



# wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 02:58 AM GMT

PDB ID : 1KMM  
Title : HISTIDYL-TRNA SYNTHETASE COMPLEXED WITH HISTIDYL-ADENYLATE  
Authors : Arnez, J.G.; Francklyn, C.S.; Moras, D.  
Deposited on : 1997-05-09  
Resolution : 2.60 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

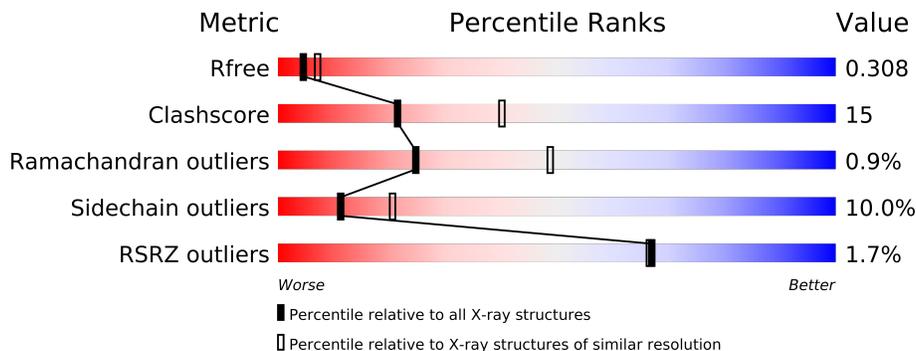
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.60 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))
$R_{free}$	66092	1718 (2.60-2.60)
Clashscore	79885	2154 (2.60-2.60)
Ramachandran outliers	78287	2113 (2.60-2.60)
Sidechain outliers	78261	2113 (2.60-2.60)
RSRZ outliers	66119	1718 (2.60-2.60)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	424	
1	B	424	
1	C	424	
1	D	424	

## 2 Entry composition i

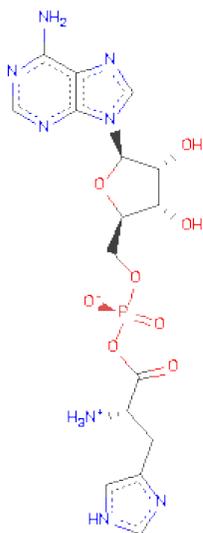
There are 3 unique types of molecules in this entry. The entry contains 14633 atoms, of which 2685 are hydrogens and 0 are deuterium.

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 HISTIDYL-TRNA SYNTHETASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	386	Total	C	H	N	O	S	0	0	0
			3675	1884	682	535	563	11			
1	B	364	Total	C	H	N	O	S	0	0	0
			3472	1777	648	507	529	11			
1	C	387	Total	C	H	N	O	S	0	0	0
			3689	1892	684	538	564	11			
1	D	362	Total	C	H	N	O	S	0	0	0
			3453	1768	643	504	527	11			

- Molecule 2 is HISTIDYL-ADENOSINE MONOPHOSPHATE (three-letter code: HAM) (formula:  $C_{16}H_{21}N_8O_8P$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	H	N	O			P
2	A	1	Total	C	H	N	O	P	0	0
			40	16	7	8	8	1		
2	B	1	Total	C	H	N	O	P	0	0
			40	16	7	8	8	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
2	C	1	Total	C	H	N	O	P	0	0
			40	16	7	8	8	1		
2	D	1	Total	C	H	N	O	P	0	0
			40	16	7	8	8	1		

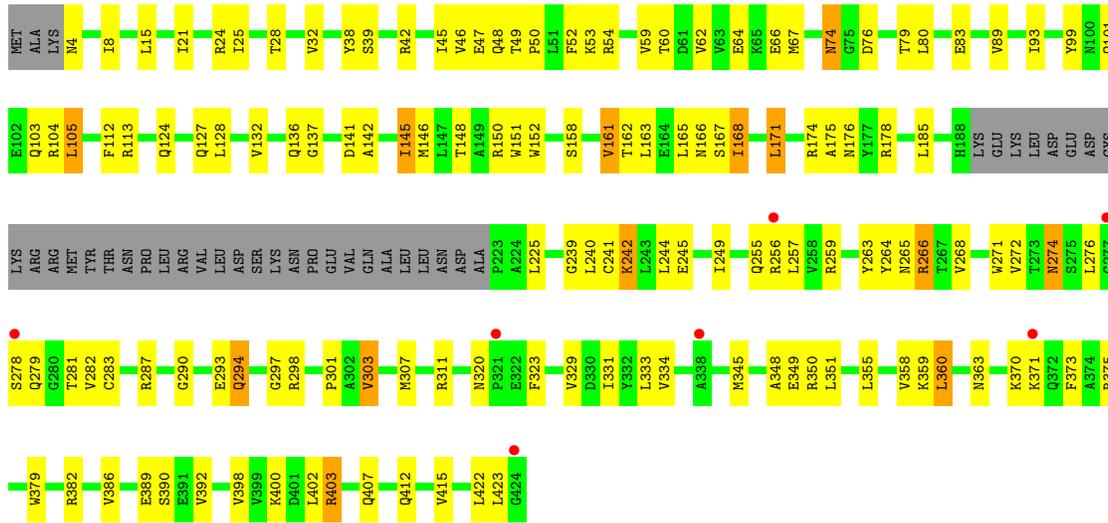
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	45	Total	O	0	0
			45	45		
3	B	50	Total	O	0	0
			50	50		
3	C	55	Total	O	0	0
			55	55		
3	D	34	Total	O	0	0
			34	34		



