W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Oct 7, 2023 – 06:13 PM EDT

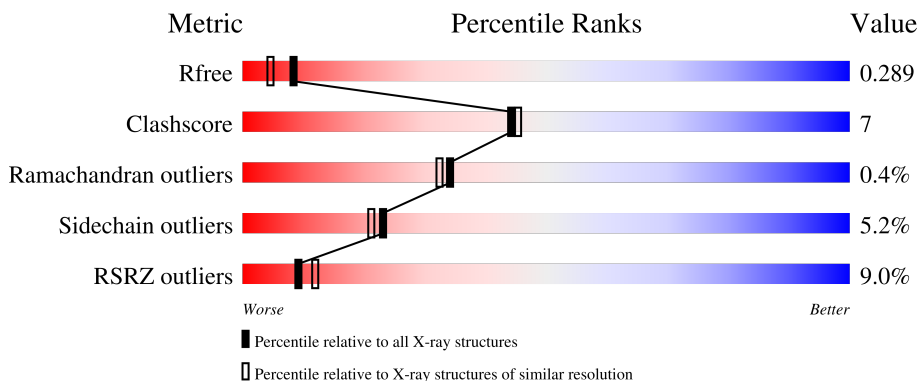|		
PDB ID	:	6DND
Title	:	Crystal structure of wild-type (WT) human Glutamate oxaloacetate transaminase 1 (GOT1)
Authors	:	Assar, Z.; Holt, M.C.; Stein, A.J.; Lairson, L.; Lyssiotis, C.A.
Deposited on	:	2018-06-06
Resolution	:	2.10 Å(reported)

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 2.10 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		0.289
Clashscore		7
Ramachandran outliers		0.4%
Sidechain outliers		5.2%
RSRZ outliers		9.0%

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$_{free}$	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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	411	13% 88% 10% •
1	B	411	5% 85% 12% •

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 6628 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 Aspartate aminotransferase, cytoplasmic.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	409	Total 3237	C 2070	N 562	O 596	S 9	0	4	0
1	B	410	Total 3244	C 2070	N 564	O 601	S 9	0	1	0

- Molecule 2 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).
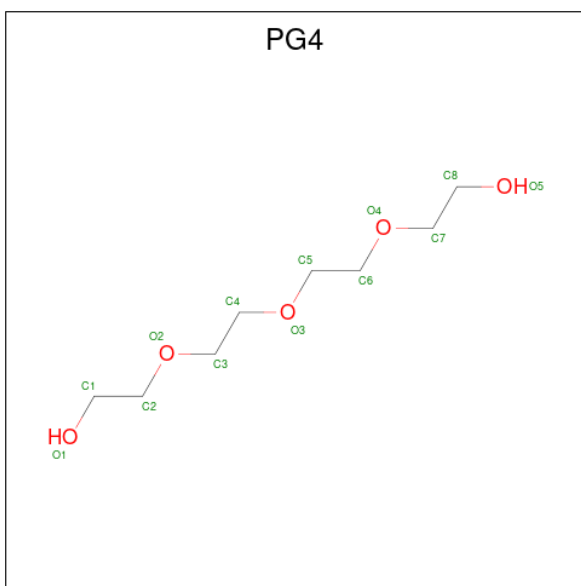
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total 7	C 4	O 3	0	0
2	B	1	Total 7	C 4	O 3	0	0

- Molecule 3 is PYRIDOXAL-5'-PHOSPHATE (three-letter code: PLP) (formula: $C_8H_{10}NO_6P$).

PLP

N1
N
C2A
C6
C2
C5A
C5
C4
C3
O3P
HO
P
O4P
C4A
OH
O3
P
O
O4A
O1P
OH
O2P

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			16	8	1	6	1		
3	B	1	Total	C	N	O	P	0	0
			16	8	1	6	1		

• Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	35	Total	O	0	0
			35	35		
4	B	66	Total	O	0	0
			66	66		

