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

wwPDB X-ray Structure Validation Summary Report (i)

Aug 22, 2020 – 06:50 AM BST

PDB ID	:	4MWA
Title	:	1.85 Angstrom Crystal Structure of GCPE Protein from Bacillus anthracis
Authors	:	Minasov, G.; Wawrzak, Z.; Brunzelle, J.S.; Xu, X.; Cui, H.; Maltseva, N.; Bishop, B.; Kwon, K.; Savchenko, A.; Joachimiak, A.; Anderson, W.F.; Center for Structural Genomics of Infectious Diseases (CSGID)
Deposited on	:	2013-09-24
Resolution	:	1.85 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 (i) symbol.

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

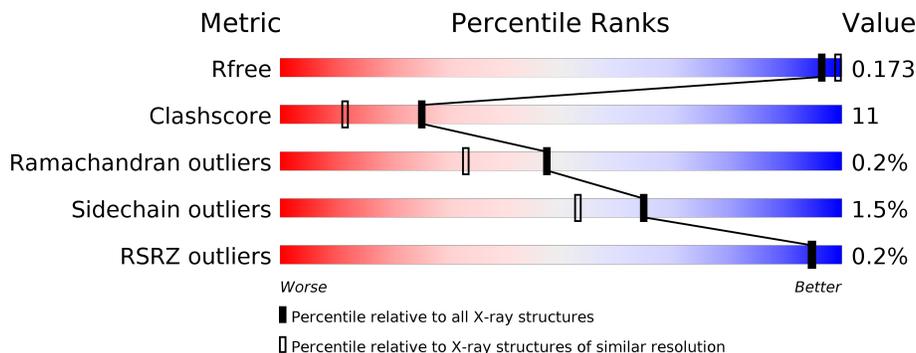|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

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 1.85 Å.

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

Metric	Percentile Ranks	Value
Rfree		0.173
Clashscore		11
Ramachandran outliers		0.2%
Sidechain outliers		1.5%
RSRZ outliers		0.2%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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 <=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	276	69% 25% • 5%
1	G	276	77% 15% • 7%
2	B	276	81% 12% • 6%
3	C	276	68% 24% • 7%
4	D	276	% 68% 23% 9%
4	F	276	73% 21% • 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	H	276	71% 22% 8%
5	E	276	79% 14% 6%

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
6	SO4	E	401	-	-	X	-
6	SO4	G	402	-	-	X	-

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 17202 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 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	261	Total 2083	C 1305	N 376	O 395	S 1	Se 6	0	10	0
1	G	258	Total 1965	C 1235	N 354	O 371	S 1	Se 4	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	EXPRESSION TAG	UNP Q81LV7
A	-1	ASN	-	EXPRESSION TAG	UNP Q81LV7
A	0	ALA	-	EXPRESSION TAG	UNP Q81LV7
A	1	MSE	-	EXPRESSION TAG	UNP Q81LV7
A	2	ASN	-	EXPRESSION TAG	UNP Q81LV7
A	3	GLU	-	EXPRESSION TAG	UNP Q81LV7
G	-2	SER	-	EXPRESSION TAG	UNP Q81LV7
G	-1	ASN	-	EXPRESSION TAG	UNP Q81LV7
G	0	ALA	-	EXPRESSION TAG	UNP Q81LV7
G	1	MSE	-	EXPRESSION TAG	UNP Q81LV7
G	2	ASN	-	EXPRESSION TAG	UNP Q81LV7
G	3	GLU	-	EXPRESSION TAG	UNP Q81LV7

- Molecule 2 is a protein called 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	259	Total 1990	C 1250	N 360	O 375	S 1	Se 4	0	2	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	SER	-	EXPRESSION TAG	UNP Q81LV7
B	-1	ASN	-	EXPRESSION TAG	UNP Q81LV7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	ALA	-	EXPRESSION TAG	UNP Q81LV7
B	1	MSE	-	EXPRESSION TAG	UNP Q81LV7
B	2	ASN	-	EXPRESSION TAG	UNP Q81LV7
B	3	GLU	-	EXPRESSION TAG	UNP Q81LV7

