Full wwPDB X-ray Structure Validation Report ⓘ

May 28, 2020 – 07:22 pm BST

PDB ID	:	1EHW
Title	:	HUMAN NUCLEOSIDE DIPHOSPHATE KINASE 4
Authors	:	Milon, L.; Meyer, P.; Chiadmi, M.; Munier, A.; Johansson, M.; Karlsson, A.; Lascu, I.; Capeau, J.; Janin, J.; Lacombe, M.-L.
Deposited on	:	2000-02-23
Resolution	:	2.40 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

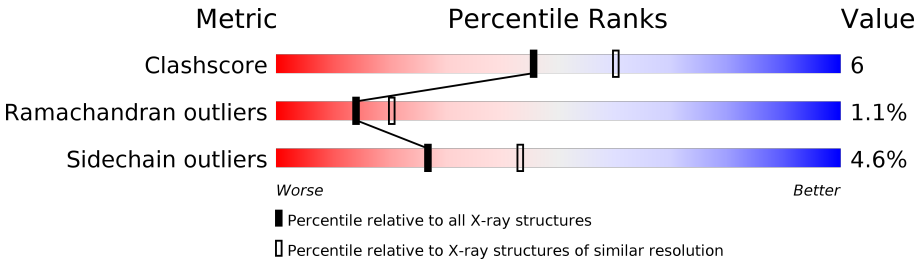
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.40 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	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)

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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	162	76% 12% • 12%
1	B	162	73% 13% • 12%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2352 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 NUCLEOSIDE DIPHOSPHATE KINASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	143	Total	C	N	O	S	0	0	0
			1131	711	211	203	6			
1	B	142	Total	C	N	O	S	0	0	0
			1123	706	210	202	5			

There are 44 discrepancies between the modelled and reference sequences:

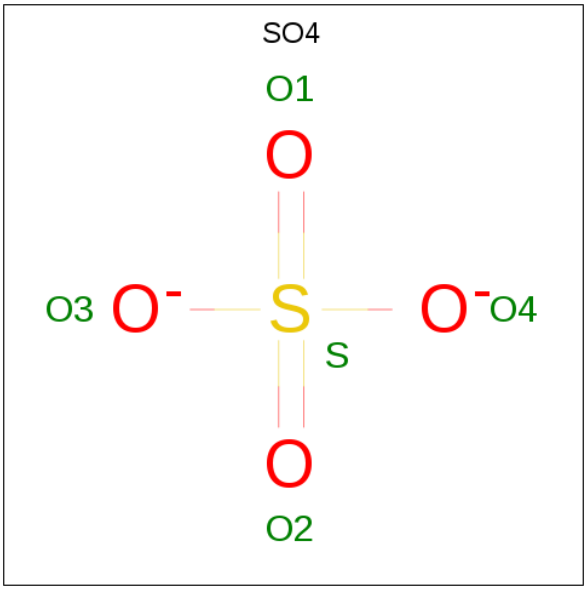
Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	ARG	SEE REMARK 999	UNP O00746
A	-18	GLY	CYS	SEE REMARK 999	UNP O00746
A	-17	SER	GLY	SEE REMARK 999	UNP O00746
A	-16	SER	PRO	SEE REMARK 999	UNP O00746
A	-15	HIS	ARG	SEE REMARK 999	UNP O00746
A	-14	HIS	ALA	SEE REMARK 999	UNP O00746
A	-13	HIS	PRO	SEE REMARK 999	UNP O00746
A	-12	HIS	GLY	SEE REMARK 999	UNP O00746
A	-11	HIS	PRO	SEE REMARK 999	UNP O00746
A	-10	HIS	SER	SEE REMARK 999	UNP O00746
A	-9	SER	LEU	SEE REMARK 999	UNP O00746
A	-8	SER	LEU	SEE REMARK 999	UNP O00746
A	-7	GLY	VAL	SEE REMARK 999	UNP O00746
A	-6	LEU	ARG	SEE REMARK 999	UNP O00746
A	-5	VAL	HIS	SEE REMARK 999	UNP O00746
A	-4	PRO	GLY	SEE REMARK 999	UNP O00746
A	-3	ARG	SER	SEE REMARK 999	UNP O00746
A	-2	GLY	GLY	SEE REMARK 999	UNP O00746
A	-1	SER	GLY	SEE REMARK 999	UNP O00746
A	0	HIS	PRO	see remark 999	UNP O00746
A	1	MET	SER	see remark 999	UNP O00746
A	2	GLY	TRP	see remark 999	UNP O00746
B	-19	MET	ARG	SEE REMARK 999	UNP O00746
B	-18	GLY	CYS	SEE REMARK 999	UNP O00746
B	-17	SER	GLY	SEE REMARK 999	UNP O00746

