WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Aug 27, 2024 – 05:01 PM JST

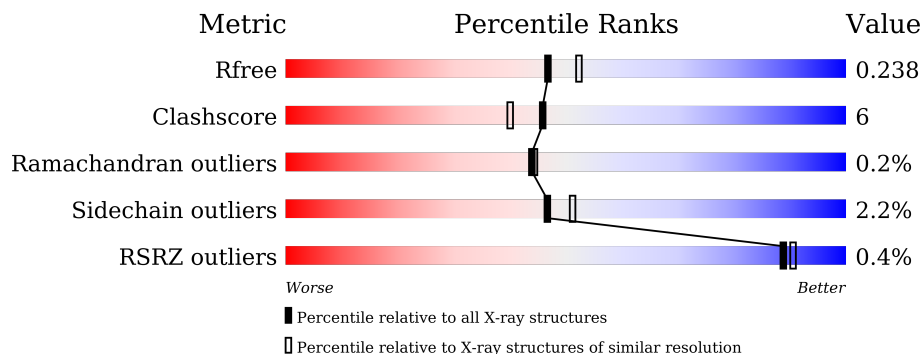|		
PDB ID	:	9J4L
Title	:	Crystal structure of GH9l Inulin fructotransferases (IFTase)
Authors	:	Chen, G.; Wang, Z.X.; Yang, Y.Q.; Li, Y.G.; Zhang, T.; Ouyang, S.Y.; Zhang, L.; Chen, Y.; Ruan, X.L.; Miao, M.
Deposited on	:	2024-08-09
Resolution	:	2.15 Å(reported)

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 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	Percentile Ranks	Value
Rfree		0.238
Clashscore		6
Ramachandran outliers		0.2%
Sidechain outliers		2.2%
RSRZ outliers		0.4%

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}	164625	1881 (2.16-2.16)
Clashscore	180529	2047 (2.16-2.16)
Ramachandran outliers	177936	2027 (2.16-2.16)
Sidechain outliers	177891	2026 (2.16-2.16)
RSRZ outliers	164620	1882 (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% 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	404	87% 12%
1	B	404	84% 15% •
1	C	404	.% 86% 13% •

2 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 9080 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 DFA-III-forming inulin fructotransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	404	Total	C	N	O	S	0	0	0
			3029	1897	534	585	13			
1	B	403	Total	C	N	O	S	0	0	0
			3028	1895	536	584	13			
1	C	403	Total	C	N	O	S	0	0	0
			3023	1891	534	585	13			

There are 3 discrepancies between the modelled and reference sequences:

| Chain	Residue	Modelled	Actual	Comment	Reference
A	412	HIS	-	expression tag	UNP F8QV43
B	412	HIS	-	expression tag	UNP F8QV43
C	412	HIS	-	expression tag	UNP F8QV43

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.
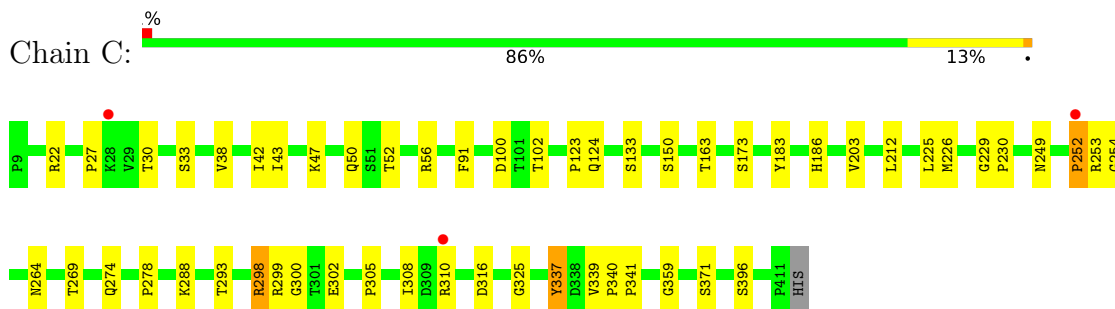
- Molecule 1: DFA-III-forming inulin fructotransferase

Chain A:

- Molecule 1: DFA-III-forming inulin fructotransferase

Chain B:

- Molecule 1: DFA-III-forming inulin fructotransferase

Chain C:
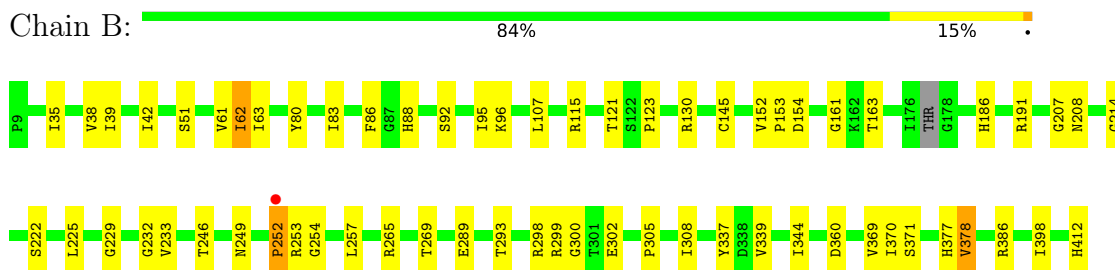
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α, β, γ	167.68Å 167.68Å 85.66Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	9.62 – 2.15 9.62 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.0 (9.62-2.15) 98.0 (9.62-2.15)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.92 (at 2.15Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.205 , 0.233 0.207 , 0.238	Depositor DCC
R_{free} test set	3679 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	23.0	Xtriage
Anisotropy	0.143	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.51 , 53.4	EDS
L-test for twinning[2]	$<|L|> = 0.46$, $<L^2> = 0.29$	Xtriage
Estimated twinning fraction	0.049 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	9080	wwPDB-VP
Average B, all atoms (Å^2)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.70% 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	0.57	1/3094 (0.0%)	0.66	1/4214 (0.0%)
1	B	0.54	1/3092 (0.0%)	0.66	1/4209 (0.0%)
1	C	0.54	0/3086	0.69	1/4202 (0.0%)
All	All	0.55	2/9272 (0.0%)	0.67	3/12625 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	252	PRO	CA-C	-5.43	1.42	1.52
1	A	252	PRO	CA-C	-5.11	1.42	1.52

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	252	PRO	N-CA-CB	10.11	115.43	103.30
1	C	252	PRO	N-CA-CB	8.32	113.28	103.30
1	A	252	PRO	N-CA-CB	7.75	112.60	103.30

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	3029	0	2957	36	0
1	B	3028	0	2959	38	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	3023	0	2954	36	0
All	All	9080	0	8870	103	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 103 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:B:299:ARG:NH1	1:B:300:GLY:O	2.08	0.85
1:A:202:MET:CE	1:C:173:SER:OG	2.25	0.83
1:A:202:MET:HE2	1:C:173:SER:OG	1.78	0.83
1:B:62:ILE:CD1	1:B:86:PHE:HE2	1.93	0.82
1:B:62:ILE:HD11	1:B:86:PHE:HE2	1.47	0.79

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	402/404 (100%)	387 (96%)	14 (4%)	1 (0%)	44	44
1	B	399/404 (99%)	386 (97%)	12 (3%)	1 (0%)	37	34
1	C	401/404 (99%)	385 (96%)	15 (4%)	1 (0%)	44	44
All	All	1202/1212 (99%)	1158 (96%)	41 (3%)	3 (0%)	44	44

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	252	PRO
1	B	252	PRO

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	252	PRO

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	331/333 (99%)	323 (98%)	8 (2%)	44	47
1	B	331/333 (99%)	322 (97%)	9 (3%)	40	42
1	C	330/333 (99%)	325 (98%)	5 (2%)	60	66
All	All	992/999 (99%)	970 (98%)	22 (2%)	47	51

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	298	ARG
1	C	133	SER
1	B	378	VAL
1	C	212	LEU
1	A	396	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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	404/404 (100%)	-0.66	1 (0%)	92	93	14, 21, 34, 57	0
1	B	403/404 (99%)	-0.64	1 (0%)	92	93	14, 21, 34, 57	0
1	C	403/404 (99%)	-0.56	3 (0%)	84	86	13, 22, 35, 54	0
All	All	1210/1212 (99%)	-0.62	5 (0%)	89	90	13, 21, 35, 57	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	252	PRO	5.2
1	B	252	PRO	5.2
1	C	252	PRO	4.5
1	C	310	ARG	2.7
1	C	28	LYS	2.2

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 ⓘ

There are no such residues in this entry.