



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 31, 2023 – 11:56 PM EDT

PDB ID : 3IBX  
Title : Crystal structure of F47Y variant of TenA (HP1287) from Helicobacter pylori  
Authors : Barison, N.; Cendron, L.; Trento, A.; Angelini, A.; Zanotti, G.  
Deposited on : 2009-07-17  
Resolution : 2.40 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 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  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

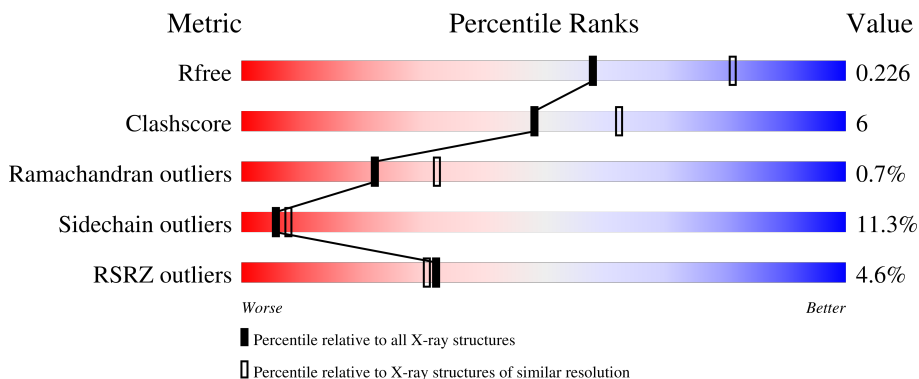
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 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	A	221	
1	D	221	

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 3617 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 Putative thiaminase II.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	221	1798	1160	290	338	10	0	0	0
1	D	216	1755	1133	284	328	10	0	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	ASP	-	expression tag	UNP A8KRL3
A	-2	PRO	-	expression tag	UNP A8KRL3
A	-1	PHE	-	expression tag	UNP A8KRL3
A	0	THR	-	expression tag	UNP A8KRL3
A	47	TYR	PHE	engineered mutation	UNP A8KRL3
D	-3	ASP	-	expression tag	UNP A8KRL3
D	-2	PRO	-	expression tag	UNP A8KRL3
D	-1	PHE	-	expression tag	UNP A8KRL3
D	0	THR	-	expression tag	UNP A8KRL3
D	47	TYR	PHE	engineered mutation	UNP A8KRL3

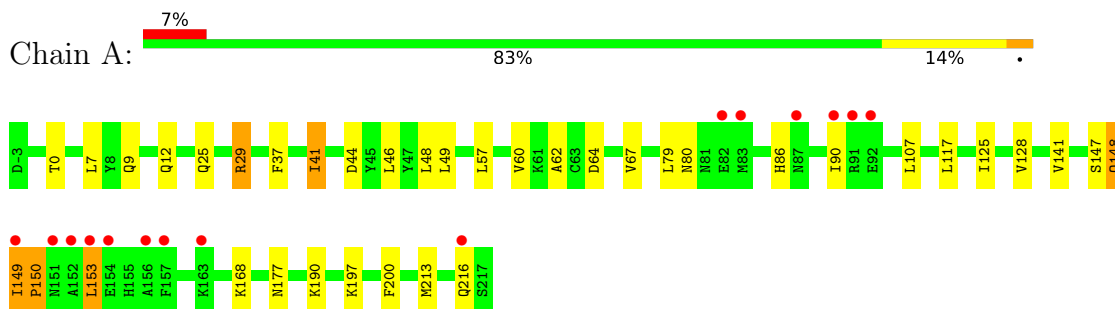
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	35	Total 35	O 35	0	0
2	D	29	Total 29	O 29	0	0

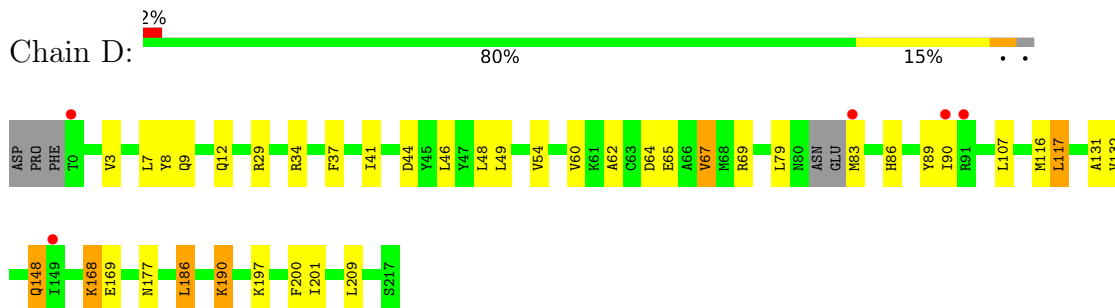
### 3 Residue-property plots [i](#)

