



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 08:31 PM 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 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
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

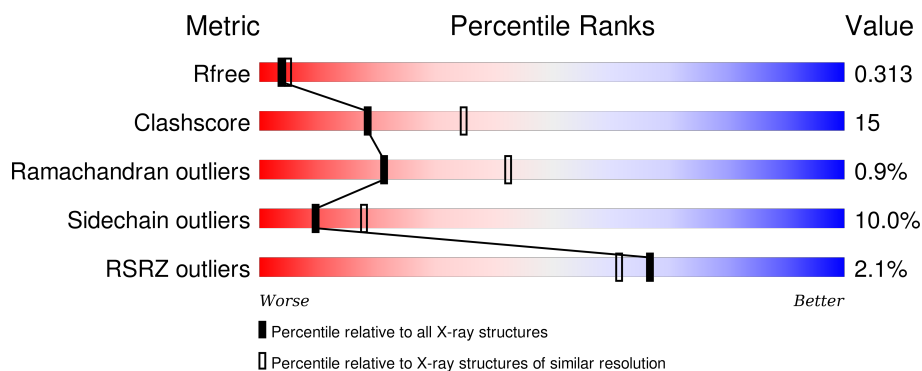
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

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}	91344	2328 (2.60-2.60)
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)
RSRZ outliers	91569	2334 (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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$. 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	A	424	<div> <div>3%</div> <div>58% 29% 9%</div> </div>
1	B	424	<div> <div>2%</div> <div>54% 27% 5% 14%</div> </div>
1	C	424	<div> <div>2%</div> <div>59% 29% 9%</div> </div>
1	D	424	<div> <div>%</div> <div>58% 24% 15%</div> </div>

2 Entry composition [i](#)

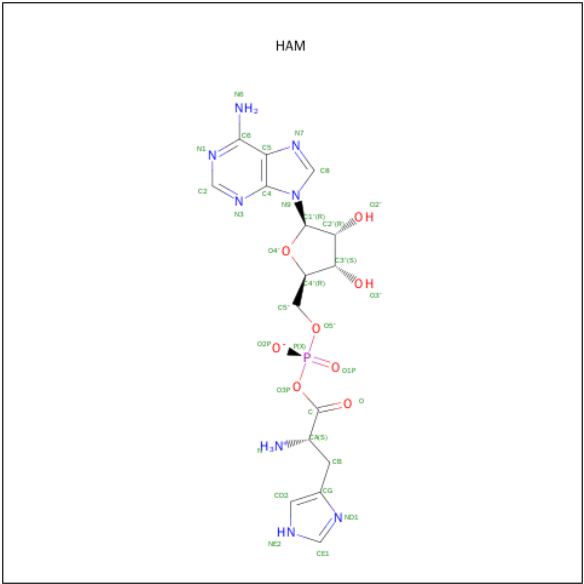
There are 3 unique types of molecules in this entry. The entry contains 14633 atoms, of which 2685 are hydrogens and 0 are deuteriums.

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 HISTIDYL-TRNA SYNTHETASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
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: C₁₆H₂₁N₈O₈P).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
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		

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ($\text{RSRZ} > 2$). 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.

Chain A:

3% 58% 29% 9%

ASP CYS LYS ARG MET THR ASN PRO LEU ARG VAL LEU ASP SER LYS ASN PRO GLU VAL GLN ALA LEU ASP ALA PRO K224 L225 G226 D227 Y228 F237 L240 L243 L249 A250 Y251 T252 Q255 R256 L257 R258 R259 G260 L261 Y264 M265 R266 T267 Y268 K271 R274 S275 L276 G277 S278 Q279 C280 T281 V282 C283 R287 Y288 V292 L295 R298 A299 T300 V303 A306 P307 R311 L312 V313 R320 P321 E322 F323 K324 I331 Y332 L333 V334 T340 G341 S342 A343 R350 E354 L360 N363 H364 K371 R372

MET ALA LYS H4 G10 D13 Y14 L15 P16 T19 A20 I21 W22 Q23 R24 I25 K30 N31 V32 Y36 E40 I41 R42 T49 P50 R54 A55 I56 G57 E58 V59 T60 E64 R65 F70 N74 T79 L80 E83 E86 Q87 R89 L98 Y99 H100 Q101 W106 M111 F112 R113 R123 Q124 F125 H126 Q127 L128 F133 G137 P138 D139 I140 D141 M146 A149 R150 R153 A154 L155 S158 E159 H160 V161 T162 L163 E164 L165 I168 A175 R176 Y177 R178 D179 V182 A183 F184 L185 E186 Q187 H188 LYS GLU LYS LEU ASP CYS LYS ARG MET THR ASN PRO LEU ARG VAL LEU ASP SER LYS ASN PRO GLU VAL GLN ALA LEU ASP ALA PRO K224 L225 G226 D227 Y228 F237 L240 L243 L249 A250 Y251 T252 Q255 R256 L257 R258 R259 G260 L261 Y264 M265 R266 T267 Y268 K271 R274 S275 L276 G277 S278 Q279 C280 T281 V282 C283 R287 Y288 V292 L295 R298 A299 T300 V303 A306 P307 R311 L312 V313 R320 P321 E322 F323 K324 I331 Y332 L333 V334 T340 G341 S342 A343 R350 E354 L360 N363 H364 K371 R372

