



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 22, 2023 – 11:19 AM EDT

PDB ID : 2QKK  
Title : Human RNase H catalytic domain mutant D210N in complex with 14-mer RNA/DNA hybrid  
Authors : Nowotny, M.; Gaidamakov, S.A.; Ghirlando, R.; Cerritelli, S.M.; Crouch, R.J.; Yang, W.  
Deposited on : 2007-07-11  
Resolution : 3.20 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](mailto: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 : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35  
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) : 2.35

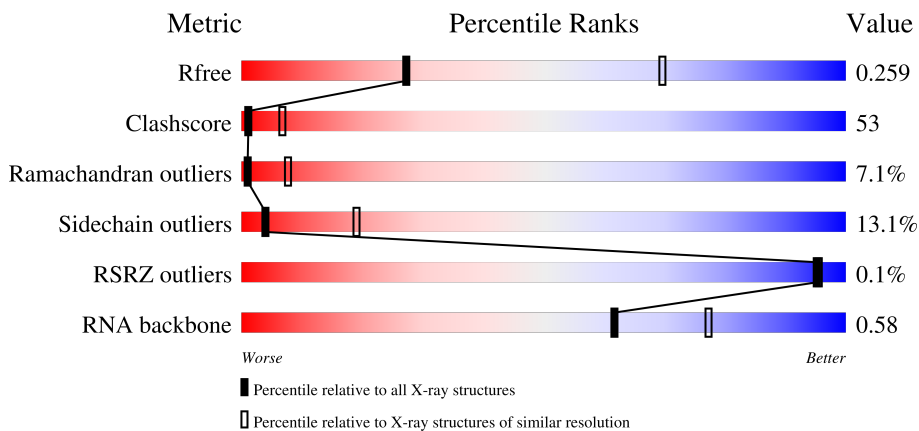
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*



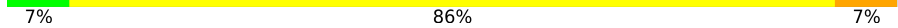
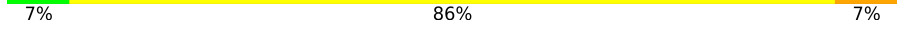
The reported resolution of this entry is 3.20 Å.

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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)
RNA backbone	3102	1010 (3.50-2.90)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	C	14	 64% 36%
1	G	14	 36% 57% 7%
1	K	14	 7% 86% 7%
1	O	14	 7% 86% 7%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	T	14	29% 71%
1	X	14	14% 64% 14% 7%
2	D	14	7% 50% 43%
2	H	14	7% 64% 29%
2	L	14	71% 29%
2	P	14	57% 43%
2	U	14	64% 36%
2	Z	14	43% 36% 21%
3	A	154	% 38% 49% 13%
3	B	154	27% 53% 14% . .
3	E	154	28% 57% 12% .
3	F	154	25% 56% 17% ..
3	I	154	29% 58% 9% .
3	J	154	36% 49% 11% ..
3	M	154	36% 47% 10% . 6%
3	N	154	% 26% 53% 16% . 5%
3	R	154	35% 50% 10% . .
3	S	154	32% 53% 11% . .
3	W	154	27% 56% 12% . .

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 16231 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 RNA chain called 5'-R(\*CP\*GP\*AP\*CP\*AP\*CP\*CP\*UP\*GP\*AP\*UP\*UP\*CP\*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	C	14	Total 289	C 131	N 49	O 96	P 13	0	0	0
1	G	14	Total 289	C 131	N 49	O 96	P 13	0	0	0
1	K	14	Total 289	C 131	N 49	O 96	P 13	0	0	0
1	O	14	Total 289	C 131	N 49	O 96	P 13	0	0	0
1	T	14	Total 289	C 131	N 49	O 96	P 13	0	0	0
1	X	13	Total 269	C 122	N 46	O 89	P 12	0	0	0

- Molecule 2 is a DNA chain called 5'-D(\*GP\*GP\*AP\*AP\*TP\*CP\*AP\*GP\*GP\*TP\*GP\*TP\*CP\*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	D	14	Total 290	C 138	N 57	O 82	P 13	0	0	0
2	H	14	Total 290	C 138	N 57	O 82	P 13	0	0	0
2	L	14	Total 290	C 138	N 57	O 82	P 13	0	0	0
2	P	14	Total 290	C 138	N 57	O 82	P 13	0	0	0
2	U	14	Total 290	C 138	N 57	O 82	P 13	0	0	0
2	Z	11	Total 229	C 109	N 47	O 63	P 10	0	0	0

