



wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 07:30 am BST

PDB ID : 5JY7
Title : Complex of Mycobacterium smegmatis trehalose synthase with maltokinase
Authors : Futterer, K.; Kermani, A.A.; Besra, G.S.
Deposited on : 2016-05-13
Resolution : 3.60 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

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

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

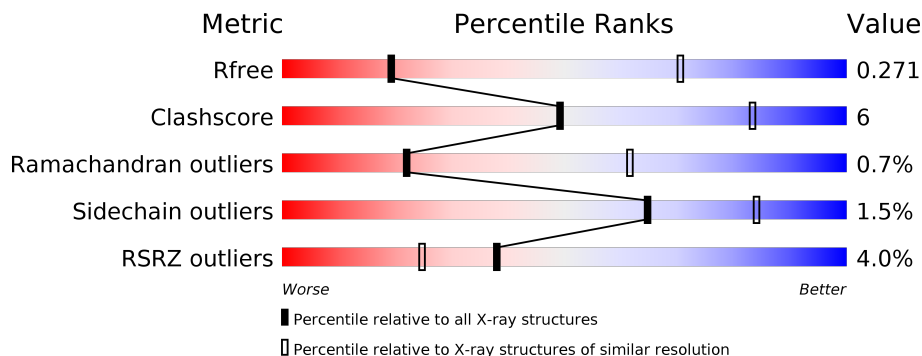
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.60 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1257 (3.70-3.50)
Clashscore	141614	1353 (3.70-3.50)
Ramachandran outliers	138981	1307 (3.70-3.50)
Sidechain outliers	138945	1307 (3.70-3.50)
RSRZ outliers	127900	1161 (3.70-3.50)

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 $\leq 5\%$. 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	593	
1	B	593	
1	C	593	
1	D	593	
1	E	593	
1	F	593	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	593	
1	H	593	
2	I	441	
2	J	441	
2	K	441	
2	L	441	
2	M	441	
2	N	441	
2	O	441	
2	P	441	

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	CA	F	601	-	-	-	X

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 56343 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 Trehalose synthase/amylase TreS.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	565	Total 4610	C 2952	N 788	O 853	S 17	0	0	0
1	B	571	Total 4661	C 2981	N 800	O 863	S 17	0	0	0
1	C	569	Total 4630	C 2964	N 792	O 857	S 17	0	0	0
1	D	571	Total 4661	C 2981	N 800	O 863	S 17	0	0	0
1	E	549	Total 4496	C 2884	N 764	O 831	S 17	0	0	0
1	F	571	Total 4661	C 2981	N 800	O 863	S 17	0	0	0
1	G	571	Total 4661	C 2981	N 800	O 863	S 17	0	0	0
1	H	544	Total 4452	C 2858	N 756	O 821	S 17	0	0	0

- Molecule 2 is a protein called Maltokinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	I	435	Total 2909	C 1811	N 523	O 569	S 6	0	0	0
2	J	435	Total 2899	C 1807	N 523	O 563	S 6	0	0	0
2	K	353	Total 2380	C 1481	N 432	O 462	S 5	0	0	0
2	L	436	Total 2913	C 1816	N 524	O 566	S 7	0	0	0
2	M	431	Total 2887	C 1799	N 519	O 563	S 6	0	0	0
2	N	349	Total 2347	C 1459	N 428	O 455	S 5	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	O	239	Total	C	N	O	S	0	0	0
			1584	974	291	315	4			
2	P	239	Total	C	N	O	S	0	0	0
			1584	974	291	315	4			

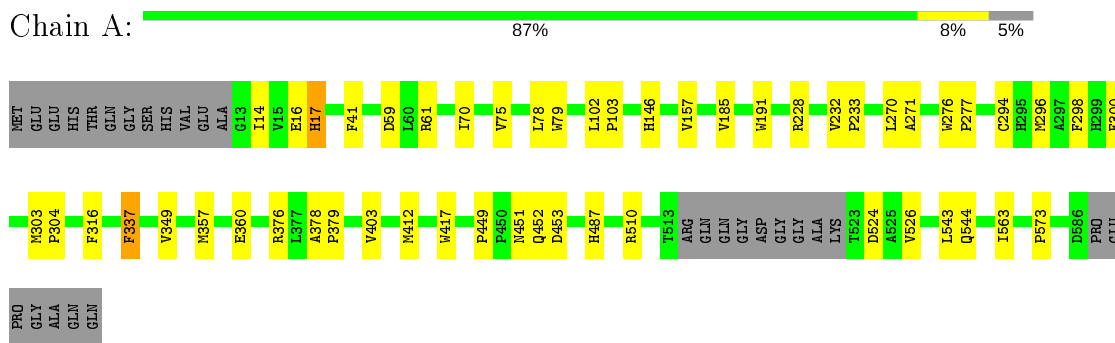
- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	G	1	Total	Ca	0	0
			1	1		
3	D	1	Total	Ca	0	0
			1	1		
3	E	1	Total	Ca	0	0
			1	1		
3	H	1	Total	Ca	0	0
			1	1		
3	B	1	Total	Ca	0	0
			1	1		
3	C	1	Total	Ca	0	0
			1	1		
3	A	1	Total	Ca	0	0
			1	1		
3	F	1	Total	Ca	0	0
			1	1		