Chain B:

2% 54% 27% 5% 14%

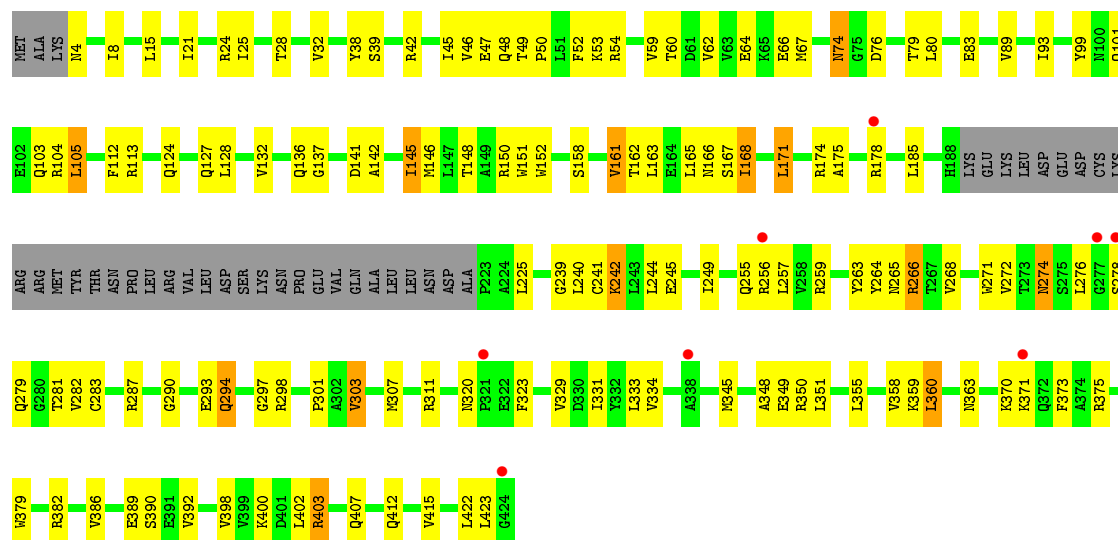
ALA
PHE
GLU
GLN
HIS
LYS
GLU
LYS
LEU
ASP
GLU
CYS
LYS
ARG
ARG
MET
THR
ASN
PRO
LEU
ARG
VAL
LEU
ASP
SER
LYS
ASN
PRO
GLU
VAL
GLN
ALA
LEU
LEU
ASN
ASP
ALA
PRO
ALA
LEU
GLY
ASP
TYR
L229
D230
E231
E232
E235
H236
F237
A238
G239
L240
C241
K242
L243

R104
L105
I108
E115
R116
P117
Q118
K119
G120
R121
Y122
R123
Q127
A128
G129
C130
F133
G134
L135
Q136
G137
P138
D139
I140
D141
A142
E143
M146
W152
S158
T159
H160
V161
T162
L165
M166
S167
G169
SER
LEU
GLU
ALA
ARG
ALA
ASN
ARG
ASP
ALA
VAL

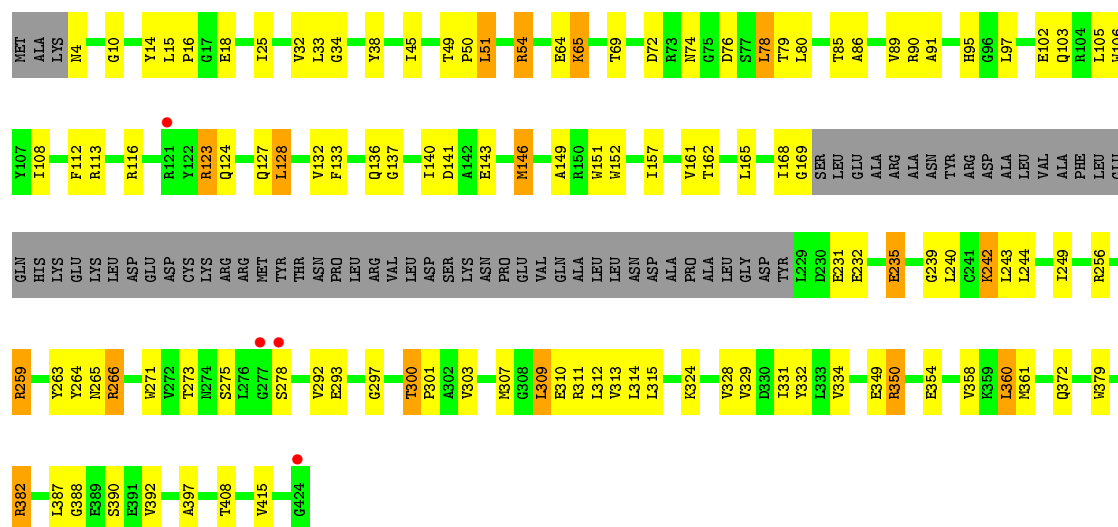
K3
A7
I8
R9
M11
Y14
L15
T19
I21
P22
Q23
R24
L33
Y38
S39
E40
I41
R42
I45
V46
E47
F52
A55
V59
T60
D61
E64
K65
M74
L80
E83
V89
R90
I93
L97
M100
Q101
E102
G103