- Molecule 3 is a protein called Ribonuclease H1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	154	1186	737	222	220	7	0	0	0
3	B	151	1175	731	223	214	7	1	0	0
3	E	149	1147	715	217	209	6	0	0	0
3	F	152	1175	733	219	216	7	3	0	0
3	I	149	1133	706	213	208	6	6	0	0
3	J	151	1166	728	218	213	7	5	0	0
3	M	145	1117	700	208	203	6	3	0	0
3	N	146	1121	701	208	206	6	0	0	0
3	R	148	1144	714	215	208	7	0	0	0
3	S	148	1126	704	210	206	6	7	0	0
3	W	148	1141	716	211	207	7	9	0	0

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	133	GLY	-	expression tag	UNP O60930
A	134	SER	-	expression tag	UNP O60930
A	135	HIS	-	expression tag	UNP O60930
A	210	ASN	ASP	engineered mutation	UNP O60930
B	133	GLY	-	expression tag	UNP O60930
B	134	SER	-	expression tag	UNP O60930
B	135	HIS	-	expression tag	UNP O60930
B	210	ASN	ASP	engineered mutation	UNP O60930
E	133	GLY	-	expression tag	UNP O60930
E	134	SER	-	expression tag	UNP O60930
E	135	HIS	-	expression tag	UNP O60930
E	210	ASN	ASP	engineered mutation	UNP O60930
F	133	GLY	-	expression tag	UNP O60930
F	134	SER	-	expression tag	UNP O60930
F	135	HIS	-	expression tag	UNP O60930
F	210	ASN	ASP	engineered mutation	UNP O60930
I	133	GLY	-	expression tag	UNP O60930
I	134	SER	-	expression tag	UNP O60930

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
I	135	HIS	-	expression tag	UNP O60930
I	210	ASN	ASP	engineered mutation	UNP O60930
J	133	GLY	-	expression tag	UNP O60930
J	134	SER	-	expression tag	UNP O60930
J	135	HIS	-	expression tag	UNP O60930
J	210	ASN	ASP	engineered mutation	UNP O60930
M	133	GLY	-	expression tag	UNP O60930
M	134	SER	-	expression tag	UNP O60930
M	135	HIS	-	expression tag	UNP O60930
M	210	ASN	ASP	engineered mutation	UNP O60930
N	133	GLY	-	expression tag	UNP O60930
N	134	SER	-	expression tag	UNP O60930
N	135	HIS	-	expression tag	UNP O60930
N	210	ASN	ASP	engineered mutation	UNP O60930
R	133	GLY	-	expression tag	UNP O60930
R	134	SER	-	expression tag	UNP O60930
R	135	HIS	-	expression tag	UNP O60930
R	210	ASN	ASP	engineered mutation	UNP O60930
S	133	GLY	-	expression tag	UNP O60930
S	134	SER	-	expression tag	UNP O60930
S	135	HIS	-	expression tag	UNP O60930
S	210	ASN	ASP	engineered mutation	UNP O60930
W	133	GLY	-	expression tag	UNP O60930
W	134	SER	-	expression tag	UNP O60930
W	135	HIS	-	expression tag	UNP O60930
W	210	ASN	ASP	engineered mutation	UNP O60930

- Molecule 4 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula: C<sub>4</sub>H<sub>12</sub>NO<sub>3</sub>).

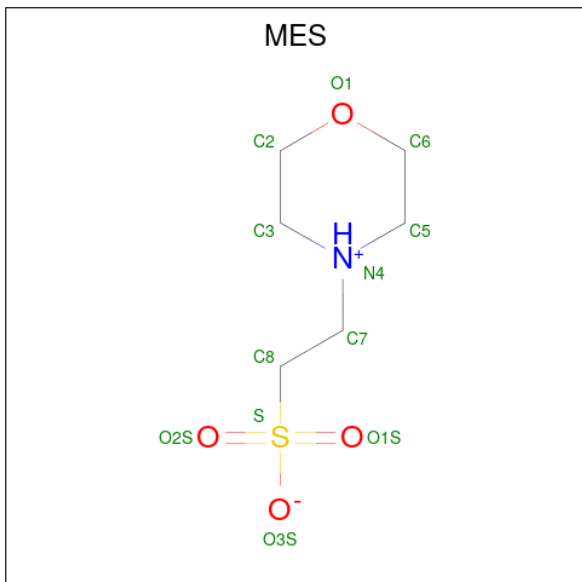


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	D	1	8	4	1	3	0	0

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
5	A	2	2	2	0	0
5	B	1	1	1	0	0
5	E	2	2	2	0	0
5	F	1	1	1	0	0
5	I	1	1	1	0	0
5	J	1	1	1	0	0
5	M	2	2	2	0	0
5	N	1	1	1	0	0
5	R	1	1	1	0	0
5	S	1	1	1	0	0
5	W	1	1	1	0	0

