



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 6, 2026 – 10:21 AM UTC

PDB ID : 1WSJ / pdb_00001wsj
Title : Crystal structure of E.coli RNase HI active site mutant (K87A/H124A)
Authors : Tsunaka, Y.; Takano, K.; Matsumura, H.; Yamagata, Y.; Kanaya, S.
Deposited on : 2004-11-07
Resolution : 2.00 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4-5-2 with Phenix2.0
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

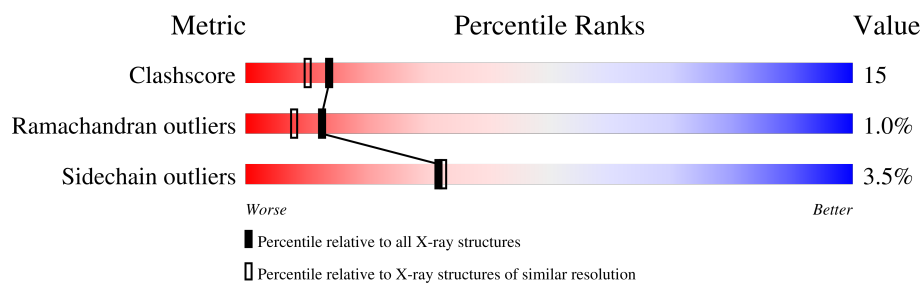
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	190562	11152 (2.00-2.00)
Ramachandran outliers	187476	11031 (2.00-2.00)
Sidechain outliers	187428	11029 (2.00-2.00)

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 $\leq 5\%$.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	155	77% 19% .
1	B	155	68% 28% ..
1	C	155	69% 25% ...
1	D	155	58% 34% . .
1	E	155	68% 28% ..
1	F	155	65% 30% ..
1	G	155	66% 26% . 5%
1	H	155	67% 26% . .

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 9931 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 Ribonuclease HI.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	155	Total	C	N	O	S	0	0	0
			1228	770	224	227	7			
1	B	153	Total	C	N	O	S	0	0	0
			1212	760	222	223	7			
1	C	149	Total	C	N	O	S	0	0	0
			1179	739	216	217	7			
1	D	149	Total	C	N	O	S	0	0	0
			1189	746	217	219	7			
1	E	153	Total	C	N	O	S	0	0	0
			1212	760	222	223	7			
1	F	152	Total	C	N	O	S	0	0	0
			1212	760	221	224	7			
1	G	148	Total	C	N	O	S	0	0	0
			1182	741	216	218	7			
1	H	151	Total	C	N	O	S	0	0	0
			1203	755	220	221	7			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	87	ALA	LYS	engineered mutation	UNP P0A7Y4
A	124	ALA	HIS	engineered mutation	UNP P0A7Y4
B	87	ALA	LYS	engineered mutation	UNP P0A7Y4
B	124	ALA	HIS	engineered mutation	UNP P0A7Y4
C	87	ALA	LYS	engineered mutation	UNP P0A7Y4
C	124	ALA	HIS	engineered mutation	UNP P0A7Y4
D	87	ALA	LYS	engineered mutation	UNP P0A7Y4
D	124	ALA	HIS	engineered mutation	UNP P0A7Y4
E	87	ALA	LYS	engineered mutation	UNP P0A7Y4
E	124	ALA	HIS	engineered mutation	UNP P0A7Y4
F	87	ALA	LYS	engineered mutation	UNP P0A7Y4
F	124	ALA	HIS	engineered mutation	UNP P0A7Y4
G	87	ALA	LYS	engineered mutation	UNP P0A7Y4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	124	ALA	HIS	engineered mutation	UNP P0A7Y4
H	87	ALA	LYS	engineered mutation	UNP P0A7Y4
H	124	ALA	HIS	engineered mutation	UNP P0A7Y4

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	47	Total O 47 47	0	0
2	B	47	Total O 47 47	0	0
2	C	44	Total O 44 44	0	0
2	D	33	Total O 33 33	0	0
2	E	26	Total O 26 26	0	0
2	F	32	Total O 32 32	0	0
2	G	35	Total O 35 35	0	0
2	H	50	Total O 50 50	0	0

3 Residue-property plots

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.

Note EDS was not executed.

