wwPDB X-ray Structure Validation Summary Report (i)

Apr 26, 2022 – 10:11 am BST

PDB ID	:	7YWE
Title	:	Monocot chimeric jacalin JAC1 from Oryza sativa: dirigent domain (crystal form 2)
Authors	:	Huwa, N.; Classen, T.; Weiergraeber, O.H.
Deposited on	:	2022-02-13
Resolution	:	2.15 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:

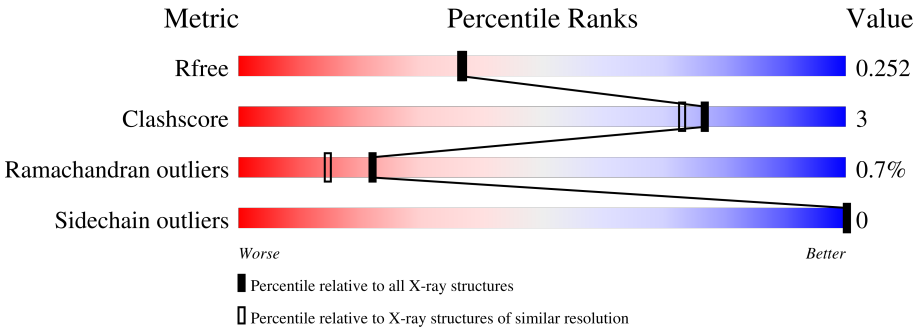
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.28
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.28

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 2.15 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	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)

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%

Mol	Chain	Length	Quality of chain
1	A	161	79% 6% • 14%
1	B	161	86% 8% 6%
1	C	161	80% 8% 12%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 3258 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.

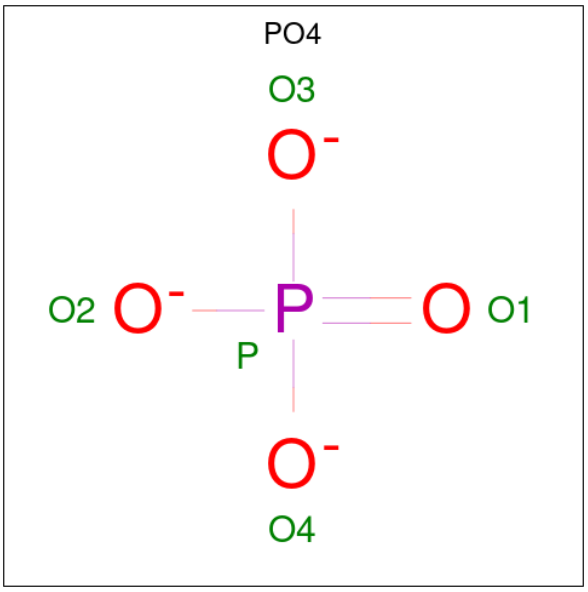
- Molecule 1 is a protein called Dirigent protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	138	Total	C	N	O	S	0	0	0
			1036	663	175	193	5			
1	B	152	Total	C	N	O	S	0	0	0
			1114	714	187	207	6			
1	C	142	Total	C	N	O	S	0	0	0
			1054	676	179	194	5			

There are 21 discrepancies between the modelled and reference sequences:

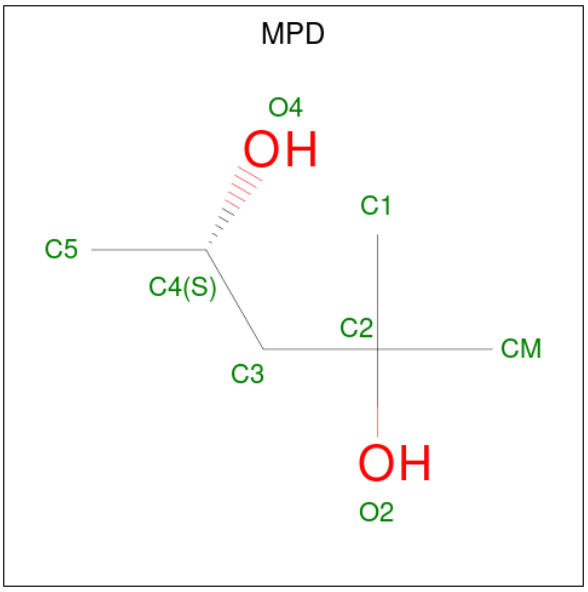
Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP Q306J3
A	0	SER	-	expression tag	UNP Q306J3
A	1	HIS	-	expression tag	UNP Q306J3
A	2	MET	-	expression tag	UNP Q306J3
A	3	LEU	-	expression tag	UNP Q306J3
A	4	GLU	-	expression tag	UNP Q306J3
A	49	ILE	THR	engineered mutation	UNP Q306J3
B	-1	GLY	-	expression tag	UNP Q306J3
B	0	SER	-	expression tag	UNP Q306J3
B	1	HIS	-	expression tag	UNP Q306J3
B	2	MET	-	expression tag	UNP Q306J3
B	3	LEU	-	expression tag	UNP Q306J3
B	4	GLU	-	expression tag	UNP Q306J3
B	49	ILE	THR	engineered mutation	UNP Q306J3
C	-1	GLY	-	expression tag	UNP Q306J3
C	0	SER	-	expression tag	UNP Q306J3
C	1	HIS	-	expression tag	UNP Q306J3
C	2	MET	-	expression tag	UNP Q306J3
C	3	LEU	-	expression tag	UNP Q306J3
C	4	GLU	-	expression tag	UNP Q306J3
C	49	ILE	THR	engineered mutation	UNP Q306J3

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O P 5 4 1	0	0
2	B	1	Total O P 5 4 1	0	0
2	C	1	Total O P 5 4 1	0	0
2	C	1	Total O P 5 4 1	0	0

- Molecule 3 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: $C_6H_{14}O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 8 6 2	0	0

- Molecule 4 is water.

