Full wwPDB X-ray Structure Validation Report ⓘ

Dec 18, 2023 – 11:30 AM EST

PDB ID	:	1VMI
Title	:	Crystal structure of Putative phosphate acetyltransferase (np_416953.1) from Escherichia coli k12 at 2.32 A resolution
Authors	:	Joint Center for Structural Genomics (JCSG)
Deposited on	:	2004-09-28
Resolution	:	2.32 Å(reported)

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

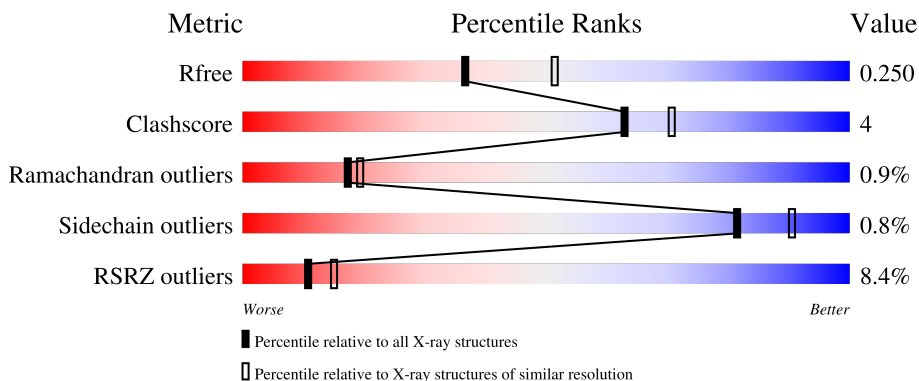|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.32 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		0.250
Clashscore		4
Ramachandran outliers		0.9%
Sidechain outliers		0.8%
RSRZ outliers		8.4%

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	5974 (2.34-2.30)
Clashscore	141614	6604 (2.34-2.30)
Ramachandran outliers	138981	6523 (2.34-2.30)
Sidechain outliers	138945	6523 (2.34-2.30)
RSRZ outliers	127900	5855 (2.34-2.30)

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	355	8% 81% 11% · 7%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 2500 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 phosphate acetyltransferase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	329	2397	1522	427	434	6	8	0	0	0

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-12	HIS	-	SEE REMARK 999	UNP P77218
A	-11	HIS	-	SEE REMARK 999	UNP P77218
A	-10	HIS	-	SEE REMARK 999	UNP P77218
A	-9	HIS	-	SEE REMARK 999	UNP P77218
A	-8	HIS	-	SEE REMARK 999	UNP P77218
A	-7	HIS	-	SEE REMARK 999	UNP P77218
A	-6	THR	-	SEE REMARK 999	UNP P77218
A	-5	ASP	-	SEE REMARK 999	UNP P77218
A	-4	PRO	-	SEE REMARK 999	UNP P77218
A	-3	ALA	-	SEE REMARK 999	UNP P77218
A	-2	LEU	-	SEE REMARK 999	UNP P77218
A	-1	ARG	-	SEE REMARK 999	UNP P77218
A	0	ALA	-	SEE REMARK 999	UNP P77218
A	62	MSE	MET	modified residue	UNP P77218
A	76	MSE	MET	modified residue	UNP P77218
A	105	MSE	MET	modified residue	UNP P77218
A	110	MSE	MET	modified residue	UNP P77218
A	153	MSE	MET	modified residue	UNP P77218
A	201	MSE	MET	modified residue	UNP P77218
A	265	MSE	MET	modified residue	UNP P77218
A	300	MSE	MET	modified residue	UNP P77218
A	339	GLY	-	SEE REMARK 999	UNP P77218
A	340	LEU	-	SEE REMARK 999	UNP P77218
A	341	CYS	-	SEE REMARK 999	UNP P77218
A	342	GLY	-	SEE REMARK 999	UNP P77218
A	343	ARG	-	SEE REMARK 999	UNP P77218

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	91	Total O 91 91		0	0

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: putative phosphate acetyltransferase
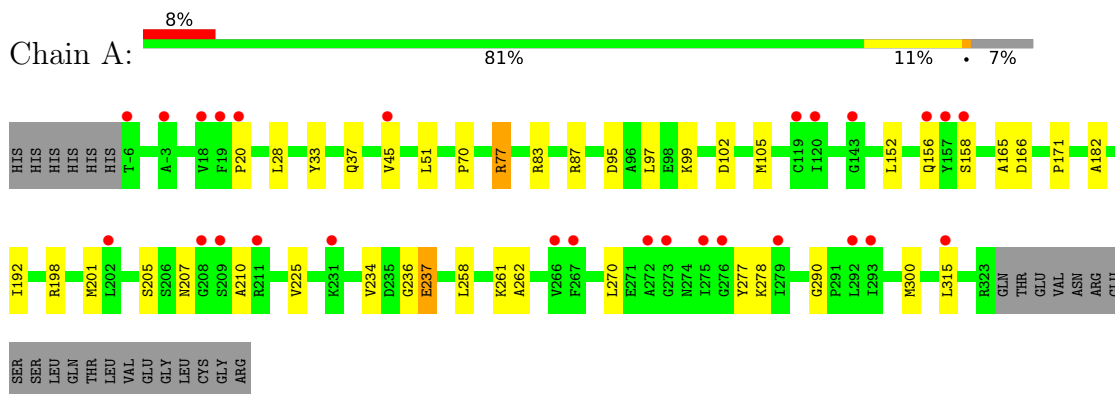
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 63 2 2	Depositor
Cell constants a, b, c, α, β, γ	95.31Å 95.31Å 175.51Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	28.26 – 2.32 28.26 – 2.32	Depositor EDS
% Data completeness (in resolution range)	99.7 (28.26-2.32) 99.8 (28.26-2.32)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	1.84 (at 2.31Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.194 , 0.247 0.201 , 0.250	Depositor DCC
R_{free} test set	1085 reflections (5.13%)	wwPDB-VP
Wilson B-factor (Å^2)	48.4	Xtriage
Anisotropy	0.209	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 47.9	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	2500	wwPDB-VP
Average B, all atoms (Å^2)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.96% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