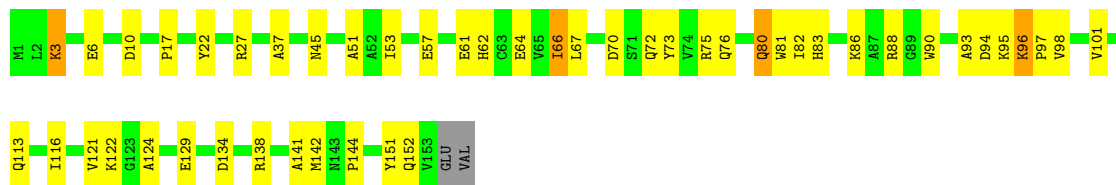
• Molecule 1: Ribonuclease HI

Chain A: 



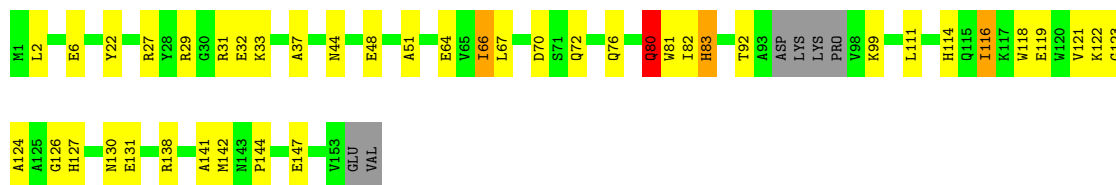
• Molecule 1: Ribonuclease HI

Chain B: 



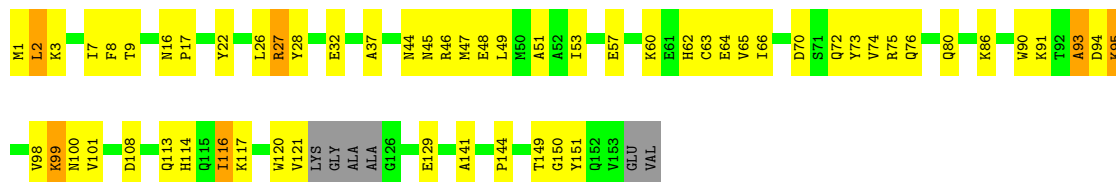
• Molecule 1: Ribonuclease HI

Chain C: 



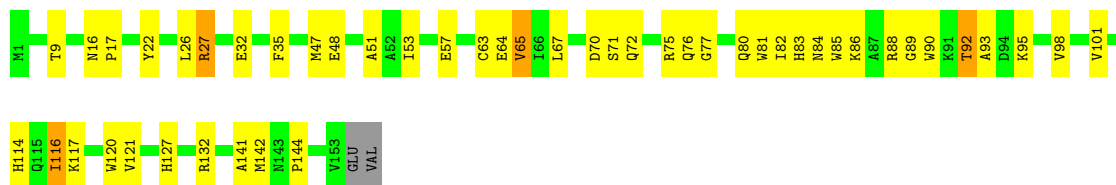
• Molecule 1: Ribonuclease HI

Chain D: 



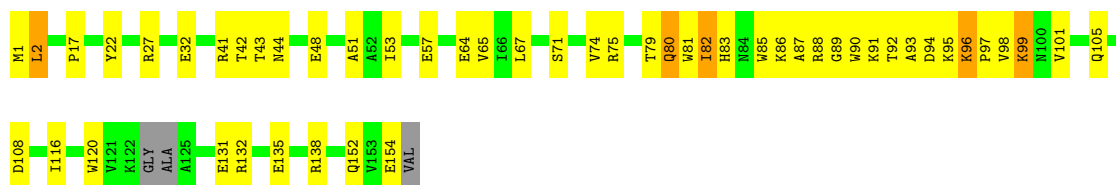
• Molecule 1: Ribonuclease HI

Chain E:  68% 28% ..



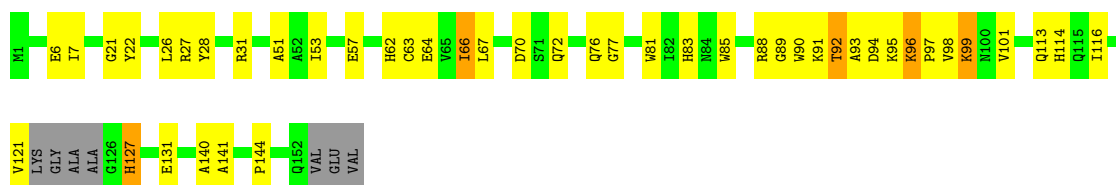
- Molecule 1: Ribonuclease HI

Chain F:  65% 30% ..



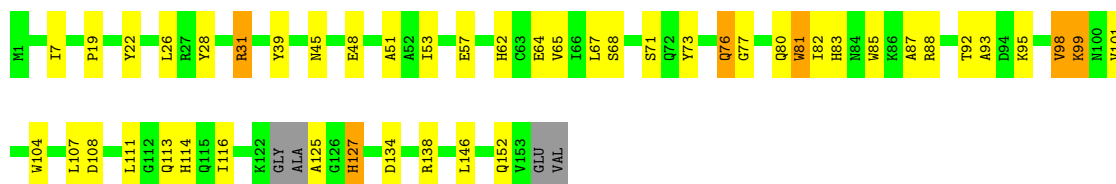
- Molecule 1: Ribonuclease HI

Chain G:  66% 26% • 5%



- Molecule 1: Ribonuclease HI

Chain H:  67% 26% ..



