



wwPDB X-ray Structure Validation Summary Report i

Mar 12, 2014 – 03:21 PM GMT

PDB ID : 4GO3
Title : Crystal structure of PnpE from Pseudomonas sp. WBC-3
Authors : Su, J.; Zhang, C.; Liu, S.; Zhu, D.; Gu, L.
Deposited on : 2012-08-18
Resolution : 2.70 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

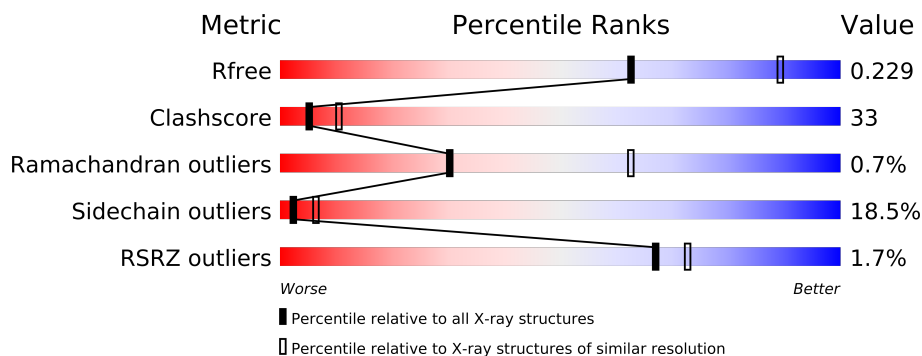
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.16 November 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	trunk22714
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk22714

1 Overall quality at a glance

The reported resolution of this entry is 2.70 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}	66092	1557 (2.70-2.70)
Clashscore	79885	1939 (2.70-2.70)
Ramachandran outliers	78287	1905 (2.70-2.70)
Sidechain outliers	78261	1905 (2.70-2.70)
RSRZ outliers	66119	1559 (2.70-2.70)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	495	
1	B	495	
1	C	495	
1	D	495	
1	E	495	
1	F	495	
1	G	495	
1	H	495	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 29686 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Putative gamma-hydroxymuconic semialdehyde dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	487	Total	C	N	O	S	0	0	0
			3685	2333	652	684	16			
1	B	487	Total	C	N	O	S	0	1	0
			3692	2338	654	684	16			
1	C	487	Total	C	N	O	S	0	1	0
			3691	2337	652	686	16			
1	D	487	Total	C	N	O	S	0	0	0
			3685	2333	652	684	16			
1	E	485	Total	C	N	O	S	0	0	0
			3665	2321	646	682	16			
1	F	487	Total	C	N	O	S	0	0	0
			3685	2333	652	684	16			
1	G	487	Total	C	N	O	S	0	0	0
			3685	2333	652	684	16			
1	H	487	Total	C	N	O	S	0	0	0
			3685	2333	652	684	16			

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	426	ASN	SER	ENGINEERED MUTATION	UNP C1I208
A	484	HIS	TYR	ENGINEERED MUTATION	UNP C1I208
A	488	LEU	-	EXPRESSION TAG	UNP C1I208
A	489	GLY	-	EXPRESSION TAG	UNP C1I208
A	490	HIS	-	EXPRESSION TAG	UNP C1I208
A	491	HIS	-	EXPRESSION TAG	UNP C1I208
A	492	HIS	-	EXPRESSION TAG	UNP C1I208
A	493	HIS	-	EXPRESSION TAG	UNP C1I208
A	494	HIS	-	EXPRESSION TAG	UNP C1I208
A	495	HIS	-	EXPRESSION TAG	UNP C1I208
B	426	ASN	SER	ENGINEERED MUTATION	UNP C1I208
B	484	HIS	TYR	ENGINEERED MUTATION	UNP C1I208
B	488	LEU	-	EXPRESSION TAG	UNP C1I208