- Molecule 6 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C<sub>6</sub>H<sub>13</sub>NO<sub>4</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
6	A	1	12	6	1	4	1	0	0

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
7	B	1	1	1	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
8	C	2	2	2	0	0
8	D	5	5	5	0	0
8	G	8	8	8	0	0
8	H	9	9	9	0	0
8	K	4	4	4	0	0
8	L	4	4	4	0	0

*Continued on next page...*



*Continued from previous page...*

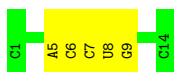
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	O	4	Total O 4 4	0	0
8	P	6	Total O 6 6	0	0
8	T	3	Total O 3 3	0	0
8	U	3	Total O 3 3	0	0
8	X	5	Total O 5 5	0	0
8	Z	3	Total O 3 3	0	0
8	A	16	Total O 16 16	0	0
8	B	15	Total O 15 15	0	0
8	E	16	Total O 16 16	0	0
8	F	13	Total O 13 13	0	0
8	I	3	Total O 3 3	0	0
8	J	12	Total O 12 12	0	0
8	M	6	Total O 6 6	0	0
8	N	7	Total O 7 7	0	0
8	R	7	Total O 7 7	0	0
8	S	11	Total O 11 11	0	0
8	W	10	Total O 10 10	0	0

### 3 Residue-property plots [i](#)

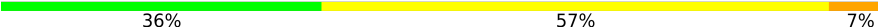
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: 5'-R(\*CP\*GP\*AP\*CP\*AP\*CP\*CP\*UP\*GP\*AP\*UP\*UP\*CP\*C)-3'

Chain C: 

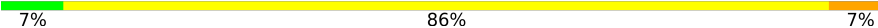


- Molecule 1: 5'-R(\*CP\*GP\*AP\*CP\*AP\*CP\*CP\*UP\*GP\*AP\*UP\*UP\*CP\*C)-3'

Chain G: 



- Molecule 1: 5'-R(\*CP\*GP\*AP\*CP\*AP\*CP\*CP\*UP\*GP\*AP\*UP\*UP\*CP\*C)-3'

Chain K: 




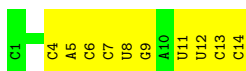
- Molecule 1: 5'-R(\*CP\*GP\*AP\*CP\*AP\*CP\*CP\*UP\*GP\*AP\*UP\*UP\*CP\*C)-3'

Chain O: 



- Molecule 1: 5'-R(\*CP\*GP\*AP\*CP\*AP\*CP\*CP\*UP\*GP\*AP\*UP\*UP\*CP\*C)-3'

Chain T: 

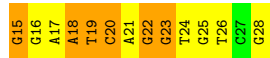


- Molecule 1: 5'-R(\*CP\*GP\*AP\*CP\*AP\*CP\*CP\*UP\*GP\*AP\*UP\*UP\*CP\*C)-3'

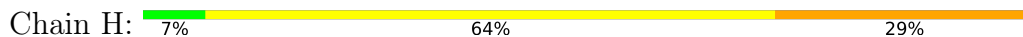
Chain X: 



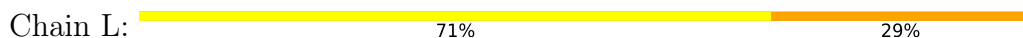
- Molecule 2: 5'-D(\*GP\*GP\*AP\*AP\*TP\*CP\*AP\*GP\*GP\*TP\*GP\*TP\*CP\*G)-3'



- Molecule 2: 5'-D(\*GP\*GP\*AP\*AP\*TP\*CP\*AP\*GP\*GP\*TP\*GP\*TP\*CP\*G)-3'



- Molecule 2: 5'-D(\*GP\*GP\*AP\*AP\*TP\*CP\*AP\*GP\*GP\*TP\*GP\*TP\*CP\*G)-3'



- Molecule 2: 5'-D(\*GP\*GP\*AP\*AP\*TP\*CP\*AP\*GP\*GP\*TP\*GP\*TP\*CP\*G)-3'



