wwPDB X-ray Structure Validation Summary Report (i)

Nov 26, 2023 – 10:28 PM JST

PDB ID	:	8KBW
Title	:	The crystal structure of syn-copalyl diphosphate synthase from Oryza sativa
Authors	:	Ma, X.L.; Xu, H.F.; Jiang, T.
Deposited on	:	2023-08-04
Resolution	:	3.49 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

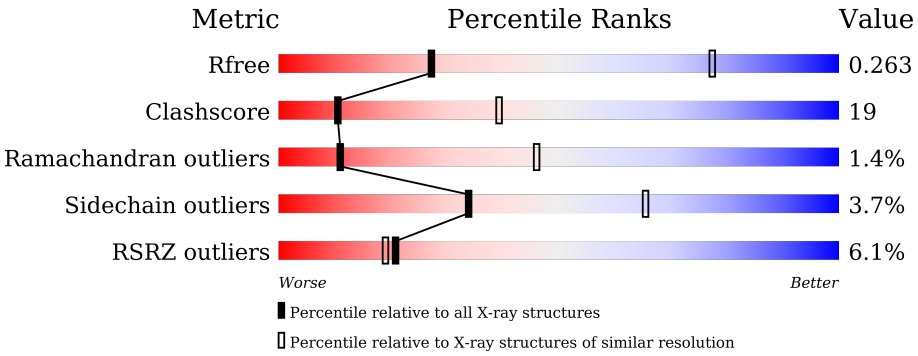
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.49 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	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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	775	51% 35% • 11%
1	B	775	5% 54% 31% • 14%
1	C	775	9% 51% 30% • 16%
1	D	775	8% 54% 33% • 11%
1	E	775	57% 30% • 11%
1	F	775	9% 52% 33% • 13%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 32467 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 Syn-copalyl diphosphate synthase, chloroplastic.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	668	Total	C	N	O	S	0	0	0
			5341	3391	916	1000	34			
1	C	648	Total	C	N	O	S	0	1	0
			5186	3297	892	962	35			
1	E	688	Total	C	N	O	S	0	0	0
			5504	3492	944	1033	35			
1	D	688	Total	C	N	O	S	0	0	0
			5504	3492	944	1033	35			
1	F	677	Total	C	N	O	S	0	0	0
			5421	3439	928	1019	35			
1	A	689	Total	C	N	O	S	0	0	0
			5511	3496	945	1035	35			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	768	GLU	-	expression tag	UNP Q0JF02
B	769	PHE	-	expression tag	UNP Q0JF02
B	770	HIS	-	expression tag	UNP Q0JF02
B	771	HIS	-	expression tag	UNP Q0JF02
B	772	HIS	-	expression tag	UNP Q0JF02
B	773	HIS	-	expression tag	UNP Q0JF02
B	774	HIS	-	expression tag	UNP Q0JF02
B	775	HIS	-	expression tag	UNP Q0JF02
C	768	GLU	-	expression tag	UNP Q0JF02
C	769	PHE	-	expression tag	UNP Q0JF02
C	770	HIS	-	expression tag	UNP Q0JF02
C	771	HIS	-	expression tag	UNP Q0JF02
C	772	HIS	-	expression tag	UNP Q0JF02
C	773	HIS	-	expression tag	UNP Q0JF02
C	774	HIS	-	expression tag	UNP Q0JF02
C	775	HIS	-	expression tag	UNP Q0JF02
E	768	GLU	-	expression tag	UNP Q0JF02

Continued on next page...

Continued from previous page...