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	489	GLY	-	EXPRESSION TAG	UNP C1I208
B	490	HIS	-	EXPRESSION TAG	UNP C1I208
B	491	HIS	-	EXPRESSION TAG	UNP C1I208
B	492	HIS	-	EXPRESSION TAG	UNP C1I208
B	493	HIS	-	EXPRESSION TAG	UNP C1I208
B	494	HIS	-	EXPRESSION TAG	UNP C1I208
B	495	HIS	-	EXPRESSION TAG	UNP C1I208
C	426	ASN	SER	ENGINEERED MUTATION	UNP C1I208
C	484	HIS	TYR	ENGINEERED MUTATION	UNP C1I208
C	488	LEU	-	EXPRESSION TAG	UNP C1I208
C	489	GLY	-	EXPRESSION TAG	UNP C1I208
C	490	HIS	-	EXPRESSION TAG	UNP C1I208
C	491	HIS	-	EXPRESSION TAG	UNP C1I208
C	492	HIS	-	EXPRESSION TAG	UNP C1I208
C	493	HIS	-	EXPRESSION TAG	UNP C1I208
C	494	HIS	-	EXPRESSION TAG	UNP C1I208
C	495	HIS	-	EXPRESSION TAG	UNP C1I208
D	426	ASN	SER	ENGINEERED MUTATION	UNP C1I208
D	484	HIS	TYR	ENGINEERED MUTATION	UNP C1I208
D	488	LEU	-	EXPRESSION TAG	UNP C1I208
D	489	GLY	-	EXPRESSION TAG	UNP C1I208
D	490	HIS	-	EXPRESSION TAG	UNP C1I208
D	491	HIS	-	EXPRESSION TAG	UNP C1I208
D	492	HIS	-	EXPRESSION TAG	UNP C1I208
D	493	HIS	-	EXPRESSION TAG	UNP C1I208
D	494	HIS	-	EXPRESSION TAG	UNP C1I208
D	495	HIS	-	EXPRESSION TAG	UNP C1I208
E	426	ASN	SER	ENGINEERED MUTATION	UNP C1I208
E	484	HIS	TYR	ENGINEERED MUTATION	UNP C1I208
E	488	LEU	-	EXPRESSION TAG	UNP C1I208
E	489	GLY	-	EXPRESSION TAG	UNP C1I208
E	490	HIS	-	EXPRESSION TAG	UNP C1I208
E	491	HIS	-	EXPRESSION TAG	UNP C1I208
E	492	HIS	-	EXPRESSION TAG	UNP C1I208
E	493	HIS	-	EXPRESSION TAG	UNP C1I208
E	494	HIS	-	EXPRESSION TAG	UNP C1I208
E	495	HIS	-	EXPRESSION TAG	UNP C1I208
F	426	ASN	SER	ENGINEERED MUTATION	UNP C1I208
F	484	HIS	TYR	ENGINEERED MUTATION	UNP C1I208
F	488	LEU	-	EXPRESSION TAG	UNP C1I208
F	489	GLY	-	EXPRESSION TAG	UNP C1I208
F	490	HIS	-	EXPRESSION TAG	UNP C1I208

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	491	HIS	-	EXPRESSION TAG	UNP C1I208
F	492	HIS	-	EXPRESSION TAG	UNP C1I208
F	493	HIS	-	EXPRESSION TAG	UNP C1I208
F	494	HIS	-	EXPRESSION TAG	UNP C1I208
F	495	HIS	-	EXPRESSION TAG	UNP C1I208
G	426	ASN	SER	ENGINEERED MUTATION	UNP C1I208
G	484	HIS	TYR	ENGINEERED MUTATION	UNP C1I208
G	488	LEU	-	EXPRESSION TAG	UNP C1I208
G	489	GLY	-	EXPRESSION TAG	UNP C1I208
G	490	HIS	-	EXPRESSION TAG	UNP C1I208
G	491	HIS	-	EXPRESSION TAG	UNP C1I208
G	492	HIS	-	EXPRESSION TAG	UNP C1I208
G	493	HIS	-	EXPRESSION TAG	UNP C1I208
G	494	HIS	-	EXPRESSION TAG	UNP C1I208
G	495	HIS	-	EXPRESSION TAG	UNP C1I208
H	426	ASN	SER	ENGINEERED MUTATION	UNP C1I208
H	484	HIS	TYR	ENGINEERED MUTATION	UNP C1I208
H	488	LEU	-	EXPRESSION TAG	UNP C1I208
H	489	GLY	-	EXPRESSION TAG	UNP C1I208
H	490	HIS	-	EXPRESSION TAG	UNP C1I208
H	491	HIS	-	EXPRESSION TAG	UNP C1I208
H	492	HIS	-	EXPRESSION TAG	UNP C1I208
H	493	HIS	-	EXPRESSION TAG	UNP C1I208
H	494	HIS	-	EXPRESSION TAG	UNP C1I208
H	495	HIS	-	EXPRESSION TAG	UNP C1I208

