



Full wwPDB X-ray Structure Validation Report

Feb 27, 2014 – 02:38 PM GMT

PDB ID : 4NYN
Title : Crystal structure of RNase H1 from halophilic archaeon Halobacterium salinarum NRC-1
Authors : You, D.J.; Angkawidjaja, C.; Koga, Y.; Kanaya, S.
Deposited on : 2013-12-11
Resolution : 1.41 Å(reported)

This is a full wwPDB validation 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/ValidationPDFNotes.html>

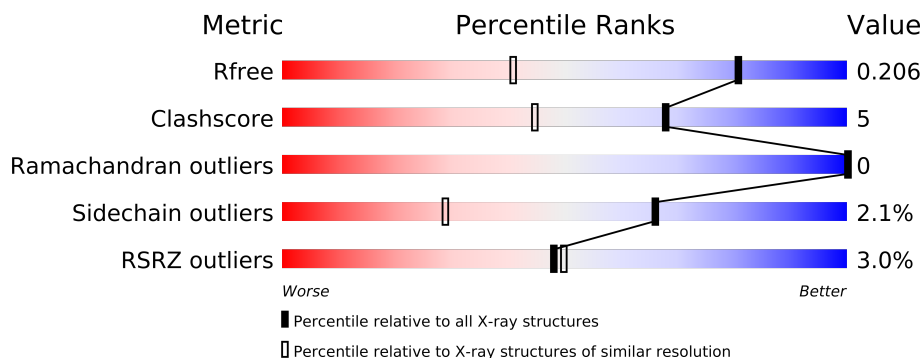
The following versions of software and data (see [references](#)) were used in the production of this report:

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 1.41 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	1110 (1.44-1.40)
Clashscore	79885	1263 (1.44-1.40)
Ramachandran outliers	78287	1226 (1.44-1.40)
Sidechain outliers	78261	1225 (1.44-1.40)
RSRZ outliers	66119	1110 (1.44-1.40)

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. 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	199	
1	B	199	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2300 atoms, of which 0 are hydrogen 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 Ribonuclease HI.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	133	Total	C	N	O	0	0	0
			999	611	184	204			
1	B	132	Total	C	N	O	0	0	0
			995	609	183	203			

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	3	Total	Mn	0	0
			3	3		
2	A	2	Total	Mn	0	0
			2	2		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	148	Total	O	0	0
			148	148		
3	B	153	Total	O	0	0
			153	153		

3 Residue-property plots

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 ($RSRZ > 2$). 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: Ribonuclease HI

Chain A: 

MET	PRO	VAL	VAL	GLN	ASP	CYS	ASP	ILE	GLN	THR	ALA	ARG	ALA	ALA	LEU	ALA	ASP	GLY	ALA	ALA	SER	PHE	SER	ASP	GLY	ASN	SER	GLU	HIS	GLU	LEU	TRP	HIS	ALA	ASP	GLY	ASP	ALA	ALA	VAL	ALA	TYR	ALA	ASP	LYS	LEU	VAL	VAL	GLN	GLY	SER	PRO	THR	ASP	ILE	THR	ALA
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

VAL	VAL	GLN	PRO	ASP	ARG	G67	G68	G69	R79	V92	S93	G94	D95	I98	Y116	G130	R166	T170	H179	L191	L196	A199
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------

• Molecule 1: Ribonuclease HI

Chain B: 

MET	PRO	VAL	VAL	GLN	ASP	CYS	ASP	ILE	GLN	THR	ALA	ARG	ALA	ALA	LEU	ALA	ASP	GLY	ALA	ALA	SER	PHE	SER	ASP	GLY	ASN	SER	GLU	HIS	GLU	LEU	TRP	HIS	ALA	ASP	GLY	ASP	ALA	ALA	VAL	ALA	TYR	ALA	ASP	LYS	LEU	VAL	VAL	GLN	GLY	SER	PRO	THR	ASP	ILE	THR	ALA
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

VAL	VAL	GLN	PRO	ASP	ARG	GLY	G68	R69	V90	S93	G94	D95	G96	D133	L142	R160	H179	R187	A199
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------

