



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 2, 2023 – 10:51 AM EDT

PDB ID : 6MUJ
Title : Formylglycine generating enzyme bound to copper
Authors : Lafrance-Vanasse, J.; Appel, M.J.; Tsai, C.-L.; Bertozzi, C.; Tainer, J.A.
Deposited on : 2018-10-23
Resolution : 2.25 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 ⓘ 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 : **FAILED**
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtrriage (Phenix) : 1.13
EDS : **FAILED**
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) : 2.35.1

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 2.25 Å.

There are no overall percentile quality scores available for this entry.

MolProbity and EDS failed to run properly - the sequence quality summary graphics cannot be shown.

2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 12000 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 Formylglycine-generating enzyme.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	289	Total 2252	C 1404	N 434	O 408	S 6	0	2	0
1	B	294	Total 2268	C 1414	N 433	O 415	S 6	0	0	0
1	C	289	Total 2236	C 1394	N 428	O 408	S 6	0	0	0
1	D	293	Total 2262	C 1410	N 432	O 414	S 6	0	0	0
1	E	291	Total 2245	C 1400	N 427	O 412	S 6	0	0	0

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	GLY	-	expression tag	UNP Q9F3C7
A	-3	ILE	-	expression tag	UNP Q9F3C7
A	-2	ASP	-	expression tag	UNP Q9F3C7
A	-1	PRO	-	expression tag	UNP Q9F3C7
A	0	PHE	-	expression tag	UNP Q9F3C7
A	1	THR	-	expression tag	UNP Q9F3C7
B	-4	GLY	-	expression tag	UNP Q9F3C7
B	-3	ILE	-	expression tag	UNP Q9F3C7
B	-2	ASP	-	expression tag	UNP Q9F3C7
B	-1	PRO	-	expression tag	UNP Q9F3C7
B	0	PHE	-	expression tag	UNP Q9F3C7
B	1	THR	-	expression tag	UNP Q9F3C7
C	-4	GLY	-	expression tag	UNP Q9F3C7
C	-3	ILE	-	expression tag	UNP Q9F3C7
C	-2	ASP	-	expression tag	UNP Q9F3C7
C	-1	PRO	-	expression tag	UNP Q9F3C7
C	0	PHE	-	expression tag	UNP Q9F3C7
C	1	THR	-	expression tag	UNP Q9F3C7
D	-4	GLY	-	expression tag	UNP Q9F3C7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-3	ILE	-	expression tag	UNP Q9F3C7
D	-2	ASP	-	expression tag	UNP Q9F3C7
D	-1	PRO	-	expression tag	UNP Q9F3C7
D	0	PHE	-	expression tag	UNP Q9F3C7
D	1	THR	-	expression tag	UNP Q9F3C7
E	-4	GLY	-	expression tag	UNP Q9F3C7
E	-3	ILE	-	expression tag	UNP Q9F3C7
E	-2	ASP	-	expression tag	UNP Q9F3C7
E	-1	PRO	-	expression tag	UNP Q9F3C7
E	0	PHE	-	expression tag	UNP Q9F3C7
E	1	THR	-	expression tag	UNP Q9F3C7

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Ca 1 1	0	0
2	B	1	Total Ca 1 1	0	0
2	C	1	Total Ca 1 1	0	0
2	D	1	Total Ca 1 1	0	0
2	E	1	Total Ca 1 1	0	0

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	4	Total Cu 4 4	0	0
3	B	3	Total Cu 3 3	0	0
3	C	4	Total Cu 4 4	0	0
3	D	5	Total Cu 5 5	0	0
3	E	4	Total Cu 4 4	0	0

- Molecule 4 is IMIDAZOLE (three-letter code: IMD) (formula: C₃H₅N₂).