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	29	Total O 29 29	0	0
2	B	29	Total O 29 29	0	0
2	C	31	Total O 31 31	0	0
2	D	24	Total O 24 24	0	0
2	E	30	Total O 30 30	0	0
2	F	34	Total O 34 34	0	0
2	G	23	Total O 23 23	0	0

Continued on next page...

Continued from previous page...

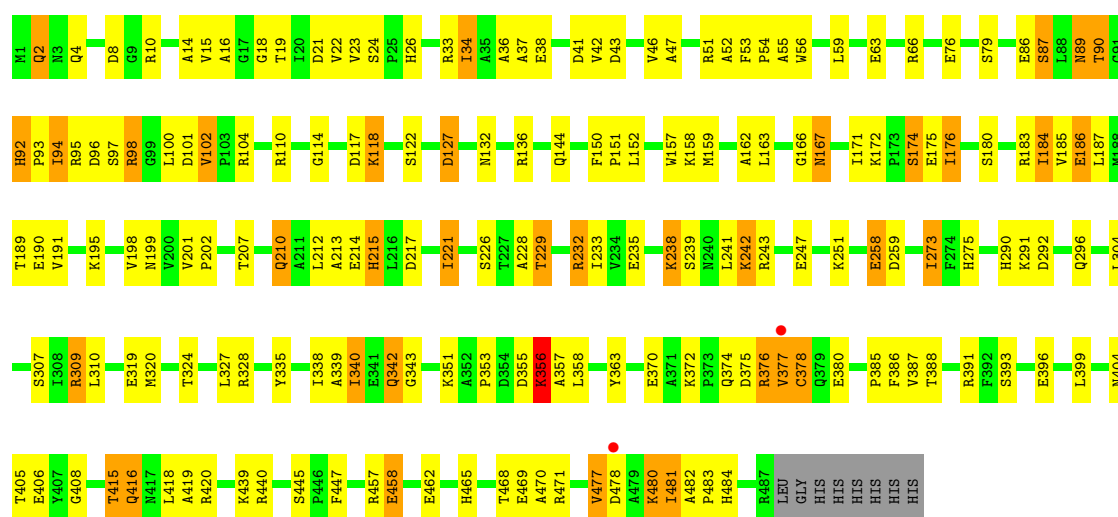
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	H	13	Total	O	0	0
			13	13		

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

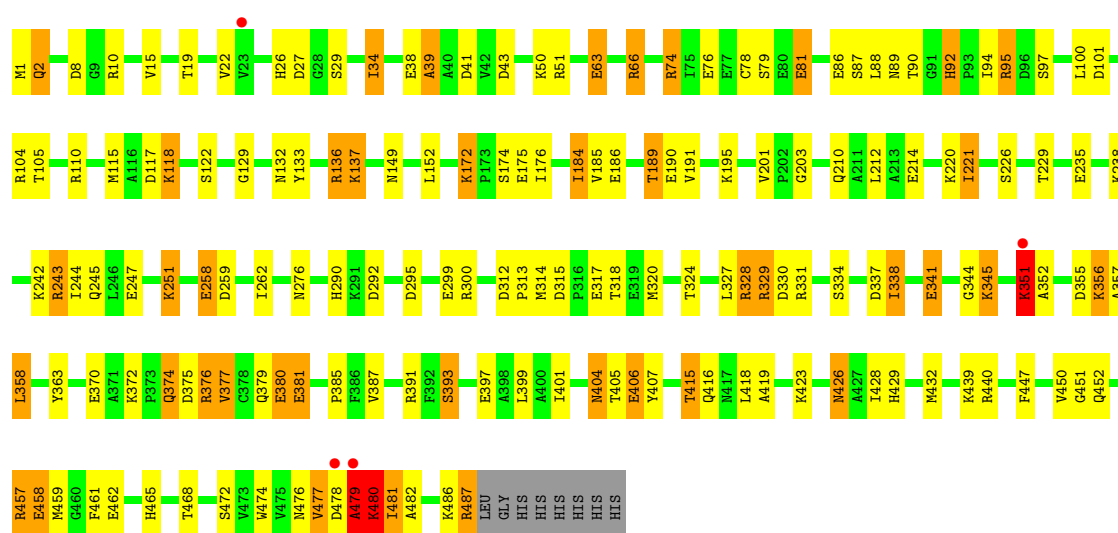
- Molecule 1: Putative gamma-hydroxymuconic semialdehyde dehydrogenase