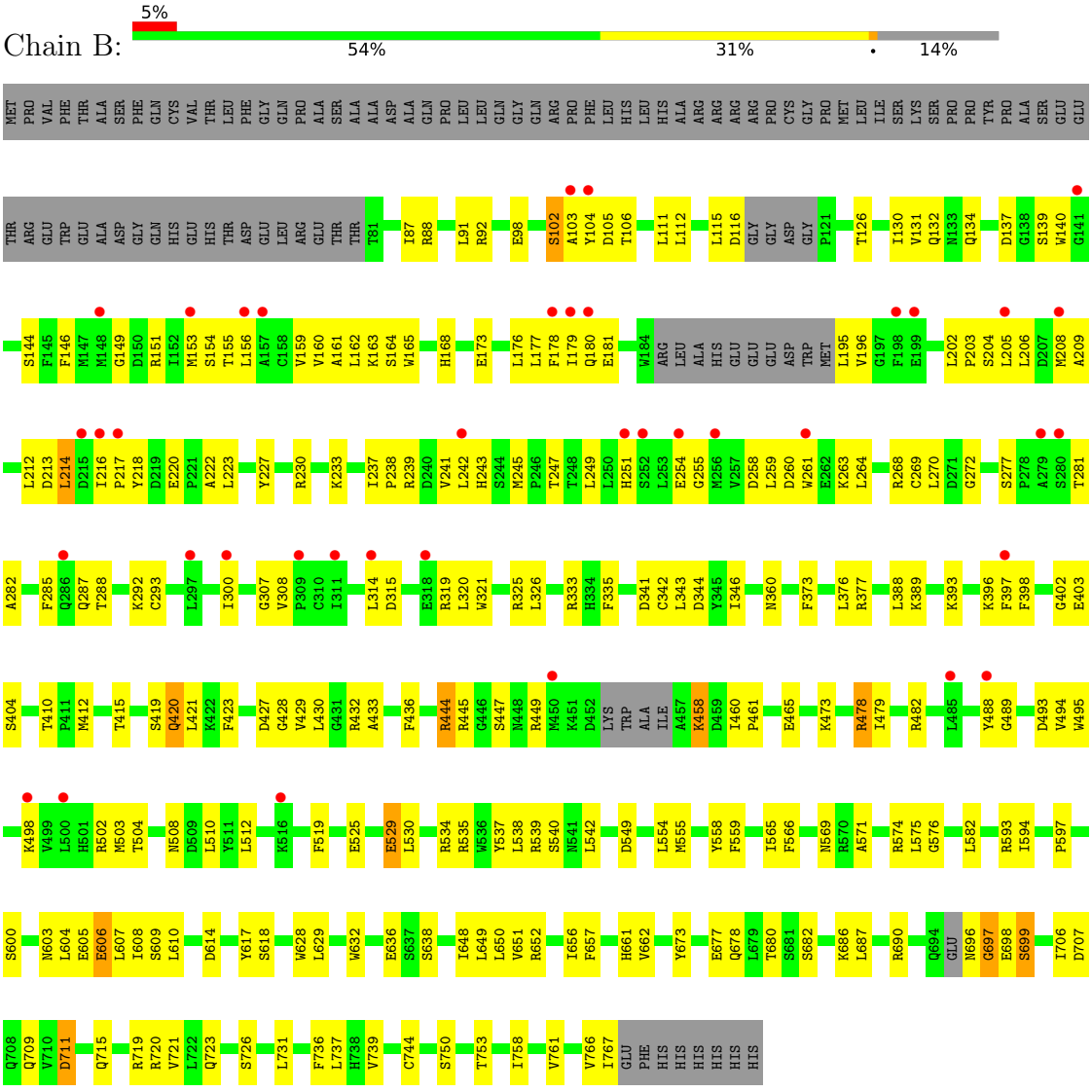
Chain	Residue	Modelled	Actual	Comment	Reference
E	769	PHE	-	expression tag	UNP Q0JF02
E	770	HIS	-	expression tag	UNP Q0JF02
E	771	HIS	-	expression tag	UNP Q0JF02
E	772	HIS	-	expression tag	UNP Q0JF02
E	773	HIS	-	expression tag	UNP Q0JF02
E	774	HIS	-	expression tag	UNP Q0JF02
E	775	HIS	-	expression tag	UNP Q0JF02
D	768	GLU	-	expression tag	UNP Q0JF02
D	769	PHE	-	expression tag	UNP Q0JF02
D	770	HIS	-	expression tag	UNP Q0JF02
D	771	HIS	-	expression tag	UNP Q0JF02
D	772	HIS	-	expression tag	UNP Q0JF02
D	773	HIS	-	expression tag	UNP Q0JF02
D	774	HIS	-	expression tag	UNP Q0JF02
D	775	HIS	-	expression tag	UNP Q0JF02
F	768	GLU	-	expression tag	UNP Q0JF02
F	769	PHE	-	expression tag	UNP Q0JF02
F	770	HIS	-	expression tag	UNP Q0JF02
F	771	HIS	-	expression tag	UNP Q0JF02
F	772	HIS	-	expression tag	UNP Q0JF02
F	773	HIS	-	expression tag	UNP Q0JF02
F	774	HIS	-	expression tag	UNP Q0JF02
F	775	HIS	-	expression tag	UNP Q0JF02
A	768	GLU	-	expression tag	UNP Q0JF02
A	769	PHE	-	expression tag	UNP Q0JF02
A	770	HIS	-	expression tag	UNP Q0JF02
A	771	HIS	-	expression tag	UNP Q0JF02
A	772	HIS	-	expression tag	UNP Q0JF02
A	773	HIS	-	expression tag	UNP Q0JF02
A	774	HIS	-	expression tag	UNP Q0JF02
A	775	HIS	-	expression tag	UNP Q0JF02

3 Residue-property plots (i)

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.