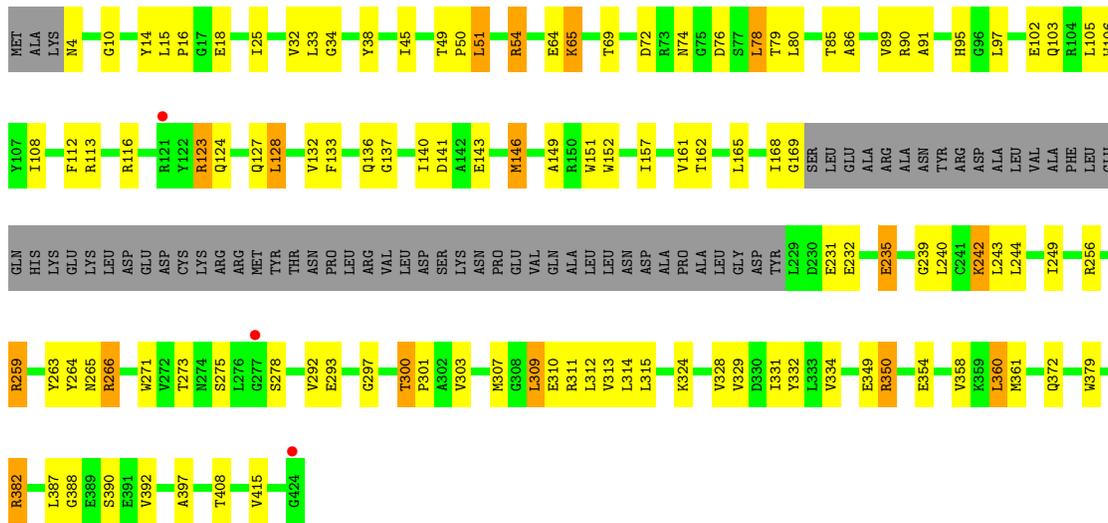
- Molecule 1: HISTIDYL-TRNA SYNTHETASE

Chain C:



- Molecule 1: HISTIDYL-TRNA SYNTHETASE

Chain D:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	61.40Å 110.70Å 108.70Å 115.00° 97.40° 90.00°	Depositor
Resolution (Å)	12.00 – 2.60 12.04 – 2.60	Depositor EDS
% Data completeness (in resolution range)	68.0 (12.00-2.60) 84.9 (12.04-2.60)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.66 (at 2.59Å)	Xtrriage
Refinement program	X-PLOR	Depositor
R, $R_{free}$	0.221 , 0.297 0.252 , 0.308	Depositor DCC
$R_{free}$ test set	1679 reflections (2.53%)	DCC
Wilson B-factor (Å <sup>2</sup> )	56.7	Xtrriage
Anisotropy	0.239	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 56.2	EDS
Estimated twinning fraction	No twinning to report.	Xtrriage
L-test for twinning	$\langle  L  \rangle = 0.54$ , $\langle L^2 \rangle = 0.39$	Xtrriage
Outliers	0 of 66471 reflections	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	14633	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	38.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.90% of the height of the origin peak. No significant pseudotranslation is detected.*

<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: HAM

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	0.59	0/3047	0.85	6/4123 (0.1%)
1	B	0.60	1/2875 (0.0%)	0.80	5/3888 (0.1%)
1	C	0.58	0/3061	0.80	4/4142 (0.1%)
1	D	0.57	0/2861	0.77	2/3870 (0.1%)
All	All	0.59	1/11844 (0.0%)	0.81	17/16023 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	130	CYS	CB-SG	-6.51	1.71	1.82

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	320	ASN	C-N-CD	-12.73	92.60	120.60
1	B	387	LEU	CA-CB-CG	7.16	131.76	115.30
1	A	323	PHE	N-CA-C	-6.18	94.31	111.00
1	C	323	PHE	N-CA-C	-5.88	95.12	111.00
1	D	360	LEU	CA-CB-CG	5.66	128.31	115.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the

chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2993	682	2269	88	0
1	B	2824	648	2152	110	0
1	C	3005	684	2279	94	0
1	D	2810	643	2139	86	0
2	A	33	7	14	2	0
2	B	33	7	14	1	0
2	C	33	7	14	2	0
2	D	33	7	14	5	0
3	A	45	0	0	1	0
3	B	50	0	0	1	0
3	C	55	0	0	1	0
3	D	34	0	0	0	0
All	All	11948	2685	8895	358	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 15.