4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	55.76Å 75.08Å 147.87Å 90.00° 90.98° 90.00°	Depositor
Resolution (Å)	50.00 – 2.00	Depositor
% Data completeness (in resolution range)	(Not available) (50.00-2.00)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.235 , 0.288	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	9931	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	0.43	0/1255	0.87	5/1698 (0.3%)
1	B	0.39	0/1239	0.85	3/1676 (0.2%)
1	C	0.40	0/1204	0.90	6/1628 (0.4%)
1	D	0.35	0/1215	0.78	2/1643 (0.1%)
1	E	0.38	0/1239	0.86	6/1676 (0.4%)
1	F	0.40	0/1238	0.87	3/1673 (0.2%)
1	G	0.37	0/1208	0.83	4/1633 (0.2%)
1	H	0.40	0/1229	0.86	4/1661 (0.2%)
All	All	0.39	0/9827	0.85	33/13288 (0.2%)

There are no bond length outliers.

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	116	ILE	N-CA-C	8.42	120.73	108.53
1	H	81	TRP	N-CA-C	8.38	120.42	111.28
1	A	64	GLU	N-CA-C	-7.55	96.93	109.24
1	E	81	TRP	N-CA-C	7.52	120.43	111.33
1	E	64	GLU	N-CA-C	-7.34	97.28	108.96

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	1228	0	1207	27	0
1	B	1212	0	1192	32	0
1	C	1179	0	1154	31	0
1	D	1189	0	1165	36	0
1	E	1212	0	1192	42	0
1	F	1212	0	1189	47	0
1	G	1182	0	1156	30	0
1	H	1203	0	1183	32	0
2	A	47	0	0	2	0
2	B	47	0	0	0	0
2	C	44	0	0	0	0
2	D	33	0	0	0	0
2	E	26	0	0	1	0
2	F	32	0	0	0	0
2	G	35	0	0	1	0
2	H	50	0	0	1	0
All	All	9931	0	9438	277	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 277 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:H:98:VAL:HG23	1:H:101:VAL:HB	1.46	0.97
1:C:127:HIS:HB3	1:C:130:ASN:HD22	1.30	0.95
1:E:98:VAL:HG23	1:E:101:VAL:HB	1.50	0.90
1:G:98:VAL:HG23	1:G:101:VAL:HB	1.53	0.90
1:B:96:LYS:HD3	1:B:97:PRO:HD2	1.56	0.87

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	153/155 (99%)	147 (96%)	4 (3%)	2 (1%)	9	5
1	B	151/155 (97%)	138 (91%)	11 (7%)	2 (1%)	9	5
1	C	145/155 (94%)	136 (94%)	8 (6%)	1 (1%)	18	14
1	D	145/155 (94%)	132 (91%)	11 (8%)	2 (1%)	9	4
1	E	151/155 (97%)	143 (95%)	7 (5%)	1 (1%)	18	14
1	F	148/155 (96%)	136 (92%)	9 (6%)	3 (2%)	6	2
1	G	144/155 (93%)	137 (95%)	6 (4%)	1 (1%)	18	14
1	H	147/155 (95%)	140 (95%)	7 (5%)	0	100	100
All	All	1184/1240 (96%)	1109 (94%)	63 (5%)	12 (1%)	12	8

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	92	THR
1	A	123	GLY
1	B	94	ASP
1	B	95	LYS
1	F	80	GLN

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	125/125 (100%)	120 (96%)	5 (4%)	28	27
1	B	123/125 (98%)	117 (95%)	6 (5%)	22	20
1	C	119/125 (95%)	114 (96%)	5 (4%)	26	25
1	D	122/125 (98%)	118 (97%)	4 (3%)	33	34
1	E	123/125 (98%)	120 (98%)	3 (2%)	43	47
1	F	124/125 (99%)	121 (98%)	3 (2%)	43	47
1	G	121/125 (97%)	117 (97%)	4 (3%)	33	34
1	H	123/125 (98%)	119 (97%)	4 (3%)	33	34

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	980/1000 (98%)	946 (96%)	34 (4%)	32 32

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	96	LYS
1	G	99	LYS
1	H	98	VAL
1	C	66	ILE
1	C	2	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 58 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	105	GLN
1	H	115	GLN
1	E	84	ASN
1	H	105	GLN
1	G	152	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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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.