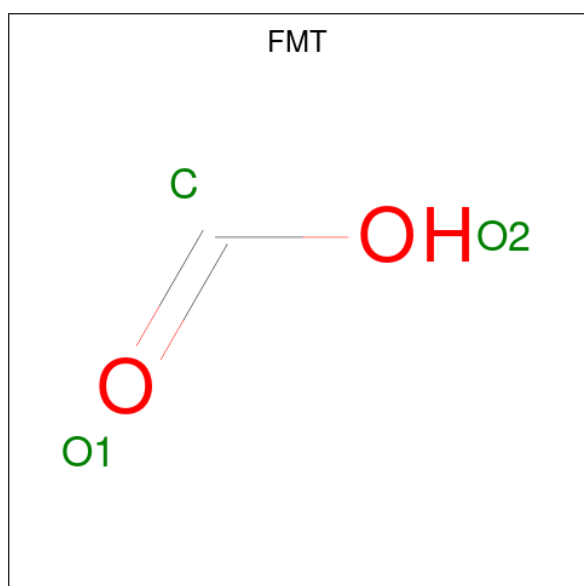
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	N	0	0
			5	3	2		
4	B	1	Total	C	N	0	0
			5	3	2		
4	D	1	Total	C	N	0	0
			5	3	2		
4	E	1	Total	C	N	0	0
			5	3	2		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



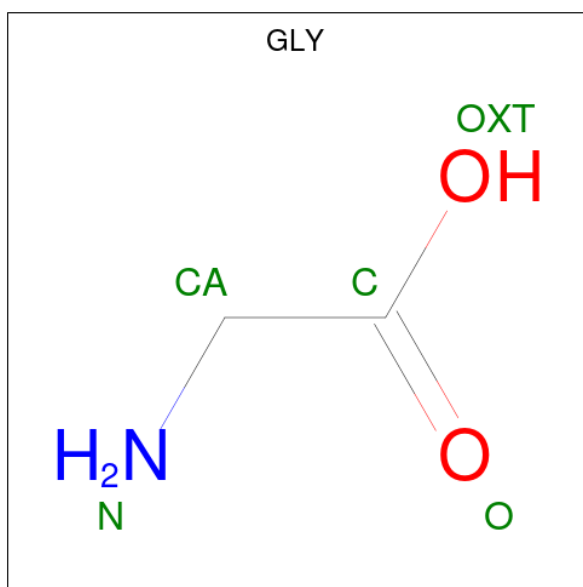
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total	C	O	0	0
			6	3	3		
5	C	1	Total	C	O	0	0
			6	3	3		
5	C	1	Total	C	O	0	0
			6	3	3		
5	D	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is FORMIC ACID (three-letter code: FMT) (formula: CH₂O₂).



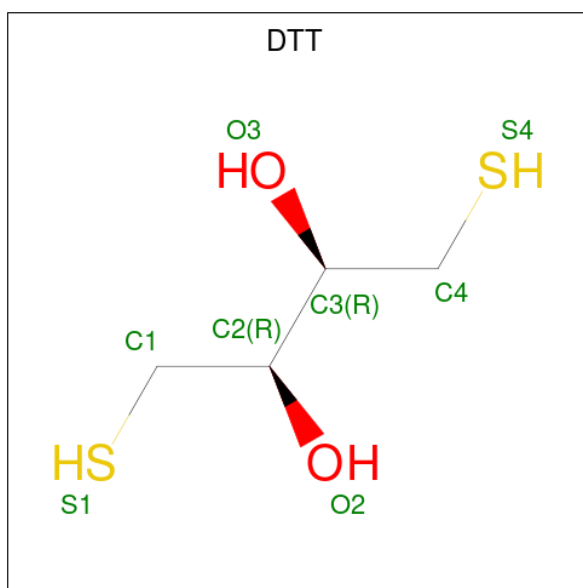
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	C	1	Total	C	O	0	0
			3	1	2		

- Molecule 7 is GLYCINE (three-letter code: GLY) (formula: C₂H₅NO₂).



