Full wwPDB X-ray Structure Validation Report (i)

Oct 18, 2023 – 02:51 AM EDT

PDB ID	:	2EIC
Title	:	Crystal Structure of Galactose Oxidase mutant W290F
Authors	:	Akyumani, N.; Tamber, S.; Firbank, S.J.; Knowles, P.F.; Phillips, S.E.; McPherson, M.J.
Deposited on	:	2007-03-12
Resolution	:	2.80 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 (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:

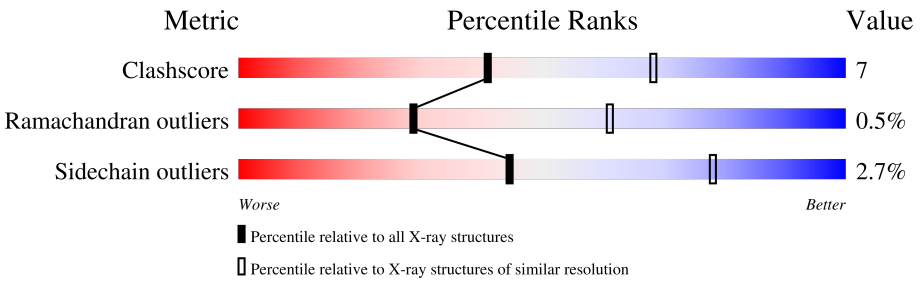
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.80 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))
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	639	83% 15% •

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5048 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 Galactose oxidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	639	4826	3015	839	953	19	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	290	PHE	TRP	engineered mutation	UNP Q01745

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Na		
2	A	1	1	1	0	0

- Molecule 3 is COPPER (I) ION (three-letter code: CU1) (formula: Cu).

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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	220	220	220	0	0

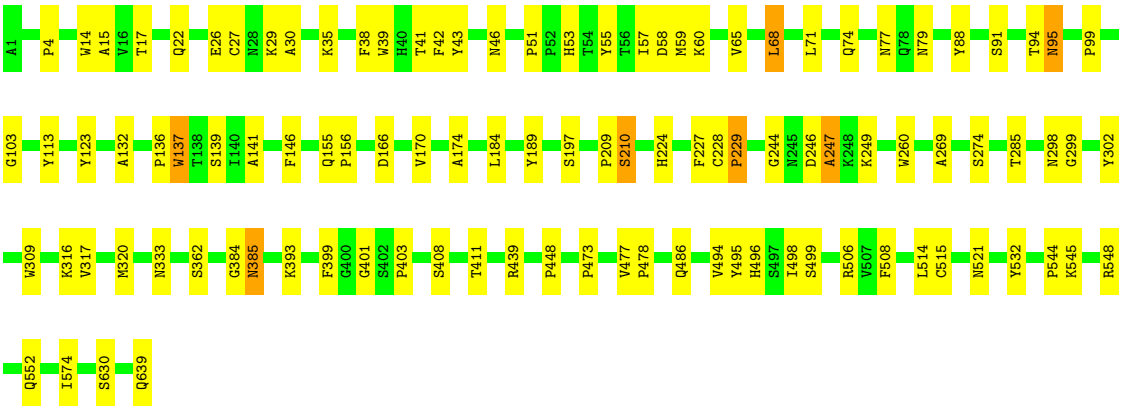
3 Residue-property plots (i)

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.

Note EDS was not executed.

- Molecule 1: Galactose oxidase

Chain A:



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	97.77Å 88.89Å 86.19Å 90.00° 117.90° 90.00°	Depositor
Resolution (Å)	30.00 – 2.80	Depositor
% Data completeness (in resolution range)	89.3 (30.00-2.80)	Depositor
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.177 , 0.233	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5048	wwPDB-VP
Average B, all atoms (Å^2)	25.0	wwPDB-VP

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: NA, CU1

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	0.39	0/4954	0.67	2/6758 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	210	SER	CB-CA-C	-6.53	97.69	110.10
1	A	210	SER	N-CA-CB	5.24	118.36	110.50

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	4826	0	4601	67	0
2	A	1	0	0	0	0
3	A	1	0	0	0	0
4	A	220	0	0	7	0
All	All	5048	0	4601	67	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 7.