Chain A:



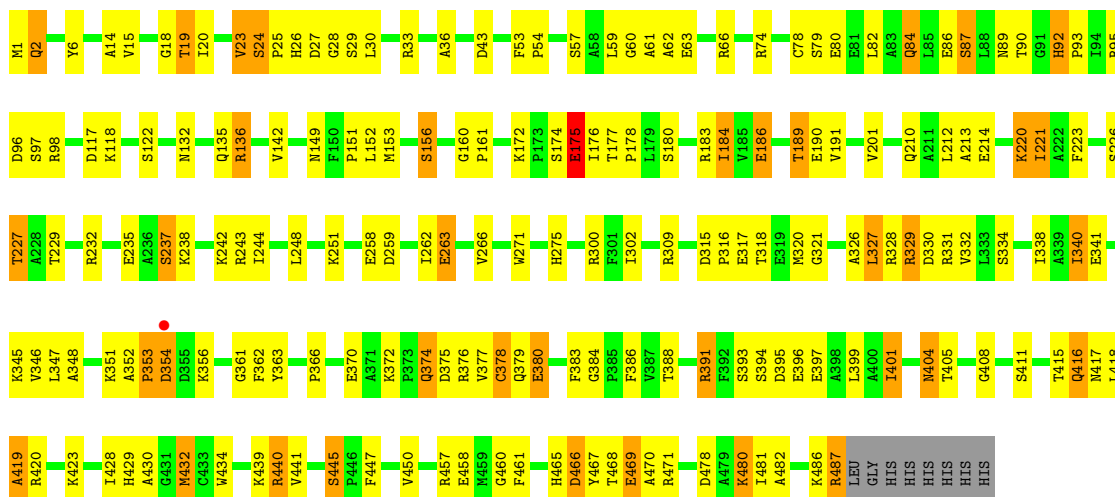
- Molecule 1: Putative gamma-hydroxymuconic semialdehyde dehydrogenase

Chain B:



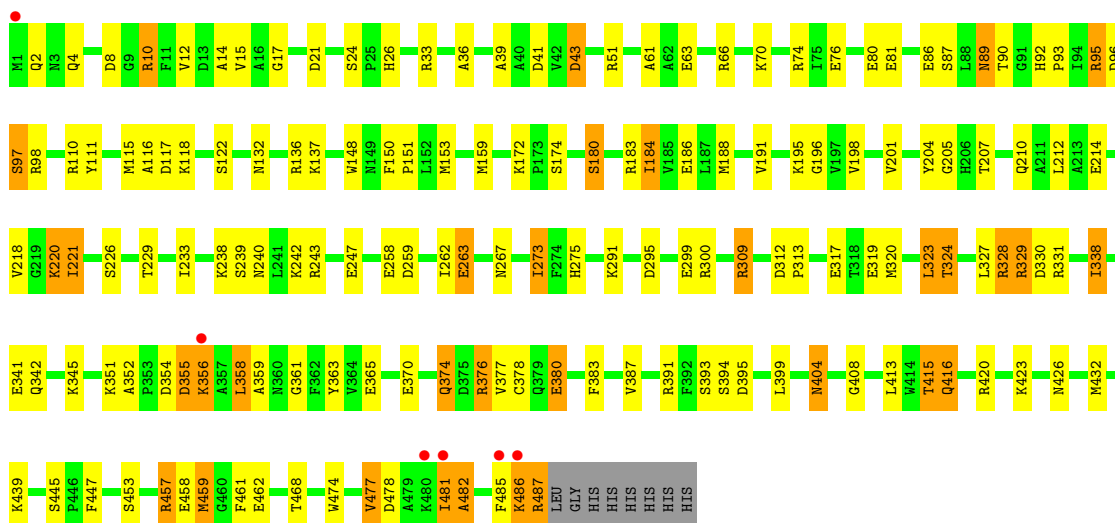
- Molecule 1: Putative gamma-hydroxymuconic semialdehyde dehydrogenase

