WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 20, 2023 – 09:22 pm GMT

PDB ID	:	7P5L
Title	:	Crystal structure of Mycobacterium hassiacum glucosyl-3-phosphoglycerate synthase at pH 7.1 - apo form
Authors	:	Silva, A.; Nunes-Costa, D.; Barbosa Pereira, P.J.; Macedo-Ribeiro, S.
Deposited on	:	2021-07-14
Resolution	:	1.22 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.32.1
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)	:	2.32.1

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 1.22 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.
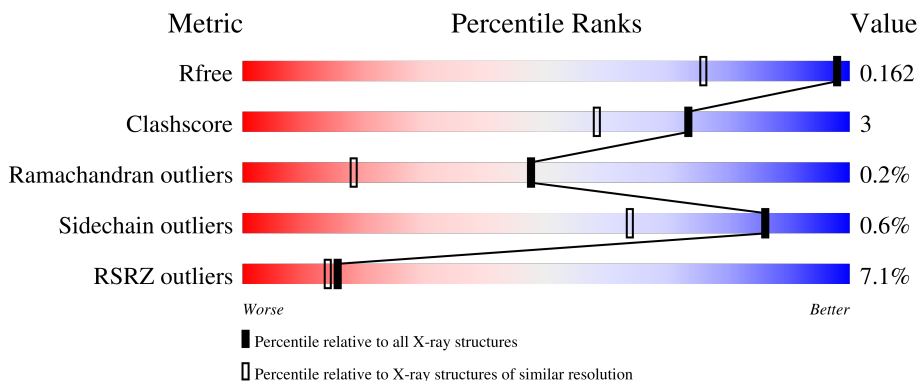
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1232 (1.24-1.20)
Clashscore	141614	1294 (1.24-1.20)
Ramachandran outliers	138981	1251 (1.24-1.20)
Sidechain outliers	138945	1250 (1.24-1.20)
RSRZ outliers	127900	1209 (1.24-1.20)

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	327	8% ... 82% ... 6% 11%
1	B	327	5% ... 83% ... 6% 11%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 10009 atoms, of which 4771 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 Glucosyl-3-phosphoglycerate synthase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	291	Total	C	H	N	O	S	9	24	0
			4731	1474	2403	414	432	8			
1	B	290	Total	C	H	N	O	S	0	19	0
			4646	1445	2368	399	425	9			

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	315	LYS	-	expression tag	UNP K5B7Z4
A	316	LEU	-	expression tag	UNP K5B7Z4
A	317	ALA	-	expression tag	UNP K5B7Z4
A	318	ALA	-	expression tag	UNP K5B7Z4
A	319	ALA	-	expression tag	UNP K5B7Z4
A	320	LEU	-	expression tag	UNP K5B7Z4
A	321	GLU	-	expression tag	UNP K5B7Z4
A	322	HIS	-	expression tag	UNP K5B7Z4
A	323	HIS	-	expression tag	UNP K5B7Z4
A	324	HIS	-	expression tag	UNP K5B7Z4
A	325	HIS	-	expression tag	UNP K5B7Z4
A	326	HIS	-	expression tag	UNP K5B7Z4
A	327	HIS	-	expression tag	UNP K5B7Z4
B	315	LYS	-	expression tag	UNP K5B7Z4
B	316	LEU	-	expression tag	UNP K5B7Z4
B	317	ALA	-	expression tag	UNP K5B7Z4
B	318	ALA	-	expression tag	UNP K5B7Z4
B	319	ALA	-	expression tag	UNP K5B7Z4
B	320	LEU	-	expression tag	UNP K5B7Z4
B	321	GLU	-	expression tag	UNP K5B7Z4
B	322	HIS	-	expression tag	UNP K5B7Z4
B	323	HIS	-	expression tag	UNP K5B7Z4
B	324	HIS	-	expression tag	UNP K5B7Z4
B	325	HIS	-	expression tag	UNP K5B7Z4
B	326	HIS	-	expression tag	UNP K5B7Z4

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	327	HIS	-	expression tag	UNP K5B7Z4

- Molecule 2 is (2S)-2-hydroxybutanedioic acid (three-letter code: LMR) (formula: $C_4H_6O_5$).