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

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	0.72	0/2436	0.80	6/3307 (0.2%)

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	3

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	77	ARG	NE-CZ-NH2	-8.24	116.18	120.30
1	A	77	ARG	NE-CZ-NH1	7.99	124.30	120.30
1	A	87	ARG	NE-CZ-NH2	-5.86	117.37	120.30
1	A	83	ARG	NE-CZ-NH2	-5.45	117.57	120.30
1	A	83	ARG	NE-CZ-NH1	5.32	122.96	120.30
1	A	95	ASP	CB-CG-OD1	5.19	122.97	118.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	165	ALA	Peptide,Mainchain
1	A	236	GLY	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	2397	0	2420	21	0
2	A	12	0	16	1	0
3	A	91	0	0	1	0
All	All	2500	0	2436	21	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 4.

All (21) 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:300:MSE:CE	1:A:300:MSE:SE	2.15	1.44
1:A:207:ASN:O	1:A:237:GLU:OE1	2.07	0.73
1:A:201:MSE:HE2	1:A:234:VAL:CG2	2.37	0.55
1:A:205:SER:HB2	1:A:210:ALA:HB3	1.91	0.52
1:A:99:LYS:HB3	1:A:105:MSE:HE3	1.91	0.52
1:A:182:ALA:HB1	1:A:201:MSE:HE1	1.90	0.52
1:A:201:MSE:HE2	1:A:234:VAL:HG21	1.95	0.49
1:A:70:PRO:O	1:A:77:ARG:HD2	2.14	0.47
1:A:28:LEU:HD21	1:A:51:LEU:HD11	1.97	0.46
1:A:171:PRO:HA	1:A:270:LEU:HB2	1.97	0.46
1:A:97:LEU:HB2	3:A:395:HOH:O	2.18	0.44
1:A:102:ASP:HB3	1:A:105:MSE:HG3	2.01	0.43
1:A:198:ARG:HD3	1:A:261:LYS:O	2.19	0.43
1:A:33:TYR:CZ	1:A:37:GLN:HG3	2.53	0.43
1:A:258:LEU:HD13	1:A:262:ALA:HB2	2.01	0.43
1:A:201:MSE:CE	1:A:234:VAL:HG21	2.50	0.42
1:A:201:MSE:CE	1:A:225:VAL:HG21	2.50	0.42
1:A:152:LEU:HD11	1:A:192:ILE:HG21	2.02	0.42
1:A:20:PRO:HA	1:A:45:VAL:HB	2.02	0.41
1:A:278:LYS:HZ1	2:A:1:GOL:H11	1.86	0.40
1:A:290:GLY:HA2	1:A:315:LEU:HD13	2.03	0.40

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	327/355 (92%)	319 (98%)	5 (2%)	3 (1%)	17	19

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	237	GLU
1	A	158	SER
1	A	166	ASP

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	238/268 (89%)	236 (99%)	2 (1%)	81	90

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	156	GLN
1	A	277	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	25	GLN

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 ⓘ

2 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	$\#|Z| > 2$
2	GOL	A	1	-	5,5,5	0.39	0	5,5,5	0.55	0
2	GOL	A	344	-	5,5,5	1.04	0	5,5,5	1.25	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	GOL	A	1	-	-	3/4/4/4	-
2	GOL	A	344	-	-	0/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	344	GOL	O3-C3-C2	2.32	121.32	110.20

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1	GOL	C1-C2-C3-O3
2	A	1	GOL	O2-C2-C3-O3
2	A	1	GOL	O1-C1-C2-C3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1	GOL	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 ⓘ

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(Å²)	Q<0.9
1	A	321/355 (90%)	0.28	27 (8%)	11	15	41, 49, 56, 65	0

All (27) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	157	TYR	7.0
1	A	158	SER	4.2
1	A	275	ILE	3.6
1	A	272	ALA	3.5
1	A	156	GLN	3.4
1	A	267	PHE	3.2
1	A	202	LEU	2.8
1	A	120	ILE	2.7
1	A	292	LEU	2.7
1	A	293	ILE	2.7
1	A	266	VAL	2.6
1	A	273	GLY	2.5
1	A	-3	ALA	2.4
1	A	-6	THR	2.4
1	A	276	GLY	2.4
1	A	45	VAL	2.4
1	A	19	PHE	2.4
1	A	211	ARG	2.3
1	A	231	LYS	2.3
1	A	143	GLY	2.3
1	A	315	LEU	2.2
1	A	20	PRO	2.2
1	A	208	GLY	2.2
1	A	18	VAL	2.1
1	A	209	SER	2.1
1	A	119	CYS	2.0
1	A	279	ILE	2.0

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	GOL	A	344	6/6	0.83	0.23	51,57,60,63	0
2	GOL	A	1	6/6	0.90	0.32	71,72,73,73	0

6.5 Other polymers ⓘ

There are no such residues in this entry.