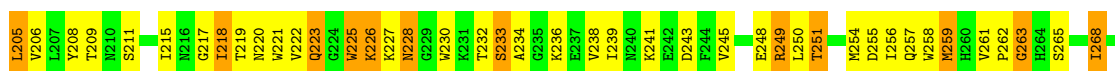
- Molecule 2: 5'-D(\*GP\*GP\*AP\*AP\*TP\*CP\*AP\*GP\*GP\*TP\*GP\*TP\*CP\*G)-3'



- Molecule 2: 5'-D(\*GP\*GP\*AP\*AP\*TP\*CP\*AP\*GP\*GP\*TP\*GP\*TP\*CP\*G)-3'

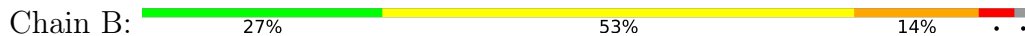


- Molecule 3: Ribonuclease H1

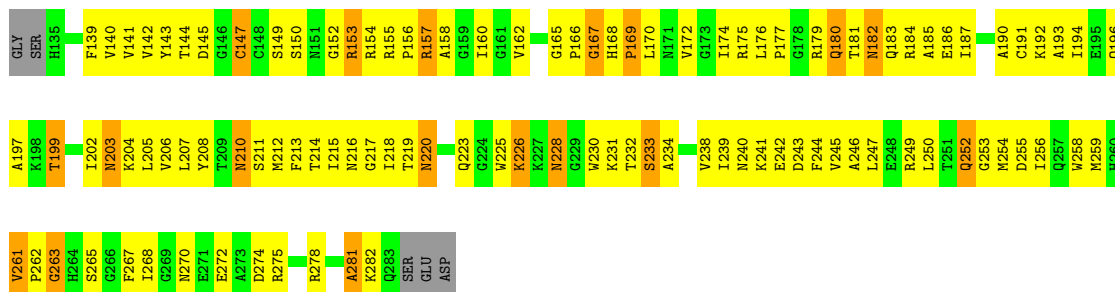
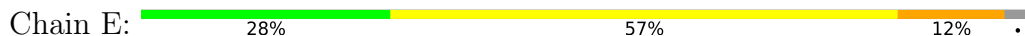




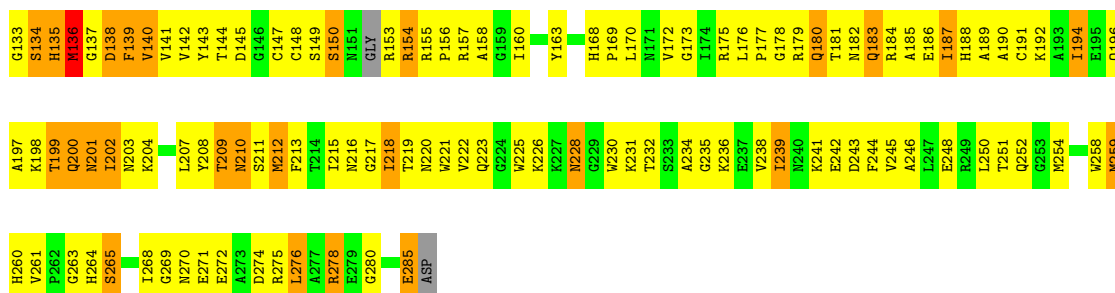
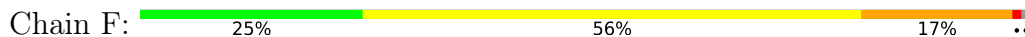
• Molecule 3: Ribonuclease H1



