



# Full wwPDB X-ray Structure Validation Report ⓘ

May 13, 2020 – 12:45 am BST

PDB ID : 3FC7  
Title : The crystal structure of a domain of HTR-like protein from *Haloarcula marismortui* ATCC 43049  
Authors : Tan, K.; Hatzos, C.; Freeman, L.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)  
Deposited on : 2008-11-21  
Resolution : 2.65 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 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 : 2.11  
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) : 2.11

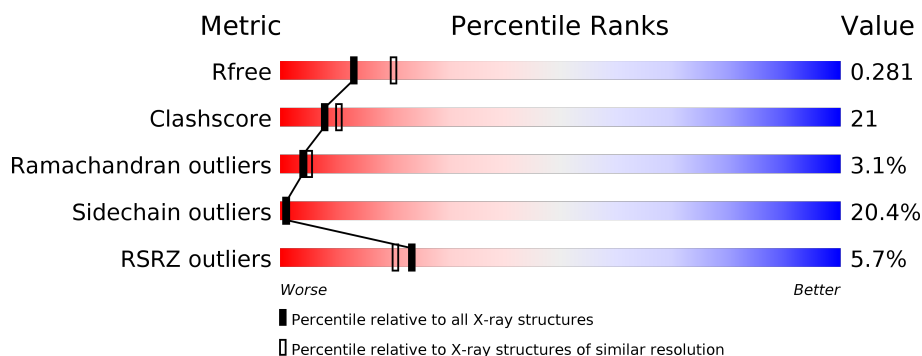
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.65 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	1332 (2.68-2.64)
Clashscore	141614	1374 (2.68-2.64)
Ramachandran outliers	138981	1349 (2.68-2.64)
Sidechain outliers	138945	1349 (2.68-2.64)
RSRZ outliers	127900	1318 (2.68-2.64)

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 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	A	125	<div> <div>3%</div> <div> <div></div> <div>48%</div> <div>24%</div> <div>7%</div> <div>20%</div> </div> </div>
1	B	125	<div> <div>6%</div> <div> <div></div> <div>38%</div> <div>34%</div> <div>7%</div> <div>22%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 1481 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 HTR-like protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	100	Total	C	N	O	Se	0	0	0
			744	446	138	158	2			
1	B	98	Total	C	N	O	Se	0	0	0
			730	439	136	153	2			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	117	SER	-	expression tag	UNP Q5V4P0
A	118	ASN	-	expression tag	UNP Q5V4P0
A	119	ALA	-	expression tag	UNP Q5V4P0
B	117	SER	-	expression tag	UNP Q5V4P0
B	118	ASN	-	expression tag	UNP Q5V4P0
B	119	ALA	-	expression tag	UNP Q5V4P0

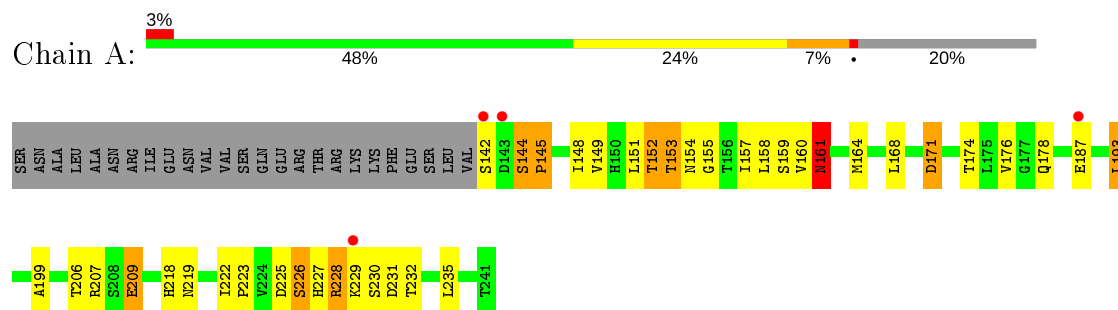
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	O	0	0
			2	2		
2	B	5	Total	O	0	0
			5	5		

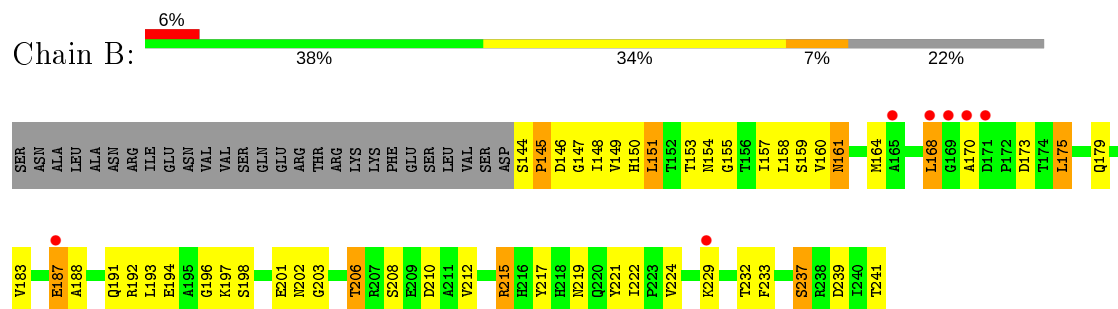
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: HTR-like protein



