wwPDB X-ray Structure Validation Summary Report (i)

May 15, 2020 – 09:51 am BST

PDB ID	:	1DKA
Title	:	DIALKYLGLYCINE DECARBOXYLASE STRUCTURE: BIFUNCTIONAL ACTIVE SITE AND ALKALI METAL BINDING SITES
Authors	:	Toney, M.D.; Hohenester, E.; Jansonius, J.N.
Deposited on	:	1993-06-18
Resolution	:	2.60 Å(reported)

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

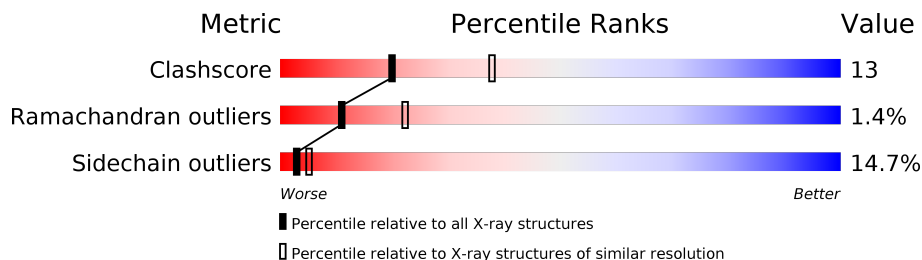|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.60 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
Clashscore		13
Ramachandran outliers		1.4%
Sidechain outliers		14.7%

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))
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)

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 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	433	62% 27% 8% ·

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 3426 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 2,2-DIALKYLGLYCINE DECARBOXYLASE (PYRUVATE).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	431	Total	C	N	O	S	0	0	0
			3252	2051	576	607	18			

There are 2 discrepancies between the modelled and reference sequences:

| Chain	Residue	Modelled	Actual	Comment	Reference
A	15	HIS	GLN	CONFLICT	UNP P16932
A	81	GLU	GLY	CONFLICT	UNP P16932

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

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

- Molecule 3 is POTASSIUM ION (three-letter code: K) (formula: K).

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

- Molecule 4 is PYRIDOXAL-5'-PHOSPHATE (three-letter code: PLP) (formula: $C_8H_{10}NO_6P$).

PLP

N1

C6

C2A

C2

C5A

C5

C4

C3

OH

O3

O3P

HO

P

O4P

C4A

O1P

O

OH

O2P

O4A
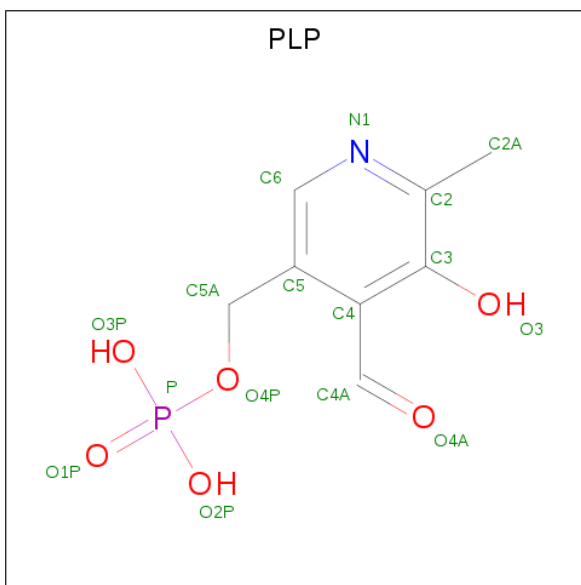
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
4	A	1	15	8	1	5	1	0	0

- Molecule 5 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: $C_6H_{13}NO_4S$).

MES

O1

C2

O

C6

C3

C5

H

N+

N4

C7

C8

S

O2S

O

S

O

O1S

O-

O3S
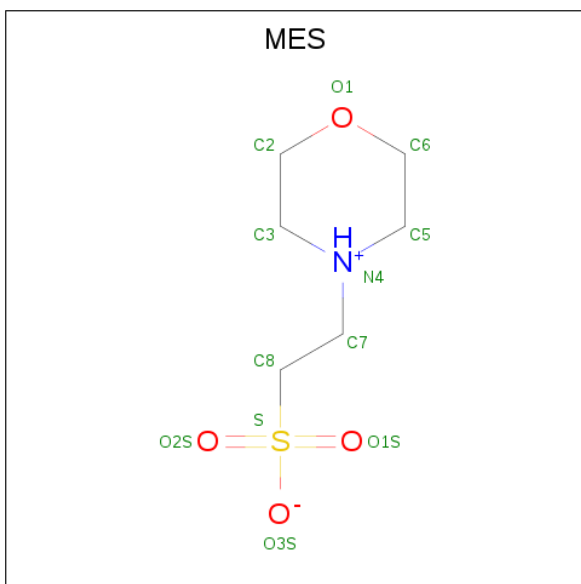
| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
5	A	1	12	6	1	4	1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	145	Total 145	O 145	0	0