• Molecule 3 is a protein called 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	257	Total 1968	C 1236	N 354	O 373	S 1	Se 4	0	1	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-2	SER	-	EXPRESSION TAG	UNP Q81LV7
C	-1	ASN	-	EXPRESSION TAG	UNP Q81LV7
C	0	ALA	-	EXPRESSION TAG	UNP Q81LV7
C	1	MSE	-	EXPRESSION TAG	UNP Q81LV7
C	2	ASN	-	EXPRESSION TAG	UNP Q81LV7
C	3	GLU	-	EXPRESSION TAG	UNP Q81LV7

• Molecule 4 is a protein called 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
4	D	251	Total 1938	C 1222	N 344	O 367	S 1	Se 4	0	4	0
4	F	262	Total 2033	C 1274	N 371	O 383	S 1	Se 4	0	5	0
4	H	255	Total 1951	C 1226	N 352	O 368	S 1	Se 4	0	1	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-2	SER	-	EXPRESSION TAG	UNP Q81LV7
D	-1	ASN	-	EXPRESSION TAG	UNP Q81LV7
D	0	ALA	-	EXPRESSION TAG	UNP Q81LV7
D	1	MSE	-	EXPRESSION TAG	UNP Q81LV7
D	2	ASN	-	EXPRESSION TAG	UNP Q81LV7
D	3	GLU	-	EXPRESSION TAG	UNP Q81LV7
F	-2	SER	-	EXPRESSION TAG	UNP Q81LV7
F	-1	ASN	-	EXPRESSION TAG	UNP Q81LV7

Continued from previous page...

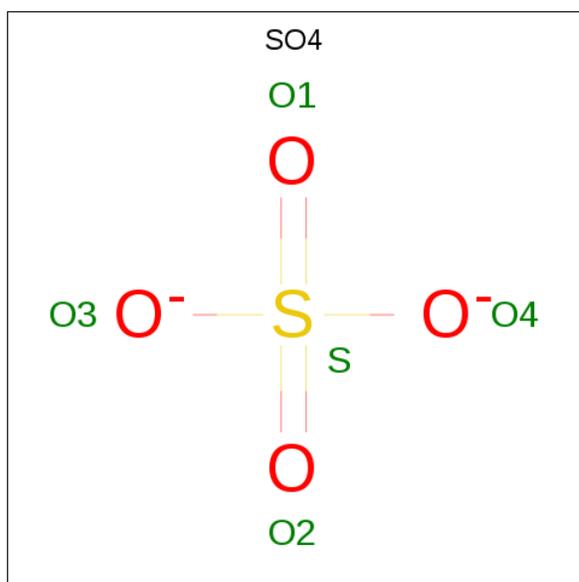| Chain	Residue	Modelled	Actual	Comment	Reference
F	0	ALA	-	EXPRESSION TAG	UNP Q81LV7
F	1	MSE	-	EXPRESSION TAG	UNP Q81LV7
F	2	ASN	-	EXPRESSION TAG	UNP Q81LV7
F	3	GLU	-	EXPRESSION TAG	UNP Q81LV7
H	-2	SER	-	EXPRESSION TAG	UNP Q81LV7
H	-1	ASN	-	EXPRESSION TAG	UNP Q81LV7
H	0	ALA	-	EXPRESSION TAG	UNP Q81LV7
H	1	MSE	-	EXPRESSION TAG	UNP Q81LV7
H	2	ASN	-	EXPRESSION TAG	UNP Q81LV7
H	3	GLU	-	EXPRESSION TAG	UNP Q81LV7

- Molecule 5 is a protein called 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
5	E	260	2003	1256	363	379	1	4	0	2	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	-2	SER	-	EXPRESSION TAG	UNP Q81LV7
E	-1	ASN	-	EXPRESSION TAG	UNP Q81LV7
E	0	ALA	-	EXPRESSION TAG	UNP Q81LV7
E	1	MSE	-	EXPRESSION TAG	UNP Q81LV7
E	2	ASN	-	EXPRESSION TAG	UNP Q81LV7
E	3	GLU	-	EXPRESSION TAG	UNP Q81LV7