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: Putative thiaminase II



- Molecule 1: Putative thiaminase II



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 41 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	148.73Å 148.73Å 233.57Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	78.00 – 2.40 62.73 – 2.40	Depositor EDS
% Data completeness (in resolution range)	97.7 (78.00-2.40) 97.7 (62.73-2.40)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.05 (at 2.40Å)	Xtrriage
Refinement program	REFMAC 5.5.0072	Depositor
R, $R_{free}$	0.218 , 0.230 0.213 , 0.226	Depositor DCC
$R_{free}$ test set	2541 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	39.7	Xtrriage
Anisotropy	0.032	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 45.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	3617	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.62% 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	0.51	1/1842 (0.1%)	0.58	0/2494
1	D	0.50	0/1796	0.63	1/2429 (0.0%)
All	All	0.51	1/3638 (0.0%)	0.60	1/4923 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	150	PRO	C-O	7.01	1.37	1.23

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	186	LEU	CA-CB-CG	5.72	128.46	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	1798	0	1745	18	0
1	D	1755	0	1712	24	0
2	A	35	0	0	1	0
2	D	29	0	0	2	0
All	All	3617	0	3457	42	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 6.

All (42) 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:168:LYS:H	1:D:168:LYS:HD2	1.41	0.85
1:A:149:ILE:H	1:A:150:PRO:HA	1.43	0.84
1:D:116:MET:HE2	1:D:132:VAL:HA	1.59	0.83
1:A:29:ARG:HG3	1:A:29:ARG:HH11	1.44	0.80
1:D:86:HIS:O	1:D:90:ILE:HG12	1.85	0.76
1:A:44:ASP:OD2	1:A:86:HIS:NE2	2.22	0.73
1:A:9:GLN:HG3	2:A:249:HOH:O	1.91	0.70
1:A:44:ASP:OD2	1:A:86:HIS:CE1	2.47	0.67
1:D:116:MET:HE3	1:D:131:ALA:O	1.95	0.67
1:D:116:MET:HE3	2:D:232:HOH:O	1.95	0.66
1:A:149:ILE:N	1:A:150:PRO:HA	2.13	0.64
1:A:37:PHE:O	1:A:41:ILE:HG23	2.00	0.62
1:A:62:ALA:HB1	1:A:67:VAL:CG2	2.32	0.60
1:A:41:ILE:HD11	1:A:90:ILE:HG13	1.85	0.59
1:A:29:ARG:HG3	1:A:29:ARG:NH1	2.17	0.55
1:A:7:LEU:HB3	1:A:200:PHE:CD1	2.42	0.55
1:D:190:LYS:N	1:D:190:LYS:HD2	2.22	0.55
1:A:62:ALA:HB1	1:A:67:VAL:HG23	1.90	0.54
1:D:116:MET:CE	1:D:131:ALA:O	2.56	0.54
1:D:7:LEU:HD22	1:D:197:LYS:HA	1.91	0.53
1:D:62:ALA:HB1	1:D:67:VAL:HG22	1.90	0.53
1:D:116:MET:CE	1:D:132:VAL:HA	2.37	0.52
1:D:65:GLU:OE2	1:D:69:ARG:NH2	2.44	0.51
1:D:168:LYS:H	1:D:168:LYS:CD	2.10	0.51
1:D:37:PHE:CE2	1:D:41:ILE:HD11	2.47	0.50
1:D:44:ASP:HB3	1:D:86:HIS:HE2	1.77	0.48
1:D:54:VAL:HG22	1:D:117:LEU:HD13	1.97	0.47
1:A:153:LEU:H	1:A:153:LEU:HD12	1.81	0.46
1:A:62:ALA:HA	1:A:125:ILE:HD12	1.98	0.46
1:D:116:MET:CE	2:D:232:HOH:O	2.58	0.45
1:D:8:TYR:CZ	1:D:12:GLN:OE1	2.69	0.45
1:D:44:ASP:OD2	1:D:86:HIS:NE2	2.49	0.44
1:A:148:GLN:HE21	1:A:148:GLN:HB3	1.67	0.44
1:D:190:LYS:HD2	1:D:190:LYS:H	1.83	0.44
1:D:41:ILE:CD1	1:D:89:TYR:CB	2.97	0.43
1:D:168:LYS:HD2	1:D:168:LYS:N	2.20	0.43
1:A:7:LEU:HD22	1:A:197:LYS:HA	2.01	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:25:GLN:O	1:A:29:ARG:HG2	2.19	0.42
1:D:148:GLN:HE21	1:D:148:GLN:HB3	1.56	0.41
1:D:41:ILE:HD13	1:D:89:TYR:HB2	2.03	0.41
1:D:7:LEU:HB3	1:D:200:PHE:CD1	2.55	0.41
1:A:48:LEU:HD11	1:A:86:HIS:CE1	2.57	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	219/221 (99%)	208 (95%)	9 (4%)	2 (1%)	17	25
1	D	212/221 (96%)	203 (96%)	8 (4%)	1 (0%)	29	41
All	All	431/442 (98%)	411 (95%)	17 (4%)	3 (1%)	22	32