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 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: 2,2-DIALKYLGLYCINE DECARBOXYLASE (PYRUVATE)
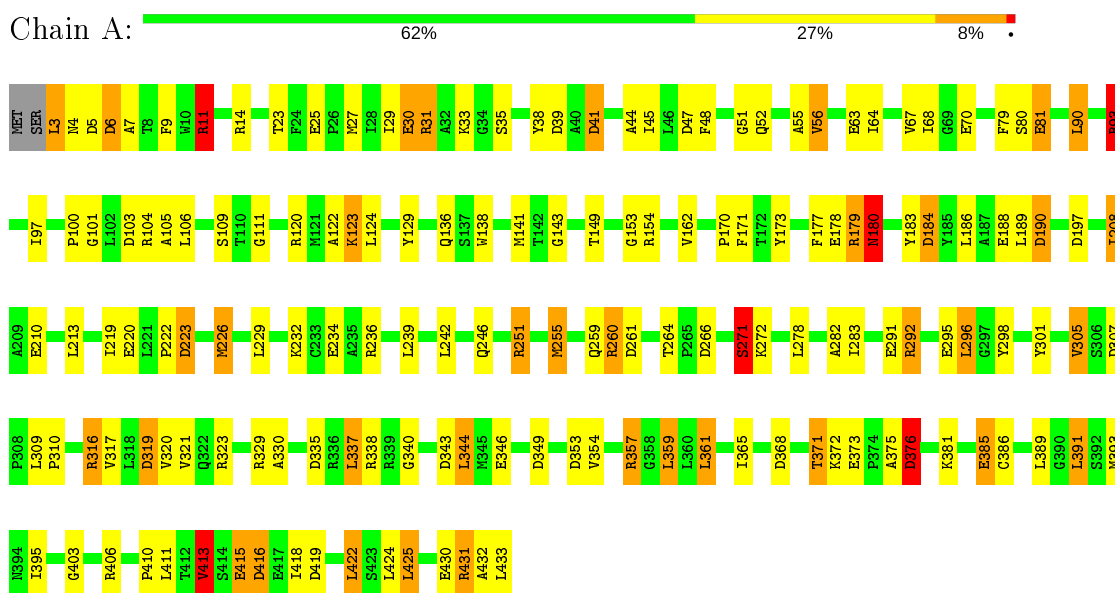
4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 64 2 2	Depositor
Cell constants a, b, c, α, β, γ	152.70Å 152.70Å 86.60Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	8.00 − 2.60	Depositor
% Data completeness (in resolution range)	(Not available) (8.00-2.60)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	TNT	Depositor
R, R_{free}	0.176 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3426	wwPDB-VP
Average B, all atoms (Å²)	21.0	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, K, MES, PLP

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	0.99	21/3309 (0.6%)	1.41	52/4478 (1.2%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	51	GLY	C-N	9.59	1.56	1.34
1	A	52	GLN	C-N	7.69	1.51	1.34
1	A	220	GLU	CD-OE1	7.37	1.33	1.25
1	A	178	GLU	CD-OE2	6.35	1.32	1.25
1	A	70	GLU	CD-OE2	6.26	1.32	1.25

The worst 5 of 52 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	6	ASP	CB-CG-OD1	8.70	126.13	118.30
1	A	307	ASP	CB-CG-OD2	-8.66	110.51	118.30
1	A	190	ASP	CB-CG-OD2	-8.44	110.70	118.30
1	A	6	ASP	CB-CG-OD2	-8.21	110.91	118.30
1	A	307	ASP	CB-CG-OD1	7.72	125.25	118.30

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	3252	0	3277	82	0
2	A	1	0	0	0	0
3	A	1	0	0	0	0
4	A	15	0	6	0	0
5	A	12	0	13	1	0
6	A	145	0	0	10	0
All	All	3426	0	3296	83	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 13.

The worst 5 of 83 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:410:PRO:O	1:A:413:VAL:HG22	1.83	0.79
1:A:338:ARG:HD3	6:A:630:HOH:O	1.85	0.75
1:A:31:ARG:HD3	6:A:609:HOH:O	1.85	0.75
1:A:309:LEU:HB3	1:A:310:PRO:HD3	1.70	0.73
1:A:3:LEU:HD12	1:A:4:ASN:H	1.55	0.72

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	429/433 (99%)	401 (94%)	22 (5%)	6 (1%)	11 22

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	55	ALA
1	A	143	GLY

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	376	ASP
1	A	6	ASP
1	A	271	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	334/336 (99%)	285 (85%)	49 (15%)	3	5

5 of 49 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	223	ASP
1	A	278	LEU
1	A	422	LEU
1	A	255	MET
1	A	292	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	15	HIS
1	A	180	ASN
1	A	259	GLN

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 carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	#$|Z| > 2$
4	PLP	A	437	1	15,15,16	1.27	1 (6%)	20,22,23	2.40	10 (50%)
5	MES	A	434	-	12,12,12	0.56	0	14,16,16	0.94	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
4	PLP	A	437	1	-	0/6/6/8	0/1/1/1
5	MES	A	434	-	-	5/6/14/14	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	437	PLP	C4A-C4	-2.97	1.45	1.51

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	437	PLP	C6-C5-C4	5.67	122.62	118.16
4	A	437	PLP	C2A-C2-C3	3.55	125.27	120.89
4	A	437	PLP	O4P-C5A-C5	3.30	115.64	109.35
4	A	437	PLP	C5-C6-N1	-3.27	118.37	123.82
4	A	437	PLP	C6-N1-C2	3.13	124.97	119.17

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	434	MES	C8-C7-N4-C3
5	A	434	MES	C7-C8-S-O2S
5	A	434	MES	N4-C7-C8-S
5	A	434	MES	C8-C7-N4-C5
5	A	434	MES	C7-C8-S-O1S

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	434	MES	1	0

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.