wwPDB X-ray Structure Validation Summary Report (i)

Jun 12, 2024 – 04:44 AM EDT

PDB ID	:	1YT5
Title	:	Crystal structure of NAD kinase from Thermotoga maritima
Authors	:	Berkeley Structural Genomics Center (BSGC)
Deposited on	:	2005-02-09
Resolution	:	2.30 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:

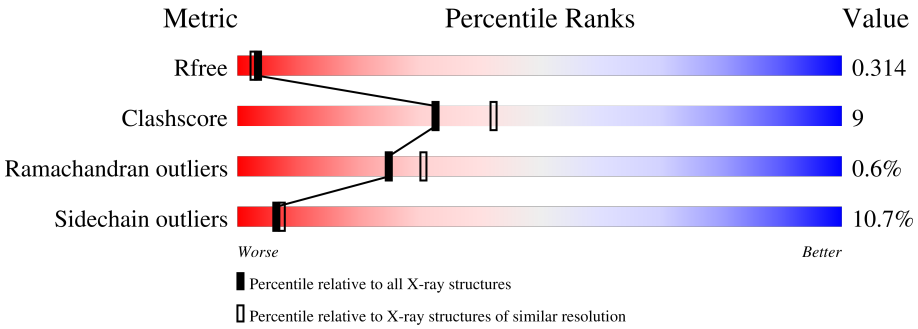
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)

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%

Mol	Chain	Length	Quality of chain
1	A	258	72% 23%
1	B	258	77% 19%
1	C	258	64% 28% 5%
1	D	258	79% 17%

2 Entry composition (i)

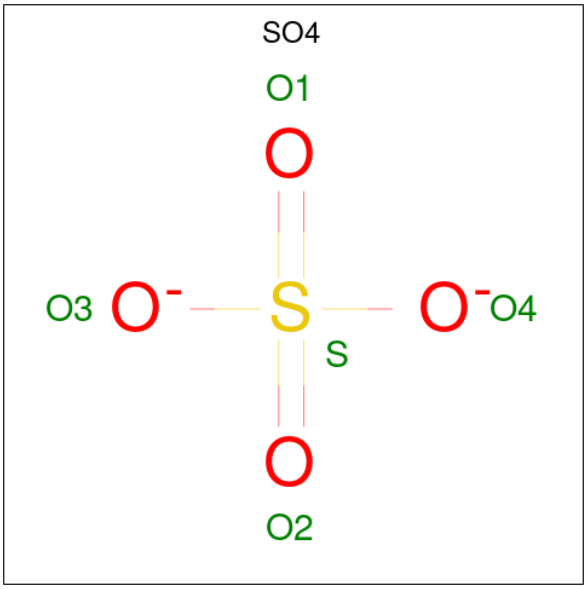
There are 3 unique types of molecules in this entry. The entry contains 8357 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 inorganic polyphosphate/ATP-NAD kinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	256	Total	C	N	O	S	0	0	0
			2044	1306	350	382	6			
1	B	256	Total	C	N	O	S	0	0	0
			2044	1306	350	382	6			
1	C	252	Total	C	N	O	S	0	0	0
			2015	1289	342	378	6			
1	D	256	Total	C	N	O	S	0	0	0
			2044	1306	350	382	6			

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



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

Continued on next page...

Continued from previous page...

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
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	C	1	Total O S 5 4 1	0	0
2	C	1	Total O S 5 4 1	0	0
2	D	1	Total O S 5 4 1	0	0
2	D	1	Total O S 5 4 1	0	0
2	D	1	Total O S 5 4 1	0	0
2	D	1	Total O S 5 4 1	0	0
2	D	1	Total O S 5 4 1	0	0
2	D	1	Total O S 5 4 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	26	Total O 26 26	0	0
3	B	43	Total O 43 43	0	0
3	C	29	Total O 29 29	0	0

Continued on next page...

Continued from previous page...

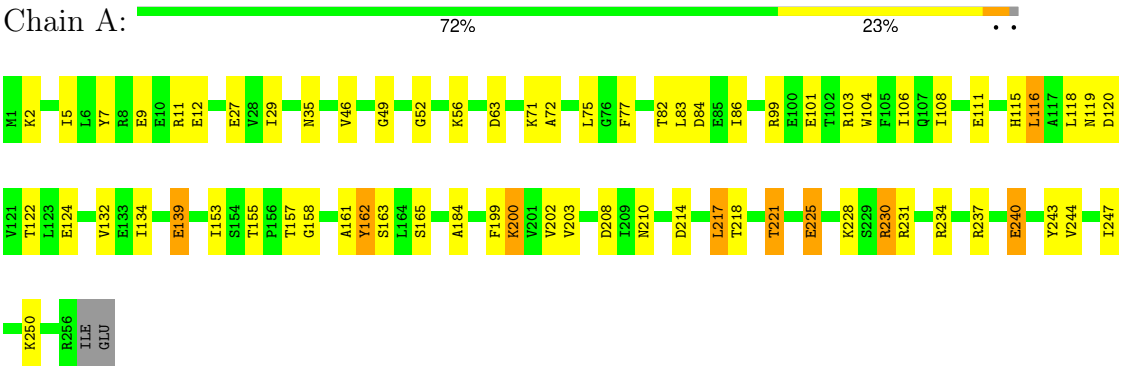
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	D	22	Total O 22 22	0	0

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.