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-16	SER	PRO	SEE REMARK 999	UNP O00746
B	-15	HIS	ARG	SEE REMARK 999	UNP O00746
B	-14	HIS	ALA	SEE REMARK 999	UNP O00746
B	-13	HIS	PRO	SEE REMARK 999	UNP O00746
B	-12	HIS	GLY	SEE REMARK 999	UNP O00746
B	-11	HIS	PRO	SEE REMARK 999	UNP O00746
B	-10	HIS	SER	SEE REMARK 999	UNP O00746
B	-9	SER	LEU	SEE REMARK 999	UNP O00746
B	-8	SER	LEU	SEE REMARK 999	UNP O00746
B	-7	GLY	VAL	SEE REMARK 999	UNP O00746
B	-6	LEU	ARG	SEE REMARK 999	UNP O00746
B	-5	VAL	HIS	SEE REMARK 999	UNP O00746
B	-4	PRO	GLY	SEE REMARK 999	UNP O00746
B	-3	ARG	SER	SEE REMARK 999	UNP O00746
B	-2	GLY	GLY	SEE REMARK 999	UNP O00746
B	-1	SER	GLY	SEE REMARK 999	UNP O00746
B	0	HIS	PRO	see remark 999	UNP O00746
B	1	MET	SER	see remark 999	UNP O00746
B	2	GLY	TRP	see remark 999	UNP O00746

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total O S 5 4 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	47	Total O 47 47	0	0
3	B	36	Total O 36 36	0	0

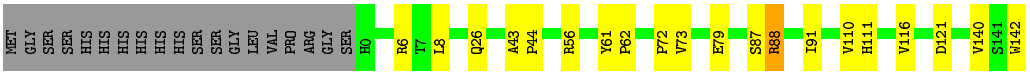
3 Residue-property plots (i)

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.

Note EDS was not executed.

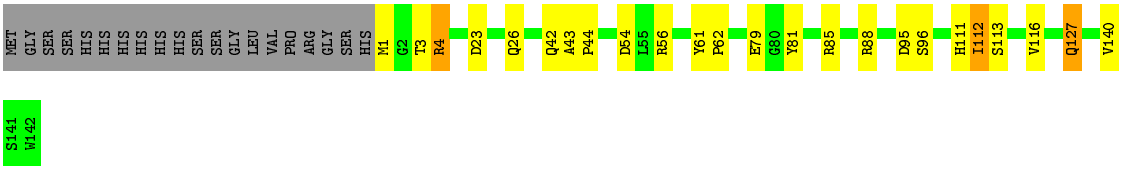
- Molecule 1: NUCLEOSIDE DIPHOSPHATE KINASE

Chain A:



- Molecule 1: NUCLEOSIDE DIPHOSPHATE KINASE

Chain B:



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 3	Depositor
Cell constants a, b, c, α, β, γ	112.96Å 112.96Å 112.96Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.40	Depositor
% Data completeness (in resolution range)	94.3 (30.00-2.40)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 0.5	Depositor
R, R_{free}	0.184 , 0.229	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2352	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

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	0.50	0/1155	0.70	0/1561
1	B	0.48	0/1147	0.70	0/1551
All	All	0.49	0/2302	0.70	0/3112

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

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	1131	0	1127	13	0
1	B	1123	0	1118	18	0
2	A	10	0	0	0	0
2	B	5	0	0	0	0
3	A	47	0	0	1	0
3	B	36	0	0	1	0
All	All	2352	0	2245	29	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 6.