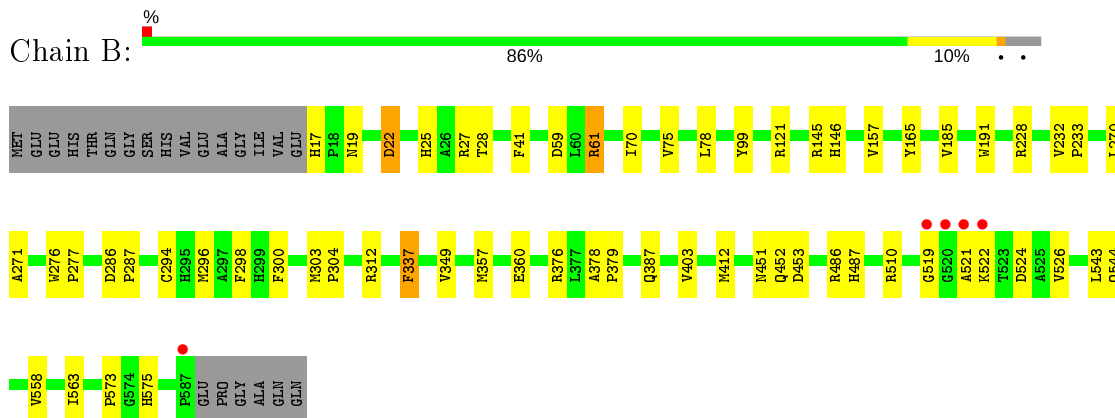
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

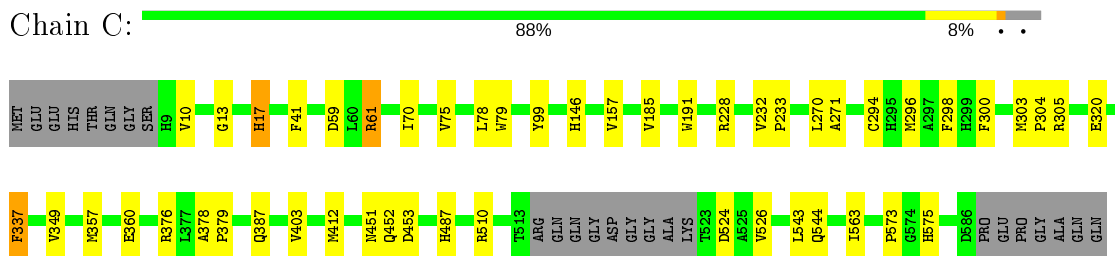
- Molecule 1: Trehalose synthase/amylase TreS



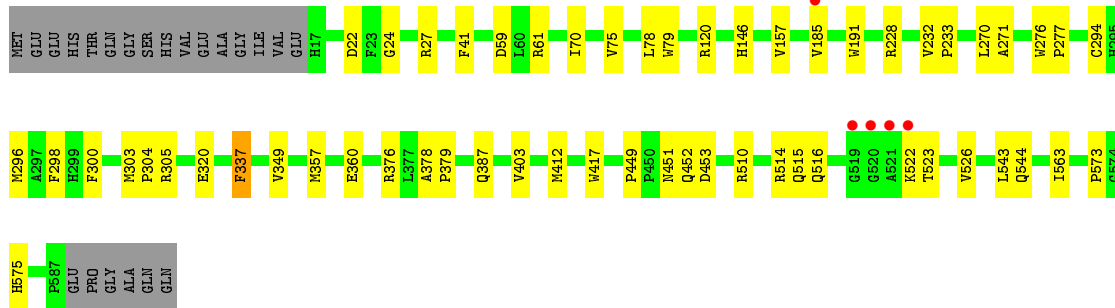
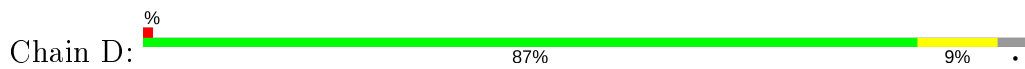
- Molecule 1: Trehalose synthase/amylase TreS