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

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

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	B	2	Total 3	Cl 3	0	1
7	A	1	Total 1	Cl 1	0	0
7	C	3	Total 3	Cl 3	0	0
7	F	2	Total 3	Cl 3	0	1

- Molecule 8 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	146	Total 147	O 147	0	2
8	B	139	Total 142	O 142	0	3
8	C	147	Total 149	O 149	0	2
8	D	116	Total 119	O 119	0	3
8	E	166	Total 172	O 172	0	6
8	F	148	Total 154	O 154	0	6
8	G	158	Total 162	O 162	0	4
8	H	148	Total 151	O 151	0	3

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.
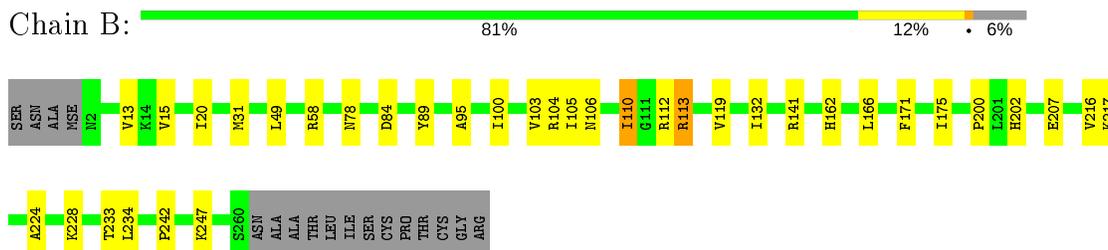
- Molecule 1: 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase

Chain A:

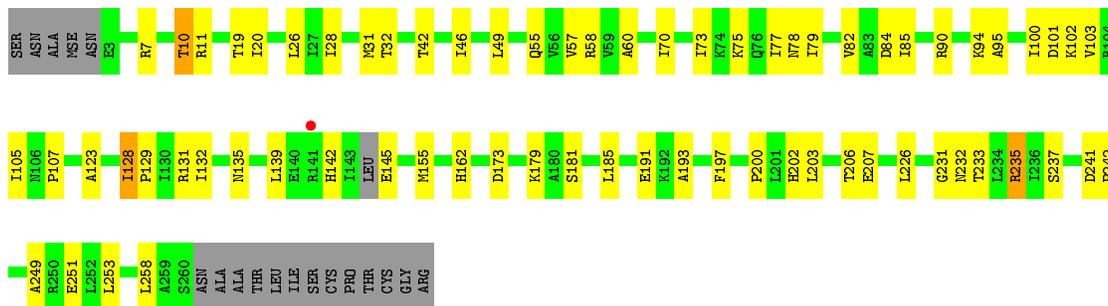- Molecule 1: 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase

Chain G:
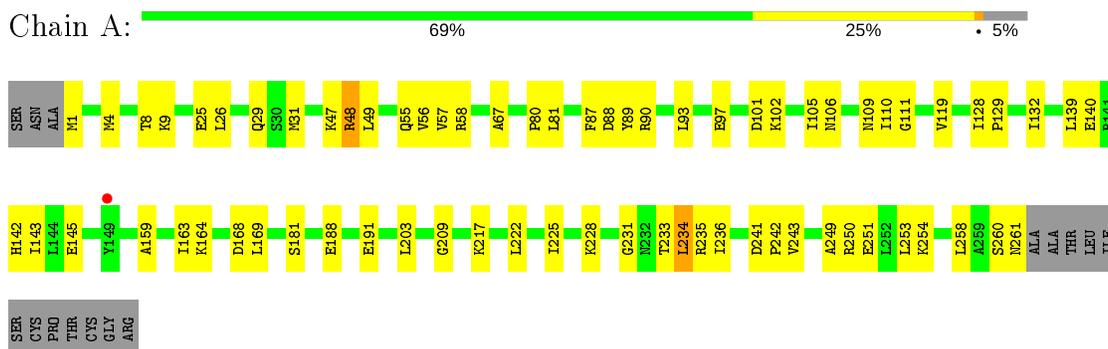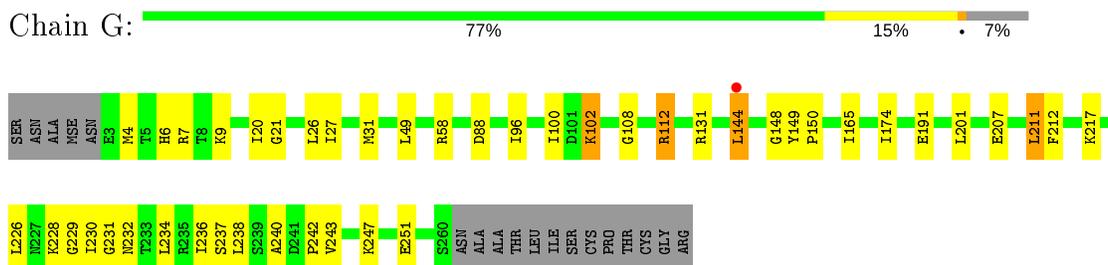
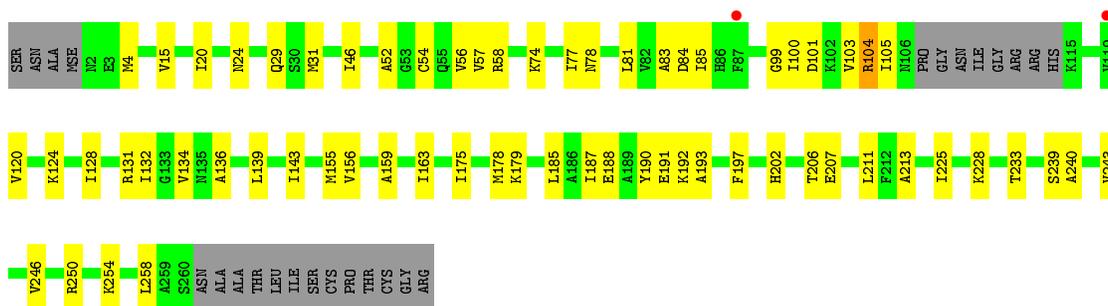- Molecule 2: 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase

Chain B:

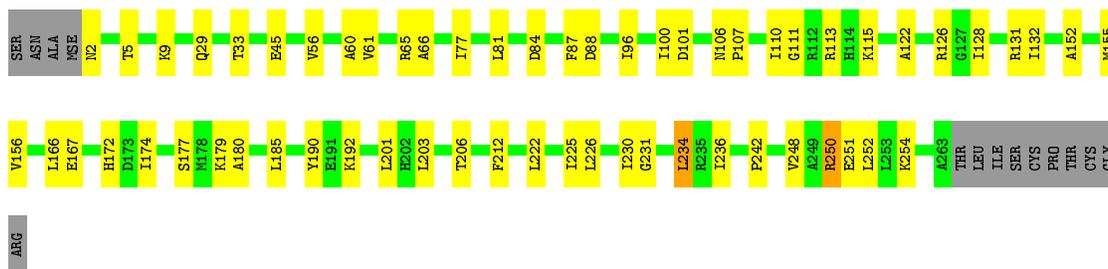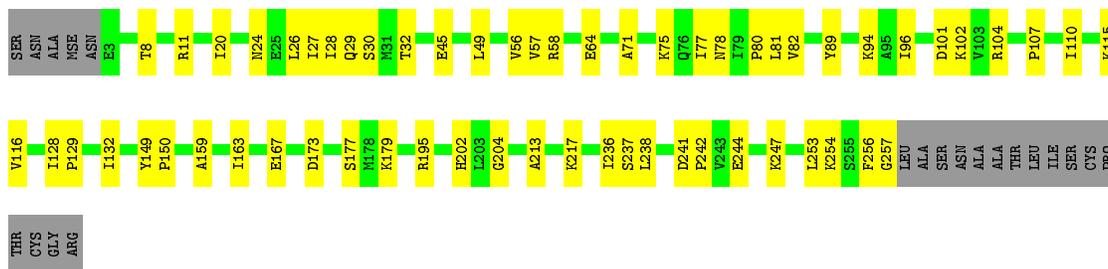- Molecule 3: 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase

Chain C:

SER ASN ALA MSE ASN E3 R7 T10 R11 T19 I20 L26 I27 I28 M31 T32 T42 I46 L49 Q55 V56 V57 R58 V59 A60 I70 I73 K74 K75 Q76 I77 N78 I79 V82 A83 D84 I85 R90 K94 A95 I100 D101 K102 V103 R104

I105 N106 P107 A123 I128 P129 I130 R131 I132 N135 L139 E140 R141 H142 I143 LEU E145 M155 H162 D173 K179 A180 S181 L185 E191 K192 A193 F197 P200 L201 H202 L203 I226 G231 N232 T233 L234 R235 I236 S237 D241 P242

A249 R250 E251 L252 L253 L258 A259 S260 ASN ALA ALA THR LEU ILE SER CYS PRO THR CYS GLY ARG

• Molecule 4: 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase

Chain D:

SER ASN ALA MSE N2 E3 M4 V15 I20 N24 Q29 S30 M31 I46 A52 G53 C54 Q55 V56 V57 R58 K74 I77 N78 L81 V82 A83 D84 I85 H86 F87 G99 I100 D101 K102 V103 R104 I105 N106 PRO GLY ASN ILE GLY ARG ARG HIS K115 V119

V120 K124 I128 R131 I132 G133 V134 N135 A136 L139 I143 M155 V156 A159 I163 I175 M178 K179 L185 A186 I187 E188 A189 Y190 E191 K192 A193 F197 H202 T206 E207 L211 F212 A213 I225 K228 T233 S239 A240 V243

V246 R250 K254 L258 A259 S260 ASN ALA ALA THR LEU ILE SER CYS PRO THR CYS GLY ARG

• Molecule 4: 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase

Chain F:

SER ASN ALA MSE N2 T5 K9 Q29 T33 E45 V56 A60 V61 R65 A66 I77 L81 D84 F87 D88 I96 I100 D101 N106 P107 I110 G111 R112 R113 H114 K115 A122 R126 G127 I128 R131 I132 A152 M155

V156 L166 E167 H172 D173 I174 S177 M178 K179 A180 L185 Y190 E191 K192 L201 H202 L203 T206 F212 L222 I225 L226 I230 G231 L234 R235 I236 P242 V248 A249 R250 E251 L252 L253 K254 A263 THR LEU ILE SER CYS PRO THR CYS GLY

ARG

• Molecule 4: 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase

Chain H:

SER ASN ALA MSE ASN E3 T8 R11 I20 N24 E25 L26 I27 I28 Q29 S30 M31 T32 E45 L49 V56 V57 R58 E64 A71 K75 Q76 I77 N78 I79 P80 L81 V82 Y89 K94 A95 I96 D101 K102 V103 R104 P107 I110 K115

V116 I128 P129 I132 Y149 P150 A159 I163 E167 D173 S177 M178 K179 R195 H202 L203 G204 A213 K217 I236 S237 L238 D241 P242 V243 E244 K247 L253 K254 S255 F256 G257 LEU ALA SER ASN ALA ALA THR LEU ILE SER CYS PRO

THR CYS GLY ARG

• Molecule 5: 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase

Chain E:

79%   14%   • 6%

SER ASN ALA MSE N2 L26 I27 I28 M31 T32 E45 I46 L49 Q55 V56 V57 R58 K94 K102 V103 R104 I105 N106 R112 R126 R131 R141 H142 I143 L144 E145 S158 A159 L160 H161 H162 I163 K164 E167 H172 D173 I174 K179 V183

