wwPDB X-ray Structure Validation Summary Report (i)

Oct 16, 2021 – 09:25 PM EDT

PDB ID	:	1PN2
Title	:	Crystal structure analysis of the selenomethionine labelled 2-enoyl-CoA hydratase 2 domain of Candida tropicalis multifunctional enzyme type 2
Authors	:	Koski, M.K.; Haapalainen, A.M.; Hiltunen, J.K.; Glumoff, T.
Deposited on	:	2003-06-12
Resolution	:	1.95 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:

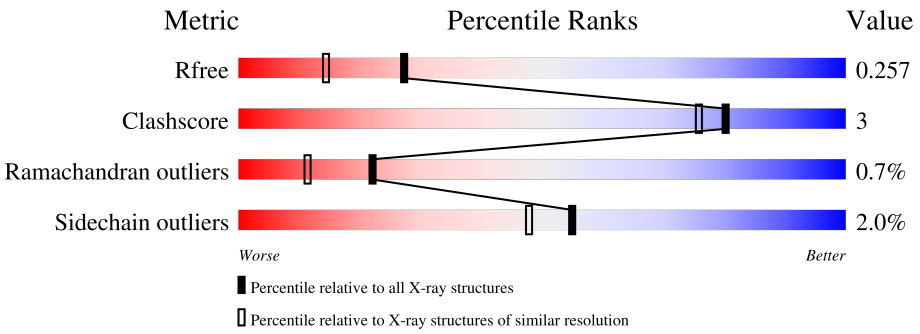
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 1.95 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	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)

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	280	83% 12% • •
1	B	280	85% 10% • •
1	C	280	81% 13% • 5%
1	D	280	86% 10% •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 9374 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.

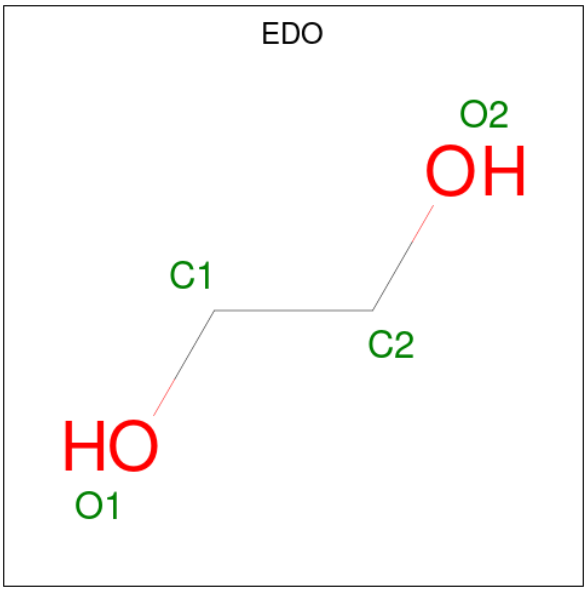
- Molecule 1 is a protein called Peroxisomal hydratase-dehydrogenase-epimerase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	272	2153	1386	367	395	2	3	98	0	0
1	B	270	2141	1380	365	391	2	3	51	0	0
1	C	267	2117	1366	360	386	2	3	36	0	0
1	D	272	2153	1386	367	395	2	3	32	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	GLU	engineered mutation	UNP P22414
A	69	MSE	MET	modified residue	UNP P22414
A	206	MSE	MET	modified residue	UNP P22414
A	222	MSE	MET	modified residue	UNP P22414
B	1	MSE	GLU	engineered mutation	UNP P22414
B	69	MSE	MET	modified residue	UNP P22414
B	206	MSE	MET	modified residue	UNP P22414
B	222	MSE	MET	modified residue	UNP P22414
C	1	MSE	GLU	engineered mutation	UNP P22414
C	69	MSE	MET	modified residue	UNP P22414
C	206	MSE	MET	modified residue	UNP P22414
C	222	MSE	MET	modified residue	UNP P22414
D	1	MSE	GLU	engineered mutation	UNP P22414
D	69	MSE	MET	modified residue	UNP P22414
D	206	MSE	MET	modified residue	UNP P22414
D	222	MSE	MET	modified residue	UNP P22414

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	4	2	2	0	0
			Total	C	O		
2	C	1	4	2	2	0	0
			Total	C	O		
2	C	1	4	2	2	0	0
			Total	C	O		
2	D	1	4	2	2	0	0
			Total	C	O		
2	D	1	4	2	2	0	0

- Molecule 3 is water.