All (29) 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:4:ARG:HA	1:B:4:ARG:HH11	1.42	0.83
1:B:127:GLN:HE21	1:B:127:GLN:HA	1.48	0.77
1:B:4:ARG:HA	1:B:4:ARG:NH1	2.02	0.74
1:B:4:ARG:HH12	1:B:79:GLU:CD	1.91	0.73
1:A:72:PRO:HB3	1:B:140:VAL:HG21	1.75	0.68
1:A:140:VAL:HG23	1:A:140:VAL:O	1.96	0.65
1:B:95:ASP:OD1	1:B:112:ILE:HG23	2.05	0.56
1:A:56:ARG:HA	1:A:61:TYR:CD1	2.43	0.53
1:A:111:HIS:HD2	3:A:552:HOH:O	1.92	0.52
1:A:61:TYR:HB3	1:A:62:PRO:HD3	1.91	0.52
1:B:127:GLN:NE2	1:B:127:GLN:HA	2.23	0.52
1:A:6:ARG:HD2	1:A:79:GLU:OE2	2.09	0.51
1:B:1:MET:N	1:B:4:ARG:HB2	2.24	0.51
1:B:61:TYR:HB3	1:B:62:PRO:HD3	1.92	0.51
1:A:88:ARG:HD2	1:A:121:ASP:HA	1.94	0.49
1:A:140:VAL:CG2	1:A:140:VAL:O	2.61	0.48
1:B:140:VAL:CG2	1:B:140:VAL:O	2.62	0.48
1:B:96:SER:HB2	1:B:111:HIS:O	2.14	0.48
1:B:23:ASP:OD1	3:B:560:HOH:O	2.20	0.46
1:B:1:MET:H2	1:B:4:ARG:HB2	1.79	0.46
1:B:85:ARG:NH1	1:B:85:ARG:HB3	2.31	0.45
1:B:140:VAL:HG23	1:B:140:VAL:O	2.16	0.45
1:A:56:ARG:HA	1:A:61:TYR:CG	2.53	0.44
1:B:43:ALA:HA	1:B:44:PRO:HD3	1.81	0.43
1:B:3:THR:HG22	1:B:81:TYR:HB2	2.01	0.43
1:A:43:ALA:HA	1:A:44:PRO:HD3	1.87	0.42
1:A:26:GLN:OE1	1:B:26:GLN:OE1	2.37	0.41
1:A:142:TRP:CD1	1:A:142:TRP:N	2.88	0.41
1:A:87:SER:O	1:A:91:ILE:HG23	2.21	0.40

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	141/162 (87%)	137 (97%)	3 (2%)	1 (1%)	22 32
1	B	140/162 (86%)	136 (97%)	2 (1%)	2 (1%)	11 15
All	All	281/324 (87%)	273 (97%)	5 (2%)	3 (1%)	14 20

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	113	SER
1	A	116	VAL
1	B	116	VAL

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	119/136 (88%)	115 (97%)	4 (3%)	37 56
1	B	118/136 (87%)	111 (94%)	7 (6%)	19 32
All	All	237/272 (87%)	226 (95%)	11 (5%)	27 43

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	8	LEU
1	A	73	VAL
1	A	88	ARG
1	A	110	VAL
1	B	4	ARG
1	B	42	GLN
1	B	54	ASP
1	B	56	ARG
1	B	88	ARG
1	B	112	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	127	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	26	GLN
1	A	127	GLN
1	A	131	GLN
1	B	42	GLN
1	B	127	GLN
1	B	131	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 carbohydrates in this entry.

5.6 Ligand geometry

3 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	SO4	A	164	-	4,4,4	0.22	0	6,6,6	0.62	0
2	SO4	B	163	-	4,4,4	0.42	0	6,6,6	0.12	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	SO4	A	160	-	4,4,4	0.36	0	6,6,6	0.35	0

There are no bond length outliers.

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

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.