P200 L203 L222 L234 R235 I236 A249 R250 E251 L258 N261 ALA ALA THR LEU ILE SER CYS PRO THR CYS GLY ARG

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 3	Depositor
Cell constants a, b, c, α, β, γ	162.68Å 162.68Å 76.48Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.77 – 1.85 29.77 – 1.85	Depositor EDS
% Data completeness (in resolution range)	99.4 (29.77-1.85) 99.4 (29.77-1.85)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.66 (at 1.85Å)	Xtriage
Refinement program	REFMAC 5.8.0046	Depositor
R, R_{free}	0.125 , 0.169 0.133 , 0.173	Depositor DCC
R_{free} test set	9677 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	19.8	Xtriage
Anisotropy	0.259	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 19.4	EDS
L-test for twinning[2]	$< |L| > = 0.41, < L^2 > = 0.23$	Xtriage
Estimated twinning fraction	0.468 for -h,-k,l 0.479 for h,-h-k,-l 0.468 for -k,-h,-l	Xtriage
Reported twinning fraction	0.241 for H, K, L 0.260 for -K, -H, -L 0.268 for -h,-k,l 0.232 for K, H, -L	Depositor
Outliers	0 of 192077 reflections	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	17202	wwPDB-VP
Average B, all atoms (Å^2)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.77% 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: CL, KCX, SO4

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	0.52	2/2093 (0.1%)	0.74	0/2817
1	G	0.53	0/1974	0.78	0/2660
2	B	0.45	0/2000	0.72	0/2696
3	C	0.46	0/1963	0.71	0/2643
4	D	0.45	0/1957	0.72	0/2636
4	F	0.50	0/2055	0.76	0/2770
4	H	0.52	0/1973	0.74	0/2659
5	E	0.52	0/2000	0.75	0/2694
All	All	0.50	2/16015 (0.0%)	0.74	0/21575

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	48[A]	ARG	CD-NE	-5.32	1.37	1.46
1	A	48[B]	ARG	CD-NE	-5.32	1.37	1.46

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

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	2083	0	2132	58	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	G	1965	0	2026	32	0
2	B	1990	0	2047	26	0
3	C	1968	0	2022	58	0
4	D	1938	0	2007	49	0
4	F	2033	0	2095	47	0
4	H	1951	0	2011	53	0
5	E	2003	0	2054	33	0
6	A	5	0	0	1	0
6	B	5	0	0	1	0
6	C	10	0	0	0	0
6	D	10	0	0	0	0
6	E	10	0	0	2	0
6	F	10	0	0	1	0
6	G	10	0	0	2	0
6	H	5	0	0	1	0
7	A	1	0	0	1	0
7	B	3	0	0	0	0
7	C	3	0	0	0	0
7	F	3	0	0	1	0
8	A	147	0	0	4	0
8	B	142	0	0	4	0
8	C	149	0	0	4	0
8	D	119	0	0	2	0
8	E	172	0	0	7	0
8	F	154	0	0	3	0
8	G	162	0	0	3	0
8	H	151	0	0	5	0
All	All	17202	0	16394	348	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 11.

The worst 5 of 348 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)
4:F:155:MSE:SE	4:F:185:LEU:HD21	2.10	1.02
4:H:236:ILE:HG22	4:H:238:LEU:CD1	1.95	0.97
4:H:29:GLN:HE22	4:H:58:ARG:NH1	1.68	0.91
3:C:241:ASP:HB3	8:C:554:HOH:O	1.73	0.87
4:H:28:ILE:HD11	4:H:253:LEU:HD12	1.61	0.82

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	268/276 (97%)	260 (97%)	7 (3%)	1 (0%)	34	19
1	G	255/276 (92%)	246 (96%)	9 (4%)	0	100	100
2	B	258/276 (94%)	245 (95%)	11 (4%)	2 (1%)	19	7
3	C	252/276 (91%)	242 (96%)	10 (4%)	0	100	100
4	D	251/276 (91%)	241 (96%)	10 (4%)	0	100	100
4	F	265/276 (96%)	260 (98%)	4 (2%)	1 (0%)	34	19
4	H	254/276 (92%)	247 (97%)	7 (3%)	0	100	100
5	E	258/276 (94%)	251 (97%)	7 (3%)	0	100	100
All	All	2061/2208 (93%)	1992 (97%)	65 (3%)	4 (0%)	47	33

