



# wwPDB X-ray Structure Validation Summary Report

Mar 12, 2014 – 04:41 PM GMT

PDB ID : 4EHS  
Title : Crystal structure of Helicobacter pylori DnaG Primase C terminal domain  
Authors : Abdul Rehman, S.A.; Gourinath, S.  
Deposited on : 2012-04-04  
Resolution : 1.78 Å(reported)

This is a wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

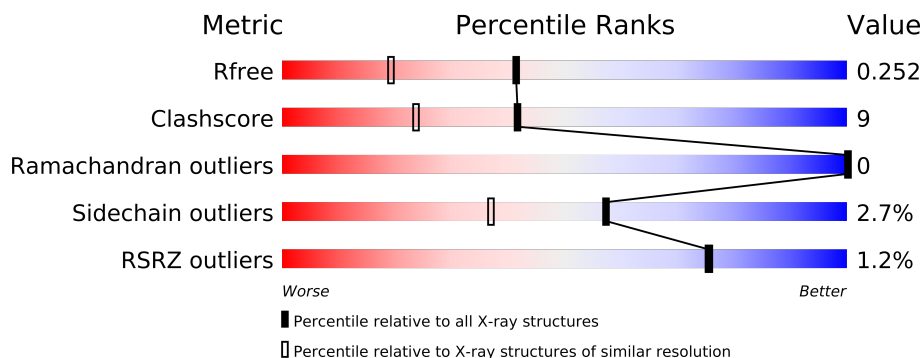
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.16 November 2013  
Xtriage (Phenix) : dev-1323  
EDS : trunk22714  
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) : trunk22714

# 1 Overall quality at a glance

The reported resolution of this entry is 1.78 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	4987 (1.80-1.76)
Clashscore	79885	6152 (1.80-1.76)
Ramachandran outliers	78287	6074 (1.80-1.76)
Sidechain outliers	78261	6073 (1.80-1.76)
RSRZ outliers	66119	4990 (1.80-1.76)

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. 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	155	
1	B	155	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	BME	A	602	-	X

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2222 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 DNA primase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	124	Total	C	N	O	S	Se	0	0	0
			1004	657	167	176	2	2			
1	B	122	Total	C	N	O	S	Se	0	0	0
			989	649	165	171	2	2			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	560	LEU	-	EXPRESSION TAG	UNP P56064
A	561	GLU	-	EXPRESSION TAG	UNP P56064
A	562	HIS	-	EXPRESSION TAG	UNP P56064
A	563	HIS	-	EXPRESSION TAG	UNP P56064
A	564	HIS	-	EXPRESSION TAG	UNP P56064
A	565	HIS	-	EXPRESSION TAG	UNP P56064
A	566	HIS	-	EXPRESSION TAG	UNP P56064
A	567	HIS	-	EXPRESSION TAG	UNP P56064
B	560	LEU	-	EXPRESSION TAG	UNP P56064
B	561	GLU	-	EXPRESSION TAG	UNP P56064
B	562	HIS	-	EXPRESSION TAG	UNP P56064
B	563	HIS	-	EXPRESSION TAG	UNP P56064
B	564	HIS	-	EXPRESSION TAG	UNP P56064
B	565	HIS	-	EXPRESSION TAG	UNP P56064
B	566	HIS	-	EXPRESSION TAG	UNP P56064
B	567	HIS	-	EXPRESSION TAG	UNP P56064

- Molecule 2 is BETA-MERCAPTOETHANOL (three-letter code: BME) (formula: C<sub>2</sub>H<sub>6</sub>OS).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	O	S	0	0
			4	2	1	1		
2	A	1	Total	C	O	S	0	0
			4	2	1	1		
2	B	1	Total	C	O	S	0	0
			4	2	1	1		
2	B	1	Total	C	O	S	0	0
			4	2	1	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	103	Total	O	0	0
			103	103		
3	B	110	Total	O	0	0
			110	110		