- Molecule 1: HTR-like protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	55.18 Å 55.18 Å 249.92 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	34.54 – 2.65 31.40 – 2.65	Depositor EDS
% Data completeness (in resolution range)	99.6 (34.54-2.65) 99.6 (31.40-2.65)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.97 (at 2.64 Å)	Xtriage
Refinement program	REFMAC 5.5.0054	Depositor
R, $R_{free}$	0.202 , 0.274 0.207 , 0.281	Depositor DCC
$R_{free}$ test set	340 reflections (4.72%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	67.5	Xtriage
Anisotropy	0.021	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 85.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	1481	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.44% 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](#)

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	0.84	0/752	0.95	1/1016 (0.1%)
1	B	0.78	0/738	0.93	3/997 (0.3%)
All	All	0.81	0/1490	0.94	4/2013 (0.2%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	192	ARG	NE-CZ-NH2	6.47	123.53	120.30
1	B	161	ASN	N-CA-CB	-5.71	100.31	110.60
1	A	161	ASN	N-CA-CB	-5.33	101.00	110.60
1	B	192	ARG	NE-CZ-NH1	-5.22	117.69	120.30

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	744	0	710	30	0
1	B	730	0	701	35	0
2	A	2	0	0	0	0
2	B	5	0	0	0	0
All	All	1481	0	1411	62	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 21.

All (62) 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:A:225:ASP:HB2	1:A:229:LYS:HB3	1.33	1.05
1:A:225:ASP:HB2	1:A:229:LYS:CB	2.03	0.89
1:B:170:ALA:HB3	1:B:175:LEU:HD13	1.55	0.87
1:B:197:LYS:O	1:B:201:GLU:HG2	1.82	0.78
1:A:152:THR:HG22	1:A:154:ASN:H	1.53	0.73
1:B:153:THR:HB	1:B:197:LYS:HG2	1.71	0.72
1:A:152:THR:CG2	1:A:154:ASN:H	2.05	0.70
1:B:160:VAL:HG23	1:B:164:MSE:HB3	1.74	0.69
1:B:170:ALA:HB3	1:B:175:LEU:CD1	2.23	0.68
1:A:152:THR:HG22	1:A:154:ASN:N	2.09	0.67
1:B:160:VAL:CG2	1:B:164:MSE:HB3	2.25	0.67
1:A:222:ILE:HD13	1:B:222:ILE:HD13	1.79	0.65
1:B:215:ARG:NH2	1:B:241:THR:HB	2.11	0.65
1:A:161:ASN:HB3	1:A:164:MSE:H	1.60	0.65
1:B:150:HIS:NE2	1:B:232:THR:OG1	2.29	0.64
1:B:144:SER:HB3	1:B:145:PRO:HD2	1.80	0.63
1:B:219:ASN:ND2	1:B:237:SER:OG	2.32	0.62
1:B:148:ILE:H	1:B:161:ASN:HD22	1.49	0.61
1:A:155:GLY:HA2	1:A:193:LEU:HD11	1.84	0.59
1:A:148:ILE:O	1:A:161:ASN:HB2	2.08	0.54
1:B:194:GLU:OE1	1:B:194:GLU:HA	2.07	0.54
1:A:157:ILE:O	1:A:176:VAL:O	2.25	0.53
1:B:144:SER:HB3	1:B:145:PRO:CD	2.38	0.53
1:A:225:ASP:HB3	1:A:228:ARG:HD2	1.91	0.52
1:B:193:LEU:HG	1:B:197:LYS:HE2	1.90	0.52
1:B:148:ILE:N	1:B:161:ASN:HD22	2.06	0.52
1:B:147:GLY:HA2	1:B:161:ASN:HD22	1.73	0.52
1:A:219:ASN:HB3	1:A:235:LEU:CD1	2.42	0.50
1:A:207:ARG:HA	1:A:219:ASN:O	2.12	0.50
1:A:222:ILE:HD13	1:B:222:ILE:CD1	2.41	0.50
1:B:215:ARG:HH22	1:B:241:THR:HB	1.77	0.50
1:A:160:VAL:CG2	1:A:164:MSE:HB3	2.42	0.49
1:B:196:GLY:HA3	1:B:233:PHE:CZ	2.48	0.49
1:A:209:GLU:OE2	1:A:218:HIS:ND1	2.46	0.48
1:B:146:ASP:O	1:B:161:ASN:ND2	2.46	0.48
1:B:154:ASN:O	1:B:179:GLN:NE2	2.47	0.48
1:A:227:HIS:H	1:A:227:HIS:CD2	2.29	0.48
1:B:224:VAL:HB	1:B:232:THR:HG23	1.97	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:210:ASP:OD2	1:B:217:TYR:HB2	2.14	0.47
1:A:153:THR:HG22	1:A:231:ASP:O	2.14	0.47
1:B:206:THR:HG22	1:B:221:TYR:HD2	1.80	0.47
1:B:157:ILE:HD11	1:B:183:VAL:HG11	1.97	0.46
1:A:207:ARG:NH1	1:A:218:HIS:NE2	2.62	0.46
1:A:222:ILE:HA	1:A:223:PRO:HD2	1.76	0.46
1:B:144:SER:C	1:B:146:ASP:H	2.20	0.46
1:A:144:SER:C	1:A:145:PRO:O	2.55	0.45
1:A:149:VAL:HG22	1:A:151:LEU:HG	1.99	0.45
1:A:229:LYS:O	1:A:231:ASP:N	2.51	0.44
1:B:187:GLU:HG3	1:B:188:ALA:H	1.82	0.43
1:A:199:ALA:HB1	1:A:223:PRO:HD3	2.01	0.43
1:B:160:VAL:HG23	1:B:164:MSE:CB	2.44	0.42
1:A:225:ASP:HA	1:B:159:SER:OG	2.20	0.42
1:A:160:VAL:HG22	1:A:164:MSE:HB3	2.01	0.42
1:B:151:LEU:HD13	1:B:155:GLY:HA2	2.01	0.42
1:B:164:MSE:HE3	1:B:168:LEU:HD21	2.01	0.42
1:A:229:LYS:HG2	1:A:232:THR:OG1	2.20	0.41
1:B:215:ARG:HD3	1:B:239:ASP:OD2	2.21	0.41
1:B:148:ILE:HB	1:B:161:ASN:HB2	2.02	0.41
1:B:147:GLY:HA2	1:B:161:ASN:ND2	2.35	0.40
1:A:209:GLU:CD	1:A:218:HIS:HD1	2.24	0.40
1:A:171:ASP:C	1:A:171:ASP:OD1	2.59	0.40
1:A:225:ASP:HB2	1:A:229:LYS:HB2	1.99	0.40

