEY	4	0
2	E	801	MAL	2	0
3	E	802	CEY	4	0
2	F	801	MAL	1	0
3	F	802	CEY	5	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	660/723 (91%)	0.94	54 (8%) 12 10	86, 118, 165, 185	0
1	B	660/723 (91%)	0.92	54 (8%) 12 10	86, 117, 162, 185	0
1	C	660/723 (91%)	0.99	66 (10%) 8 7	88, 119, 167, 201	0
1	D	660/723 (91%)	0.90	54 (8%) 12 10	87, 117, 162, 192	0
1	E	660/723 (91%)	0.88	49 (7%) 15 12	85, 119, 165, 191	0
1	F	660/723 (91%)	0.90	41 (6%) 21 16	84, 117, 167, 190	0
All	All	3960/4338 (91%)	0.92	318 (8%) 13 11	84, 118, 165, 201	0

The worst 5 of 318 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	356	ALA	4.8
1	C	377	PRO	4.5
1	D	449	PHE	3.9
1	B	180	ALA	3.8
1	A	347	LEU	3.7

### 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 [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. 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
2	MAL	E	801	23/23	0.80	0.38	-0.34	121,144,152,156	0
3	CEY	A	802	67/67	0.89	0.35	-0.41	105,125,132,139	0
2	MAL	A	801	23/23	0.75	0.35	-0.55	111,139,149,153	0
2	MAL	B	801	23/23	0.79	0.33	-0.57	117,144,154,155	0
2	MAL	D	801	23/23	0.77	0.34	-0.58	127,147,156,160	0
2	MAL	F	801	23/23	0.77	0.36	-0.62	122,139,148,156	0
3	CEY	D	802	67/67	0.88	0.33	-0.62	109,124,130,134	0
3	CEY	F	802	67/67	0.90	0.33	-0.65	108,121,132,148	0
2	MAL	C	801	23/23	0.76	0.35	-0.83	110,140,153,163	0
3	CEY	B	802	67/67	0.85	0.31	-0.86	106,128,160,171	0
3	CEY	C	802	67/67	0.89	0.32	-0.94	110,125,132,141	0
3	CEY	E	802	67/67	0.87	0.30	-0.96	110,127,134,137	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.