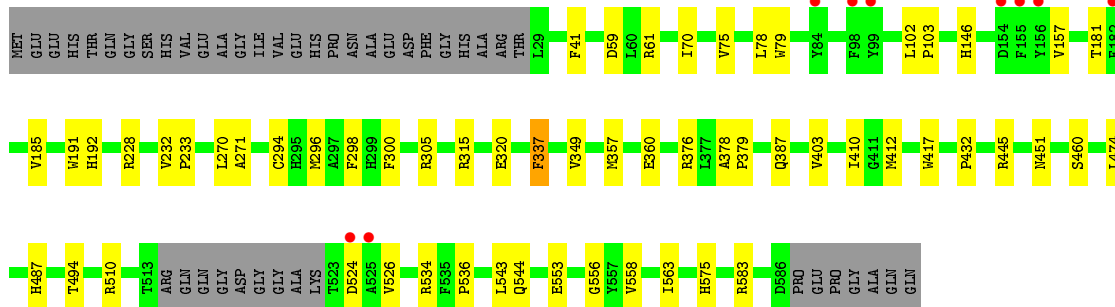
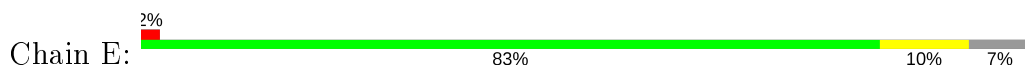
- Molecule 1: Trehalose synthase/amylase TreS



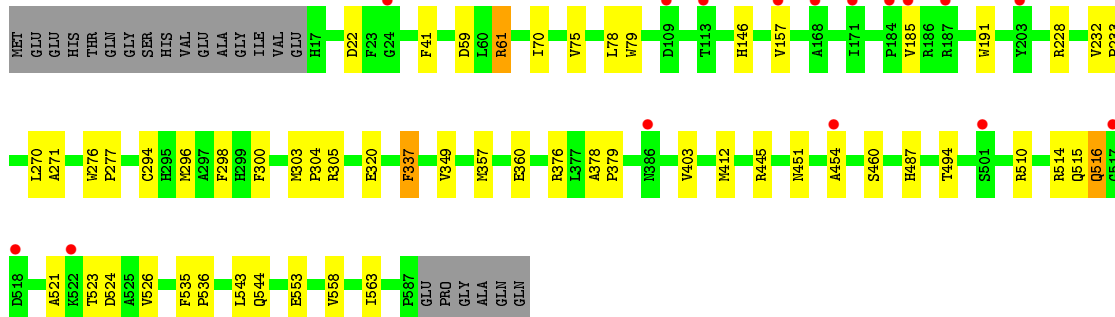
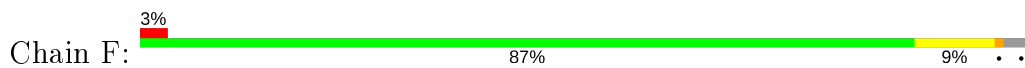
- Molecule 1: Trehalose synthase/amylase TreS



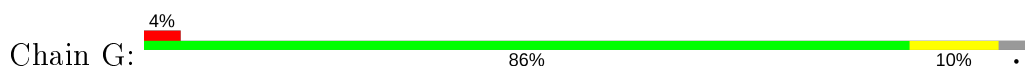
• Molecule 1: Trehalose synthase/amylase TreS

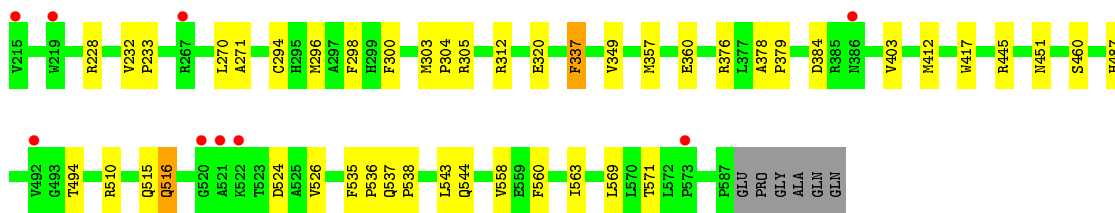


• Molecule 1: Trehalose synthase/amylase TreS

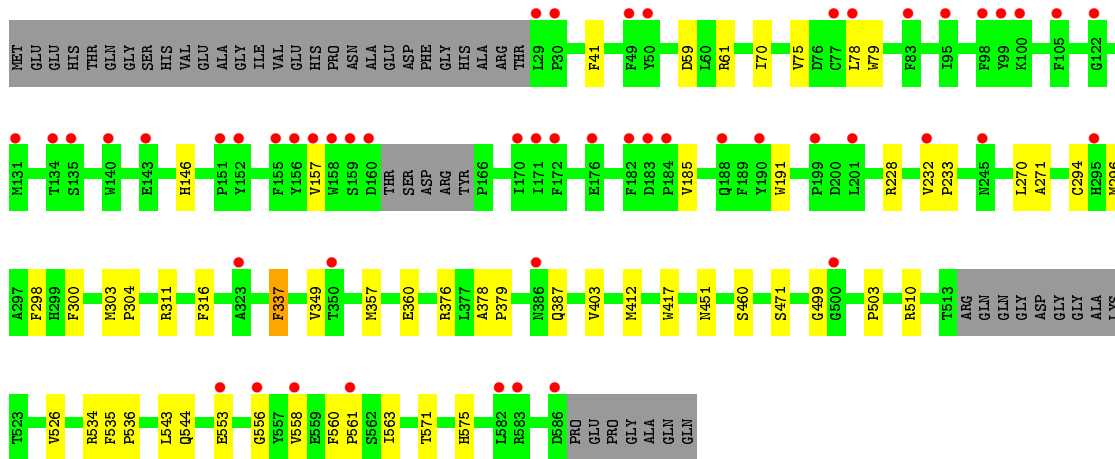
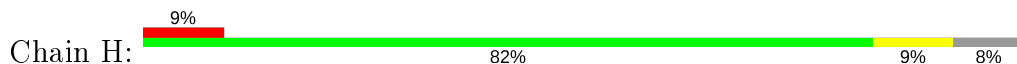