LMR
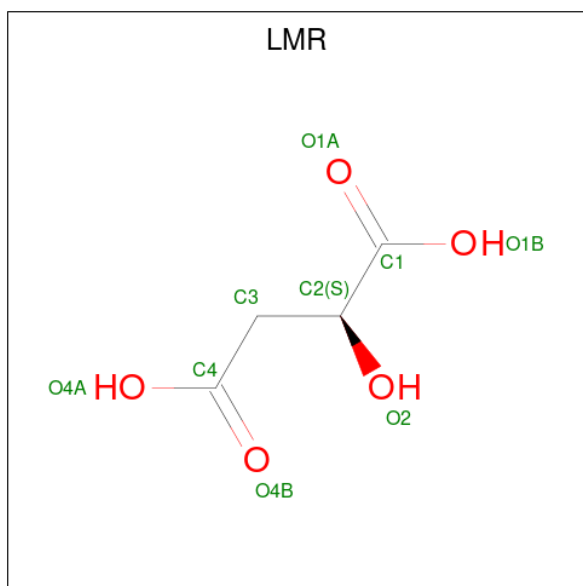
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	9	4	5	0	0

- Molecule 3 is D-MALATE (three-letter code: MLT) (formula: $C_4H_6O_5$).

MLT
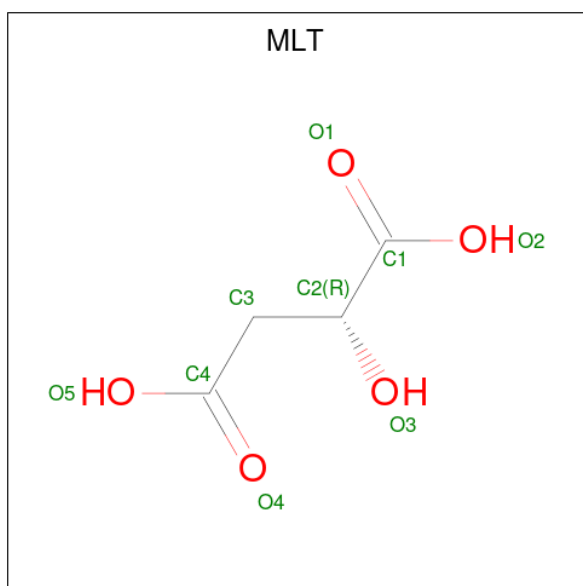
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total C O 9 4 5			0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	307	Total O 322 322		0	19
4	B	280	Total O 292 292		0	16

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: Glucosyl-3-phosphoglycerate synthase

Chain A:

• Molecule 1: Glucosyl-3-phosphoglycerate synthase

Chain B:
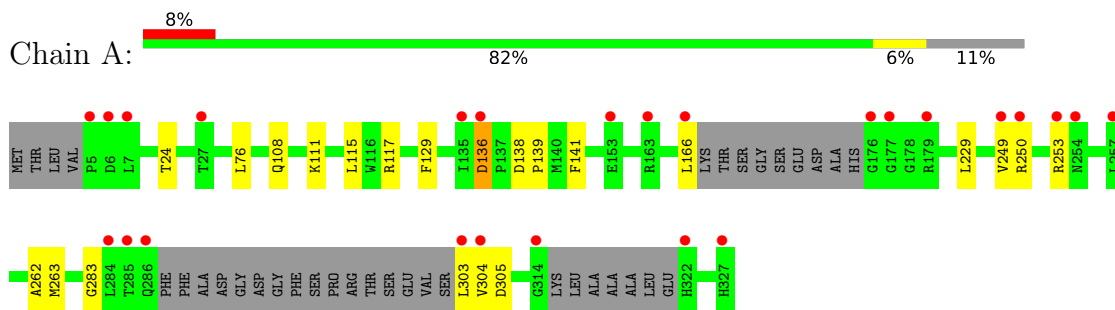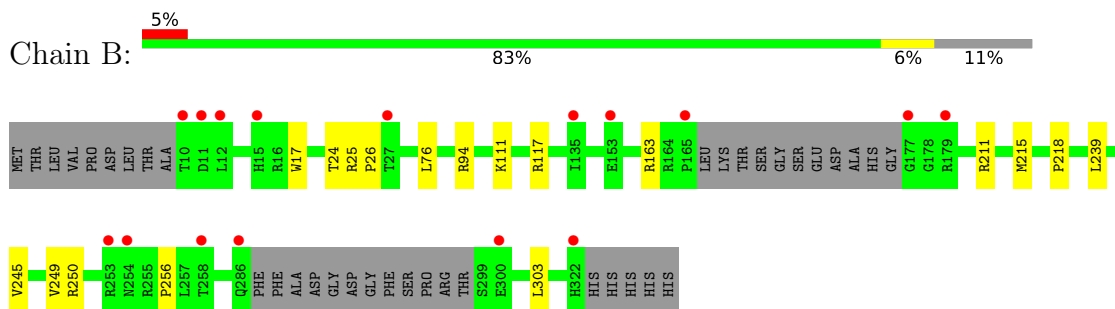
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	71.10Å 90.61Å 95.26Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.23 – 1.22 48.23 – 1.22	Depositor EDS
% Data completeness (in resolution range)	99.7 (48.23-1.22) 99.7 (48.23-1.22)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.87 (at 1.22Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.143 , 0.164 0.142 , 0.162	Depositor DCC
R_{free} test set	9129 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	12.2	Xtriage
Anisotropy	0.237	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.44 , 49.2	EDS
L-test for twinning[2]	$< |L| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.015 for -h,l,k	Xtriage
F_o,F_c correlation	0.98	EDS
Total number of atoms	10009	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 4.18% 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 (i)

