

# wwPDB X-ray Structure Validation Summary Report (i)

May 21, 2020 – 08:06 pm BST

PDB ID	:	1T3H
$\operatorname{Title}$	:	X-ray Structure of Dephospho-CoA Kinase from E. coli Norteast Structural
		Genomics Consortium Target ER57
Authors	:	Kuzin, A.P.; Chen, Y.; Forouhar, F.; Edstrom, W.; Benach, J.; Vorobiev, S.;
		Acton, T.; Shastry, R.; Ma, LC.; Xia, R.; Montelione, G.; Tong, L.; Hunt, J.;
		Northeast Structural Genomics Consortium (NESG)
Deposited on	:	2004-04-26
Resolution	:	2.50  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 following versions of software and data (see references (1)) were used in the production of this report:

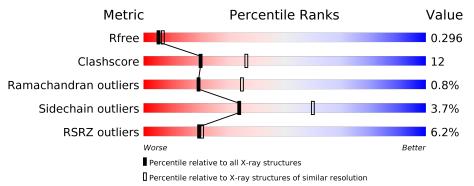
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
$\mathrm{EDS}$	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
$\operatorname{Refmac}$	:	5.8.0158
CCP4	:	7.0.044   (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber $(2001)$
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

# 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 2.50 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	$egin{array}{c} { m Whole \ archive} \ (\#{ m Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries},{ m resolution\ range}({ m \AA}))$
$R_{free}$	130704	4661(2.50-2.50)
Clashscore	141614	$5346 \ (2.50-2.50)$
Ramachandran outliers	138981	5231(2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559(2.50-2.50)

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% 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	А	214	7%	23%					
1	В	214	3% 75%	19%					
1	С	214	7%           59%         25°	% •	13%				

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:



Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	SO4	В	302	-	-	-	Х



# 2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4803 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	А	207	Total	С	Ν	Ο	Se	0	0	Ο
L	11	201	1603	1004	293	303	3	0		0
1	В	207	Total	С	Ν	Ο	Se	0	0	0
	D	201	1603	1004	293	303	3	0	0	U
1	C	187	Total	С	Ν	Ο	Se	0	0	0
L		C 187	1444	906	259	276	3	U	U	0

• Molecule 1 is a protein called Dephospho-CoA kinase.

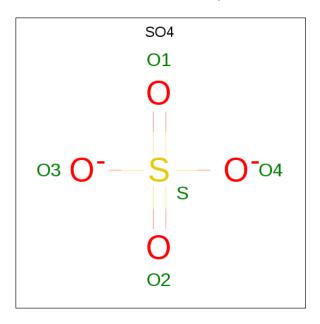
Chain	Residue	Modelled	Actual	Comment	Reference
А	1	MSE	MET	MODIFIED RESIDUE	UNP P0A6I9
A	58	MSE	MET	MODIFIED RESIDUE	UNP P0A6I9
A	142	MSE	MET	MODIFIED RESIDUE	UNP P0A6I9
A	207	GLU	-	EXPRESSION TAG	UNP P0A6I9
A	208	LEU	-	EXPRESSION TAG	UNP P0A6I9
A	209	HIS	-	EXPRESSION TAG	UNP P0A6I9
A	210	HIS	-	EXPRESSION TAG	UNP P0A6I9
A	211	HIS	-	EXPRESSION TAG	UNP P0A6I9
A	212	HIS	_	EXPRESSION TAG	UNP P0A6I9
A	213	HIS	-	EXPRESSION TAG	UNP P0A6I9
А	214	HIS	_	EXPRESSION TAG	UNP P0A6I9
В	1	MSE	MET	MODIFIED RESIDUE	UNP P0A6I9
В	58	MSE	MET	MODIFIED RESIDUE	UNP P0A6I9
В	142	MSE	MET	MODIFIED RESIDUE	UNP P0A6I9
В	207	GLU	-	EXPRESSION TAG	UNP P0A6I9
В	208	LEU	-	EXPRESSION TAG	UNP P0A6I9
В	209	HIS	-	EXPRESSION TAG	UNP P0A6I9
В	210	HIS	-	EXPRESSION TAG	UNP P0A6I9
В	211	HIS	-	EXPRESSION TAG	UNP P0A6I9
В	212	HIS	-	EXPRESSION TAG	UNP P0A6I9
В	213	HIS	-	EXPRESSION TAG	UNP P0A6I9
В	214	HIS	-	EXPRESSION TAG	UNP P0A6I9
С	1	MSE	MET	MODIFIED RESIDUE	UNP P0A6I9



