



Full wwPDB X-ray Structure Validation Report ⓘ

Jun 27, 2024 – 04:03 PM EDT

PDB ID : 8TX6
Title : Crystal structure of an engineered variant of galactose oxidase, GOaseRd7BB, from *Fusarium graminearum*
Authors : Selvaraj, B.; Orth, P.
Deposited on : 2023-08-22
Resolution : 1.56 Å (reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ 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 ⓘ](#)) 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 : 2.37.1
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) : 2.37.1

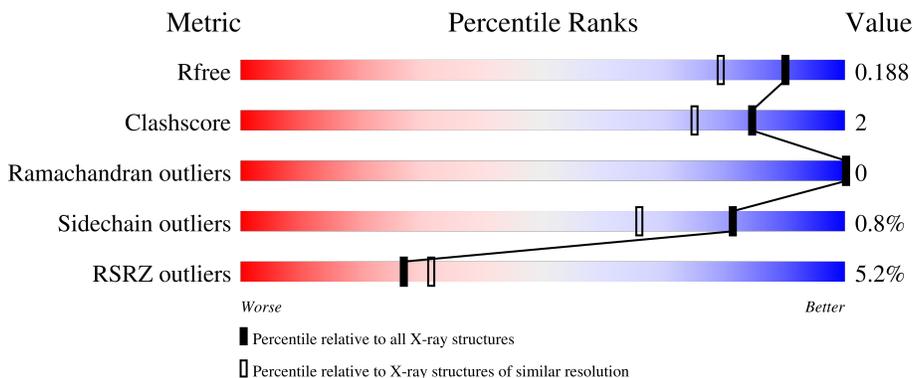
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

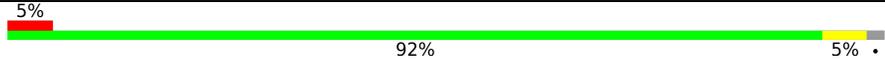
The reported resolution of this entry is 1.56 Å.

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	1483 (1.56-1.56)
Clashscore	141614	1529 (1.56-1.56)
Ramachandran outliers	138981	1498 (1.56-1.56)
Sidechain outliers	138945	1495 (1.56-1.56)
RSRZ outliers	127900	1465 (1.56-1.56)

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 $\leq 5\%$. 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	653	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 5252 atoms, of which 3 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 Galactose oxidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	638	4819	3013	832	955	19	0	1	0

There are 40 discrepancies between the modelled and reference sequences:

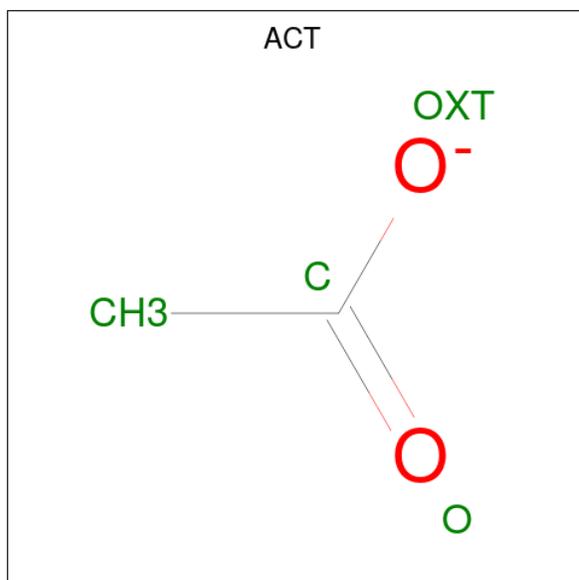
Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP P0CS93
A	7	VAL	SER	conflict	UNP P0CS93
A	10	PRO	SER	conflict	UNP P0CS93
A	42	GLN	PHE	conflict	UNP P0CS93
A	45	VAL	ALA	conflict	UNP P0CS93
A	55	ILE	TYR	conflict	UNP P0CS93
A	62	VAL	THR	conflict	UNP P0CS93
A	70	VAL	MET	conflict	UNP P0CS93
A	94	VAL	THR	conflict	UNP P0CS93
A	172	SER	ALA	conflict	UNP P0CS93
A	191	GLN	ASN	conflict	UNP P0CS93
A	195	GLU	GLY	conflict	UNP P0CS93
A	219	SER	VAL	conflict	UNP P0CS93
A	242	SER	THR	conflict	UNP P0CS93
A	290	PHE	TRP	conflict	UNP P0CS93
A	294	GLN	VAL	conflict	UNP P0CS93
A	295	VAL	PHE	conflict	UNP P0CS93
A	406	GLU	GLN	conflict	UNP P0CS93
A	459	GLN	ARG	conflict	UNP P0CS93
A	464	THR	PHE	conflict	UNP P0CS93
A	494	ALA	VAL	conflict	UNP P0CS93
A	535	ASP	ASN	conflict	UNP P0CS93
A	555	VAL	LYS	conflict	UNP P0CS93
A	559	TRP	ARG	conflict	UNP P0CS93
A	563	TRP	SER	conflict	UNP P0CS93
A	566	MET	SER	conflict	UNP P0CS93
A	570	ALA	LYS	conflict	UNP P0CS93