• Molecule 1: Trehalose synthase/amylase TreS

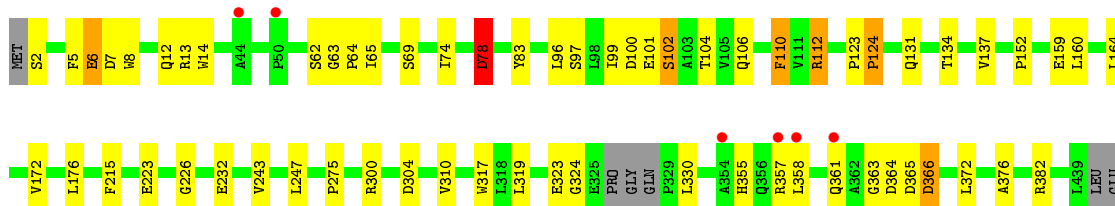
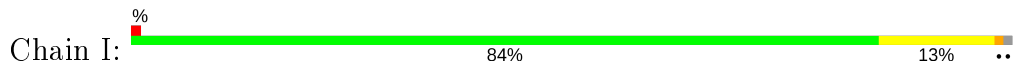




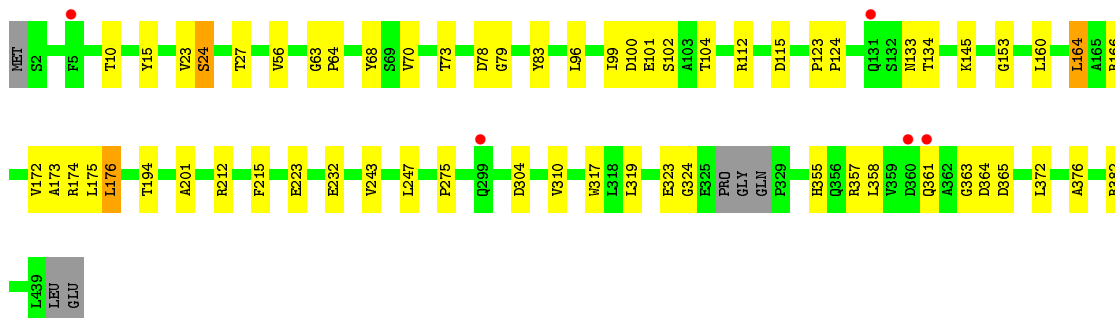
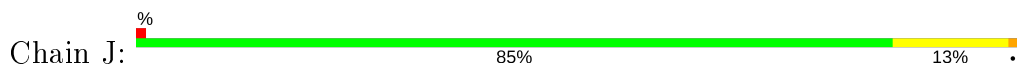
• Molecule 1: Trehalose synthase/amylase TreS