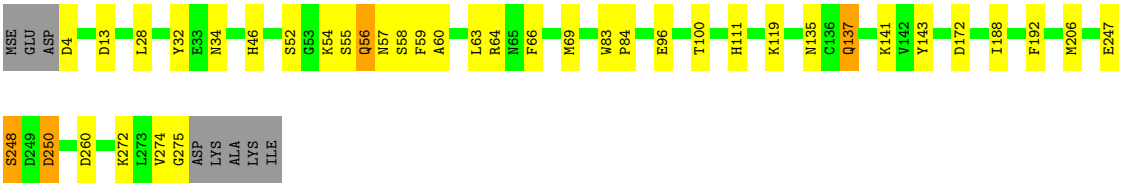
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	153	153	153	0	0
			Total	O		
3	B	147	147	147	0	0
			Total	O		
3	C	236	236	236	0	0
			Total	O		
3	D	254	254	254	0	0

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.

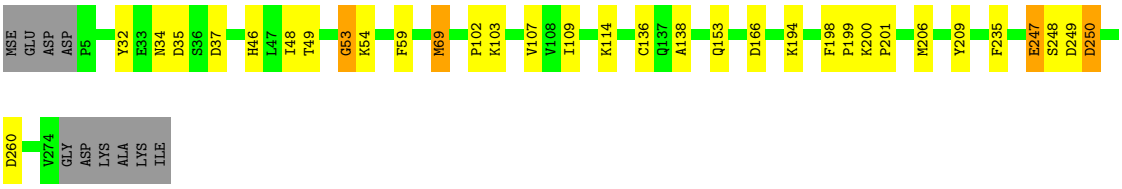
• Molecule 1: Peroxisomal hydratase-dehydrogenase-epimerase

Chain A:



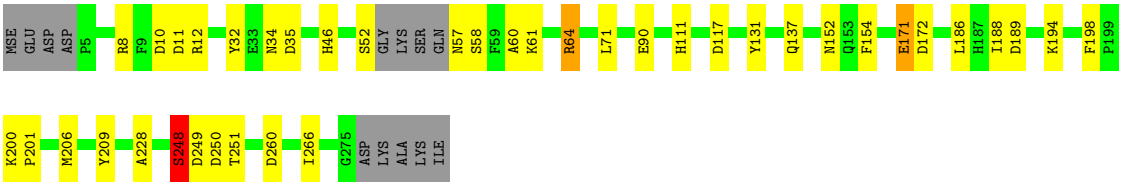
• Molecule 1: Peroxisomal hydratase-dehydrogenase-epimerase

Chain B:



• Molecule 1: Peroxisomal hydratase-dehydrogenase-epimerase

Chain C:



• Molecule 1: Peroxisomal hydratase-dehydrogenase-epimerase

Chain D:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	178.80Å 60.65Å 131.12Å 90.00° 94.57° 90.00°	Depositor
Resolution (Å)	30.00 – 1.95 29.57 – 1.95	Depositor EDS
% Data completeness (in resolution range)	95.2 (30.00-1.95) 95.2 (29.57-1.95)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.28 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.1	Depositor
R, R_{free}	0.179 , 0.214 0.235 , 0.257	Depositor DCC
R_{free} test set	9761 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å^2)	26.4	Xtriage
Anisotropy	0.619	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 54.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	9374	wwPDB-VP
Average B, all atoms (Å^2)	11.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.11% 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 (i)

5.1 Standard geometry (i)

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

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	0.94	9/2207 (0.4%)	1.08	17/2996 (0.6%)
1	B	0.68	3/2195 (0.1%)	0.85	6/2979 (0.2%)
1	C	1.00	6/2170 (0.3%)	1.02	13/2945 (0.4%)
1	D	0.81	3/2207 (0.1%)	0.92	12/2995 (0.4%)
All	All	0.87	21/8779 (0.2%)	0.97	48/11915 (0.4%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	1
1	C	0	1
All	All	0	3