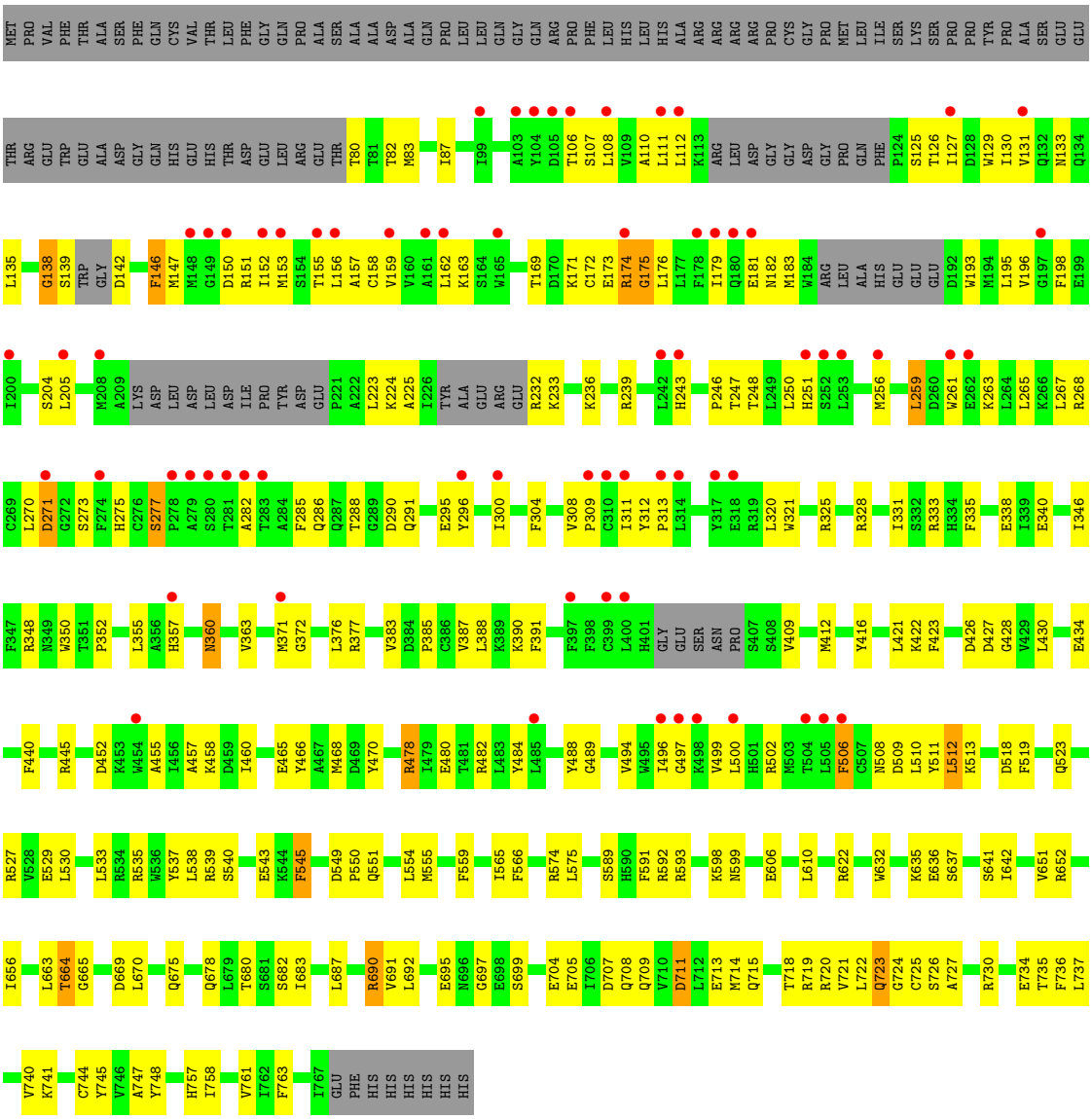
• Molecule 1: Syn-copalyl diphosphate synthase, chloroplastic

Chain B:



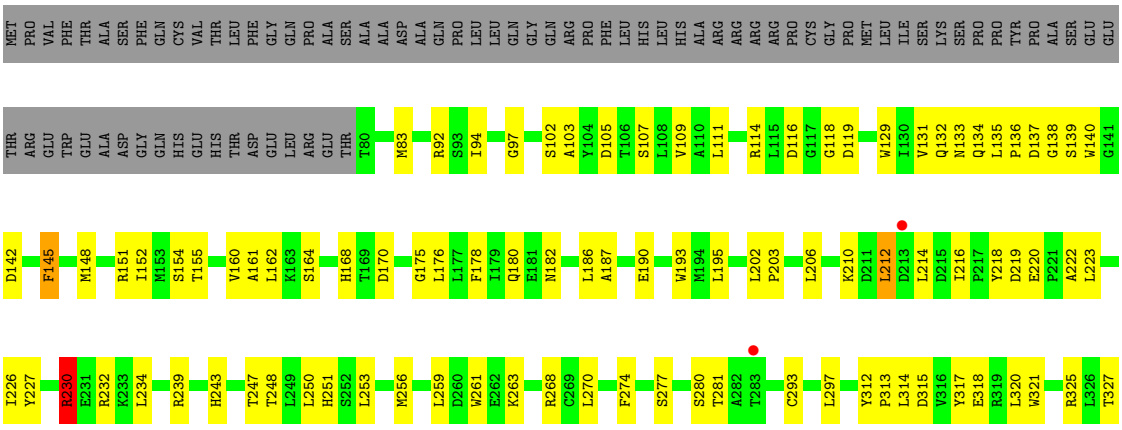
• Molecule 1: Syn-copalyl diphosphate synthase, chloroplastic

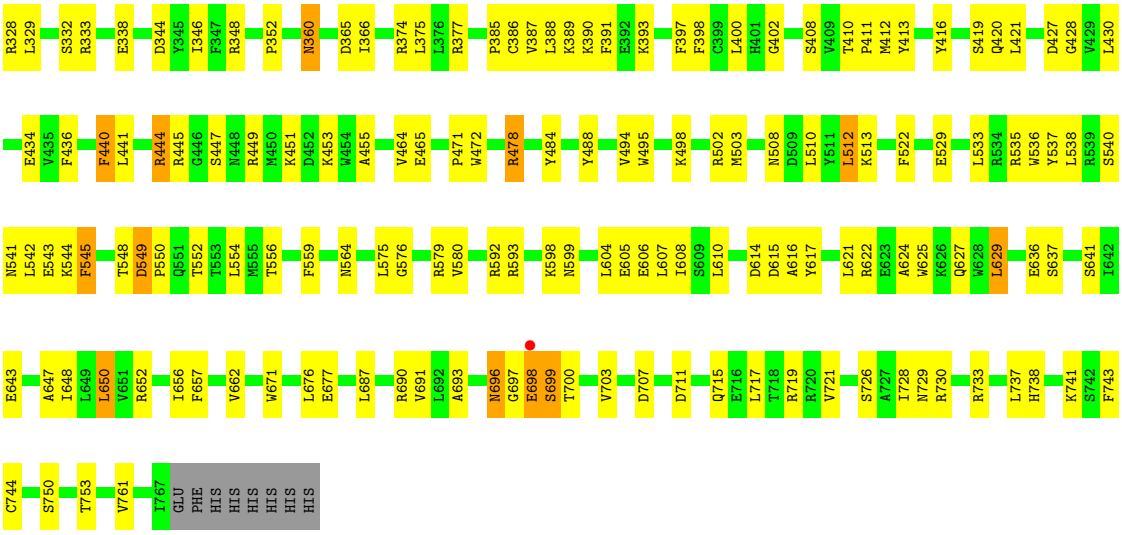
Chain C:



• Molecule 1: Syn-copalyl diphosphate synthase, chloroplastic

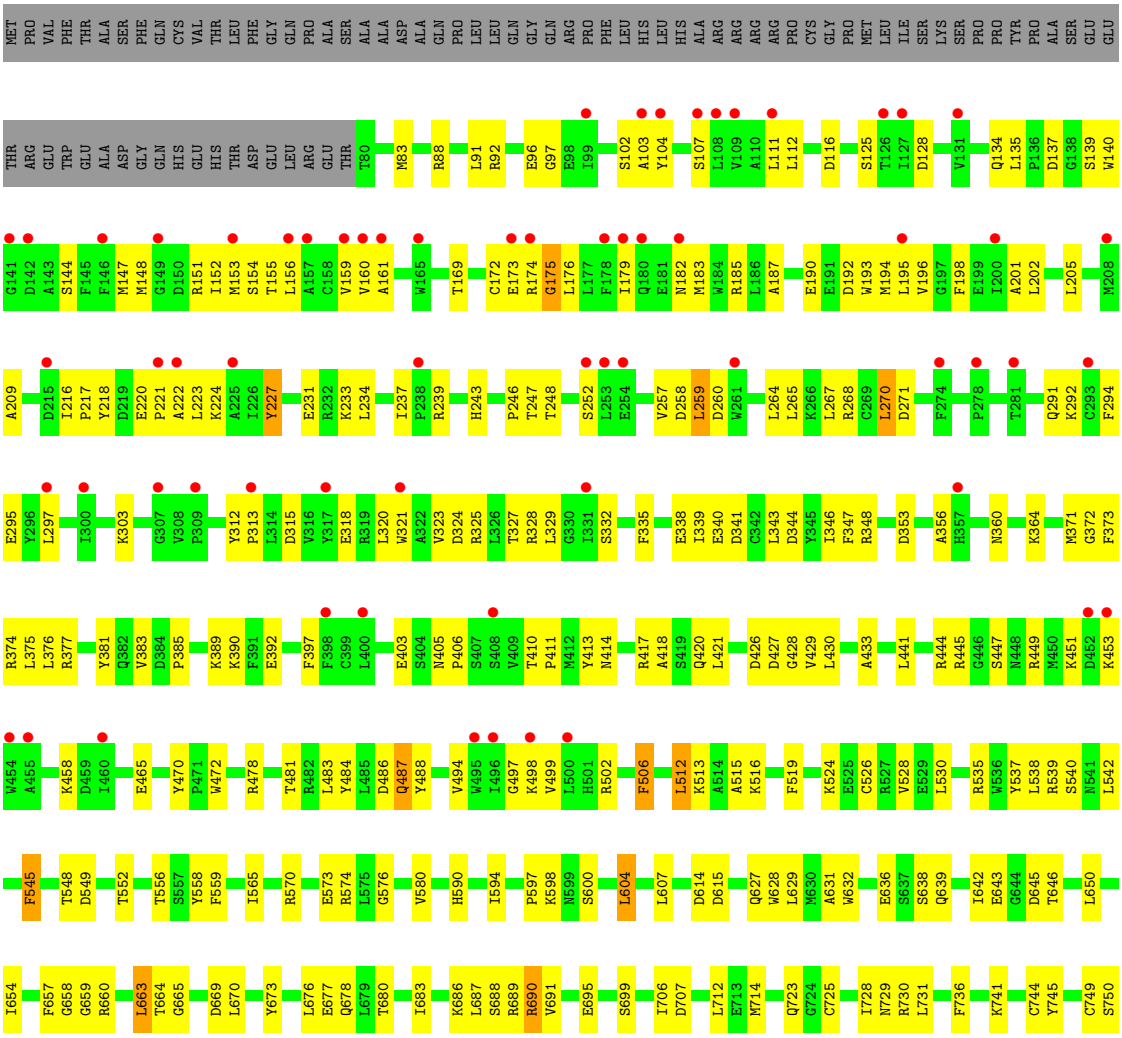
Chain E:

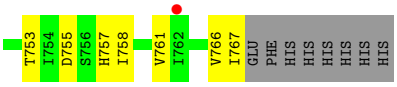




• Molecule 1: Syn-copalyl diphosphate synthase, chloroplastic

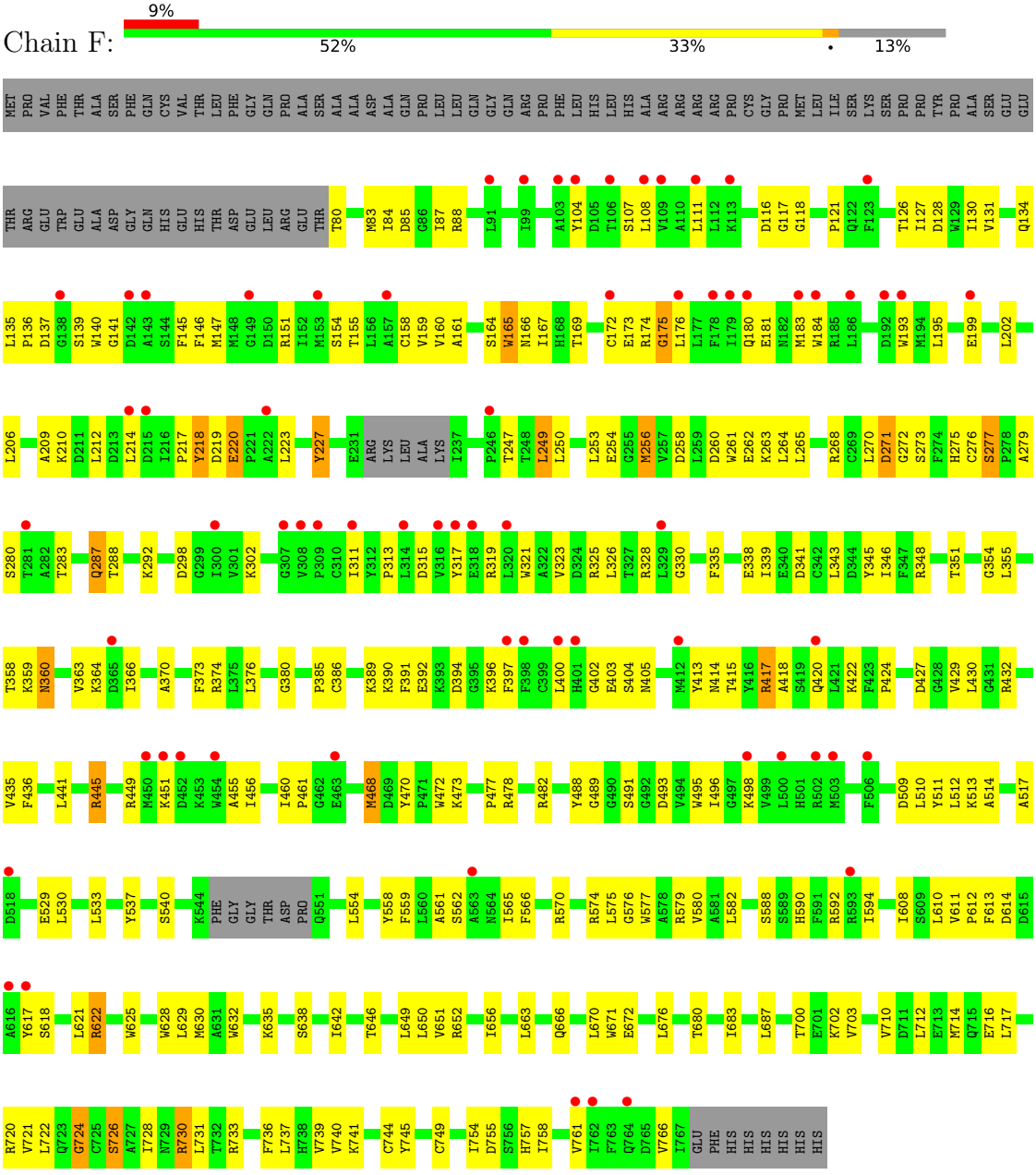
Chain D:





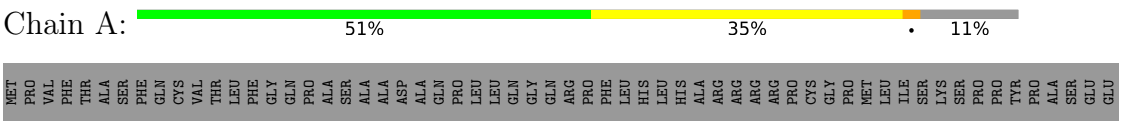
• Molecule 1: Syn-copalyl diphosphate synthase, chloroplastic

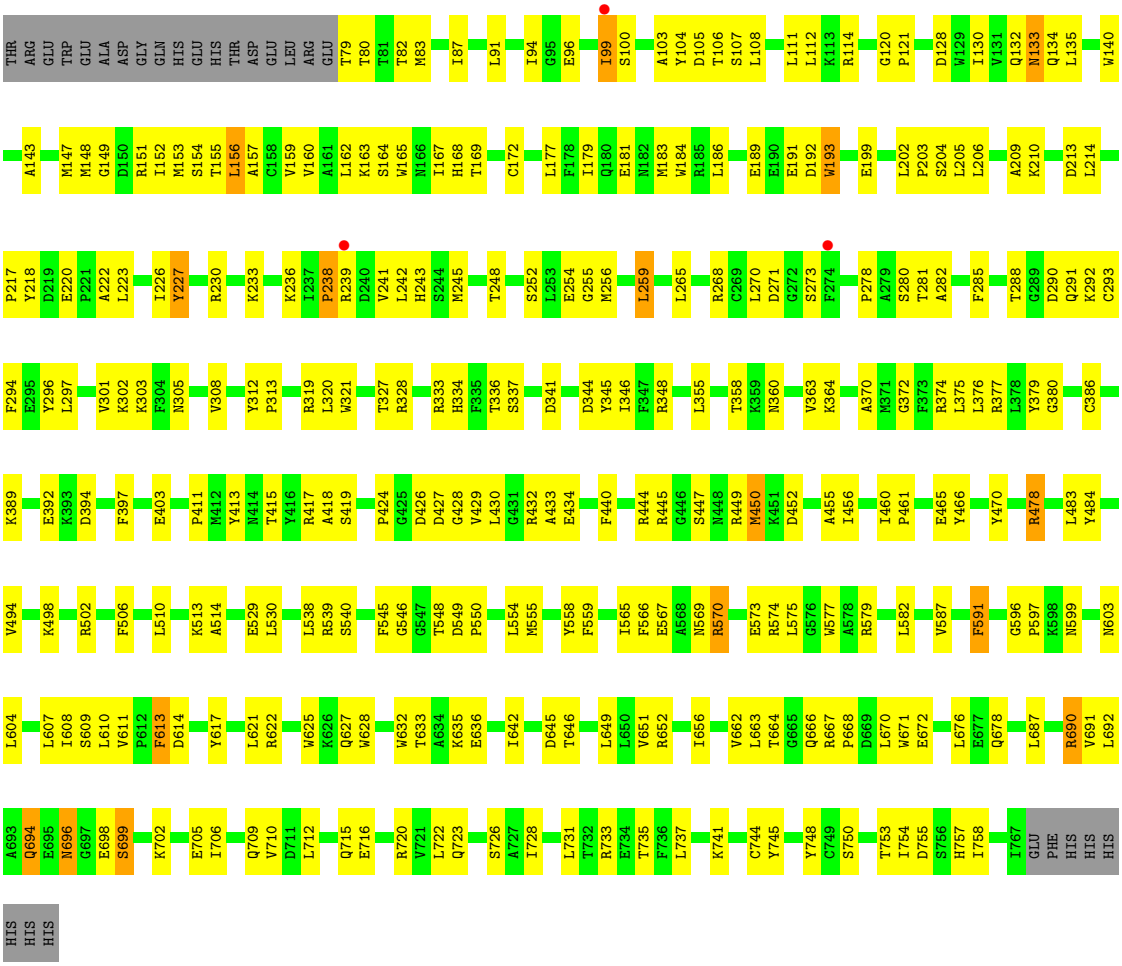
Chain F:



• Molecule 1: Syn-copalyl diphosphate synthase, chloroplastic

Chain A:




THR ARG GLU TRP GLU ALA ASP GLY GLN HIS GLU HIS THR ASP GLU LEU ARG GLU T79 T80 T81 T82 M83 I87 L91 I94 G95 E96 I99 S100 A103 Y104 D105 T106 S107 L108 L111 L112 K113 R114 G120 P121 D128 W129 I130 V131 Q132 N133 Q134 L135 W140
A143 M147 M148 G149 D150 R151 I152 M153 S154 T155 L156 A157 C158 V159 V160 A161 L162 K163 S164 W165 N166 I167 H168 T169 C172 L177 F178 I179 Q180 E181 N182 M183 W184 R185 L186 E189 E190 E191 D192 W193 E199 L202 P203 S204 L205 L206 A209 K210 D213 L214
P217 Y218 D219 E220 P221 A222 L223 I226 Y227 R230 K233 K236 I237 P238 R239 D240 V241 L242 H243 S244 M245 T248 S252 L253 E254 G255 M256 L259 L265 R268 C269 L270 D271 G272 S273 F274 P278 A279 S280 T281 A282 F285 T288 G289 D290 Q291 K292 C293
F294 E295 Y296 L297 V301 K302 K303 F304 N305 V308 Y312 P313 R319 L320 W321 T327 R328 R333 H334 F335 T336 S337 D341 D344 Y345 I346 F347 R348 L355 T358 K359 N360 V363 K364 A370 M371 G372 F373 R374 L375 L376 R377 L378 Y379 G380 C386
K389 E392 K393 D394 F397 E403 P411 W412 Y413 N414 T415 Y416 R417 A418 S419 P424 G425 D426 D427 G428 V429 L430 G431 R432 A433 E434 F440 R444 R445 G446 S447 N448 R449 M450 K451 D452 A455 I456 I460 P461 E465 Y466 Y470 R478 L483 Y484
V494 K498 R502 P506 L510 K513 A514 E529 L530 L538 R539 S540 F545 G546 G547 T548 D549 P550 L554 M555 Y558 F559 I565 F566 E567 A568 N569 R570 E573 R574 L575 G576 W577 A578 R579 L582 V587 P591 G596 P597 K598 N599 N603
L604 L607 I608 S609 L610 V611 P612 F613 D614 Y617 L621 R622 W625 K626 Q627 W628 W632 T633 A634 K635 E636 I642 D645 T646 L649 L650 V651 R652 I656 V662 L663 T664 G665 Q666 R667 P668 D669 L670 W671 E672 L676 E677 Q678 L687 R690 V691 L692
A693 Q694 E695 N696 G697 E698 S699 K702 E705 I706 Q709 V710 D711 L712 Q715 E716 R720 V721 L722 Q723 S726 A727 I728 L731 T732 R733 E734 T735 F736 L737 K741 C744 Y745 Y748 C749 S750 T753 I754 D755 S756 H757 I758 I767 GLU PHE HIS HIS HIS
HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	129.70Å 174.34Å 296.55Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	97.62 – 3.49 98.19 – 3.49	Depositor EDS
% Data completeness (in resolution range)	96.5 (97.62-3.49) 95.4 (98.19-3.49)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.63 (at 3.49Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.201 , 0.263 0.201 , 0.263	Depositor DCC
R_{free} test set	2000 reflections (2.33%)	wwPDB-VP
Wilson B-factor (Å^2)	86.5	Xtriage
Anisotropy	0.366	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 56.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	32467	wwPDB-VP
Average B, all atoms (Å^2)	81.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.38% 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

