wwPDB X-ray Structure Validation Summary Report (i)

Jun 15, 2024 – 10:01 AM EDT

PDB ID	:	2BG5
Title	:	Crystal Structure of the Phosphoenolpyruvate-binding Enzyme I-Domain from the Thermoanaerobacter tengcongensis PEP: Sugar Phosphotransferase System (PTS)
Authors	:	Oberholzer, A.E.; Bumann, M.; Schneider, P.; Baechler, C.; Siebold, C.; Baumann, U.; Erni, B.
Deposited on	:	2004-12-17
Resolution	:	1.82 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:

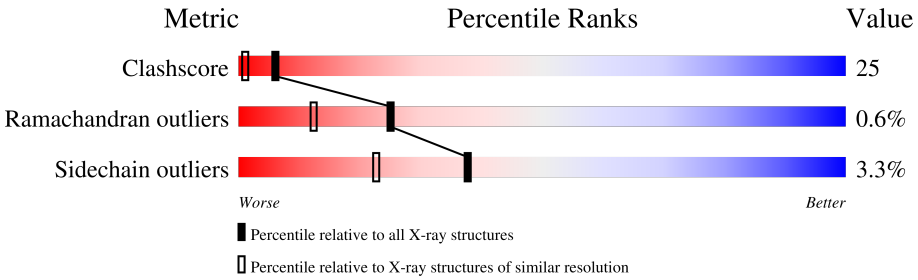
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 1.82 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	8401 (1.84-1.80)
Ramachandran outliers	138981	8290 (1.84-1.80)
Sidechain outliers	138945	8290 (1.84-1.80)

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	324	76% 22% •
1	B	324	77% 19% •
1	C	324	70% 26% •
1	D	324	75% 21% • •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 10932 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 PHOSPHOENOLPYRUVATE-PROTEIN KINASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	324	Total	C	N	O	S	Se	0	0	1
			2523	1602	419	483	2	17			
1	B	323	Total	C	N	O	S	Se	0	0	1
			2515	1597	418	482	2	16			
1	C	324	Total	C	N	O	S	Se	0	0	1
			2523	1602	419	483	2	17			
1	D	324	Total	C	N	O	S	Se	0	0	1
			2523	1602	419	483	2	17			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	205	Total	O	0	0
			205	205		
2	B	227	Total	O	0	0
			227	227		
2	C	196	Total	O	0	0
			196	196		
2	D	220	Total	O	0	0
			220	220		

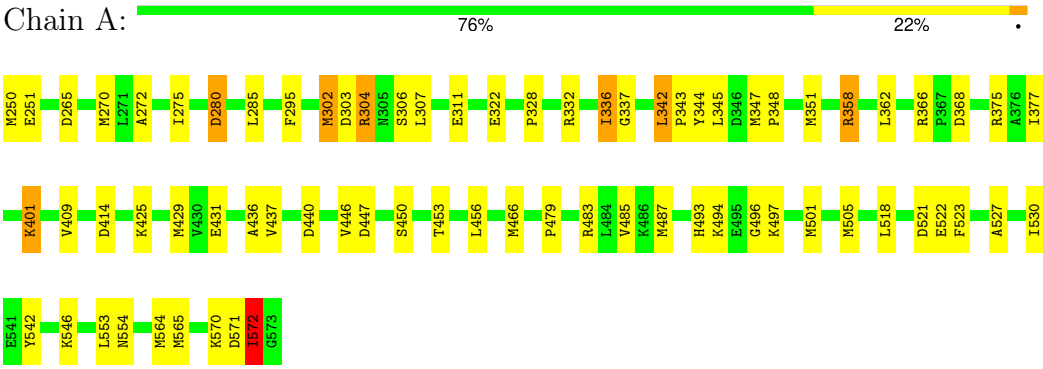
3 Residue-property plots (i)

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.

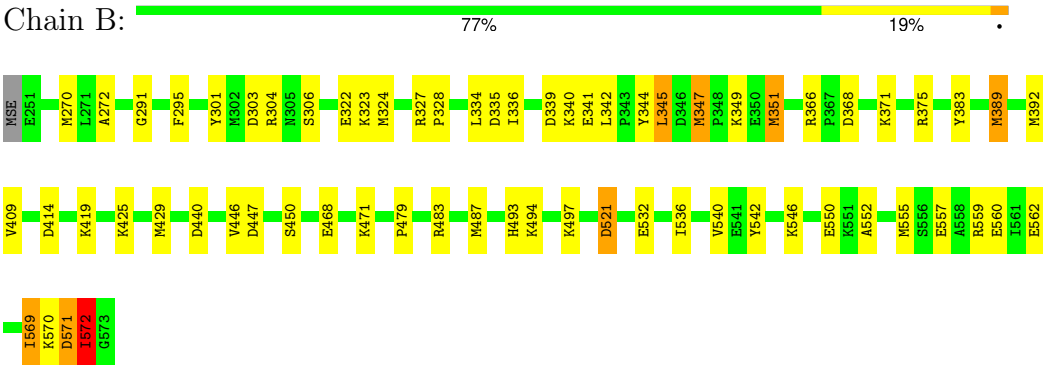
• Molecule 1: PHOSPHOENOLPYRUVATE-PROTEIN KINASE

Chain A:



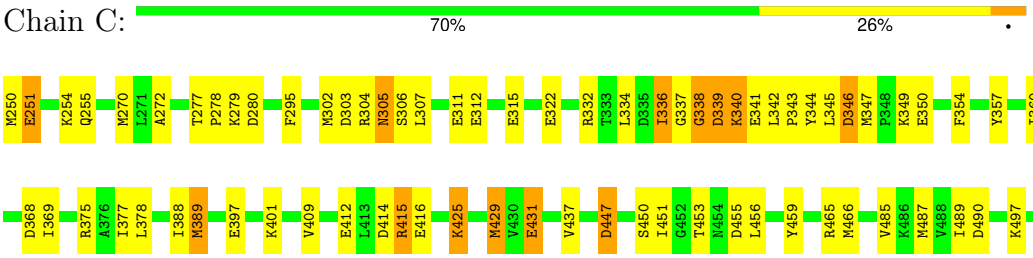
• Molecule 1: PHOSPHOENOLPYRUVATE-PROTEIN KINASE

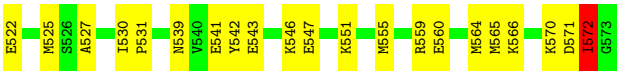
Chain B:



• Molecule 1: PHOSPHOENOLPYRUVATE-PROTEIN KINASE

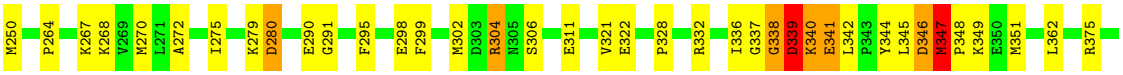
Chain C:

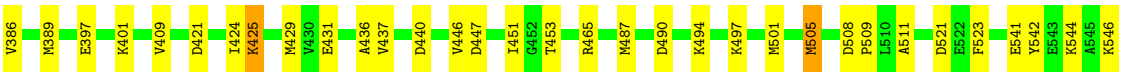


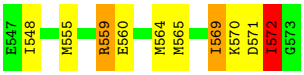


• Molecule 1: PHOSPHOENOLPYRUVATE-PROTEIN KINASE

Chain D:







4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	82.09Å 91.44Å 181.85Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 1.82	Depositor
% Data completeness (in resolution range)	99.8 (30.00-1.82)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	REFMAC 5.2.0001	Depositor
R, R_{free}	0.208 , 0.250	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	10932	wwPDB-VP
Average B, all atoms (Å^2)	6.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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	0.65	1/2545 (0.0%)	1.03	12/3400 (0.4%)
1	B	0.60	1/2537 (0.0%)	0.97	8/3390 (0.2%)
1	C	0.58	1/2545 (0.0%)	1.02	9/3400 (0.3%)
1	D	0.59	1/2545 (0.0%)	1.03	11/3400 (0.3%)
All	All	0.61	4/10172 (0.0%)	1.01	40/13590 (0.3%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	389	MSE	SE-CE	-6.71	1.55	1.95
1	C	572	ILE	C-N	-5.19	1.23	1.33
1	B	572	ILE	C-N	-5.10	1.23	1.33
1	D	572	ILE	C-N	-5.06	1.24	1.33

The worst 5 of 40 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	304	ARG	NE-CZ-NH1	10.58	125.59	120.30
1	D	304	ARG	NE-CZ-NH2	-10.33	115.14	120.30
1	C	455	ASP	CB-CG-OD2	10.14	127.42	118.30
1	D	347	MSE	CA-CB-CG	-8.17	99.41	113.30
1	B	389	MSE	CG-SE-CE	-8.06	81.18	98.90

There are no chirality outliers.

There are no planarity outliers.

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	2523	0	2575	129	1
1	B	2515	0	2566	78	1
1	C	2523	0	2575	150	1
1	D	2523	0	2575	158	1
2	A	205	0	0	24	0
2	B	227	0	0	27	0
2	C	196	0	0	29	0
2	D	220	0	0	26	0
All	All	10932	0	10291	503	2

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 25.

The worst 5 of 503 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:564:MSE:CE	1:A:565:MSE:HE3	1.24	1.56
1:A:564:MSE:HE3	1:A:565:MSE:CE	1.34	1.54
1:D:564:MSE:HE2	1:D:565:MSE:CE	1.50	1.42
1:D:564:MSE:CE	1:D:565:MSE:HE1	1.54	1.36
1:C:489:ILE:HG13	2:C:2149:HOH:O	1.33	1.28

All (2) 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)
1:C:322:GLU:OE2	1:D:322:GLU:OE2[1_655]	2.04	0.16
1:A:322:GLU:OE2	1:B:322:GLU:OE2[1_655]	2.12	0.08

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	322/324 (99%)	312 (97%)	9 (3%)	1 (0%)	41 27
1	B	321/324 (99%)	314 (98%)	6 (2%)	1 (0%)	41 27
1	C	322/324 (99%)	308 (96%)	11 (3%)	3 (1%)	17 6
1	D	322/324 (99%)	308 (96%)	11 (3%)	3 (1%)	17 6
All	All	1287/1296 (99%)	1242 (96%)	37 (3%)	8 (1%)	25 12

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	572	ILE
1	B	572	ILE
1	C	338	GLY
1	C	572	ILE
1	D	339	ASP

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	271/254 (107%)	262 (97%)	9 (3%)	38 23
1	B	270/254 (106%)	262 (97%)	8 (3%)	41 26
1	C	271/254 (107%)	261 (96%)	10 (4%)	34 19
1	D	271/254 (107%)	262 (97%)	9 (3%)	38 23
All	All	1083/1016 (107%)	1047 (97%)	36 (3%)	38 23

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	339	ASP
1	D	569	ILE
1	D	340	LYS
1	D	505	MSE
1	B	347	MSE

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

Mol	Chain	Res	Type
1	A	469	HIS
1	A	493	HIS
1	A	554	ASN
1	B	493	HIS

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

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.