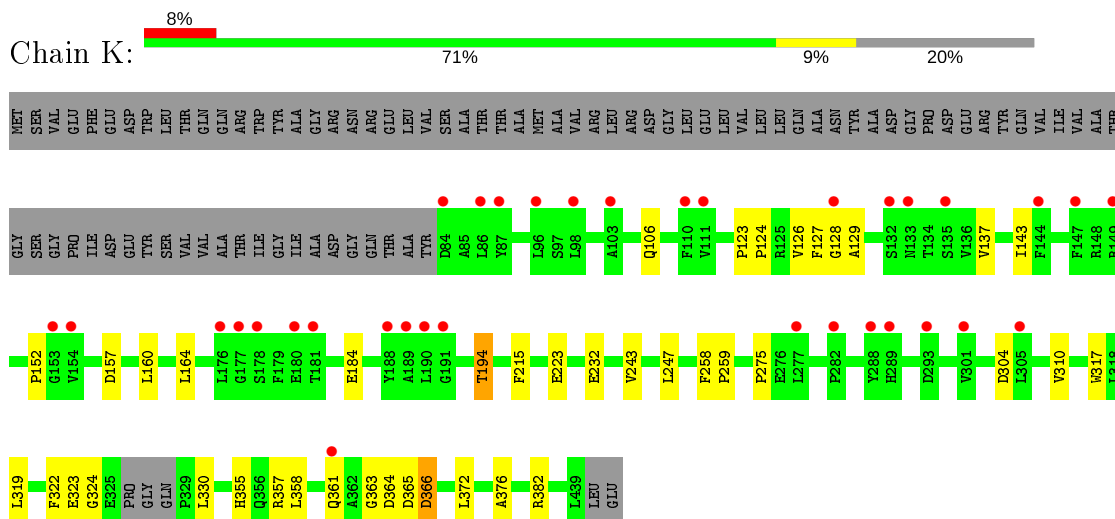
• Molecule 2: Maltokinase



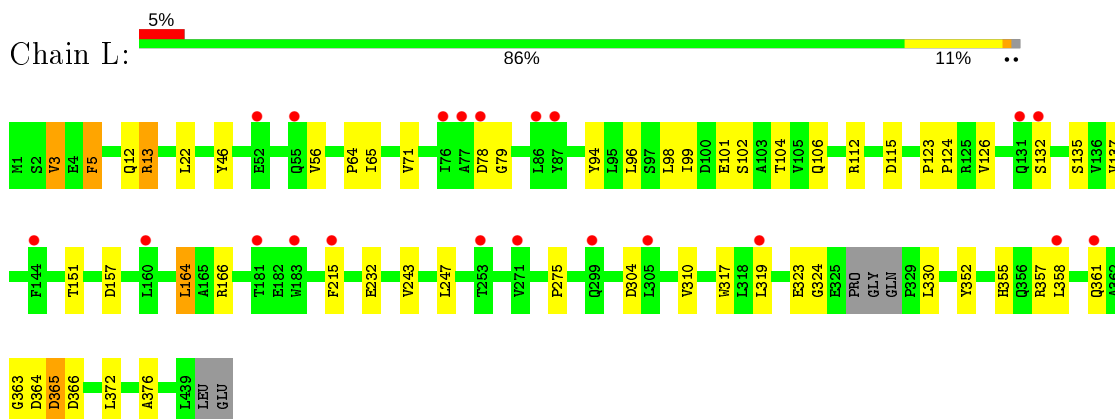
• Molecule 2: Maltokinase



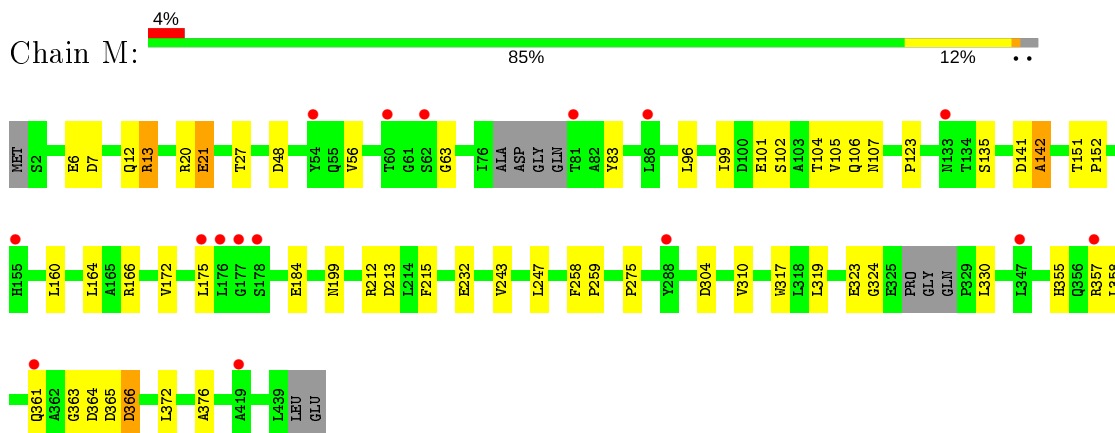
- Molecule 2: Maltokinase



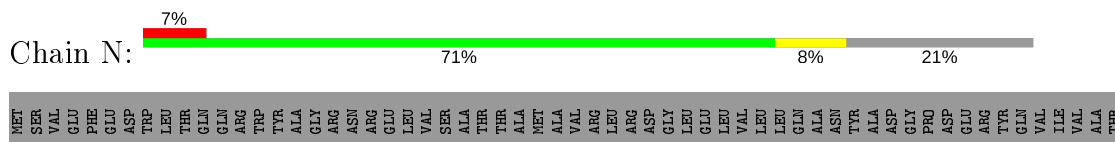
- Molecule 2: Maltokinase

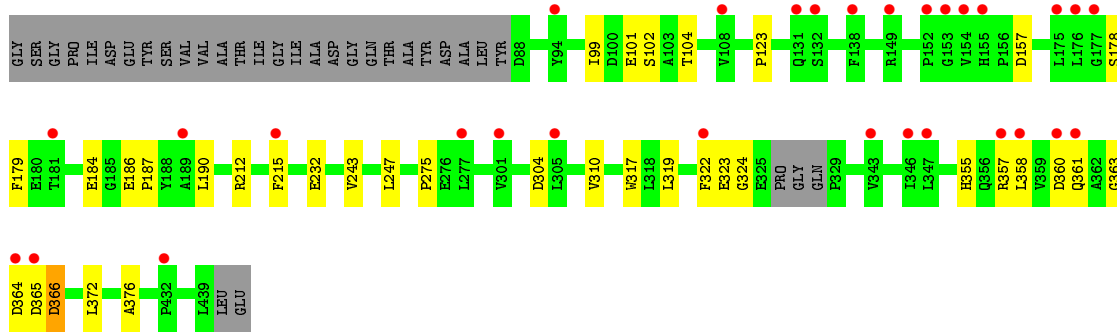


- Molecule 2: Maltokinase

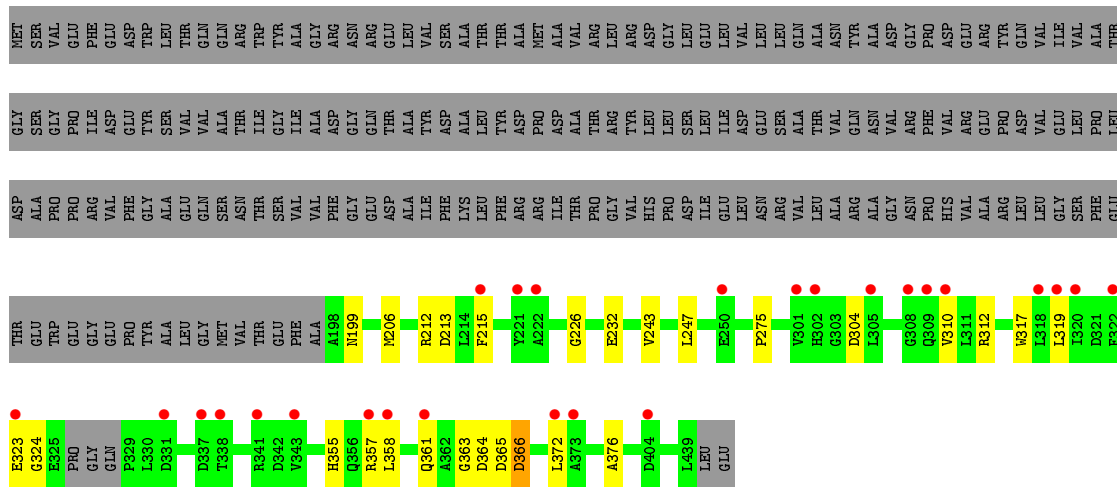


- Molecule 2: Maltokinase

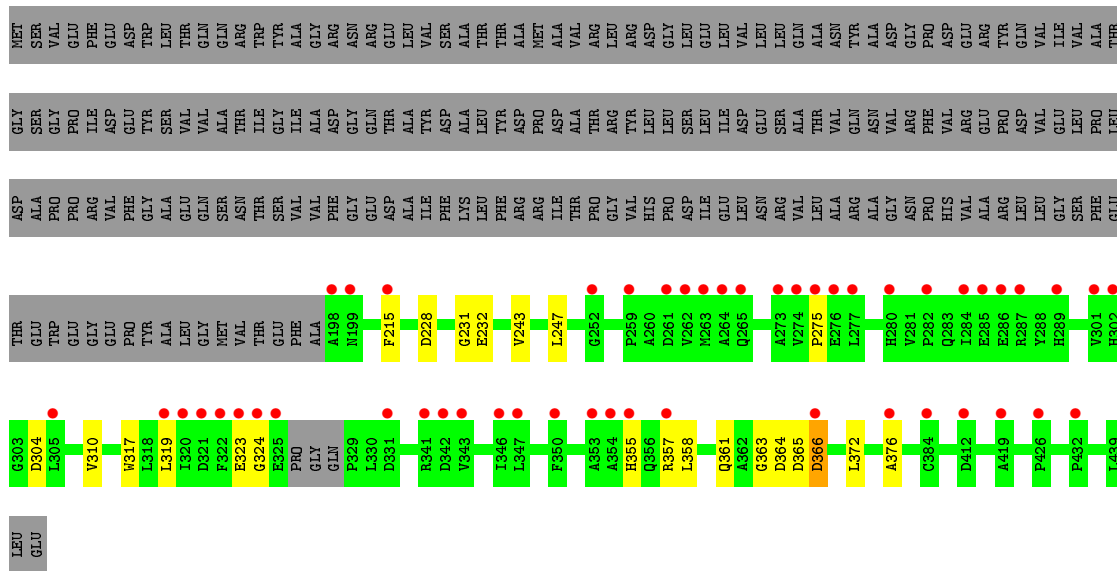




• Molecule 2: Maltokinase



• Molecule 2: Maltokinase



4 Data and refinement statistics

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, α , β , γ	315.68Å 315.68Å 124.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.90 – 3.60 49.91 – 3.60	Depositor EDS
% Data completeness (in resolution range)	99.8 (49.90-3.60) 99.9 (49.91-3.60)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	0.17	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.30 (at 3.57Å)	Xtrriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.261 , 0.281 0.254 , 0.271	Depositor DCC
R_{free} test set	7119 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	125.2	Xtrriage
Anisotropy	0.002	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 76.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	0.036 for h,-k,-l	Xtrriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	56343	wwPDB-VP
Average B, all atoms (Å ²)	111.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.57% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality i