WORLDWIDE
wwPDB
PROTEIN DATA BANK

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: Aspartate aminotransferase, cytoplasmic
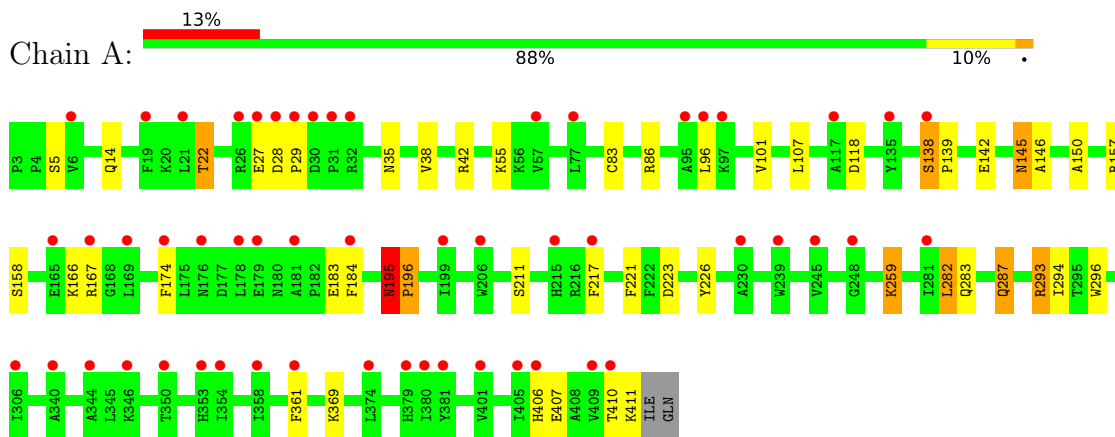
• Molecule 1: Aspartate aminotransferase, cytoplasmic
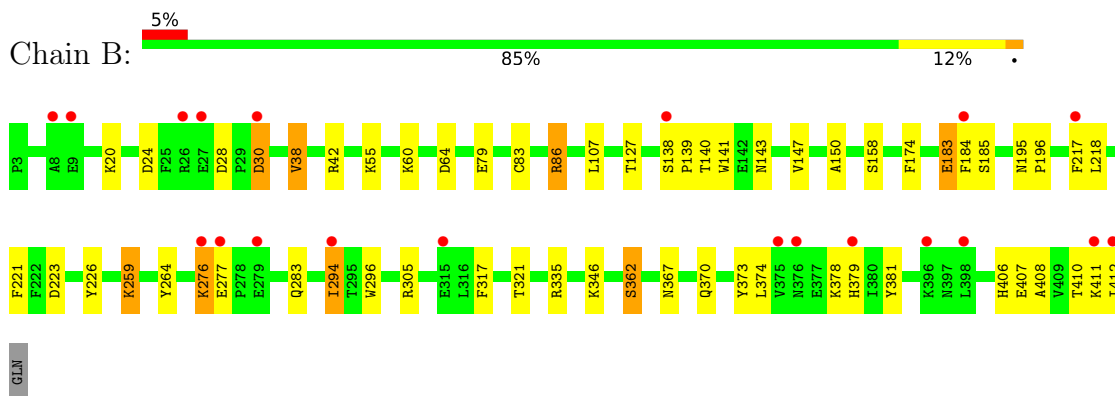
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	91.04Å 96.86Å 97.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	68.85 – 2.10 48.43 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.9 (68.85-2.10) 99.9 (48.43-2.10)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.22 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.208 , 0.269 0.231 , 0.289	Depositor DCC
R_{free} test set	2570 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	50.6	Xtriage
Anisotropy	0.049	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 52.3	EDS
L-test for twinning[2]	$< \lvert L \rvert > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.016 for -h,l,k	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	6628	wwPDB-VP
Average B, all atoms (Å^2)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.71% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< \lvert L \rvert >$, $< L^2 >$ 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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: PLP, PG4

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	0.68	0/3332	0.79	6/4533 (0.1%)
1	B	0.75	1/3328 (0.0%)	0.79	4/4525 (0.1%)
All	All	0.72	1/6660 (0.0%)	0.79	10/9058 (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	A	0	1
1	B	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	79	GLU	CD-OE1	6.00	1.32	1.25

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	195	ASN	C-N-CD	-6.66	105.96	120.60
1	A	287[A]	GLN	CA-CB-CG	-5.96	100.28	113.40
1	A	287[B]	GLN	CA-CB-CG	-5.96	100.28	113.40
1	B	86	ARG	NE-CZ-NH2	5.45	123.02	120.30
1	A	118	ASP	CB-CG-OD1	5.41	123.17	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	195	ASN	Peptide
1	B	195	ASN	Peptide

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	3237	0	3147	43	0
1	B	3244	0	3179	47	0
2	A	7	0	8	0	0
2	B	7	0	9	0	0
3	A	16	0	7	3	0
3	B	16	0	7	4	0
4	A	35	0	0	2	0
4	B	66	0	0	2	0
All	All	6628	0	6357	85	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.