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

- Molecule 8 is 2,3-DIHYDROXY-1,4-DITHIOBUTANE (three-letter code: DTT) (formula: C₄H₁₀O₂S₂).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	S		
8	D	1	8	4	2	2	0	0

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	104	Total 104	O 104	0	0
9	B	125	Total 125	O 125	0	0
9	C	155	Total 155	O 155	0	0
9	D	134	Total 134	O 134	0	0
9	E	135	Total 135	O 135	0	0

MolProbity and EDS failed to run properly - this section is therefore empty.

3 Data and refinement statistics i

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	140.01Å 140.01Å 217.51Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	38.74 – 2.25	Depositor
% Data completeness (in resolution range)	84.4 (38.74-2.25)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.60 (at 2.24Å)	Xtrriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, R_{free}	0.202 , 0.246	Depositor
Wilson B-factor (Å ²)	37.9	Xtrriage
Anisotropy	0.539	Xtrriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.006 for -h,-k,l	Xtrriage
Total number of atoms	12000	wwPDB-VP
Average B, all atoms (Å ²)	48.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 29.16 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.6271e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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.

4 Model quality [i](#)

4.1 Standard geometry [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.2 Too-close contacts [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3 Torsion angles [i](#)

4.3.1 Protein backbone [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3.2 Protein sidechains [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3.3 RNA [i](#)

MolProbity failed to run properly - this section is therefore empty.

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

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

4.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

4.6 Ligand geometry [i](#)

Of 36 ligands modelled in this entry, 25 are monoatomic - leaving 11 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
7	GLY	D	401	-	3,3,4	0.81	0	0,2,4	-	-
5	GOL	D	402	-	5,5,5	0.54	0	5,5,5	0.58	0
4	IMD	E	406	-	3,5,5	0.50	0	4,5,5	0.51	0
4	IMD	B	501	-	3,5,5	0.40	0	4,5,5	0.61	0
5	GOL	C	509	-	5,5,5	0.53	0	5,5,5	0.58	0
5	GOL	C	508	-	5,5,5	0.69	0	5,5,5	0.82	0
6	FMT	C	507	-	2,2,2	0.63	0	1,1,1	0.17	0
4	IMD	D	410	-	3,5,5	0.36	0	4,5,5	0.76	0
8	DTT	D	409	-	7,7,7	1.09	0	4,8,8	1.62	1 (25%)
4	IMD	A	506	-	3,5,5	0.27	0	4,5,5	0.88	0
5	GOL	C	501	-	5,5,5	0.41	0	5,5,5	0.41	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
7	GLY	D	401	-	-	0/0/1/2	-
5	GOL	D	402	-	-	1/4/4/4	-
4	IMD	E	406	-	-	-	0/1/1/1
5	GOL	C	509	-	-	4/4/4/4	-
4	IMD	B	501	-	-	-	0/1/1/1
5	GOL	C	508	-	-	3/4/4/4	-
4	IMD	D	410	-	-	-	0/1/1/1
8	DTT	D	409	-	-	4/8/8/8	-
4	IMD	A	506	-	-	-	0/1/1/1
5	GOL	C	501	-	-	4/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	D	409	DTT	C3-C4-S4	2.26	121.04	114.47

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	C	501	GOL	O1-C1-C2-O2
5	C	501	GOL	O1-C1-C2-C3
5	C	508	GOL	O1-C1-C2-O2
5	C	508	GOL	O1-C1-C2-C3
5	C	509	GOL	C1-C2-C3-O3

There are no ring outliers.

No monomer is involved in short contacts.

4.7 Other polymers [i](#)

There are no such residues in this entry.

4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

5 Fit of model and data [i](#)

5.1 Protein, DNA and RNA chains [i](#)

EDS failed to run properly - this section is therefore empty.

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

EDS failed to run properly - this section is therefore empty.

5.3 Carbohydrates [i](#)

EDS failed to run properly - this section is therefore empty.

5.4 Ligands [i](#)

EDS failed to run properly - this section is therefore empty.

5.5 Other polymers [i](#)

EDS failed to run properly - this section is therefore empty.