wwPDB X-ray Structure Validation Summary Report (i)

Oct 18, 2023 – 05:06 AM EDT

PDB ID	:	2CVO
Title	:	Crystal structure of putative N-acetyl-gamma-glutamyl-phosphate reductase (AK071544) from rice (Oryza sativa)
Authors	:	Nonaka, T.; Kita, A.; Miura-Ohnuma, J.; Katoh, E.; Inagaki, N.; Yamazaki, T.; Miki, K.
Deposited on	:	2005-06-10
Resolution	:	2.20 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

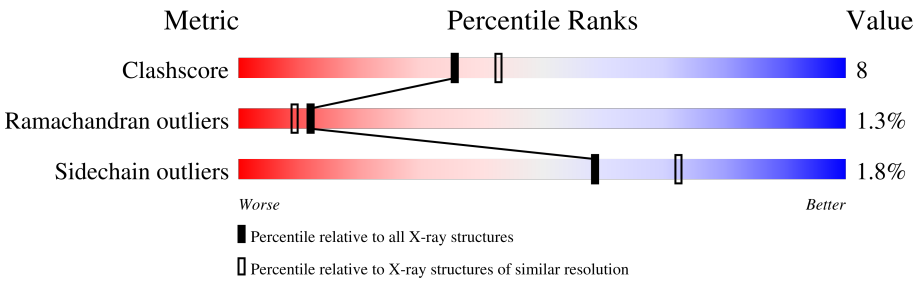
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.20 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))
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	366	80% 13% • 5%
1	B	366	78% 15% • 5%
1	C	366	80% 14% • 5%
1	D	366	77% 17% • 6%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 11554 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 putative Semialdehyde dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	348	2685	1700	469	503	13	0	0	0
1	B	348	2685	1700	469	503	13	0	0	0
1	C	348	2685	1700	469	503	13	0	0	0
1	D	345	2666	1689	465	499	13	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	145	CSO	CYS	modified residue	UNP Q6AV34
B	145	CSO	CYS	modified residue	UNP Q6AV34
C	145	CSO	CYS	modified residue	UNP Q6AV34
D	145	CSO	CYS	modified residue	UNP Q6AV34

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	186	186	186	0	0
2	B	204	204	204	0	0
2	C	245	245	245	0	0
2	D	198	198	198	0	0

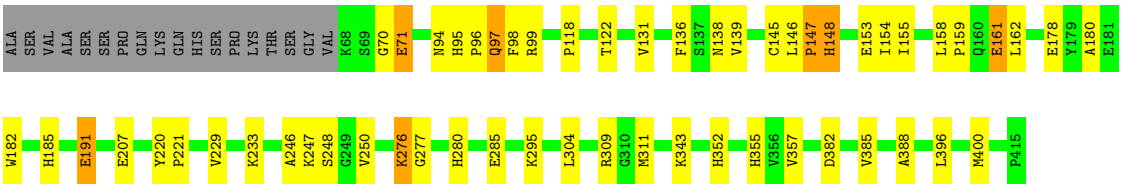
3 Residue-property plots (i)

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.

Note EDS was not executed.

• Molecule 1: putative Semialdehyde dehydrogenase

Chain A:



• Molecule 1: putative Semialdehyde dehydrogenase

Chain B:

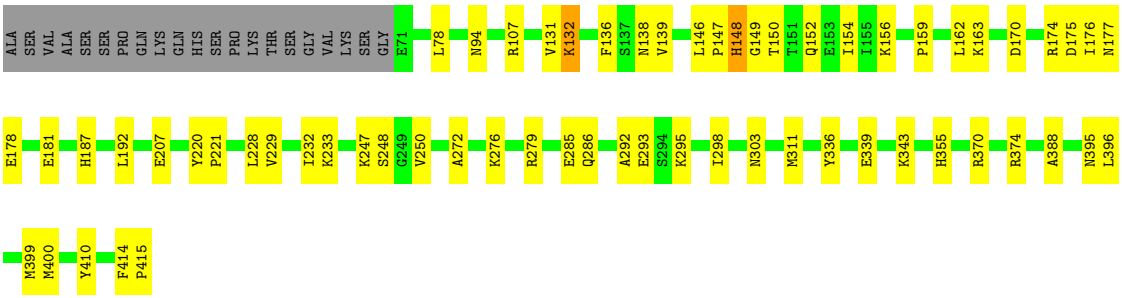
• Molecule 1: putative Semialdehyde dehydrogenase