5.1 Standard geometry i

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

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	0.37	0/4754	0.53	1/6485 (0.0%)
1	B	0.41	0/4807	0.56	1/6556 (0.0%)
1	C	0.38	0/4774	0.54	1/6513 (0.0%)
1	D	0.34	0/4807	0.52	1/6556 (0.0%)
1	E	0.31	0/4636	0.51	1/6324 (0.0%)
1	F	0.36	2/4807 (0.0%)	0.51	1/6556 (0.0%)
1	G	0.29	0/4807	0.51	1/6556 (0.0%)
1	H	0.29	0/4590	0.51	1/6259 (0.0%)
2	I	0.42	0/2973	0.60	0/4091
2	J	0.38	0/2963	0.57	0/4078
2	K	0.32	0/2432	0.52	0/3343
2	L	0.39	2/2977 (0.1%)	0.55	0/4095
2	M	0.33	0/2950	0.54	0/4058
2	N	0.32	0/2398	0.53	0/3296
2	O	0.31	0/1615	0.53	0/2219
2	P	0.32	0/1615	0.53	0/2219
All	All	0.35	4/57905 (0.0%)	0.53	8/79204 (0.0%)

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
2	L	0	2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	61	ARG	CZ-NH1	11.18	1.47	1.33
1	F	61	ARG	CD-NE	8.59	1.61	1.46

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	L	352	TYR	C-N	8.42	1.53	1.34
2	L	366	ASP	C-N	7.62	1.51	1.34

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	61	ARG	CG-CD-NE	8.76	130.19	111.80
1	A	61	ARG	CG-CD-NE	7.92	128.42	111.80
1	G	61	ARG	CG-CD-NE	7.85	128.28	111.80
1	H	61	ARG	CG-CD-NE	7.82	128.23	111.80
1	E	61	ARG	CG-CD-NE	7.82	128.23	111.80

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	L	365	ASP	Mainchain,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	4610	0	4348	32	0
1	B	4661	0	4407	39	0
1	C	4630	0	4359	31	0
1	D	4661	0	4407	36	0
1	E	4496	0	4260	59	0
1	F	4661	0	4407	47	0
1	G	4661	0	4407	74	0
1	H	4452	0	4222	72	0
2	I	2909	0	2344	53	0
2	J	2899	0	2336	51	0
2	K	2380	0	1950	39	0
2	L	2913	0	2356	39	0
2	M	2887	0	2329	51	0
2	N	2347	0	1921	37	0
2	O	1584	0	1255	43	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	P	1584	0	1255	30	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
3	G	1	0	0	0	0
3	H	1	0	0	0	0
All	All	56343	0	50563	625	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 6.

The worst 5 of 625 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:I:215:PHE:CB	2:I:361:GLN:HE22	1.62	1.11
2:O:215:PHE:CB	2:O:361:GLN:HE22	1.64	1.11
2:L:215:PHE:CB	2:L:361:GLN:HE22	1.64	1.10
2:N:215:PHE:CB	2:N:361:GLN:HE22	1.64	1.10
2:K:215:PHE:CB	2:K:361:GLN:HE22	1.64	1.10

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	561/593 (95%)	546 (97%)	13 (2%)	2 (0%)	34 71
1	B	569/593 (96%)	551 (97%)	16 (3%)	2 (0%)	34 71

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	565/593 (95%)	549 (97%)	14 (2%)	2 (0%)	34	71
1	D	569/593 (96%)	550 (97%)	16 (3%)	3 (0%)	29	68
1	E	545/593 (92%)	533 (98%)	12 (2%)	0	100	100
1	F	569/593 (96%)	553 (97%)	14 (2%)	2 (0%)	34	71
1	G	569/593 (96%)	548 (96%)	20 (4%)	1 (0%)	47	79
1	H	538/593 (91%)	528 (98%)	10 (2%)	0	100	100
2	I	431/441 (98%)	382 (89%)	43 (10%)	6 (1%)	11	48
2	J	431/441 (98%)	382 (89%)	41 (10%)	8 (2%)	8	42
2	K	349/441 (79%)	318 (91%)	25 (7%)	6 (2%)	9	45
2	L	432/441 (98%)	388 (90%)	35 (8%)	9 (2%)	7	40
2	M	425/441 (96%)	381 (90%)	35 (8%)	9 (2%)	7	40
2	N	345/441 (78%)	318 (92%)	24 (7%)	3 (1%)	17	57
2	O	235/441 (53%)	221 (94%)	13 (6%)	1 (0%)	34	71
2	P	235/441 (53%)	220 (94%)	14 (6%)	1 (0%)	34	71
All	All	7368/8272 (89%)	6968 (95%)	345 (5%)	55 (1%)	22	61