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	640	GLY	-	expression tag	UNP P0CS93
A	641	GLY	-	expression tag	UNP P0CS93
A	642	GLY	-	expression tag	UNP P0CS93
A	643	GLY	-	expression tag	UNP P0CS93
A	644	SER	-	expression tag	UNP P0CS93
A	645	TRP	-	expression tag	UNP P0CS93
A	646	SER	-	expression tag	UNP P0CS93
A	647	HIS	-	expression tag	UNP P0CS93
A	648	PRO	-	expression tag	UNP P0CS93
A	649	GLN	-	expression tag	UNP P0CS93
A	650	PHE	-	expression tag	UNP P0CS93
A	651	GLU	-	expression tag	UNP P0CS93
A	652	LYS	-	expression tag	UNP P0CS93

- Molecule 2 is ACETATE ION (three-letter code: ACT) (formula: C₂H₃O₂).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
2	A	1	7	2	3	2	0	0

- Molecule 3 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cu		
3	A	1	1	1	0	0

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total 1	Ca 1	0	0

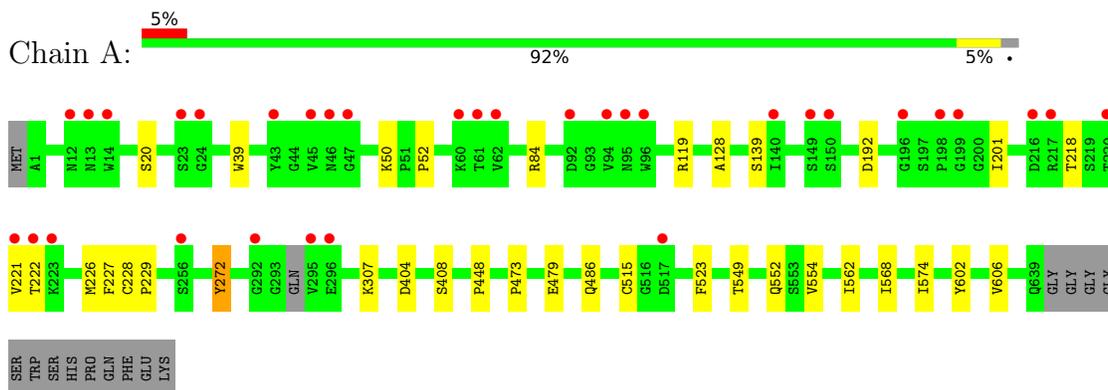
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	424	Total 424	O 424	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 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: Galactose oxidase



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	97.62Å 88.92Å 86.17Å 90.00° 117.82° 90.00°	Depositor
Resolution (Å)	22.75 – 1.56 22.75 – 1.56	Depositor EDS
% Data completeness (in resolution range)	85.3 (22.75-1.56) 85.3 (22.75-1.56)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.37 (at 1.56Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.164 , 0.188 0.164 , 0.188	Depositor DCC
R_{free} test set	3891 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å ²)	22.6	Xtrriage
Anisotropy	0.008	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 38.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	5252	wwPDB-VP
Average B, all atoms (Å ²)	27.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.49% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: CA, CU, ACT

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	0.56	0/4950	0.71	2/6760 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	272	TYR	CD1-CE1-CZ	8.25	127.22	119.80
1	A	272	TYR	CG-CD1-CE1	-5.38	117.00	121.30

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	4819	0	4590	22	0
2	A	4	3	3	0	0
3	A	1	0	0	0	0
4	A	1	0	0	0	0
5	A	424	0	0	4	0
All	All	5249	3	4593	22	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 2.