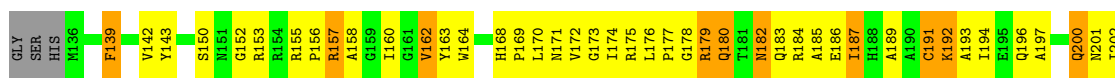
• Molecule 3: Ribonuclease H1

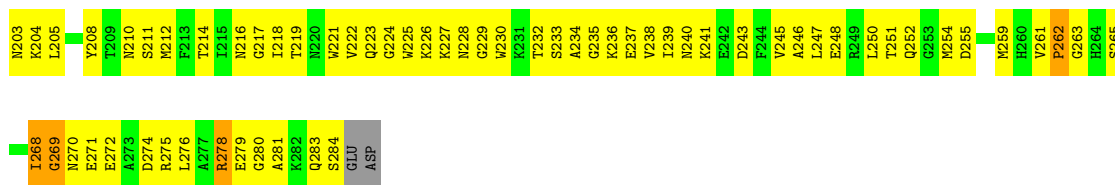


• Molecule 3: Ribonuclease H1

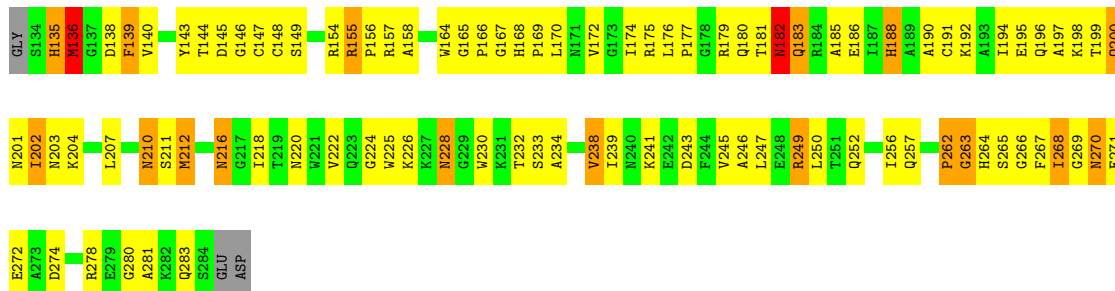


• Molecule 3: Ribonuclease H1

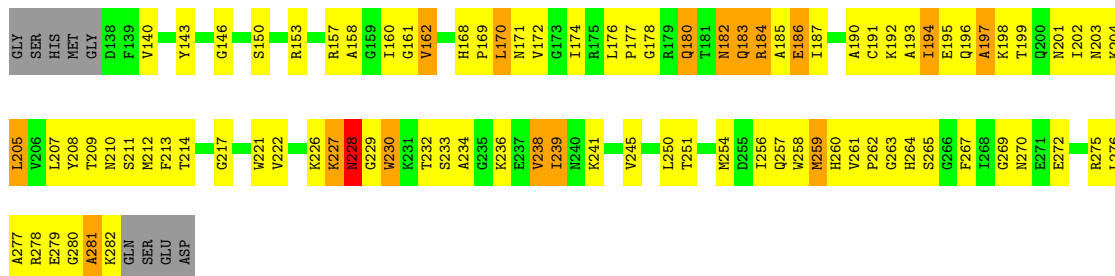
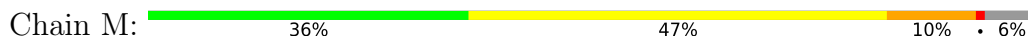