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	110	ILE
1	A	87	PHE
2	B	112	ARG
4	F	242	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	219/215 (102%)	215 (98%)	4 (2%)	59	45
1	G	206/215 (96%)	201 (98%)	5 (2%)	49	33
2	B	209/215 (97%)	206 (99%)	3 (1%)	67	55
3	C	205/214 (96%)	201 (98%)	4 (2%)	55	40
4	D	206/216 (95%)	205 (100%)	1 (0%)	88	86
4	F	214/216 (99%)	211 (99%)	3 (1%)	67	55
4	H	206/216 (95%)	205 (100%)	1 (0%)	88	86
5	E	209/214 (98%)	204 (98%)	5 (2%)	49	33
All	All	1674/1721 (97%)	1648 (98%)	26 (2%)	65	49

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	D	104	ARG
5	E	104	ARG
1	G	234	LEU
5	E	94[A]	LYS
5	E	94[B]	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	162	HIS
5	E	2	ASN
4	F	37	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)

7 non-standard protein/DNA/RNA residues 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	#$|Z| > 2$
1	KCX	A	102	1	7,11,12	0.54	0	4,12,14	0.68	0
5	KCX	E	35	5	7,11,12	0.88	0	4,12,14	1.21	0
5	KCX	E	102	5	7,11,12	0.82	0	4,12,14	0.40	0
1	KCX	G	102	1	7,11,12	0.64	0	4,12,14	1.10	1 (25%)
3	KCX	C	247	3	7,11,12	0.52	0	4,12,14	0.67	0
3	KCX	C	102	3	7,11,12	0.59	0	4,12,14	0.30	0
2	KCX	B	217	2	7,11,12	0.50	0	4,12,14	0.66	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
1	KCX	A	102	1	-	0/7/10/12	-
5	KCX	E	35	5	-	1/7/10/12	-
5	KCX	E	102	5	-	1/7/10/12	-
1	KCX	G	102	1	-	1/7/10/12	-
3	KCX	C	247	3	-	1/7/10/12	-
3	KCX	C	102	3	-	1/7/10/12	-
2	KCX	B	217	2	-	1/7/10/12	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	G	102	KCX	CE-NZ-CX	2.08	126.46	122.95

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	E	102	KCX	O-C-CA-CB
3	C	102	KCX	O-C-CA-CB
2	B	217	KCX	CG-CD-CE-NZ
5	E	35	KCX	CG-CD-CE-NZ
1	G	102	KCX	CD-CE-NZ-CX

There are no ring outliers.

5 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	102	KCX	1	0
5	E	102	KCX	1	0
1	G	102	KCX	1	0
3	C	102	KCX	2	0
2	B	217	KCX	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 23 ligands modelled in this entry, 10 are monoatomic - leaving 13 for Mogul analysis.

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	SO4	F	402	-	4,4,4	0.35	0	6,6,6	0.22	0
6	SO4	D	401	-	4,4,4	0.34	0	6,6,6	0.19	0
6	SO4	C	402	-	4,4,4	0.41	0	6,6,6	0.17	0
6	SO4	E	402	-	4,4,4	0.30	0	6,6,6	0.15	0
6	SO4	G	401	-	4,4,4	0.24	0	6,6,6	0.40	0
6	SO4	E	401	-	4,4,4	0.30	0	6,6,6	0.42	0
6	SO4	F	401	-	4,4,4	0.29	0	6,6,6	0.29	0
6	SO4	C	401	-	4,4,4	0.26	0	6,6,6	0.25	0
6	SO4	G	402	-	4,4,4	0.34	0	6,6,6	0.09	0
6	SO4	A	401	-	4,4,4	0.31	0	6,6,6	0.19	0
6	SO4	B	401	-	4,4,4	0.33	0	6,6,6	0.26	0
6	SO4	D	402	-	4,4,4	0.30	0	6,6,6	0.05	0
6	SO4	H	401	-	4,4,4	0.40	0	6,6,6	0.17	0

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.