Chain C:

• Molecule 1: putative Semialdehyde dehydrogenase

Chain D:

77%
17%
6%


ALA SER VAL ALA SER SER PRO GLN LYS GLN HIS SER PRO LYS THR SER GLY VAL LYS SER GLY E71 L78 N94 R107 V131 K132 F136 S137 N138 V139 L146 P147 H148 G149 T150 T151 Q152 E153 I154 I155 K156 P159 L162 K163 D170 R174 D175 I176 N177
E178 E181 H187 L192 E207 Y220 P221 L228 V229 I232 K233 K247 S248 G249 V250 A272 K276 R279 E285 Q286 A292 E293 S294 K295 I298 N303 M311 Y336 E339 K343 H355 R370 R374 A388 N395 L396
M399 M400 Y410 F414 P415

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α, β, γ	86.11Å 86.11Å 316.33Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.93 – 2.20	Depositor
% Data completeness (in resolution range)	98.2 (19.93-2.20)	Depositor
R_{merge}	0.04	Depositor
R_{sym}	0.04	Depositor
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.171 , 0.213	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	11554	wwPDB-VP
Average B, all atoms (Å^2)	15.0	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: CSO

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	0.30	0/2730	0.59	0/3697
1	B	0.31	0/2730	0.61	0/3697
1	C	0.31	0/2730	0.60	0/3697
1	D	0.31	0/2711	0.59	0/3673
All	All	0.31	0/10901	0.60	0/14764

There are no bond length outliers.

There are no bond angle outliers.

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	2685	0	2697	41	0
1	B	2685	0	2697	51	0
1	C	2685	0	2697	42	0
1	D	2666	0	2676	53	0
2	A	186	0	0	6	0
2	B	204	0	0	7	0
2	C	245	0	0	3	0
2	D	198	0	0	4	0
All	All	11554	0	10767	174	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 8.

The worst 5 of 174 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:239:VAL:HG21	1:B:296:VAL:HG22	1.54	0.89
1:D:107:ARG:HB2	1:D:107:ARG:NH1	1.89	0.88
1:B:146:LEU:HD13	1:B:154:ILE:HD12	1.60	0.83
1:C:241:ASN:HD22	1:D:374:ARG:HE	1.24	0.81
1:A:146:LEU:HB3	1:A:147:PRO:HD2	1.64	0.80

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	345/366 (94%)	323 (94%)	16 (5%)	6 (2%)	9 6
1	B	345/366 (94%)	329 (95%)	9 (3%)	7 (2%)	7 4
1	C	345/366 (94%)	333 (96%)	10 (3%)	2 (1%)	25 26
1	D	342/366 (93%)	325 (95%)	14 (4%)	3 (1%)	17 16
All	All	1377/1464 (94%)	1310 (95%)	49 (4%)	18 (1%)	12 9

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	148	HIS
1	B	71	GLU
1	B	147	PRO
1	A	70	GLY
1	B	70	GLY

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	289/304 (95%)	282 (98%)	7 (2%)	49	62
1	B	289/304 (95%)	285 (99%)	4 (1%)	67	80
1	C	289/304 (95%)	282 (98%)	7 (2%)	49	62
1	D	287/304 (94%)	284 (99%)	3 (1%)	76	86
All	All	1154/1216 (95%)	1133 (98%)	21 (2%)	59	72

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	210	ASN
1	C	255	ARG
1	D	148	HIS
1	D	132	LYS
1	C	247	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	312	GLN
1	D	286	GLN
1	C	97	GLN
1	D	346	ASN
1	C	411	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)

4 non-standard protein/DNA/RNA residues 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	$\#\|Z\| > 2$
1	CSO	B	145	1	3,6,7	1.30	1 (33%)	0,6,8	-	-
1	CSO	A	145	1	3,6,7	1.36	1 (33%)	0,6,8	-	-
1	CSO	C	145	1	3,6,7	1.35	1 (33%)	0,6,8	-	-
1	CSO	D	145	1	3,6,7	1.20	0	0,6,8	-	-

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
1	CSO	B	145	1	-	0/1/5/7	-
1	CSO	A	145	1	-	0/1/5/7	-
1	CSO	C	145	1	-	0/1/5/7	-
1	CSO	D	145	1	-	0/1/5/7	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	145	CSO	CB-CA	2.13	1.58	1.53
1	A	145	CSO	CB-CA	2.12	1.58	1.53
1	B	145	CSO	CB-CA	2.00	1.58	1.53

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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.