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	64	ASP
1	D	64	ASP
1	A	149	ILE

### 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	
1	A	193/193 (100%)	171 (89%)	22 (11%)	5	7
1	D	188/193 (97%)	167 (89%)	21 (11%)	6	8
All	All	381/386 (99%)	338 (89%)	43 (11%)	6	8

All (43) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	0	THR
1	A	12	GLN
1	A	29	ARG
1	A	41	ILE
1	A	46	LEU
1	A	49	LEU
1	A	57	LEU
1	A	60	VAL
1	A	79	LEU
1	A	80	ASN
1	A	107	LEU
1	A	117	LEU
1	A	128	VAL
1	A	141	VAL
1	A	147	SER
1	A	148	GLN
1	A	153	LEU
1	A	168	LYS
1	A	177	ASN
1	A	190	LYS
1	A	213	MET
1	A	216	GLN
1	D	3	VAL
1	D	9	GLN
1	D	29	ARG
1	D	34	ARG
1	D	46	LEU
1	D	48	LEU
1	D	49	LEU
1	D	60	VAL
1	D	67	VAL
1	D	79	LEU
1	D	83	MET
1	D	107	LEU
1	D	117	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	148	GLN
1	D	168	LYS
1	D	169	GLU
1	D	177	ASN
1	D	186	LEU
1	D	190	LYS
1	D	201	ILE
1	D	209	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	73	ASN
1	A	80	ASN
1	A	102	ASN
1	A	148	GLN
1	A	171	GLN
1	A	216	GLN
1	D	73	ASN
1	D	87	ASN
1	D	102	ASN
1	D	144	GLN
1	D	148	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

### 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](#)

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

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	221/221 (100%)	0.31	15 (6%) <span style="border: 1px solid red; padding: 2px;">17</span> <span style="border: 1px solid red; padding: 2px;">15</span>	8, 20, 38, 41	0
1	D	216/221 (97%)	0.06	5 (2%) <span style="border: 1px solid blue; padding: 2px;">60</span> <span style="border: 1px solid blue; padding: 2px;">58</span>	10, 17, 35, 40	0
All	All	437/442 (98%)	0.19	20 (4%) <span style="border: 1px solid red; padding: 2px;">32</span> <span style="border: 1px solid red; padding: 2px;">31</span>	8, 19, 37, 41	0

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	83	MET	4.8
1	A	149	ILE	4.0
1	A	152	ALA	3.5
1	A	87	ASN	3.2
1	D	83	MET	3.2
1	A	82	GLU	3.1
1	D	0	THR	3.0
1	A	157	PHE	3.0
1	A	151	ASN	2.9
1	A	91	ARG	2.8
1	A	153	LEU	2.6
1	A	154	GLU	2.5
1	A	156	ALA	2.4
1	D	149	ILE	2.3
1	A	92	GLU	2.2
1	A	216	GLN	2.1
1	A	163	LYS	2.1
1	D	91	ARG	2.1
1	A	90	ILE	2.0
1	D	90	ILE	2.0

## 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 monosaccharides 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.