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	98/125 (78%)	91 (93%)	3 (3%)	4 (4%)	<b>3</b> <b>3</b>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	96/125 (77%)	88 (92%)	6 (6%)	2 (2%)	7	10
All	All	194/250 (78%)	179 (92%)	9 (5%)	6 (3%)	4	5

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	226	SER
1	B	203	GLY
1	A	230	SER
1	A	161	ASN
1	A	145	PRO
1	B	145	PRO

### 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	82/103 (80%)	65 (79%)	17 (21%)	1	1
1	B	80/103 (78%)	64 (80%)	16 (20%)	1	1
All	All	162/206 (79%)	129 (80%)	33 (20%)	1	1

All (33) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	142	SER
1	A	144	SER
1	A	152	THR
1	A	153	THR
1	A	158	LEU
1	A	159	SER
1	A	161	ASN
1	A	168	LEU
1	A	171	ASP
1	A	174	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	178	GLN
1	A	187	GLU
1	A	193	LEU
1	A	206	THR
1	A	209	GLU
1	A	226	SER
1	A	228	ARG
1	B	149	VAL
1	B	151	LEU
1	B	158	LEU
1	B	168	LEU
1	B	173	ASP
1	B	175	LEU
1	B	187	GLU
1	B	191	GLN
1	B	198	SER
1	B	202	ASN
1	B	206	THR
1	B	208	SER
1	B	212	VAL
1	B	215	ARG
1	B	229	LYS
1	B	237	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	219	ASN
1	B	161	ASN
1	B	219	ASN
1	B	220	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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

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 [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	98/125 (78%)	0.14	4 (4%) 37 33	30, 39, 58, 77	0
1	B	96/125 (76%)	0.21	7 (7%) 15 12	29, 39, 50, 55	0
All	All	194/250 (77%)	0.17	11 (5%) 23 20	29, 39, 52, 77	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	170	ALA	3.9
1	B	169	GLY	3.6
1	B	168	LEU	3.3
1	A	142	SER	3.1
1	B	171	ASP	3.0
1	B	187	GLU	3.0
1	A	187	GLU	2.8
1	B	229	LYS	2.7
1	A	229	LYS	2.5
1	A	143	ASP	2.3
1	B	165	ALA	2.2

### 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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.