Chain C:



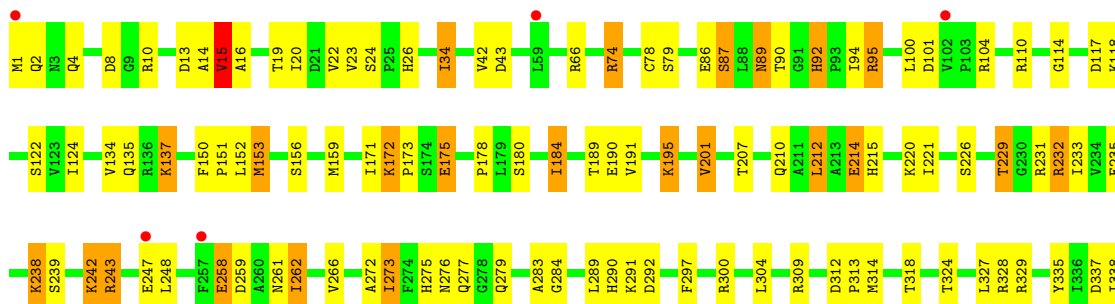
• Molecule 1: Putative gamma-hydroxymuconic semialdehyde dehydrogenase

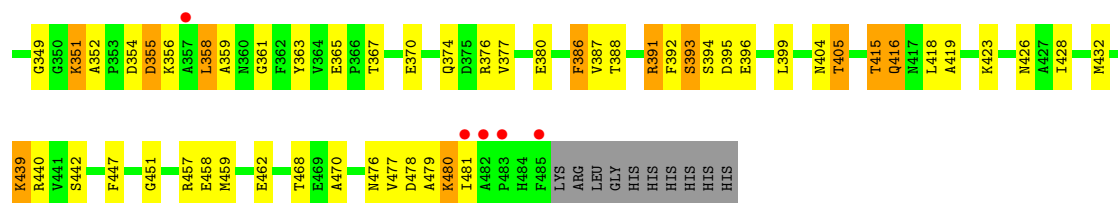
Chain D:



• Molecule 1: Putative gamma-hydroxymuconic semialdehyde dehydrogenase

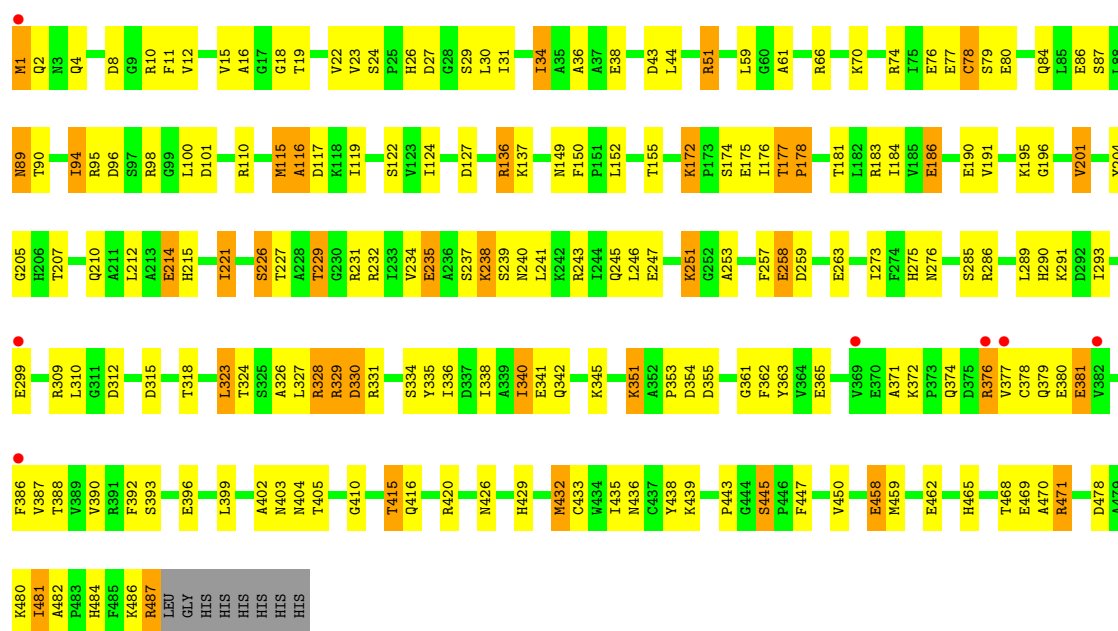
Chain E:





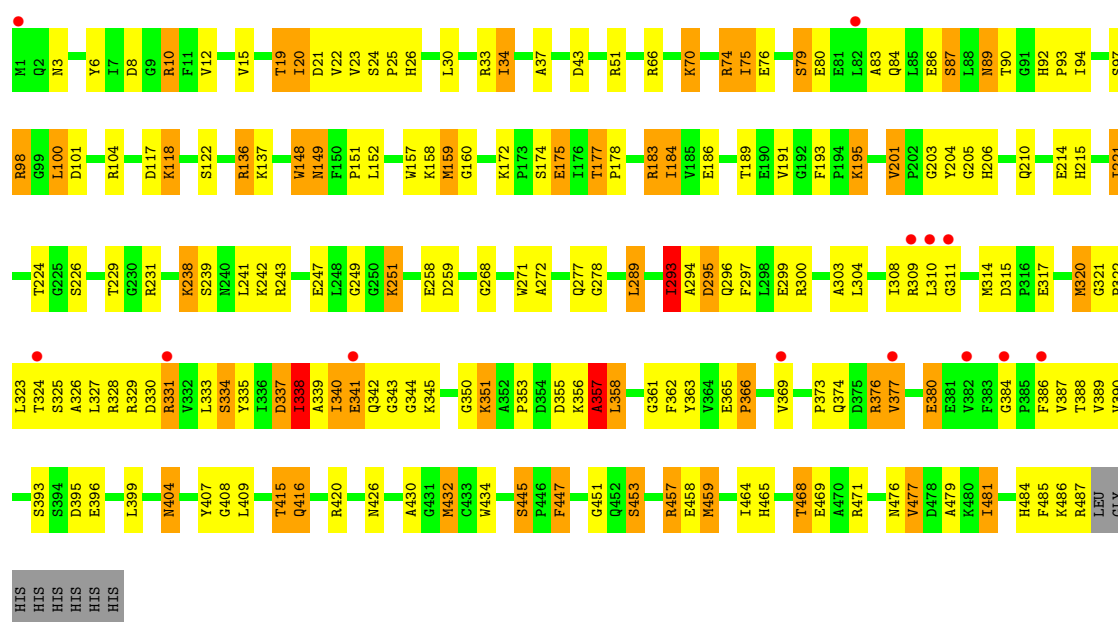
● Molecule 1: Putative gamma-hydroxymuconic semialdehyde dehydrogenase

Chain F:



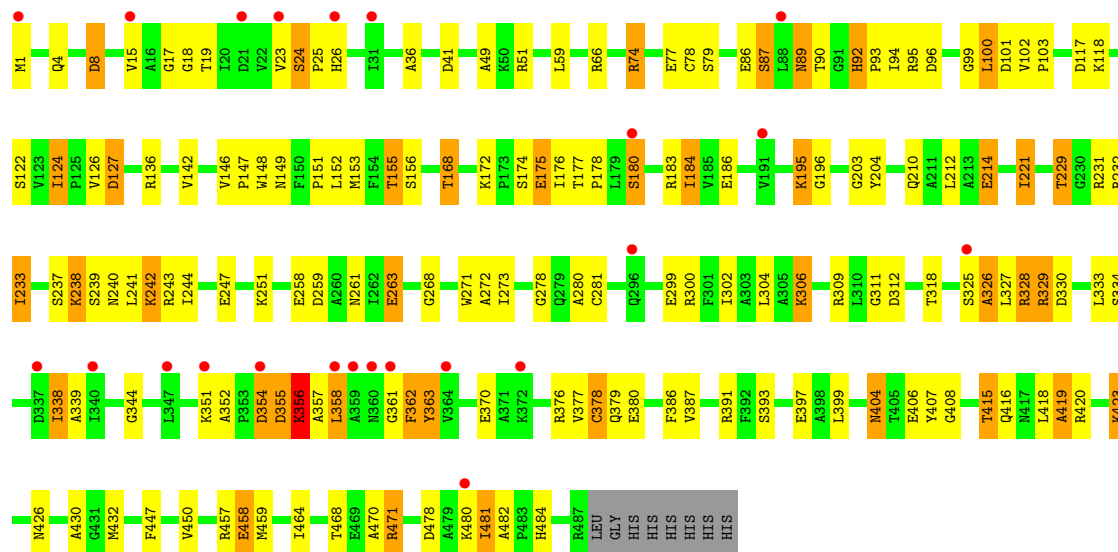
● Molecule 1: Putative gamma-hydroxymuconic semialdehyde dehydrogenase

Chain G:



● Molecule 1: Putative gamma-hydroxymuconic semialdehyde dehydrogenase

Chain H: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	83.53Å 144.25Å 138.29Å 90.00° 93.59° 90.00°	Depositor
Resolution (Å)	41.68 – 2.70 50.03 – 2.70	Depositor EDS
% Data completeness (in resolution range)	93.5 (41.68-2.70) 96.4 (50.03-2.70)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	6.12 (at 2.69Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.4_486)	Depositor
R, R_{free}	0.223 , 0.237 0.217 , 0.229	Depositor DCC
R_{free} test set	4293 reflections (4.99%)	DCC
Wilson B-factor (Å ²)	38.6	Xtriage
Anisotropy	0.313	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 16.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 86063 reflections	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	29686	wwPDB-VP
Average B, all atoms (Å ²)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.50% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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	0.97	2/3761 (0.1%)	0.68	1/5101 (0.0%)
1	B	0.95	1/3772 (0.0%)	0.72	4/5116 (0.1%)
1	C	1.01	2/3770 (0.1%)	0.70	0/5113
1	D	0.98	1/3761 (0.0%)	0.74	5/5101 (0.1%)
1	E	0.98	1/3741 (0.0%)	0.69	1/5076 (0.0%)
1	F	0.90	1/3761 (0.0%)	0.76	7/5101 (0.1%)
1	G	0.90	1/3761 (0.0%)	0.85	8/5101 (0.2%)
1	H	0.92	3/3761 (0.1%)	0.69	1/5101 (0.0%)
All	All	0.95	12/30088 (0.0%)	0.73	27/40810 (0.1%)

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	B	0	2
1	G	0	1
All	All	0	3

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	78	CYS	CB-SG	-10.06	1.65	1.82
1	E	78	CYS	CB-SG	-10.06	1.65	1.82
1	H	78	CYS	CB-SG	-10.06	1.65	1.82
1	F	78	CYS	CB-SG	-10.05	1.65	1.82
1	D	378	CYS	CB-SG	-5.91	1.72	1.81

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	482	ALA	CB-CA-C	16.20	134.40	110.10

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	479	ALA	CB-CA-C	13.03	129.65	110.10
1	G	477	VAL	CB-CA-C	13.02	136.14	111.40
1	G	315	ASP	CB-CA-C	11.93	134.27	110.40
1	F	115	MET	CB-CA-C	-10.98	88.44	110.40

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	479	ALA	Peptide
1	B	480	LYS	Peptide
1	G	477	VAL	Peptide

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3685	0	0	121	4
1	B	3692	0	0	134	18
1	C	3691	0	0	139	2
1	D	3685	0	0	98	2
1	E	3665	0	0	123	2
1	F	3685	0	0	148	10
1	G	3685	0	0	159	0
1	H	3685	0	0	114	15
2	A	29	0	0	17	1
2	B	29	0	0	26	0
2	C	31	0	0	29	0
2	D	24	0	0	15	0
2	E	30	0	0	28	0
2	F	34	0	0	21	0
2	G	23	0	0	27	0
2	H	13	0	0	9	0
All	All	29686	0	0	974	27

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 33.