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	48.88Å 61.45Å 82.42Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.27 – 1.78 49.27 – 1.78	Depositor EDS
% Data completeness (in resolution range)	99.6 (49.27-1.78) 99.6 (49.27-1.78)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.51 (at 1.77Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.195 , 0.249 0.199 , 0.252	Depositor DCC
$R_{free}$ test set	1250 reflections (5.39%)	DCC
Wilson B-factor (Å <sup>2</sup> )	22.2	Xtriage
Anisotropy	0.385	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 32.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	1 of 24450 reflections (0.004%)	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	2222	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.28% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities 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: BME

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	0.94	0/1017	1.01	2/1358 (0.1%)
1	B	0.98	1/1002 (0.1%)	1.25	12/1338 (0.9%)
All	All	0.96	1/2019 (0.0%)	1.13	14/2696 (0.5%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	521	GLU	CD-OE1	5.30	1.31	1.25

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	545	ARG	NE-CZ-NH2	-13.39	113.61	120.30
1	B	545	ARG	NE-CZ-NH1	11.38	125.99	120.30
1	B	513	ARG	NE-CZ-NH1	9.72	125.16	120.30
1	B	513	ARG	NE-CZ-NH2	-8.91	115.84	120.30
1	A	545	ARG	NE-CZ-NH1	-7.82	116.39	120.30

There are no chirality outliers.

There are no planarity outliers.

### 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	1004	0	1079	26	0
1	B	989	0	1068	15	0
2	A	8	0	10	3	0
2	B	8	0	10	0	0
3	A	103	0	0	12	0
3	B	110	0	0	5	0
All	All	2222	0	2167	38	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 9.

The worst 5 of 38 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:550:LYS:HE2	3:B:777:HOH:O	1.60	1.02
1:A:510:LYS:HD2	3:A:799:HOH:O	1.59	1.00
1:A:455:GLU:CG	3:A:801:HOH:O	2.20	0.89
1:A:455:GLU:HG2	3:A:801:HOH:O	1.73	0.84
1:A:484:GLN:HB3	3:A:732:HOH:O	1.79	0.81

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	122/155 (79%)	118 (97%)	4 (3%)	0	100	100
1	B	120/155 (77%)	120 (100%)	0	0	100	100
All	All	242/310 (78%)	238 (98%)	4 (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	111/139 (80%)	109 (98%)	2 (2%)	71	56
1	B	109/139 (78%)	105 (96%)	4 (4%)	45	24
All	All	220/278 (79%)	214 (97%)	6 (3%)	57	37

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	443	GLU
1	B	550	LYS
1	B	454	LEU
1	A	561	GLU
1	B	469	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	484	GLN
1	B	474	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 ⓘ

4 ligands are modelled in this entry.

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	$\# Z  > 2$
2	BME	A	601	1	3,3,3	0.70	0	2,2,2	0.28	0
2	BME	A	602	1	3,3,3	1.22	1 (33%)	2,2,2	1.51	1 (50%)
2	BME	B	601	1	3,3,3	0.31	0	2,2,2	0.29	0
2	BME	B	602	1	3,3,3	1.02	0	2,2,2	1.07	0

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
2	BME	A	601	1	-	0/1/1/1	0/0/0/0
2	BME	A	602	1	-	0/1/1/1	0/0/0/0
2	BME	B	601	1	-	0/1/1/1	0/0/0/0
2	BME	B	602	1	-	0/1/1/1	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	602	BME	C2-S2	-2.02	1.73	1.80

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	602	BME	O1-C1-C2	-2.11	101.68	110.80

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, 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	124/155 (80%)	-0.12	0 100 100	13, 22, 35, 51	0
1	B	122/155 (78%)	-0.17	2 (1%) 68 68	13, 21, 32, 62	0
All	All	246/310 (79%)	-0.15	2 (0%) 75 83	13, 22, 34, 62	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	439	ALA	2.8
1	B	549	MSE	2.7

### 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, 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	BME	A	602	4/4	0.44	5.24	38,42,43,44	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	BME	A	601	4/4	0.16	1.77	36,42,44,49	0
2	BME	B	601	4/4	0.15	1.63	32,38,39,52	0
2	BME	B	602	4/4	0.14	1.63	35,39,39,43	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.