• Molecule 3: Ribonuclease H1



• Molecule 3: Ribonuclease H1



• Molecule 3: Ribonuclease H1



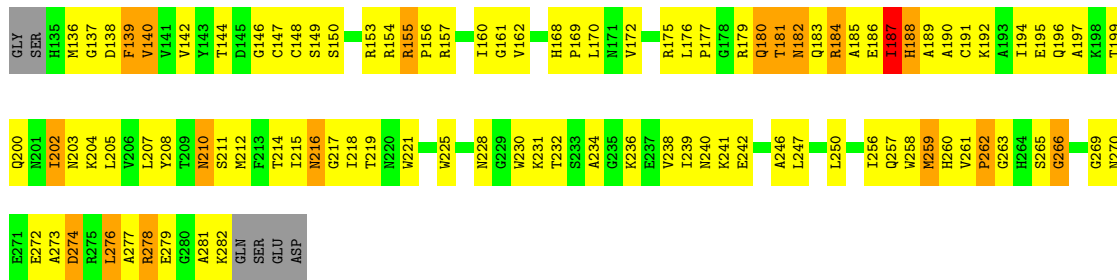
• Molecule 3: Ribonuclease H1





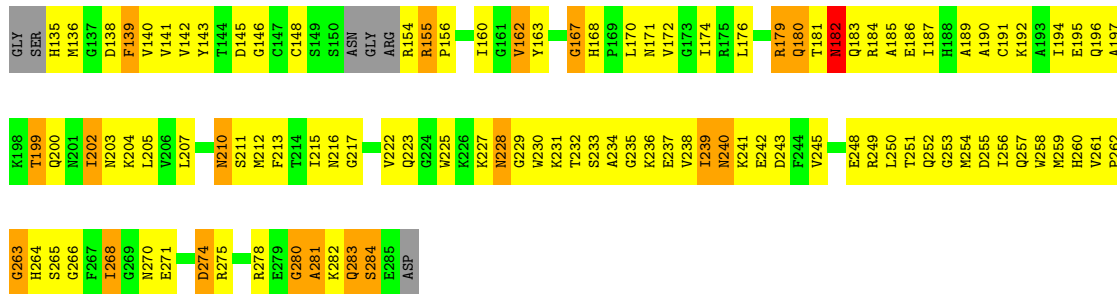
• Molecule 3: Ribonuclease H1

Chain S: 32% 53% 11%



• Molecule 3: Ribonuclease H1

Chain W: 27% 56% 12%



## 4 Data and refinement statistics i

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	151.06Å 176.20Å 125.84Å 90.00° 90.22° 90.00°	Depositor
Resolution (Å)	30.00 – 3.20 29.78 – 3.00	Depositor EDS
% Data completeness (in resolution range)	92.8 (30.00-3.20) 84.0 (29.78-3.00)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.27 (at 3.00Å)	Xtrriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.212 , 0.269 0.208 , 0.259	Depositor DCC
$R_{free}$ test set	1752 reflections (3.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	55.8	Xtrriage
Anisotropy	0.816	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , 21.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	0.177 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	16231	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	62.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.32% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<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.

## 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: MES, CA, CL, TRS

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	C	0.60	0/321	0.80	0/497
1	G	0.61	0/321	0.77	0/497
1	K	0.40	0/321	0.69	0/497
1	O	0.37	0/321	0.66	0/497
1	T	0.40	0/321	0.66	0/497
1	X	0.29	0/299	0.66	0/463
2	D	0.67	0/326	1.08	0/503
2	H	0.61	0/326	0.87	0/503
2	L	0.48	0/326	1.02	0/503
2	P	0.54	0/326	1.04	0/503
2	U	0.42	0/326	1.02	0/503
2	Z	0.42	0/258	1.02	0/398
3	A	0.49	0/1212	0.74	0/1637
3	B	0.44	0/1201	0.91	4/1622 (0.2%)
3	E	0.46	0/1172	0.68	0/1586
3	F	0.46	0/1200	0.71	1/1620 (0.1%)
3	I	0.35	0/1158	0.60	1/1571 (0.1%)
3	J	0.39	0/1192	0.67	1/1611 (0.1%)
3	M	0.41	0/1142	0.67	0/1546
3	N	0.45	0/1145	0.77	3/1550 (0.2%)
3	R	0.42	0/1170	0.68	0/1583
3	S	0.39	0/1151	0.63	0/1560
3	W	0.35	0/1166	0.62	0/1576
All	All	0.44	0/16701	0.75	10/23323 (0.0%)

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
2	D	0	6
2	H	0	4
2	L	0	6
2	P	0	6
2	U	0	5
2	Z	0	6
All	All	0	33

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	155	ARG	C-N-CD	-16.28	84.79	120.60
3	B	156	PRO	CA-N-CD	-6.56	102.32	111.50
3	B	155	ARG	C-N-CA	6.44	149.06	122.00
3	N	262	PRO	N-CA-C	5.74	127.02	112.10
3	I	152	GLY	N-CA-C	-5.61	99.08	113.10

There are no chirality outliers.

5 of 33 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	15	DG	Sidechain
2	D	18	DA	Sidechain
2	D	19	DT	Sidechain
2	D	20	DC	Sidechain
2	D	22	DG	Sidechain

## 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	C	289	0	153	3	0
1	G	289	0	153	13	0
1	K	289	0	153	16	0
1	O	289	0	153	14	0
1	T	289	0	153	13	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	X	269	0	142	9	0
2	D	290	0	159	27	0
2	H	290	0	159	33	1
2	L	290	0	159	47	0
2	P	290	0	159	35	0
2	U	290	0	159	41	0
2	Z	229	0	125	36	0
3	A	1186	0	1137	115	0
3	B	1175	0	1144	120	0
3	E	1147	0	1105	124	0
3	F	1175	0	1135	156	0
3	I	1133	0	1074	128	0
3	J	1166	0	1130	108	0
3	M	1117	0	1083	106	0
3	N	1121	0	1075	131	0
3	R	1144	0	1104	112	0
3	S	1126	0	1077	139	0
3	W	1141	0	1101	141	0
4	D	8	0	12	0	0
5	A	2	0	0	0	0
5	B	1	0	0	0	0
5	E	2	0	0	0	0
5	F	1	0	0	0	0
5	I	1	0	0	0	0
5	J	1	0	0	0	0
5	M	2	0	0	0	0
5	N	1	0	0	0	0
5	R	1	0	0	0	0
5	S	1	0	0	0	0
5	W	1	0	0	0	0
6	A	12	0	13	1	0
7	B	1	0	0	0	0
8	A	16	0	0	1	0
8	B	15	0	0	4	0
8	C	2	0	0	0	0
8	D	5	0	0	1	0
8	E	16	0	0	4	0
8	F	13	0	0	1	0
8	G	8	0	0	0	0
8	H	9	0	0	0	0
8	I	3	0	0	1	0
8	J	12	0	0	4	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	K	4	0	0	1	0
8	L	4	0	0	2	0
8	M	6	0	0	1	0
8	N	7	0	0	0	0
8	O	4	0	0	0	0
8	P	6	0	0	3	0
8	R	7	0	0	2	0
8	S	11	0	0	2	0
8	T	3	0	0	0	0
8	U	3	0	0	0	0
8	W	10	0	0	2	0
8	X	5	0	0	1	0
8	Z	3	0	0	0	0
All	All	16231	0	14017	1574	1