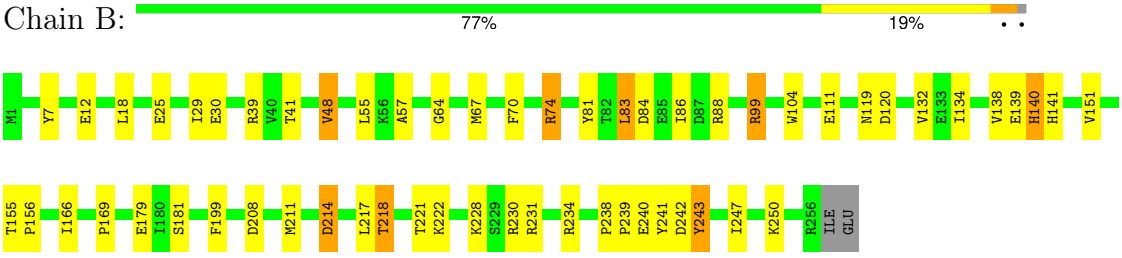
• Molecule 1: inorganic polyphosphate/ATP-NAD kinase

Chain A:



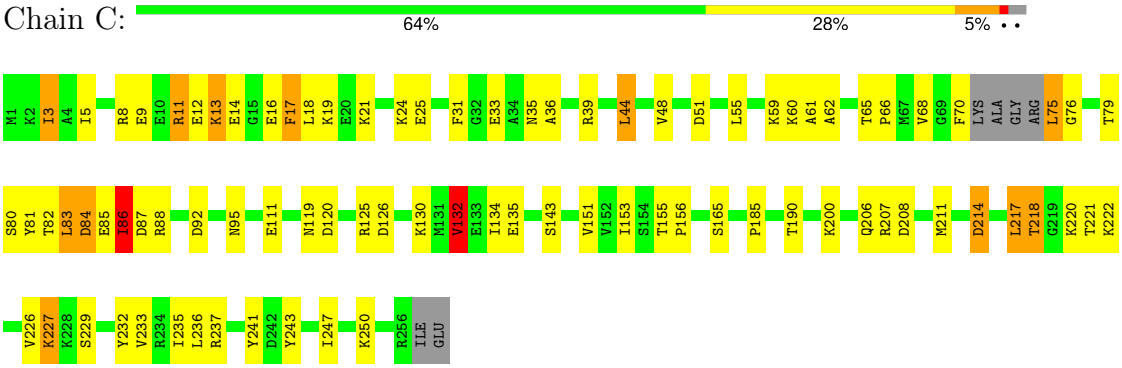
• Molecule 1: inorganic polyphosphate/ATP-NAD kinase

Chain B:



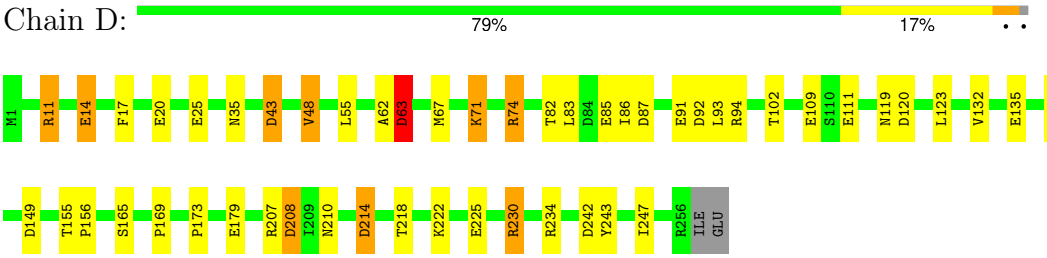
• Molecule 1: inorganic polyphosphate/ATP-NAD kinase

Chain C:



• Molecule 1: inorganic polyphosphate/ATP-NAD kinase

Chain D:


79%
17%
M1
R11
E14
F17
E20
E25
N35
D43
V48
L55
A62
D63
M67
K71
R74
T82
L83
D84
E85
I86
D87
E91
D92
L93
R94
T102
E109
S110
E111
N119
D120
L123
V132
E135
E139
H140
W145
D149
T155
P156
S165
P169
P173
E179
R207
D208
I209
N210
D214
T218
K222
E225
R230
R234
D242
Y243
I247
R256
ILE
GLU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	131.45Å 137.15Å 58.25Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	12.00 – 2.30 49.65 – 2.19	Depositor EDS
% Data completeness (in resolution range)	5.1 (12.00-2.30) 95.7 (49.65-2.19)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	3.13 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.0	Depositor
R, R_{free}	0.213 , 0.280 0.256 , 0.314	Depositor DCC
R_{free} test set	2676 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å^2)	38.7	Xtriage
Anisotropy	0.593	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 42.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	8357	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 21.17 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 7.4248e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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

