WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jan 30, 2024 – 01:45 PM EST

PDB ID	:	1HKG
Title	:	STRUCTURAL DYNAMICS OF YEAST HEXOKINASE DURING CATAL-YSIS
Authors	:	Bennett Jr., W.S.; Steitz, T.A.
Deposited on	:	1980-12-22
Resolution	:	3.50 Å(reported)

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

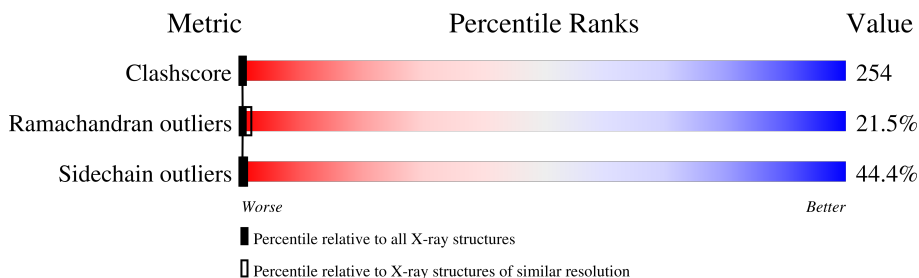|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	*NOT EXECUTED*
EDS	:	*NOT EXECUTED*
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	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 3.50 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	Percentile Ranks	Value
Clashscore		254
Ramachandran outliers		21.5%
Sidechain outliers		44.4%

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))
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)

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	457	· 35% 37% 25%

2 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 3298 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 HEXOKINASE A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	457	Total 3298	C 2131	N 543	O 611	S 13	0	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. 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: HEXOKINASE A

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	144.80Å 78.80Å 61.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 3.50	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-3.50)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	PROLSQ	Depositor
R, R_{free}	(Not available) , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3298	wwPDB-VP
Average B, all atoms (Å²)	0.0	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	3.74	297/2729 (10.9%)	2.67	224/3600 (6.2%)

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	2

The worst 5 of 297 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	51	TRP	NE1-CE2	50.00	2.02	1.37
1	A	109	TRP	CZ3-CH2	-32.93	0.87	1.40
1	A	109	TRP	NE1-CE2	-29.84	0.98	1.37
1	A	199	TYR	CE2-CZ	29.05	1.76	1.38
1	A	51	TRP	CZ3-CH2	-27.46	0.96	1.40

The worst 5 of 224 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	109	TRP	CD1-NE1-CE2	22.37	129.13	109.00
1	A	414	TRP	CD1-NE1-CE2	-19.17	91.75	109.00
1	A	51	TRP	CD1-CG-CD2	17.07	119.95	106.30
1	A	333	ARG	NE-CZ-NH2	14.62	127.61	120.30
1	A	34	ARG	NE-CZ-NH2	14.51	127.56	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	296	TYR	Sidechain
1	A	84	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	3298	0	3129	1631	1
All	All	3298	0	3129	1631	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 254.

The worst 5 of 1631 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:414:TRP:CH2	1:A:414:TRP:CZ3	1.82	1.66
1:A:148:LYS:HZ1	1:A:186:ASP:CA	1.08	1.62
1:A:199:TYR:CE2	1:A:199:TYR:CZ	1.76	1.62
1:A:276:PHE:CZ	1:A:276:PHE:CE2	1.83	1.61
1:A:312:ALA:HB3	1:A:315:VAL:CB	1.24	1.61

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)
1:A:106:SER:OG	1:A:223:THR:OG1[3_555]	1.98	0.22

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	372/457 (81%)	214 (58%)	78 (21%)	80 (22%)	0	1

5 of 80 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	4	SER
1	A	39	ASP
1	A	42	ALA
1	A	43	ALA
1	A	46	PRO

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	284/284 (100%)	158 (56%)	126 (44%)	0	0

5 of 126 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	208	ILE
1	A	388	ILE
1	A	259	ILE
1	A	385	SER
1	A	417	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	387	HIS
1	A	379	GLN
1	A	279	ASN
1	A	325	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	246	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 ⓘ

There are no ligands in this entry.

5.7   Other polymers ⓘ

There are no such residues in this entry.

5.8   Polymer linkage issues ⓘ

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	101:THR	C	102:THR	N	1.20
1	A	38:UNK	C	39:ASP	N	1.15

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.