



wwPDB X-ray Structure Validation Summary Report ⓘ

May 21, 2020 – 10:19 am BST

PDB ID : 2IOF
Title : Crystal structure of phosphonoacetaldehyde hydrolase with sodium borohydride-reduced substrate intermediate
Authors : Allen, K.A.; Lahiri, S.D.; Zhang, G.; Dunaway-Mariano, D.
Deposited on : 2006-10-10
Resolution : 2.50 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.11

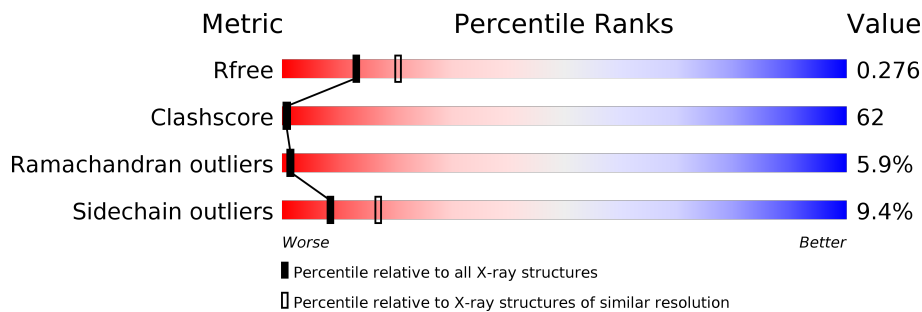
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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}	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	267	
2	K	267	

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 4441 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 Phosphonoacetaldehyde hydrolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	256	2039	1296	339	387	17	0	0	0

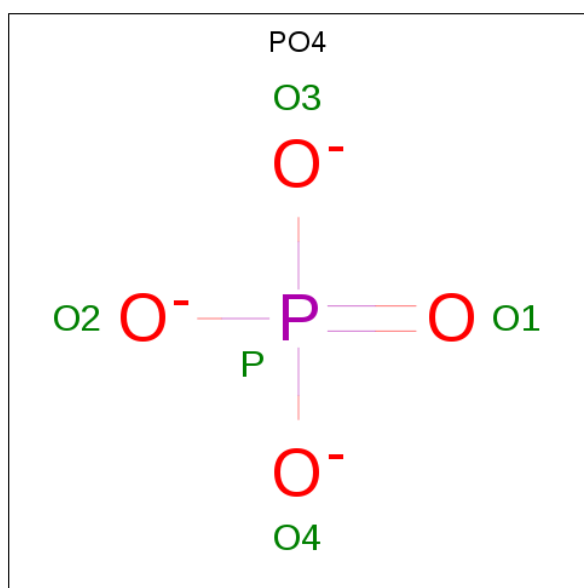
- Molecule 2 is a protein called Phosphonoacetaldehyde hydrolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	K	256	2041	1298	339	387	17	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	53	LDH	LYS	MODIFIED RESIDUE	UNP O31156

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O P 5 4 1	0	0
3	K	1	Total O P 5 4 1	0	0

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Mg 1 1	0	0
4	K	1	Total Mg 1 1	0	0

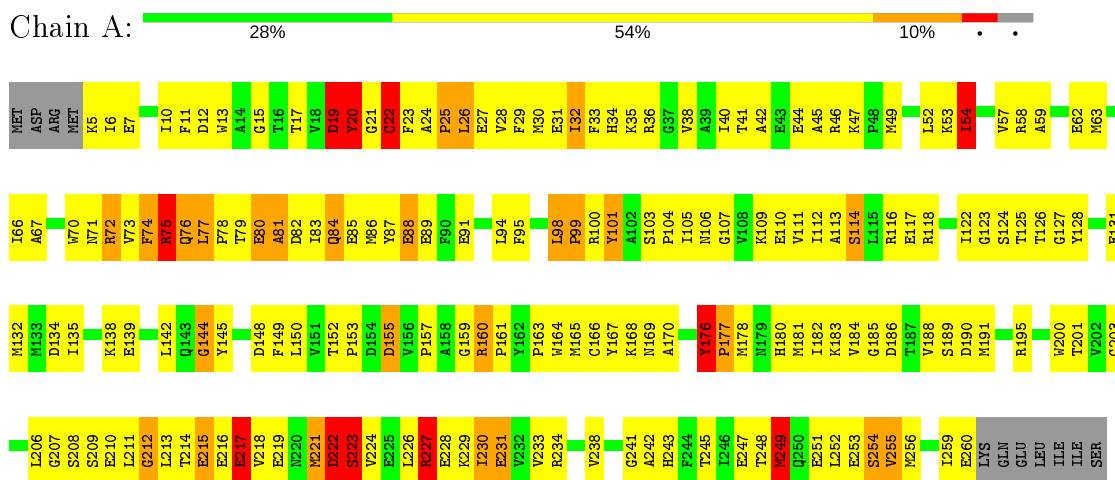
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	162	Total O 162 162	0	0
5	K	187	Total O 187 187	0	0