6 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	F	402	SO4	1	0
6	E	401	SO4	2	0
6	G	402	SO4	2	0
6	A	401	SO4	1	0
6	B	401	SO4	1	0
6	H	401	SO4	1	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, 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	255/276 (92%)	-0.43	1 (0%)	92	92	12, 21, 33, 42	0
1	G	253/276 (91%)	-0.53	1 (0%)	92	92	11, 17, 26, 32	0
2	B	254/276 (92%)	-0.37	0	100	100	15, 24, 40, 49	0
3	C	251/276 (90%)	-0.40	1 (0%)	92	92	14, 22, 32, 54	0
4	D	247/276 (89%)	-0.28	2 (0%)	86	86	13, 24, 41, 51	0
4	F	258/276 (93%)	-0.45	0	100	100	12, 19, 29, 43	0
4	H	251/276 (90%)	-0.48	0	100	100	13, 19, 29, 37	0
5	E	254/276 (92%)	-0.53	0	100	100	10, 17, 24, 30	0
All	All	2023/2208 (91%)	-0.43	5 (0%)	95	94	10, 20, 33, 54	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	87	PHE	3.1
1	A	149	TYR	2.6
4	D	119	VAL	2.3
3	C	141	ARG	2.2
1	G	144	LEU	2.1

6.2 Non-standard residues in protein, DNA, RNA chains (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
1	KCX	G	102	12/13	0.94	0.13	16,20,28,32	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	KCX	B	217	12/13	0.94	0.09	18,24,36,37	0
1	KCX	A	102	12/13	0.95	0.08	15,16,18,18	0
3	KCX	C	247	12/13	0.95	0.09	16,23,32,35	0
3	KCX	C	102	12/13	0.95	0.08	17,18,19,20	0
5	KCX	E	35	12/13	0.95	0.08	15,17,27,28	0
5	KCX	E	102	12/13	0.96	0.07	10,11,13,14	0

6.3 Carbohydrates (i)

There are no monosaccharides 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
7	CL	C	405	1/1	0.89	0.08	50,50,50,50	0
6	SO4	G	402	5/5	0.94	0.14	34,43,45,46	0
7	CL	B	402[A]	1/1	0.95	0.09	32,32,32,32	1
7	CL	A	402	1/1	0.95	0.14	53,53,53,53	0
7	CL	B	402[B]	1/1	0.95	0.09	21,21,21,21	1
7	CL	C	404	1/1	0.96	0.09	44,44,44,44	0
6	SO4	E	402	5/5	0.96	0.11	38,42,43,45	0
6	SO4	B	401	5/5	0.97	0.07	29,31,35,35	0
7	CL	F	404[A]	1/1	0.97	0.14	31,31,31,31	1
7	CL	F	404[B]	1/1	0.97	0.14	23,23,23,23	1
6	SO4	C	402	5/5	0.97	0.06	29,31,35,38	0
6	SO4	F	402	5/5	0.97	0.06	37,37,39,39	0
6	SO4	D	401	5/5	0.97	0.08	32,34,37,38	0
6	SO4	F	401	5/5	0.97	0.10	33,36,41,42	0
6	SO4	D	402	5/5	0.98	0.07	39,39,39,40	5
7	CL	C	403	1/1	0.98	0.12	37,37,37,37	0
6	SO4	H	401	5/5	0.98	0.07	25,30,31,36	0
6	SO4	A	401	5/5	0.99	0.07	18,21,23,23	0
6	SO4	C	401	5/5	0.99	0.06	25,26,27,27	0
7	CL	F	403	1/1	0.99	0.10	30,30,30,30	0
6	SO4	E	401	5/5	0.99	0.06	18,19,21,21	0
6	SO4	G	401	5/5	0.99	0.07	16,19,20,21	0

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
7	CL	B	403	1/1	0.99	0.15	42,42,42,42	1

6.5 Other polymers (i)

There are no such residues in this entry.