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 53.

The worst 5 of 1574 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)
2:L:15:DG:H4'	2:L:16:DG:H5'	1.16	1.15
2:L:15:DG:H4'	2:L:16:DG:C5'	1.77	1.14
3:W:199:THR:HG23	3:W:200:GLN:HE21	1.12	1.14
3:A:183:GLN:HE21	3:A:183:GLN:HA	1.10	1.11
3:R:187:ILE:HD11	3:R:243:ASP:HB3	1.37	1.07

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:15:DG:C5'	2:H:15:DG:C5'[2_455]	1.91	0.29

## 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	
3	A	152/154 (99%)	112 (74%)	34 (22%)	6 (4%)	3	22
3	B	149/154 (97%)	106 (71%)	28 (19%)	15 (10%)	0	3
3	E	147/154 (96%)	111 (76%)	29 (20%)	7 (5%)	2	17
3	F	148/154 (96%)	112 (76%)	23 (16%)	13 (9%)	1	4
3	I	147/154 (96%)	104 (71%)	30 (20%)	13 (9%)	1	4
3	J	149/154 (97%)	111 (74%)	29 (20%)	9 (6%)	1	12
3	M	143/154 (93%)	103 (72%)	34 (24%)	6 (4%)	3	20
3	N	142/154 (92%)	98 (69%)	31 (22%)	13 (9%)	1	3
3	R	146/154 (95%)	107 (73%)	30 (20%)	9 (6%)	1	11
3	S	146/154 (95%)	98 (67%)	38 (26%)	10 (7%)	1	9
3	W	144/154 (94%)	101 (70%)	29 (20%)	14 (10%)	0	3
All	All	1613/1694 (95%)	1163 (72%)	335 (21%)	115 (7%)	1	8

5 of 115 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	228	ASN
3	B	135	HIS
3	B	150	SER
3	B	155	ARG
3	B	156	PRO

### 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	
3	A	120/124 (97%)	102 (85%)	18 (15%)	3	14
3	B	121/124 (98%)	99 (82%)	22 (18%)	1	8

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	E	115/124 (93%)	103 (90%)	12 (10%)	7	28
3	F	120/124 (97%)	101 (84%)	19 (16%)	2	12
3	I	112/124 (90%)	104 (93%)	8 (7%)	14	47
3	J	119/124 (96%)	103 (87%)	16 (13%)	4	18
3	M	113/124 (91%)	99 (88%)	14 (12%)	4	21
3	N	113/124 (91%)	97 (86%)	16 (14%)	3	15
3	R	116/124 (94%)	100 (86%)	16 (14%)	3	16
3	S	112/124 (90%)	97 (87%)	15 (13%)	4	18
3	W	115/124 (93%)	104 (90%)	11 (10%)	8	32
All	All	1276/1364 (94%)	1109 (87%)	167 (13%)	4	19

5 of 167 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	N	212	MET
3	S	180	GLN
3	N	245	VAL
3	R	183	GLN
3	S	210	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 69 such sidechains are listed below:

Mol	Chain	Res	Type
3	R	260	HIS
3	S	183	GLN
3	W	200	GLN
3	I	183	GLN
3	I	182	ASN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	C	13/14 (92%)	0	0
1	G	13/14 (92%)	1 (7%)	0
1	K	13/14 (92%)	2 (15%)	0
1	O	13/14 (92%)	1 (7%)	0
1	T	13/14 (92%)	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	X	12/14 (85%)	2 (16%)	0
All	All	77/84 (91%)	6 (7%)	0

5 of 6 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	G	7	C
1	K	7	C
1	K	14	C
1	O	7	C
1	X	7	C

There are no RNA pucker outliers to report.

## 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](#)

Of 17 ligands modelled in this entry, 15 are monoatomic - leaving 2 for Mogul analysis.