4 Data and refinement statistics

Property	Value	Source
Space group	I 41	Depositor
Cell constants a, b, c, α , β , γ	115.80Å 115.80Å 38.28Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.94 – 1.41 40.94 – 1.41	Depositor EDS
% Data completeness (in resolution range)	99.6 (40.94-1.41) 99.7 (40.94-1.41)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.01 (at 1.41Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.175 , 0.206 0.175 , 0.206	Depositor DCC
R_{free} test set	2472 reflections (5.33%)	DCC
Wilson B-factor (Å ²)	14.2	Xtriage
Anisotropy	0.512	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 36.7	EDS
Estimated twinning fraction	0.019 for -k,-h,-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 48988 reflections	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	2300	wwPDB-VP
Average B, all atoms (Å ²)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.72% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MN

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	0.48	0/1016	0.56	0/1379
1	B	0.42	0/1012	0.49	0/1374
All	All	0.45	0/2028	0.52	0/2753

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	B	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	95	ASP	Peptide

5.2 Close contacts

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	999	0	936	12	0
1	B	995	0	933	11	0
2	A	2	0	0	0	0
2	B	3	0	0	0	0
3	A	148	0	0	2	0
3	B	153	0	0	3	0
All	All	2300	0	1869	19	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 5.

All (19) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:93:SER:HB3	3:A:514:HOH:O	1.70	0.91
1:A:92:VAL:HG12	1:A:93:SER:O	1.83	0.79
1:A:95:ASP:OD1	1:A:95:ASP:N	2.19	0.74
1:A:166:ARG:O	1:A:170:THR:HG23	1.95	0.67
1:B:179:HIS:HE1	3:B:376:HOH:O	1.82	0.62
1:A:191:LEU:HD21	1:B:142:LEU:CD1	2.30	0.61
1:B:160:ARG:HD3	3:B:411:HOH:O	2.03	0.58
1:A:67:GLY:N	1:A:130:GLY:O	2.42	0.52
1:A:191:LEU:HD21	1:B:142:LEU:HD12	1.91	0.51
1:B:69:ARG:HD2	1:B:133:ASP:CG	2.31	0.50
1:A:79:ARG:HG2	1:A:196:LEU:HD21	1.94	0.49
1:A:179:HIS:HD2	3:A:538:HOH:O	1.97	0.47
1:B:90:VAL:HG21	1:B:187:ARG:HG2	1.98	0.45
1:A:191:LEU:CD2	1:B:142:LEU:HD12	2.46	0.45
1:B:93:SER:C	1:B:95:ASP:H	2.21	0.44
1:A:191:LEU:CD2	1:B:142:LEU:CD1	2.96	0.44
1:B:93:SER:C	1:B:95:ASP:N	2.74	0.41
1:A:98:ILE:HD12	1:A:98:ILE:N	2.35	0.41
1:B:179:HIS:CE1	3:B:376:HOH:O	2.67	0.40

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	131/199 (66%)	129 (98%)	2 (2%)	0	100	100
1	B	130/199 (65%)	126 (97%)	4 (3%)	0	100	100
All	All	261/398 (66%)	255 (98%)	6 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	96/145 (66%)	93 (97%)	3 (3%)	52	15
1	B	96/145 (66%)	95 (99%)	1 (1%)	85	63
All	All	192/290 (66%)	188 (98%)	4 (2%)	66	29

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	69	ARG
1	A	95	ASP
1	A	116	TYR
1	B	95	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	179	HIS

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 ⓘ

Of 5 ligands modelled in this entry, 5 are monoatomic - leaving 0 for Mogul analysis.

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.

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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	133/199 (66%)	-0.08	3 (2%) 57 60	10, 15, 26, 46	0
1	B	132/199 (66%)	-0.15	5 (3%) 38 39	11, 16, 30, 52	0
All	All	265/398 (66%)	-0.11	8 (3%) 48 50	10, 15, 28, 52	0

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	94	GLY	3.9
1	A	94	GLY	3.7
1	A	93	SER	3.6
1	B	96	GLY	3.0
1	A	95	ASP	2.4
1	B	68	GLY	2.1
1	B	69	ARG	2.1
1	B	199	ALA	2.0

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

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

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

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	RSR	LLDF	B-factors(Å ²)	Q<0.9
2	MN	A	302	1/1	0.06	-0.48	11,11,11,11	0
2	MN	A	301	1/1	0.05	-1.43	11,11,11,11	0
2	MN	B	202	1/1	0.05	-3.18	12,12,12,12	0
2	MN	B	201	1/1	0.04	-5.39	11,11,11,11	0
2	MN	B	203	1/1	0.03	-6.12	15,15,15,15	0

6.5 Other polymers ⓘ

There are no such residues in this entry.