The worst 5 of 358 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:382:ARG:HH11	1:D:382:ARG:HG2	1.36	0.89
1:B:3:LYS:NZ	1:B:3:LYS:HB3	1.87	0.89
1:B:89:VAL:HG21	1:B:291:LEU:HD22	1.55	0.89
1:D:311:ARG:NH1	2:D:450:HAM:H2'	1.88	0.89
1:A:161:VAL:HG12	1:A:271:TRP:HE3	1.38	0.88

There are no symmetry-related clashes.

## 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	382/424 (90%)	360 (94%)	16 (4%)	6 (2%)	14 28
1	B	360/424 (85%)	340 (94%)	17 (5%)	3 (1%)	27 53

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	C	383/424 (90%)	358 (94%)	23 (6%)	2 (0%)	38 67
1	D	358/424 (84%)	342 (96%)	14 (4%)	2 (1%)	33 63
All	All	1483/1696 (87%)	1400 (94%)	70 (5%)	13 (1%)	25 49

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	321	PRO
1	A	137	GLY
1	B	137	GLY
1	C	137	GLY
1	A	266	ARG

### 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	304/342 (89%)	275 (90%)	29 (10%)	12 23
1	B	289/342 (84%)	259 (90%)	30 (10%)	10 18
1	C	306/342 (90%)	274 (90%)	32 (10%)	10 18
1	D	288/342 (84%)	260 (90%)	28 (10%)	12 22
All	All	1187/1368 (87%)	1068 (90%)	119 (10%)	11 20

5 of 119 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	333	LEU
1	C	161	VAL
1	D	275	SER
1	B	387	LEU
1	C	39	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	23	GLN
1	B	418	HIS
1	D	236	HIS
1	B	166	ASN
1	B	341	GLN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

4 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	HAM	A	450	-	36,36,36	1.31	4 (11%)	53,53,53	1.79	6 (11%)
2	HAM	B	450	-	36,36,36	1.54	4 (11%)	53,53,53	1.36	5 (9%)
2	HAM	C	450	-	36,36,36	1.40	3 (8%)	53,53,53	1.68	5 (9%)
2	HAM	D	450	-	36,36,36	1.34	4 (11%)	53,53,53	1.25	3 (5%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	HAM	A	450	-	-	0/21/39/39	0/2/4/4
2	HAM	B	450	-	-	0/21/39/39	0/2/4/4
2	HAM	C	450	-	-	0/21/39/39	0/2/4/4
2	HAM	D	450	-	-	0/21/39/39	0/2/4/4

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	450	HAM	CA-C	5.67	1.59	1.52
2	C	450	HAM	P-O3P	5.15	1.70	1.61
2	B	450	HAM	P-O3P	4.88	1.69	1.61
2	A	450	HAM	P-O3P	4.45	1.68	1.61
2	C	450	HAM	CA-C	4.44	1.58	1.52

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	450	HAM	O3P-C-O	9.65	129.75	122.11
2	C	450	HAM	O3P-C-O	8.75	129.04	122.11
2	D	450	HAM	O3P-C-O	5.36	126.35	122.11
2	B	450	HAM	O3P-C-O	4.83	125.93	122.11
2	C	450	HAM	O-C-CA	-4.51	108.59	121.64

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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 [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 <sup>2</sup> )	Q<0.9
1	A	386/424 (91%)	-0.16	8 (2%) 60 58	19, 43, 78, 94	0
1	B	364/424 (85%)	-0.17	8 (2%) 59 56	20, 45, 81, 108	0
1	C	387/424 (91%)	-0.15	7 (1%) 65 64	19, 45, 77, 98	0
1	D	362/424 (85%)	-0.12	3 (0%) 83 85	20, 44, 81, 107	0
All	All	1499/1696 (88%)	-0.15	26 (1%) 67 66	19, 44, 80, 108	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	178	ARG	3.7
1	C	338	ALA	3.4
1	C	424	GLY	3.3
1	A	278	SER	3.0
1	D	277	GLY	2.9

### 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	HAM	C	450	33/33	0.15	0.70	0,31,35,37	0
2	HAM	A	450	33/33	0.16	0.66	0,28,35,37	0
2	HAM	D	450	33/33	0.18	0.34	0,49,54,54	0
2	HAM	B	450	33/33	0.17	0.27	0,42,49,50	0

## 6.5 Other polymers

There are no such residues in this entry.