The worst 5 of 974 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:H:263:GLU:CG	1:H:300:ARG:NH2	1.68	1.51
1:G:10:ARG:NH1	1:G:12:VAL:CG1	1.73	1.50
1:B:238:LYS:NZ	1:F:235:GLU:CG	1.87	1.38
1:A:159:MET:CE	1:A:198:VAL:CG1	2.01	1.36
1:E:74:ARG:NH2	1:E:190:GLU:OE1	1.56	1.35

The worst 5 of 27 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:393:SER:CB	1:H:354:ASP:O[2_545]	0.79	1.41
1:B:356:LYS:NZ	1:F:341:GLU:CB[2_454]	0.88	1.32
1:B:356:LYS:NZ	1:F:341:GLU:CA[2_454]	0.92	1.28
1:A:478:ASP:OD2	1:H:423:LYS:CE[2_445]	1.10	1.10
1:B:393:SER:OG	1:H:354:ASP:O[2_545]	1.14	1.06

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	485/495 (98%)	452 (93%)	31 (6%)	2 (0%)	43	76
1	B	486/495 (98%)	452 (93%)	30 (6%)	4 (1%)	27	58
1	C	486/495 (98%)	440 (90%)	41 (8%)	5 (1%)	22	51
1	D	485/495 (98%)	453 (93%)	31 (6%)	1 (0%)	56	86
1	E	483/495 (98%)	446 (92%)	35 (7%)	2 (0%)	43	76
1	F	485/495 (98%)	443 (91%)	40 (8%)	2 (0%)	43	76
1	G	485/495 (98%)	432 (89%)	45 (9%)	8 (2%)	14	35
1	H	485/495 (98%)	453 (93%)	29 (6%)	3 (1%)	33	66
All	All	3880/3960 (98%)	3571 (92%)	282 (7%)	27 (1%)	30	62

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	178	PRO
1	A	356	LYS
1	C	178	PRO
1	G	341	GLU
1	G	357	ALA

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	376/383 (98%)	308 (82%)	68 (18%)	2	7
1	B	377/383 (98%)	308 (82%)	69 (18%)	2	6
1	C	377/383 (98%)	307 (81%)	70 (19%)	2	6
1	D	376/383 (98%)	308 (82%)	68 (18%)	2	7
1	E	374/383 (98%)	306 (82%)	68 (18%)	2	6
1	F	376/383 (98%)	304 (81%)	72 (19%)	2	6
1	G	376/383 (98%)	303 (81%)	73 (19%)	2	5
1	H	376/383 (98%)	307 (82%)	69 (18%)	2	6
All	All	3008/3064 (98%)	2451 (82%)	557 (18%)	2	6

5 of 557 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	331	ARG
1	E	247	GLU
1	H	212	LEU
1	D	380	GLU
1	E	79	SER

Some 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 chains 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.

5.5 Carbohydrates ⓘ

There are no carbohydrates 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(Å ²)	Q<0.9
1	A	487/495 (98%)	-0.11	2 (0%) 90 93	14, 34, 61, 93	0
1	B	487/495 (98%)	-0.08	4 (0%) 83 87	16, 36, 61, 102	0
1	C	487/495 (98%)	-0.15	1 (0%) 93 96	12, 37, 62, 145	0
1	D	487/495 (98%)	0.04	6 (1%) 75 81	14, 35, 59, 165	0
1	E	485/495 (97%)	0.05	10 (2%) 60 67	16, 37, 61, 113	0
1	F	487/495 (98%)	0.16	7 (1%) 72 77	11, 44, 76, 131	0
1	G	487/495 (98%)	0.22	13 (2%) 52 57	18, 47, 94, 137	0
1	H	487/495 (98%)	0.36	23 (4%) 30 34	16, 45, 79, 148	0
All	All	3894/3960 (98%)	0.06	66 (1%) 67 73	11, 39, 73, 165	0

The worst 5 of 66 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	358	LEU	4.8
1	E	482	ALA	4.6
1	F	376	ARG	4.2
1	F	382	VAL	3.8
1	H	340	ILE	3.8

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 carbohydrates 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.