5.1   Standard geometry (i)

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

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	0.45	1/2448 (0.0%)	0.69	0/3333
1	B	0.52	2/2375 (0.1%)	0.70	0/3235
All	All	0.49	3/4823 (0.1%)	0.69	0/6568

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
All	All	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	256	PRO	N-CD	-13.05	1.29	1.47
1	A	139	PRO	N-CD	8.17	1.59	1.47
1	B	218	PRO	N-CD	5.98	1.56	1.47

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	117	ARG	Sidechain
1	B	117	ARG	Sidechain

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	2328	2403	2357	16	0
1	B	2278	2368	2333	15	0
2	A	9	0	4	3	0
3	B	9	0	4	0	0
4	A	322	0	0	4	0
4	B	292	0	0	3	0
All	All	5238	4771	4698	31	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 3.

The worst 5 of 31 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:108:GLN:OE1	1:A:303:LEU:HD22	1.70	0.91
1:B:163:ARG:HB3	1:B:163:ARG:NH1	2.01	0.76
1:B:163:ARG:HB3	1:B:163:ARG:HH11	1.51	0.75
1:A:249:VAL:O	1:A:250:ARG:HD3	1.92	0.69
1:B:303:LEU:HD22	4:B:677:HOH:O	2.01	0.61

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	309/327 (94%)	300 (97%)	7 (2%)	2 (1%)	25	5

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	303/327 (93%)	295 (97%)	8 (3%)	0	100	100
All	All	612/654 (94%)	595 (97%)	15 (2%)	2 (0%)	47	16

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	283[A]	GLY
1	A	283[B]	GLY

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	260/265 (98%)	257 (99%)	3 (1%)	71	37
1	B	253/265 (96%)	252 (100%)	1 (0%)	91	76
All	All	513/530 (97%)	509 (99%)	4 (1%)	86	56

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	76	LEU
1	A	136[A]	ASP
1	A	136[B]	ASP
1	B	76	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	266	GLN

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 monosaccharides in this entry.

5.6   Ligand geometry ⓘ

2 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	LMR	A	401	-	8,8,8	1.14	1 (12%)	10,10,10	1.43	2 (20%)
3	MLT	B	401	-	8,8,8	1.08	0	10,10,10	1.29	2 (20%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	LMR	A	401	-	-	5/8/8/8	-
3	MLT	B	401	-	-	3/8/8/8	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	401	LMR	O1B-C1	-2.04	1.23	1.30

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	401	LMR	O1A-C1-C2	-3.09	116.50	122.54
3	B	401	MLT	O1-C1-C2	-2.69	117.29	122.54
2	A	401	LMR	O1B-C1-C2	2.56	118.34	112.72
3	B	401	MLT	O2-C1-C2	2.07	117.26	112.72

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	401	LMR	O1A-C1-C2-C3
2	A	401	LMR	O1B-C1-C2-C3
3	B	401	MLT	O1-C1-C2-C3
3	B	401	MLT	O2-C1-C2-C3
2	A	401	LMR	C2-C3-C4-O4A

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	401	LMR	3	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(Å²)	Q<0.9
1	A	291/327 (88%)	0.01	25 (8%)	10	10	9, 14, 40, 61	1 (0%)
1	B	290/327 (88%)	0.00	16 (5%)	25	23	9, 14, 38, 59	0
All	All	581/654 (88%)	0.01	41 (7%)	16	14	9, 14, 39, 61	1 (0%)

The worst 5 of 41 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	12	LEU	10.7
1	A	166	LEU	10.0
1	B	177	GLY	7.5
1	A	284[A]	LEU	7.1
1	A	254	ASN	6.1

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

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

6.3   Carbohydrates ⓘ

There are no monosaccharides 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. 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	LMR	A	401	9/9	0.84	0.20	28,36,41,42	0
3	MLT	B	401	9/9	0.93	0.14	25,56,81,81	0

6.5 Other polymers ⓘ

There are no such residues in this entry.