Chain	Residue	Modelled	Actual	Comment	Reference
С	58	MSE	MET	MODIFIED RESIDUE	UNP P0A6I9
С	142	MSE	MET	MODIFIED RESIDUE	UNP P0A6I9
С	207	GLU	-	EXPRESSION TAG	UNP P0A6I9
С	208	LEU	-	EXPRESSION TAG	UNP P0A6I9
С	209	HIS	-	EXPRESSION TAG	UNP P0A6I9
С	210	HIS	-	EXPRESSION TAG	UNP P0A6I9
С	211	HIS	-	EXPRESSION TAG	UNP P0A6I9
С	212	HIS	-	EXPRESSION TAG	UNP P0A6I9
С	213	HIS	-	EXPRESSION TAG	UNP P0A6I9
С	214	HIS	-	EXPRESSION TAG	UNP P0A6I9

Continued from previous page...

• Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	А	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{S} \\ 5 & 4 & 1 \end{array}$	0	0
2	В	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{S} \\ 5 & 4 & 1 \end{array}$	0	0

• Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	А	41	Total         O           41         41	0	0
3	В	59	Total O 59 59	0	0



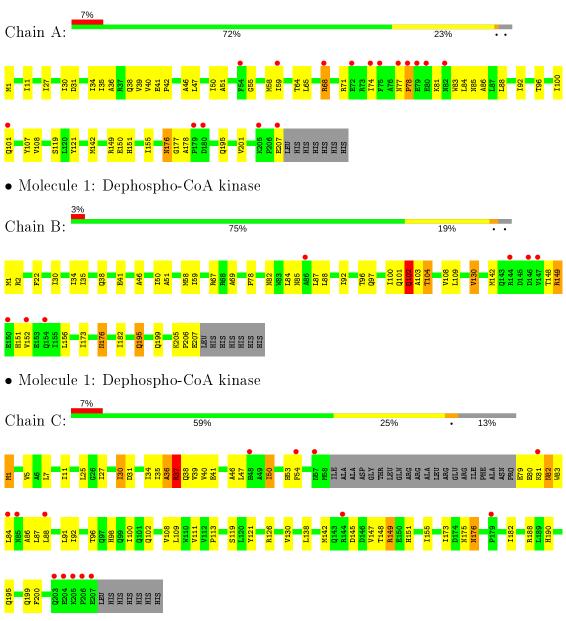
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	С	43	Total         O           43         43	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 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: Dephospho-CoA kinase



## 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants	55.63Å $80.91$ Å $75.45$ Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $93.95^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	19.53 - 2.50	Depositor
Resolution (A)	29.92 - 2.49	EDS
% Data completeness	91.2(19.53-2.50)	Depositor
(in resolution range)	97.9(29.92-2.49)	EDS
R <sub>merge</sub>	0.08	Depositor
R <sub>sym</sub>	0.08	Depositor
$< I/\sigma(I) > 1$	$3.17 (at 2.51 \text{\AA})$	Xtriage
Refinement program	CNS 1.1	Depositor
D D.	0.238 , $0.284$	Depositor
$R, R_{free}$	0.250 , $0.296$	DCC
$R_{free}$ test set	2268 reflections $(5.02\%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	27.1	Xtriage
Anisotropy	0.164	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.31 , $42.1$	EDS
L-test for twinning <sup>2</sup>	$ \langle L  \rangle = 0.46, \langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	4803	wwPDB-VP
Average B, all atoms $(Å^2)$	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 4.76% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 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: 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 Bo		nd lengths	Bond angles	
	Unam	RMSZ	# Z  > 5	RMSZ	# Z  > 5
1	А	0.40	0/1628	0.63	0/2214
1	В	1.01	4/1628~(0.2%)	1.00	9/2214~(0.4%)
1	С	0.55	3/1466~(0.2%)	0.87	7/1994~(0.4%)
All	All	0.71	7/4722~(0.1%)	0.85	16/6422~(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	В	1	2
1	С	0	1
All	All	1	3

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(A)	Ideal(Å)
1	В	102	GLN	CG-CD	-25.72	0.91	1.51
1	В	104	THR	N-CA	-20.23	1.05	1.46
1	В	102	GLN	CB-CG	-15.53	1.10	1.52
1	С	37	ARG	CA-C	-8.30	1.31	1.52
1	С	36	ALA	C-N	-8.11	1.15	1.34

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\mathbf{Observed}(^{o})$	$Ideal(^{o})$
1	В	102	GLN	CA-CB-CG	21.58	160.87	113.40
1	С	37	ARG	CB-CA-C	-16.30	77.81	110.40
1	В	102	GLN	CB-CA-C	15.35	141.10	110.40



Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
1	В	104	THR	N-CA-CB	12.95	134.91	110.30
1	В	103	ALA	C-N-CA	10.99	149.18	121.70

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	В	102	GLN	CA

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	В	102	GLN	Sidechain
1	В	104	THR	Mainchain
1	С	37	ARG	Mainchain

#### 5.2 Too-close contacts (i)

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	А	1603	0	1609	37	0
1	В	1603	0	1608	31	0
1	С	1444	0	1441	48	0
2	А	5	0	0	0	0
2	В	5	0	0	0	0
3	А	41	0	0	4	0
3	В	59	0	0	5	0
3	С	43	0	0	3	0
All	All	4803	0	4658	116	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 12.

The worst 5 of 116 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-1 Atom-2		Clash overlap (Å)	
1:B:30:ILE:HD12	1:B:100:ILE:HD13	1.44	0.95	



Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:30:ILE:HD13	1:A:100:ILE:HD13	1.62	0.79
1:C:38:GLN:HA	1:C:41:GLU:HG3	1.63	0.78
1:A:47:LEU:HD22	1:A:59:ILE:HD12	1.69	0.74
1:C:36:ALA:HA	1:C:92:ILE:CD1	2.19	0.71

Continued from previous page...

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	Perce	ntiles
1	А	205/214~(96%)	189~(92%)	12~(6%)	4 (2%)	7	12
1	В	205/214~(96%)	192 (94%)	13~(6%)	0	100	100
1	С	183/214~(86%)	166 (91%)	16 (9%)	1 (0%)	29	48
All	All	593/642~(92%)	547 (92%)	41 (7%)	5 (1%)	19	35

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	А	84	LEU
1	А	178	ALA
1	С	50	ILE
1	А	78	PRO
1	А	177	GLY

#### 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



Mol	Chain	Analysed	Rotameric Outliers		Percentiles		
1	А	168/172~(98%)	163~(97%)	5(3%)	41 68		
1	В	168/172~(98%)	163~(97%)	5(3%)	41 68		
1	С	153/172~(89%)	145~(95%)	8 (5%)	23 44		
All	All	489/516~(95%)	471 (96%)	18 (4%)	34 60		

analysed, and the total number of residues.

5 of 18 residues with a non-rotameric side chain are listed below:

Mo	1	Chain	$\mathbf{Res}$	Type
1		В	176	ASN
1		В	195	GLN
1		С	82	ASN
1		В	102	GLN
1		В	130	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
1	В	151	HIS
1	В	176	ASN
1	С	94	GLN
1	В	94	GLN
1	В	98	HIS

#### 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 carbohydrates in this entry.



#### 5.6 Ligand geometry (i)

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 C		Chain Res	Res Link	Link	Bond lengths			Bond angles		
	Type	Chain	nes		Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	SO4	А	301	-	4,4,4	0.20	0	$^{6,6,6}$	0.11	0
2	SO4	В	302	-	4,4,4	0.29	0	$^{6,6,6}$	0.17	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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	С	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	С	37:ARG	С	38:GLN	Ν	1.18
1	С	36:ALA	С	37:ARG	Ν	1.15



## 6 Fit of model and data (i)

## 6.1 Protein, DNA and RNA chains (i)

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, 95<sup>th</sup> 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 $>$	$\# RSRZ {>}2$	$OWAB(Å^2)$	$\mathbf{Q}{<}0.9$
1	А	204/214~(95%)	0.36	16 (7%) 13 13	13, 33, 67, 78	1 (0%)
1	В	204/214~(95%)	0.04	7 (3%) 45 48	6, 24, 56, 73	2(0%)
1	С	184/214~(85%)	0.35	14 (7%) 13 14	7, 28, 65, 80	1 (0%)
All	All	592/642~(92%)	0.25	37 (6%) 20 21	6, 27, 66, 80	4 (0%)

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	А	179	PRO	7.1
1	С	206	PRO	7.0
1	С	207	GLU	6.6
1	С	205	LYS	5.7
1	А	180	ASP	5.2

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

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

### 6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

### 6.4 Ligands (i)

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,  $95^{th}$  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	$\mathbf{B} ext{-factors}(\mathbf{A}^2)$	$\mathbf{Q}{<}0.9$
2	SO4	В	302	5/5	0.47	0.73	$114,\!114,\!115,\!115$	0
2	SO4	А	301	5/5	0.97	0.13	42,42,44,44	0

## 6.5 Other polymers (i)

There are no such residues in this entry.