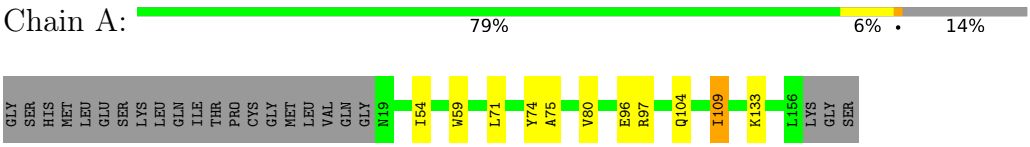
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	10	Total O 10 10	0	0
4	B	10	Total O 10 10	0	0
4	C	6	Total O 6 6	0	0

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. 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. 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.

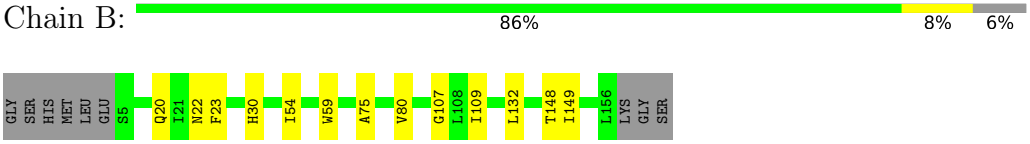
• Molecule 1: Dirigent protein

Chain A:



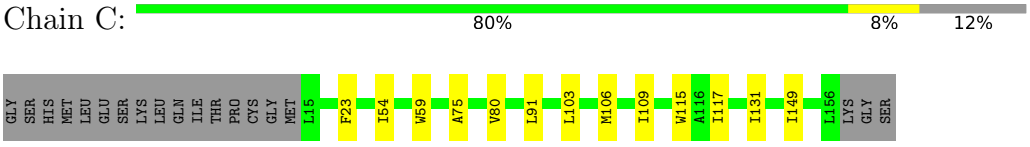
• Molecule 1: Dirigent protein

Chain B:



• Molecule 1: Dirigent protein

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α, β, γ	158.90Å 158.90Å 47.85Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.87 – 2.15 45.87 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.4 (45.87-2.15) 99.7 (45.87-2.15)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.01 (at 2.16Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.218 , 0.254 0.215 , 0.252	Depositor DCC
R_{free} test set	1222 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	59.3	Xtriage
Anisotropy	0.476	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.018 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	3258	wwPDB-VP
Average B, all atoms (Å^2)	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.52% 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

Bond lengths and bond angles in the following residue types are not validated in this section: PO4, MPD

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	0.25	0/1059	0.46	0/1438
1	B	0.25	0/1138	0.45	0/1549
1	C	0.25	0/1077	0.46	0/1464
All	All	0.25	0/3274	0.46	0/4451

There are no bond length outliers.

There are no bond angle outliers.

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	1036	0	993	6	0
1	B	1114	0	1061	7	0
1	C	1054	0	1009	7	0
2	A	5	0	0	0	0
2	B	5	0	0	1	0
2	C	10	0	0	0	0
3	A	8	0	14	0	0
4	A	10	0	0	0	0
4	B	10	0	0	0	0
4	C	6	0	0	0	0
All	All	3258	0	3077	19	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 19 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:59:TRP:HB2	1:A:75:ALA:HB3	1.74	0.69
1:C:23:PHE:HB2	1:C:149:ILE:HB	1.77	0.67
1:B:59:TRP:HB2	1:B:75:ALA:HB3	1.78	0.64
1:C:54:ILE:HG12	1:C:80:VAL:HG12	1.85	0.59
1:C:59:TRP:HB2	1:C:75:ALA:HB3	1.88	0.55

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	136/161 (84%)	130 (96%)	5 (4%)	1 (1%)	22	15
1	B	150/161 (93%)	144 (96%)	5 (3%)	1 (1%)	22	15
1	C	140/161 (87%)	136 (97%)	3 (2%)	1 (1%)	22	15
All	All	426/483 (88%)	410 (96%)	13 (3%)	3 (1%)	22	15

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	109	ILE
1	B	109	ILE
1	C	109	ILE

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	105/131 (80%)	105 (100%)	0	100	100
1	B	111/131 (85%)	111 (100%)	0	100	100
1	C	105/131 (80%)	105 (100%)	0	100	100
All	All	321/393 (82%)	321 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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

Mol	Chain	Res	Type
1	A	20	GLN
1	B	20	GLN
1	B	22	ASN

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)

5 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	PO4	C	202	-	4,4,4	0.91	0	6,6,6	0.42	0
2	PO4	A	201	-	4,4,4	0.91	0	6,6,6	0.42	0
3	MPD	A	202	-	7,7,7	0.14	0	9,10,10	0.49	0
2	PO4	C	201	-	4,4,4	0.97	0	6,6,6	0.41	0
2	PO4	B	201	-	4,4,4	0.92	0	6,6,6	0.42	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
3	MPD	A	202	-	-	3/5/5/5	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	202	MPD	C1-C2-C3-C4
3	A	202	MPD	O2-C2-C3-C4
3	A	202	MPD	C2-C3-C4-C5

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	201	PO4	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

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.