All (22) 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:228:CYS:SG	1:A:272:TYR:HE1	1.20	1.63
1:A:228:CYS:SG	1:A:272:TYR:CE1	2.10	0.99
1:A:552:GLN:OE1	5:A:801:HOH:O	1.81	0.98
1:A:307:LYS:NZ	5:A:802:HOH:O	2.29	0.66
1:A:228:CYS:N	1:A:229:PRO:HD3	2.22	0.53
1:A:201:ILE:HD11	1:A:218:THR:CG2	2.39	0.53
1:A:479:GLU:HB2	5:A:995:HOH:O	2.10	0.51
1:A:549:THR:HG22	1:A:562:ILE:HG22	1.94	0.49
1:A:448:PRO:HA	1:A:574:ILE:HD11	1.95	0.48
1:A:84:ARG:HD2	5:A:1065:HOH:O	2.17	0.44
1:A:473:PRO:HG2	1:A:486:GLN:HB2	1.99	0.44
1:A:554:VAL:HG11	1:A:606:VAL:HG21	2.00	0.43
1:A:568:ILE:HG12	1:A:602:TYR:CE1	2.53	0.43
1:A:226:MET:O	1:A:229:PRO:HG3	2.18	0.43
1:A:228:CYS:SG	1:A:272:TYR:CZ	2.88	0.43
1:A:227:PHE:CE2	1:A:228:CYS:SG	3.12	0.43
1:A:39:TRP:O	1:A:139:SER:HA	2.18	0.42
1:A:192:ASP:HB3	1:A:523:PHE:HZ	1.85	0.42
1:A:221:VAL:HG13	1:A:222:THR:HG23	2.01	0.42
1:A:404:ASP:HB2	1:A:408:SER:HB2	2.02	0.41
1:A:201:ILE:HD11	1:A:218:THR:HG23	2.02	0.40
1:A:52:PRO:HA	1:A:128:ALA:O	2.21	0.40

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	635/653 (97%)	613 (96%)	22 (4%)	0	100 100

There are no Ramachandran outliers to report.

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	527/537 (98%)	523 (99%)	4 (1%)	81 66

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

Mol	Chain	Res	Type
1	A	20	SER
1	A	50	LYS
1	A	119	ARG
1	A	515	CYS

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	42	GLN
1	A	78	GLN
1	A	135	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](#)

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 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	$\# Z > 2$
2	ACT	A	701	-	3,3,3	1.09	0	3,3,3	1.25	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.

No monomer is involved in short contacts.

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, 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	638/653 (97%)	0.02	33 (5%) 27 31	16, 25, 46, 71	0

All (33) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	45	VAL	10.0
1	A	46	ASN	5.9
1	A	220	THR	5.4
1	A	221	VAL	5.3
1	A	295	VAL	5.2
1	A	94	VAL	5.1
1	A	222	THR	3.9
1	A	292	GLY	3.5
1	A	196	GLY	3.2
1	A	23	SER	3.2
1	A	24	GLY	3.1
1	A	217	ARG	3.0
1	A	223	LYS	3.0
1	A	61	THR	2.8
1	A	216	ASP	2.8
1	A	140	ILE	2.7
1	A	92	ASP	2.7
1	A	62	VAL	2.6
1	A	198	PRO	2.6
1	A	12	ASN	2.6
1	A	13	ASN	2.5
1	A	95	ASN	2.5
1	A	150	SER	2.4
1	A	296	GLU	2.4
1	A	149	SER	2.3
1	A	14	TRP	2.3
1	A	199	GLY	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	60	LYS	2.2
1	A	96	TRP	2.2
1	A	517	ASP	2.1
1	A	256	SER	2.1
1	A	43	TYR	2.0
1	A	47	GLY	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

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
4	CA	A	703	1/1	0.93	0.05	37,37,37,37	0
2	ACT	A	701	4/4	0.98	0.07	20,22,27,27	0
3	CU	A	702	1/1	1.00	0.02	27,27,27,27	0

6.5 Other polymers [i](#)

There are no such residues in this entry.