The worst 5 of 85 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:30:ASP:CG	1:B:379:HIS:CD2	1.97	1.37
1:A:293:ARG:HD3	4:A:631:HOH:O	1.24	1.29
1:B:30:ASP:OD2	1:B:379:HIS:CD2	1.94	1.19
1:A:22:THR:HG22	1:A:35:ASN:HD21	1.04	1.15
1:B:259:LYS:HE2	3:B:502:PLP:O4A	1.46	1.13

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	411/411 (100%)	393 (96%)	16 (4%)	2 (0%)	29	26
1	B	409/411 (100%)	395 (97%)	13 (3%)	1 (0%)	47	49
All	All	820/822 (100%)	788 (96%)	29 (4%)	3 (0%)	34	32

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	196	PRO
1	B	196	PRO
1	A	138	SER

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	343/350 (98%)	321 (94%)	22 (6%)	17	14
1	B	348/350 (99%)	332 (95%)	16 (5%)	27	26
All	All	691/700 (99%)	653 (94%)	38 (6%)	23	19

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	183	GLU
1	B	346	LYS
1	B	221	PHE
1	B	277	GLU
1	B	374	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	376	ASN
1	B	379	HIS
1	B	406	HIS
1	A	195	ASN
1	B	128	ASN

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.

5.5   Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6   Ligand geometry ⓘ

4 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	#$|Z| > 2$
3	PLP	B	502	-	16,16,16	2.89	6 (37%)	20,23,23	1.97	6 (30%)
2	PG4	B	501	-	6,6,12	0.66	0	5,5,11	0.88	0
3	PLP	A	502	-	16,16,16	2.97	3 (18%)	20,23,23	2.11	7 (35%)
2	PG4	A	501	-	6,6,12	0.61	0	5,5,11	0.68	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
3	PLP	B	502	-	-	4/8/8/8	0/1/1/1
2	PG4	B	501	-	-	3/4/4/10	-
3	PLP	A	502	-	-	3/8/8/8	0/1/1/1
2	PG4	A	501	-	-	2/4/4/10	-

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	502	PLP	C4-C5	7.55	1.51	1.42
3	B	502	PLP	C3-C2	7.23	1.48	1.40
3	A	502	PLP	C3-C2	6.16	1.47	1.40
3	A	502	PLP	C4-C3	5.95	1.49	1.40
3	B	502	PLP	C4-C5	5.55	1.49	1.42

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	502	PLP	O4P-C5A-C5	5.32	119.48	109.35
3	A	502	PLP	O4P-C5A-C5	4.62	118.16	109.35
3	A	502	PLP	C4-C3-C2	-4.58	117.36	120.19
3	A	502	PLP	C6-N1-C2	3.49	125.63	119.17
3	B	502	PLP	C3-C4-C5	-3.13	115.86	118.26

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	502	PLP	C4-C5-C5A-O4P
3	A	502	PLP	C6-C5-C5A-O4P
3	B	502	PLP	C4-C5-C5A-O4P
3	B	502	PLP	C6-C5-C5A-O4P
3	B	502	PLP	C3-C4-C4A-O4A

There are no ring outliers.

2 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	502	PLP	4	0
3	A	502	PLP	3	0

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	409/411 (99%)	0.86	54 (13%)	3	4	40, 60, 82, 99	0
1	B	410/411 (99%)	0.46	20 (4%)	29	35	36, 52, 74, 91	0
All	All	819/822 (99%)	0.66	74 (9%)	9	12	36, 56, 81, 99	0

The worst 5 of 74 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	179	GLU	5.7
1	A	217	PHE	5.5
1	A	27	GLU	4.8
1	B	138	SER	4.6
1	B	412	ILE	4.3

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 ⓘ

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
2	PG4	A	501	7/13	0.70	0.16	67,70,74,76	0
2	PG4	B	501	7/13	0.80	0.21	51,59,74,83	0
3	PLP	A	502	16/16	0.97	0.22	39,73,97,105	10
3	PLP	B	502	16/16	0.97	0.16	37,62,71,71	10

6.5 Other polymers ⓘ

There are no such residues in this entry.