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	0.78	0/2083	0.93	4/2812 (0.1%)
1	B	0.75	0/2083	0.93	5/2812 (0.2%)
1	C	0.69	0/2053	0.85	7/2772 (0.3%)
1	D	0.71	0/2083	0.90	10/2812 (0.4%)
All	All	0.73	0/8302	0.90	26/11208 (0.2%)

There are no bond length outliers.

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	208	ASP	CB-CG-OD2	8.52	125.96	118.30
1	D	92	ASP	CB-CG-OD2	7.69	125.22	118.30
1	D	87	ASP	CB-CG-OD2	7.32	124.89	118.30
1	B	214	ASP	CB-CG-OD2	7.05	124.65	118.30
1	D	208	ASP	CB-CG-OD2	6.96	124.56	118.30

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	2044	0	2069	45	0
1	B	2044	0	2069	30	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2015	0	2034	47	0
1	D	2044	0	2069	24	0
2	A	30	0	0	0	0
2	B	15	0	0	0	0
2	C	10	0	0	0	0
2	D	35	0	0	0	0
3	A	26	0	0	0	0
3	B	43	0	0	0	0
3	C	29	0	0	1	0
3	D	22	0	0	1	0
All	All	8357	0	8241	144	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 9.

The worst 5 of 144 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:D:71:LYS:NZ	1:D:74:ARG:O	1.73	1.20
1:A:118:LEU:O	1:A:155:THR:HG21	1.68	0.91
1:C:211:MET:CE	1:C:218:THR:HG21	2.09	0.81
1:C:82:THR:H	1:C:85:GLU:HG3	1.46	0.81
1:C:211:MET:HE3	1:C:218:THR:HG21	1.64	0.79

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	254/258 (98%)	235 (92%)	17 (7%)	2 (1%)	19	23
1	B	254/258 (98%)	240 (94%)	13 (5%)	1 (0%)	34	42

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	C	248/258 (96%)	237 (96%)	10 (4%)	1 (0%)	34 42
1	D	254/258 (98%)	241 (95%)	11 (4%)	2 (1%)	19 23
All	All	1010/1032 (98%)	953 (94%)	51 (5%)	6 (1%)	25 31

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	162	TYR
1	D	139	GLU
1	B	141	HIS
1	D	63	ASP
1	A	139	GLU

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	223/225 (99%)	207 (93%)	16 (7%)	14 18
1	B	223/225 (99%)	204 (92%)	19 (8%)	10 13
1	C	221/225 (98%)	183 (83%)	38 (17%)	2 2
1	D	223/225 (99%)	201 (90%)	22 (10%)	8 9
All	All	890/900 (99%)	795 (89%)	95 (11%)	6 7

5 of 95 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	87	ASP
1	C	235	ILE
1	C	92	ASP
1	C	217	LEU
1	D	17	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	107	GLN
1	A	206	GLN
1	C	95	ASN
1	C	186	GLN
1	D	107	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

18 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	317	-	4,4,4	0.18	0	6,6,6	0.21	0
2	SO4	D	315	-	4,4,4	0.22	0	6,6,6	0.25	0
2	SO4	B	318	-	4,4,4	0.23	0	6,6,6	0.24	0
2	SO4	B	310	-	4,4,4	0.21	0	6,6,6	0.28	0
2	SO4	A	313	-	4,4,4	0.24	0	6,6,6	0.21	0
2	SO4	D	311	-	4,4,4	0.22	0	6,6,6	0.16	0
2	SO4	D	303	-	4,4,4	0.26	0	6,6,6	0.24	0
2	SO4	B	302	-	4,4,4	0.26	0	6,6,6	0.27	0
2	SO4	A	309	-	4,4,4	0.25	0	6,6,6	0.13	0
2	SO4	A	301	-	4,4,4	0.23	0	6,6,6	0.39	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	SO4	C	307	-	4,4,4	0.25	0	6,6,6	0.16	0
2	SO4	D	312	-	4,4,4	0.21	0	6,6,6	0.09	0
2	SO4	D	306	-	4,4,4	0.23	0	6,6,6	0.10	0
2	SO4	A	316	-	4,4,4	0.22	0	6,6,6	0.16	0
2	SO4	D	308	-	4,4,4	0.24	0	6,6,6	0.16	0
2	SO4	C	305	-	4,4,4	0.30	0	6,6,6	0.43	0
2	SO4	A	304	-	4,4,4	0.25	0	6,6,6	0.42	0
2	SO4	D	314	-	4,4,4	0.23	0	6,6,6	0.10	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

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.