5.1 Standard geometry

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	0.63	0/5634	0.84	4/7625 (0.1%)
1	B	0.54	0/5455	0.74	1/7375 (0.0%)
1	C	0.54	0/5297	0.75	1/7155 (0.0%)
1	D	0.51	0/5627	0.70	1/7615 (0.0%)
1	E	0.57	2/5627 (0.0%)	0.77	4/7615 (0.1%)
1	F	0.49	0/5540	0.75	3/7496 (0.0%)
All	All	0.55	2/33180 (0.0%)	0.76	14/44881 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	744	CYS	CB-SG	-6.91	1.70	1.82
1	E	698	GLU	CB-CG	5.77	1.63	1.52

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	177	LEU	CA-CB-CG	6.72	130.76	115.30
1	E	234	LEU	CA-CB-CG	6.52	130.29	115.30
1	F	249	LEU	CA-CB-CG	6.07	129.27	115.30
1	E	629	LEU	CA-CB-CG	-6.05	101.38	115.30
1	F	530	LEU	CA-CB-CG	5.90	128.87	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5511	0	5420	228	0
1	B	5341	0	5268	196	0
1	C	5186	0	5140	199	0
1	D	5504	0	5413	199	0
1	E	5504	0	5413	181	0
1	F	5421	0	5323	224	0
All	All	32467	0	31977	1193	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 19.

The worst 5 of 1193 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:E:608:ILE:HD11	1:E:621:LEU:HD22	1.34	1.08
1:A:156:LEU:HD11	1:A:205:LEU:HG	1.38	1.04
1:F:460:ILE:HG23	1:F:461:PRO:HD3	1.49	0.94
1:B:606:GLU:O	1:B:608:ILE:N	2.01	0.92
1:B:156:LEU:HD11	1:B:205:LEU:HG	1.52	0.91

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	687/775 (89%)	621 (90%)	56 (8%)	10 (2%)	10 45
1	B	658/775 (85%)	611 (93%)	37 (6%)	10 (2%)	10 45
1	C	635/775 (82%)	571 (90%)	52 (8%)	12 (2%)	8 40
1	D	686/775 (88%)	637 (93%)	41 (6%)	8 (1%)	13 50
1	E	686/775 (88%)	630 (92%)	49 (7%)	7 (1%)	15 54

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	671/775 (87%)	605 (90%)	57 (8%)	9 (1%)	12	48
All	All	4023/4650 (86%)	3675 (91%)	292 (7%)	56 (1%)	11	46

5 of 56 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	403	GLU
1	B	607	LEU
1	B	697	GLY
1	C	183	MET
1	C	664	THR

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	598/672 (89%)	571 (96%)	27 (4%)	27	61
1	B	582/672 (87%)	563 (97%)	19 (3%)	38	68
1	C	565/672 (84%)	547 (97%)	18 (3%)	39	69
1	D	597/672 (89%)	580 (97%)	17 (3%)	43	72
1	E	597/672 (89%)	568 (95%)	29 (5%)	25	59
1	F	589/672 (88%)	569 (97%)	20 (3%)	37	68
All	All	3528/4032 (88%)	3398 (96%)	130 (4%)	34	65

5 of 130 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	444	ARG
1	A	545	PHE
1	E	239	ARG
1	E	232	ARG
1	A	569	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	420	GLN
1	C	251	HIS
1	E	420	GLN
1	D	420	GLN
1	F	420	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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

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, 95^{th} 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	689/775 (88%)	0.33	3 (0%) 92 90	32, 56, 100, 127	0
1	B	668/775 (86%)	0.56	39 (5%) 23 20	34, 76, 118, 132	0
1	C	648/775 (83%)	0.72	69 (10%) 6 7	39, 85, 150, 177	0
1	D	688/775 (88%)	0.67	65 (9%) 8 9	41, 84, 138, 160	0
1	E	688/775 (88%)	0.31	3 (0%) 92 90	41, 67, 111, 131	0
1	F	677/775 (87%)	0.75	68 (10%) 7 8	50, 95, 145, 169	0
All	All	4058/4650 (87%)	0.55	247 (6%) 21 19	32, 76, 135, 177	0

The worst 5 of 247 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	496	ILE	5.6
1	D	454	TRP	5.2
1	F	192	ASP	5.1
1	C	200	ILE	5.1
1	C	104	TYR	4.6

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

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.