• Molecule 1: HISTIDYL-TRNA SYNTHETASE



• Molecule 1: HISTIDYL-TRNA SYNTHETASE



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	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) 74.7 (12.04-2.60)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.66 (at 2.59Å)	Xtriage
Refinement program	X-PLOR	Depositor
R, R_{free}	0.221 , 0.297 0.258 , 0.313	Depositor DCC
R_{free} test set	1684 reflections (2.60%)	DCC
Wilson B-factor (Å ²)	56.7	Xtriage
Anisotropy	0.239	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 89.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.54$, $\langle L^2 \rangle = 0.39$	Xtriage
Outliers	0 of 66471 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	14633	wwPDB-VP
Average B, all atoms (Å ²)	38.0	wwPDB-VP

Xtriage'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.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

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 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	A	2993	682	2949	88	0
1	B	2824	648	2798	113	0
1	C	3005	684	2962	95	0
1	D	2810	643	2780	85	0
2	A	33	7	21	2	0
2	B	33	7	21	1	0
2	C	33	7	21	2	0
2	D	33	7	21	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	11573	361	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 15.

The worst 5 of 361 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:382:ARG:HH11	1:D:382:ARG:HG2	1.35	0.90
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 [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	382/424 (90%)	360 (94%)	16 (4%)	6 (2%)	12 24

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	360/424 (85%)	340 (94%)	17 (5%)	3 (1%)	24	46
1	C	383/424 (90%)	358 (94%)	23 (6%)	2 (0%)	34	60
1	D	358/424 (84%)	342 (96%)	14 (4%)	2 (1%)	30	56
All	All	1483/1696 (87%)	1400 (94%)	70 (5%)	13 (1%)	21	42

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%)	11	20
1	B	289/342 (84%)	259 (90%)	30 (10%)	9	16
1	C	306/342 (90%)	274 (90%)	32 (10%)	8	16
1	D	288/342 (84%)	260 (90%)	28 (10%)	10	19
All	All	1187/1368 (87%)	1068 (90%)	119 (10%)	9	18

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 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 ⓘ

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	-	27,36,36	1.35	3 (11%)	26,53,53	1.48	4 (15%)
2	HAM	B	450	-	27,36,36	1.46	2 (7%)	26,53,53	1.24	2 (7%)
2	HAM	C	450	-	27,36,36	1.40	2 (7%)	26,53,53	1.34	5 (19%)
2	HAM	D	450	-	27,36,36	1.23	2 (7%)	26,53,53	1.01	1 (3%)

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	-	-	0/17/39/39	0/4/4/4
2	HAM	B	450	-	-	0/17/39/39	0/4/4/4
2	HAM	C	450	-	-	0/17/39/39	0/4/4/4
2	HAM	D	450	-	-	0/17/39/39	0/4/4/4

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	450	HAM	O-C	2.20	1.26	1.21
2	D	450	HAM	P-O3P	3.38	1.67	1.60
2	A	450	HAM	CA-C	3.54	1.57	1.53
2	D	450	HAM	CA-C	3.85	1.58	1.53
2	C	450	HAM	CA-C	3.93	1.58	1.53

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	450	HAM	O-C-CA	-4.17	108.59	121.03
2	A	450	HAM	O-C-CA	-3.93	109.31	121.03
2	B	450	HAM	O-C-CA	-3.92	109.33	121.03
2	D	450	HAM	O-C-CA	-3.54	110.46	121.03
2	C	450	HAM	O5'-P-O1P	-2.08	101.55	109.62

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	450	HAM	2	0
2	B	450	HAM	1	0
2	C	450	HAM	2	0
2	D	450	HAM	5	0

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, 95th 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)	Q<0.9
1	A	386/424 (91%)	-0.18	11 (2%) 56 49	19, 43, 78, 94	0
1	B	364/424 (85%)	-0.17	9 (2%) 61 54	20, 45, 81, 108	0
1	C	387/424 (91%)	-0.16	8 (2%) 67 61	19, 45, 77, 98	0
1	D	362/424 (85%)	-0.13	4 (1%) 82 79	20, 44, 81, 107	0
All	All	1499/1696 (88%)	-0.16	32 (2%) 67 61	19, 44, 80, 108	0

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	178	ARG	4.0
1	D	424	GLY	3.8
1	C	338	ALA	3.3
1	C	424	GLY	3.3
1	B	74	ASN	3.0

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, 95th 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	LLDF	B-factors(\AA^2)	Q<0.9
2	HAM	C	450	33/33	0.95	0.15	0.63	0,31,35,37	0
2	HAM	A	450	33/33	0.91	0.16	0.57	0,28,35,37	0
2	HAM	D	450	33/33	0.90	0.18	0.39	0,49,54,54	0
2	HAM	B	450	33/33	0.90	0.17	0.28	0,42,49,50	0

6.5 Other polymers [i](#)

There are no such residues in this entry.