5 of 55 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	14	ILE
1	B	25	HIS
1	C	10	VAL
2	I	275	PRO
2	J	275	PRO

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	489/512 (96%)	486 (99%)	3 (1%)	86	94
1	B	495/512 (97%)	488 (99%)	7 (1%)	67	85

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	489/512 (96%)	484 (99%)	5 (1%)	76	88
1	D	495/512 (97%)	491 (99%)	4 (1%)	81	91
1	E	480/512 (94%)	477 (99%)	3 (1%)	86	94
1	F	495/512 (97%)	491 (99%)	4 (1%)	81	91
1	G	495/512 (97%)	491 (99%)	4 (1%)	81	91
1	H	475/512 (93%)	472 (99%)	3 (1%)	86	94
2	I	211/351 (60%)	198 (94%)	13 (6%)	18	53
2	J	208/351 (59%)	198 (95%)	10 (5%)	25	60
2	K	178/351 (51%)	175 (98%)	3 (2%)	60	82
2	L	210/351 (60%)	202 (96%)	8 (4%)	33	66
2	M	210/351 (60%)	202 (96%)	8 (4%)	33	66
2	N	175/351 (50%)	172 (98%)	3 (2%)	60	82
2	O	110/351 (31%)	108 (98%)	2 (2%)	59	81
2	P	110/351 (31%)	108 (98%)	2 (2%)	59	81
All	All	5325/6904 (77%)	5243 (98%)	82 (2%)	65	84

5 of 82 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	I	100	ASP
2	I	366	ASP
2	N	104	THR
2	I	102	SER
2	I	112	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	452	GLN
1	D	452	GLN
2	L	11	GLN
1	D	146	HIS
1	D	197	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](#)

Of 8 ligands modelled in this entry, 8 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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, 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	565/593 (95%)	-0.21	0 100 100	34, 53, 96, 118	0
1	B	571/593 (96%)	-0.25	5 (0%) 84 73	28, 43, 64, 79	0
1	C	569/593 (95%)	-0.21	0 100 100	35, 51, 79, 106	0
1	D	571/593 (96%)	-0.13	5 (0%) 84 73	44, 77, 111, 129	0
1	E	549/593 (92%)	0.06	9 (1%) 72 57	67, 102, 176, 208	0
1	F	571/593 (96%)	0.16	16 (2%) 53 37	68, 119, 249, 314	0
1	G	571/593 (96%)	0.24	21 (3%) 41 27	80, 142, 218, 251	0
1	H	544/593 (91%)	0.60	51 (9%) 8 4	89, 159, 248, 316	0
2	I	435/441 (98%)	-0.32	6 (1%) 75 61	38, 62, 111, 129	0
2	J	435/441 (98%)	-0.09	5 (1%) 80 68	55, 92, 152, 171	0
2	K	353/441 (80%)	0.30	34 (9%) 8 4	106, 160, 228, 246	0
2	L	436/441 (98%)	0.10	21 (4%) 30 19	88, 119, 163, 181	0
2	M	431/441 (97%)	0.09	16 (3%) 41 27	99, 131, 188, 228	0
2	N	349/441 (79%)	0.37	30 (8%) 10 6	116, 177, 233, 254	0
2	O	239/441 (54%)	0.38	26 (10%) 5 3	168, 210, 253, 266	0
2	P	239/441 (54%)	1.04	50 (20%) 1 0	150, 195, 247, 256	0
All	All	7428/8272 (89%)	0.08	295 (3%) 38 25	28, 106, 220, 316	0

The worst 5 of 295 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	O	301	VAL	6.8
2	L	361	GLN	6.6
2	M	177	GLY	6.3
2	K	132	SER	6.2
2	P	342	ASP	6.1

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. 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
3	CA	H	601	1/1	0.53	0.33	20,20,20,20	0
3	CA	F	601	1/1	0.79	0.81	20,20,20,20	0
3	CA	E	601	1/1	0.93	0.21	20,20,20,20	0
3	CA	D	601	1/1	0.93	0.06	20,20,20,20	0
3	CA	B	601	1/1	0.97	0.13	20,20,20,20	0
3	CA	G	601	1/1	0.98	0.16	20,20,20,20	0
3	CA	A	601	1/1	0.98	0.05	20,20,20,20	0
3	CA	C	601	1/1	0.99	0.11	20,20,20,20	0

6.5 Other polymers [i](#)

There are no such residues in this entry.