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	96	GLU	CD-OE1	-22.05	1.01	1.25
1	C	61	LYS	CA-CB	-14.98	1.21	1.53
1	C	171	GLU	CD-OE2	-14.30	1.09	1.25
1	A	56	GLN	C-N	-10.74	1.09	1.34
1	C	171	GLU	CD-OE1	10.44	1.37	1.25

The worst 5 of 48 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	249	ASP	CB-CG-OD2	14.50	131.35	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	60	ALA	C-N-CA	14.40	157.70	121.70
1	A	96	GLU	OE1-CD-OE2	14.06	140.18	123.30
1	A	56	GLN	O-C-N	-13.59	100.96	122.70
1	C	60	ALA	N-CA-CB	-10.41	95.52	110.10

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	56	GLN	Mainchain
1	B	250	ASP	Peptide
1	C	248	SER	Peptide

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	2153	0	2131	11	0
1	B	2141	0	2127	13	0
1	C	2117	0	2100	13	0
1	D	2153	0	2134	9	0
2	A	4	0	6	1	0
2	C	8	0	12	0	0
2	D	8	0	12	0	0
3	A	153	0	0	0	0
3	B	147	0	0	1	0
3	C	236	0	0	0	0
3	D	254	0	0	0	0
All	All	9374	0	8522	45	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 45 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:54:LYS:O	1:A:58:SER:HB3	1.89	0.73
1:B:247:GLU:OE1	1:B:247:GLU:HA	1.91	0.69
1:C:58:SER:HB2	1:C:111:HIS:HE1	1.63	0.63
1:C:64:ARG:HB2	1:C:137:GLN:HB3	1.84	0.60
1:A:247:GLU:O	1:A:248:SER:HB3	2.02	0.59

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	270/280 (96%)	263 (97%)	5 (2%)	2 (1%)	22 11
1	B	268/280 (96%)	259 (97%)	7 (3%)	2 (1%)	22 11
1	C	263/280 (94%)	257 (98%)	4 (2%)	2 (1%)	19 9
1	D	270/280 (96%)	265 (98%)	4 (2%)	1 (0%)	34 22
All	All	1071/1120 (96%)	1044 (98%)	20 (2%)	7 (1%)	22 11

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	248	SER
1	B	53	GLY
1	B	54	LYS
1	C	248	SER
1	D	55	SER

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	233/236 (99%)	229 (98%)	4 (2%)	60	55
1	B	232/236 (98%)	226 (97%)	6 (3%)	46	36
1	C	229/236 (97%)	225 (98%)	4 (2%)	60	55
1	D	233/236 (99%)	228 (98%)	5 (2%)	53	46
All	All	927/944 (98%)	908 (98%)	19 (2%)	55	48

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	56	GLN
1	D	243	LEU
1	D	260	ASP
1	D	197	LYS
1	B	250	ASP

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	111	HIS
1	C	111	HIS
1	C	140	ASN
1	D	111	HIS
1	D	140	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	EDO	C	803	-	3,3,3	0.19	0	2,2,2	0.47	0
2	EDO	D	802	-	3,3,3	0.50	0	2,2,2	0.28	0
2	EDO	A	805	-	3,3,3	0.41	0	2,2,2	0.08	0
2	EDO	D	804	-	3,3,3	0.29	0	2,2,2	0.70	0
2	EDO	C	801	-	3,3,3	0.25	0	2,2,2	1.12	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
2	EDO	C	803	-	-	0/1/1/1	-
2	EDO	D	802	-	-	1/1/1/1	-
2	EDO	A	805	-	-	1/1/1/1	-
2	EDO	D	804	-	-	1/1/1/1	-
2	EDO	C	801	-	-	0/1/1/1	-

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
2	A	805	EDO	O1-C1-C2-O2
2	D	804	EDO	O1-C1-C2-O2
2	D	802	EDO	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	805	EDO	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	60:ALA	C	61:LYS	N	1.12
1	A	56:GLN	C	57:ASN	N	1.09

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

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

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

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

6.3 Carbohydrates (i)

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

6.4 Ligands (i)

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

6.5 Other polymers (i)

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