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	$\# Z  > 2$
4	TRS	D	2003	-	7,7,7	0.75	0	9,9,9	0.41	0
6	MES	A	2002	-	12,12,12	1.07	1 (8%)	14,16,16	1.18	2 (14%)

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
4	TRS	D	2003	-	-	0/9/9/9	-
6	MES	A	2002	-	-	0/6/14/14	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	2002	MES	C8-S	2.59	1.81	1.77

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	2002	MES	O1S-S-C8	2.27	109.65	106.92
6	A	2002	MES	O3S-S-C8	2.20	109.33	105.77

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	2002	MES	1	0

## 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 [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	C	14/14 (100%)	-0.49	0 100 100	17, 24, 34, 40	0
1	G	14/14 (100%)	-0.51	0 100 100	24, 31, 38, 41	0
1	K	14/14 (100%)	-0.64	0 100 100	56, 62, 72, 75	0
1	O	14/14 (100%)	-0.52	0 100 100	44, 52, 74, 76	0
1	T	14/14 (100%)	-0.48	0 100 100	44, 57, 76, 85	0
1	X	13/14 (92%)	-0.56	0 100 100	68, 89, 144, 152	0
2	D	14/14 (100%)	-0.42	0 100 100	22, 32, 43, 50	0
2	H	14/14 (100%)	-0.57	0 100 100	28, 36, 43, 48	0
2	L	14/14 (100%)	-0.51	0 100 100	55, 74, 92, 123	0
2	P	14/14 (100%)	-0.45	0 100 100	42, 56, 81, 103	0
2	U	14/14 (100%)	-0.43	0 100 100	54, 69, 83, 91	0
2	Z	11/14 (78%)	-0.45	0 100 100	82, 88, 117, 131	0
3	A	154/154 (100%)	-0.48	1 (0%) 89 83	18, 36, 73, 96	0
3	B	151/154 (98%)	-0.54	0 100 100	30, 51, 82, 108	1 (0%)
3	E	149/154 (96%)	-0.50	0 100 100	32, 51, 77, 93	0
3	F	152/154 (98%)	-0.47	0 100 100	30, 51, 83, 101	1 (0%)
3	I	149/154 (96%)	-0.30	0 100 100	56, 93, 117, 129	1 (0%)
3	J	151/154 (98%)	-0.39	0 100 100	51, 65, 87, 96	2 (1%)
3	M	145/154 (94%)	-0.43	0 100 100	35, 59, 80, 102	1 (0%)
3	N	146/154 (94%)	-0.32	1 (0%) 87 81	50, 68, 87, 96	1 (0%)
3	R	148/154 (96%)	-0.46	0 100 100	34, 57, 80, 87	1 (0%)
3	S	148/154 (96%)	-0.29	0 100 100	52, 70, 98, 109	3 (2%)
3	W	148/154 (96%)	-0.47	0 100 100	57, 75, 96, 102	3 (2%)
All	All	1805/1862 (96%)	-0.43	2 (0%) 95 95	17, 63, 98, 152	14 (0%)



All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	286	ASP	4.5
3	N	151	ASN	2.3

## 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 monosaccharides 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<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	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
6	MES	A	2002	12/12	0.74	0.37	136,136,138,138	0
4	TRS	D	2003	8/8	0.82	0.14	94,96,97,97	0
7	CL	B	2001	1/1	0.88	0.09	49,49,49,49	0
5	CA	M	1014	1/1	0.92	0.08	106,106,106,106	0
5	CA	E	1003	1/1	0.95	0.14	44,44,44,44	0
5	CA	J	1008	1/1	0.96	0.15	53,53,53,53	0
5	CA	F	1013	1/1	0.96	0.15	32,32,32,32	0
5	CA	R	1011	1/1	0.97	0.14	64,64,64,64	0
5	CA	S	1007	1/1	0.97	0.15	77,77,77,77	0
5	CA	W	1012	1/1	0.97	0.10	81,81,81,81	0
5	CA	M	1010	1/1	0.97	0.08	38,38,38,38	0
5	CA	E	1004	1/1	0.97	0.05	68,68,68,68	0
5	CA	N	1005	1/1	0.98	0.16	56,56,56,56	0
5	CA	I	1006	1/1	0.98	0.09	78,78,78,78	0
5	CA	B	1009	1/1	0.98	0.17	38,38,38,38	0
5	CA	A	1001	1/1	0.99	0.08	44,44,44,44	0
5	CA	A	1002	1/1	0.99	0.05	23,23,23,23	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.