All (67) 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:210:SER:HA	4:A:847:HOH:O	1.70	0.90
1:A:35:LYS:HE2	1:A:74:GLN:HG3	1.53	0.90
1:A:22:GLN:HE21	1:A:46:ASN:HB2	1.44	0.82
1:A:35:LYS:HE3	1:A:71:LEU:HD21	1.67	0.75
1:A:403:PRO:HD2	1:A:408:SER:OG	1.95	0.67
1:A:210:SER:HB3	4:A:849:HOH:O	1.95	0.66
1:A:22:GLN:CG	1:A:42:PHE:HA	2.28	0.64
1:A:298:ASN:HD22	1:A:317:VAL:H	1.48	0.60
1:A:22:GLN:HG3	1:A:42:PHE:HA	1.83	0.60
1:A:506:ARG:HD3	4:A:861:HOH:O	2.01	0.59
1:A:95:ASN:HD22	1:A:95:ASN:N	1.99	0.59
1:A:473:PRO:HG2	1:A:486:GLN:HB2	1.84	0.59
1:A:298:ASN:ND2	1:A:317:VAL:H	2.00	0.59
1:A:317:VAL:O	1:A:320:MET:HG2	2.03	0.58
1:A:22:GLN:NE2	1:A:46:ASN:HB2	2.16	0.57
1:A:41:THR:HG21	1:A:53:HIS:CG	2.40	0.56
1:A:170:VAL:HB	1:A:514:LEU:HD13	1.86	0.56
1:A:495:TYR:O	1:A:496:HIS:HB2	2.06	0.56
1:A:4:PRO:HD3	1:A:113:TYR:CE2	2.42	0.55
1:A:401:GLY:O	1:A:411:THR:HG22	2.07	0.55
1:A:43:TYR:HB2	1:A:137:TRP:CZ3	2.42	0.54
1:A:51:PRO:HD3	1:A:136:PRO:HA	1.89	0.54
1:A:174:ALA:HB3	1:A:498:ILE:HD13	1.91	0.53
1:A:30:ALA:HB3	1:A:55:TYR:OH	2.09	0.52
1:A:448:PRO:HA	1:A:574:ILE:HD11	1.91	0.51
1:A:22:GLN:HG2	1:A:42:PHE:HA	1.92	0.50
1:A:224:HIS:HD2	1:A:246:ASP:HB2	1.76	0.50
1:A:14:TRP:CD1	1:A:59:MET:HA	2.46	0.50
1:A:384:GLY:C	1:A:385:ASN:HD22	2.16	0.49
1:A:298:ASN:ND2	1:A:316:LYS:HA	2.27	0.49
1:A:227:PHE:O	1:A:244:GLY:HA3	2.13	0.49
1:A:95:ASN:HD22	1:A:95:ASN:H	1.60	0.49
1:A:95:ASN:N	1:A:95:ASN:ND2	2.61	0.48
1:A:39:TRP:O	1:A:139:SER:HA	2.12	0.48
1:A:385:ASN:HB2	1:A:399:PHE:CE2	2.48	0.48
1:A:38:PHE:HB3	1:A:141:ALA:HA	1.96	0.48
1:A:385:ASN:HD22	1:A:385:ASN:N	2.10	0.48
1:A:184:LEU:HD21	1:A:260:TRP:HZ2	1.77	0.48
1:A:68:LEU:C	1:A:68:LEU:HD12	2.34	0.47
1:A:210:SER:N	4:A:847:HOH:O	2.46	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:57:ILE:O	1:A:123:TYR:HA	2.14	0.47
1:A:88:TYR:CZ	1:A:99:PRO:HG3	2.50	0.47
1:A:184:LEU:HD21	1:A:260:TRP:CZ2	2.50	0.47
1:A:209:PRO:C	4:A:847:HOH:O	2.54	0.47
1:A:17:THR:C	1:A:27:CYS:SG	2.94	0.46
1:A:362:SER:N	4:A:668:HOH:O	2.49	0.46
1:A:285:THR:O	1:A:299:GLY:HA2	2.16	0.45
1:A:393:LYS:HE3	4:A:709:HOH:O	2.17	0.45
1:A:544:PRO:HG2	1:A:630:SER:HB3	1.99	0.45
1:A:499:SER:HA	1:A:508:PHE:O	2.16	0.45
1:A:15:ALA:HB2	1:A:60:LYS:NZ	2.32	0.45
1:A:228:CYS:N	1:A:229:PRO:HD3	2.31	0.45
1:A:65:VAL:HA	1:A:146:PHE:O	2.18	0.44
1:A:26:GLU:OE1	1:A:29:LYS:HE2	2.18	0.44
1:A:477:VAL:HA	1:A:478:PRO:HD2	1.86	0.43
1:A:77:ASN:OD1	1:A:79:ASN:HB2	2.19	0.42
1:A:155:GLN:HA	1:A:156:PRO:HD3	1.93	0.42
1:A:189:TYR:CD1	1:A:197:SER:HB2	2.55	0.42
1:A:15:ALA:HB3	1:A:58:ASP:HB3	2.02	0.41
1:A:514:LEU:N	1:A:521:ASN:OD1	2.53	0.41
1:A:103:GLY:HA3	1:A:166:ASP:O	2.21	0.41
1:A:247:ALA:O	1:A:269:ALA:HA	2.21	0.41
1:A:302:TYR:HB2	1:A:309:TRP:CE2	2.56	0.41
1:A:94:THR:HB	1:A:95:ASN:HD22	1.86	0.40
1:A:224:HIS:CD2	1:A:249:LYS:HB2	2.57	0.40
1:A:302:TYR:HB2	1:A:309:TRP:CD2	2.57	0.40
1:A:333:ASN:N	1:A:333:ASN:HD22	2.20	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	637/639 (100%)	604 (95%)	30 (5%)	3 (0%)	29 61

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	132	ALA
1	A	494	VAL
1	A	247	ALA

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	526/526 (100%)	512 (97%)	14 (3%)	44 78

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	68	LEU
1	A	91	SER
1	A	95	ASN
1	A	137	TRP
1	A	229	PRO
1	A	274	SER
1	A	385	ASN
1	A	439	ARG
1	A	515	CYS
1	A	532	TYR
1	A	545	LYS
1	A	548	ARG
1	A	552	GLN
1	A	639	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	22	GLN
1	A	46	ASN
1	A	78	GLN
1	A	95	ASN
1	A	267	GLN
1	A	298	ASN
1	A	333	ASN
1	A	385	ASN
1	A	413	ASN
1	A	552	GLN
1	A	597	ASN
1	A	600	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)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

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 (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.