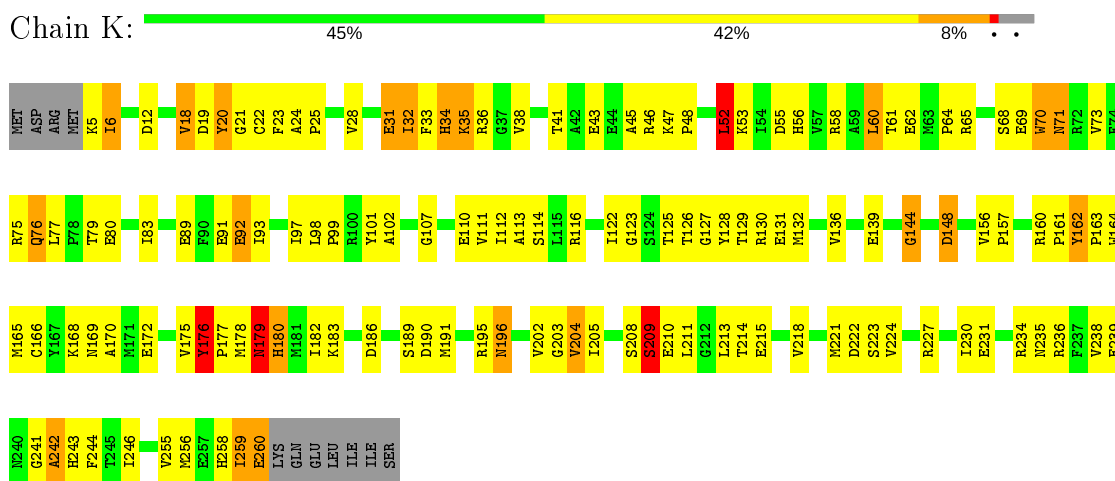
3 Residue-property plots i

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.

- Molecule 1: Phosphonoacetaldehyde hydrolase



- Molecule 2: Phosphonoacetaldehyde hydrolase



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	159.82Å 64.02Å 98.33Å 90.00° 123.52° 90.00°	Depositor
Resolution (Å)	100.00 – 2.50 81.98 – 2.46	Depositor EDS
% Data completeness (in resolution range)	68.4 (100.00-2.50) 86.4 (81.98-2.46)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.17 (at 2.45Å)	Xtrriage
Refinement program	CNS 1.1, REFMAC	Depositor
R, R_{free}	0.231 , 0.298 0.272 , 0.276	Depositor DCC
R_{free} test set	2775 reflections (9.87%)	wwPDB-VP
Wilson B-factor (Å ²)	18.1	Xtrriage
Anisotropy	0.751	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 29.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.84	EDS
Total number of atoms	4441	wwPDB-VP
Average B, all atoms (Å ²)	28.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.53% of the height of the origin peak. No significant pseudotranslation is detected.*

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.

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: PO4, LDH, MG

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	0.82	0/2081	1.11	15/2813 (0.5%)
2	K	0.80	0/2071	1.05	6/2799 (0.2%)
All	All	0.81	0/4152	1.08	21/5612 (0.4%)

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
2	K	0	2
All	All	0	3

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	176	TYR	CA-CB-CG	7.24	127.16	113.40
1	A	77	LEU	N-CA-C	-6.60	93.18	111.00
1	A	176	TYR	CB-CG-CD1	-6.37	117.18	121.00
1	A	249	MET	CG-SD-CE	-6.33	90.07	100.20
1	A	75	ARG	N-CA-C	-6.29	94.02	111.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	128	TYR	Sidechain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
2	K	128	TYR	Sidechain
2	K	162	TYR	Sidechain

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	2039	0	2013	300	0
2	K	2041	0	2015	208	1
3	A	5	0	0	0	0
3	K	5	0	0	0	0
4	A	1	0	0	0	0
4	K	1	0	0	0	0
5	A	162	0	0	137	0
5	K	187	0	0	105	1
All	All	4441	0	4028	507	1

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 62.

The worst 5 of 507 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:160:ARG:HA	5:A:611:HOH:O	1.35	1.26
2:K:161:PRO:HD2	5:K:642:HOH:O	1.36	1.24
1:A:177:PRO:HD2	5:A:566:HOH:O	1.42	1.18
1:A:191:MET:HB3	5:A:579:HOH:O	1.54	1.08
1:A:10:ILE:HG23	5:A:570:HOH:O	1.51	1.07

All (1) 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	Interatomic distance (Å)	Clash overlap (Å)
2:K:144:GLY:O	5:K:539:HOH:O[4_455]	2.16	0.04

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	254/267 (95%)	205 (81%)	29 (11%)	20 (8%)	1	1
2	K	253/267 (95%)	215 (85%)	28 (11%)	10 (4%)	3	3
All	All	507/534 (95%)	420 (83%)	57 (11%)	30 (6%)	1	1

5 of 30 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	19	ASP
1	A	20	TYR
1	A	54	ILE
1	A	75	ARG
1	A	76	GLN

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	219/230 (95%)	197 (90%)	22 (10%)	7	15
2	K	218/229 (95%)	199 (91%)	19 (9%)	10	20
All	All	437/459 (95%)	396 (91%)	41 (9%)	8	17

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	230	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	K	6	ILE
2	K	196	ASN
1	A	231	GLU
1	A	249	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	84	GLN
2	K	179	ASN
2	K	196	ASN
2	K	243	HIS

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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
2	LDH	K	53	2	9,10,11	0.49	0	5,10,12	1.06	1 (20%)

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
2	LDH	K	53	2	-	3/8/9/11	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
2	K	53	LDH	CE-NZ-CH	-2.34	108.44	113.43

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	K	53	LDH	N-CA-CB-CG
2	K	53	LDH	C-CA-CB-CG
2	K	53	LDH	CE-CD-CG-CB

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	K	53	LDH	2	0

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	$\# Z > 2$
3	PO4	K	361	4	4,4,4	1.68	1 (25%)	6,6,6	0.48	0
3	PO4	A	361	4	4,4,4	1.69	1 (25%)	6,6,6	0.45	0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	361	PO4	P-O3	-2.20	1.48	1.54
3	K	361	PO4	P-O4	-2.05	1.48	1.54

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.

5.7 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

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.