Tables

Table ST1

List of primers used for mutagenesis and EMSA experiments. The locations of mutations are shown in lowercase	
HpDnaBWTFw BamH1	GCGGATCCATGGATCATTTAAAGCATTTGC
HpDnaBWTRv BamH1	GCGGATCCTCAAGTTGTAACTATATCATAA
HpDnaGFwSalI	ACGGTCGACATGATTCTTAAAAGTTCCATTG
HpDnaGRvSalI	ACGGTCGACATATGGCGACTAATTCTCCTTG
HpDnaG-CTD Fw Nhe1	CTAGCTAGCGAGCGAGTCTCTTTTCAGCCTTT
HpDnaG-CTD Rv Xho1	CGGCTCGAGTATGGCGACTAATTCTCCTTGTTTT
FwHpDnaG F534A Fw	5’CCTAAAAGCTCGCTCCCTgctAGCGAAAAAATGATCTGT 3
RvHpDnaG F534A Rv	5’ACAGATCATTTTTTCGCTagcAGGGAGCGAGCTTTTAGG 3’
Hpb70NZ (EMSA)	CTGCCACACACCACCAAAGCTAACTGACAGGAGAATCCAGATGGAT ATTTAAAGCATTTGCAGCAATTG

Table ST2

Structure based alignment of *H. pylori*DnaG-CTD with other structures using Rapido (63).			
	Bst (X-ray)	*Bst* (NMR)	*E.coli*
#aligned	106	98	66
RMSD rigid	8.77	8.81	15.69
#aligned (rigid bodies)	74	67	62
RMSD flex	1.11	0.83	0.54
# rigid bodies	4	5	5

Table ST3.

Average B-factors of different regions of DnaG C-terminal domain			
Protein	Globe region	Linker	Hairpin helix
*Hp*DnaG (4EHS)	18.4 (442-494)	24.3 (495-506)	22.7 (507-551)
*Ec*DnaG (1T3W)	61.7 (451-519)	89.6 (520-526)	66.10 (527-577)
*Bst*DnaG (2R6A)*	98.92 (456-532)	103.06 (533-555)	92.64 (556-592)

Residue range of the respective regions has been shown in parenthesis. Terminal residues were left out to keep the bias level at minimum. The linker regions clearly show higher temperature factors compared to other regions, indicating flexibility of this region.

*The *Bst*DnaG (2R6A) structure is in complex with helicase, where the HHR is bound to helicase, thus may have less B- factors compared to other parts of the structure.

Table ST4.

Interaction energies of the *Hp*DnaB (helicase) with native *Hp*DnaG-CTD and mutant *Hp*DnaG-CTD F534A (primase). All energies are in kcal/mol.

Complex	ΔE_{Vdw}	ΔE_{ele}	ΔE_{gb}	$\Delta E_{surface}$	ΔG_{gb}	ΔG_{gb_sol}	$\Delta \Delta G_{bind}$
HpDnaG-CTD Native	-85.34	-1426.25	1515.60	-13.18	-1511.61	1502.42	-9.18
*HpDnaG -CTD Mutant (*F534A)	-89.66	-1446.78	1546.81	-13.30	-1536.44	1533.51	-2.93

Supplementary Figures

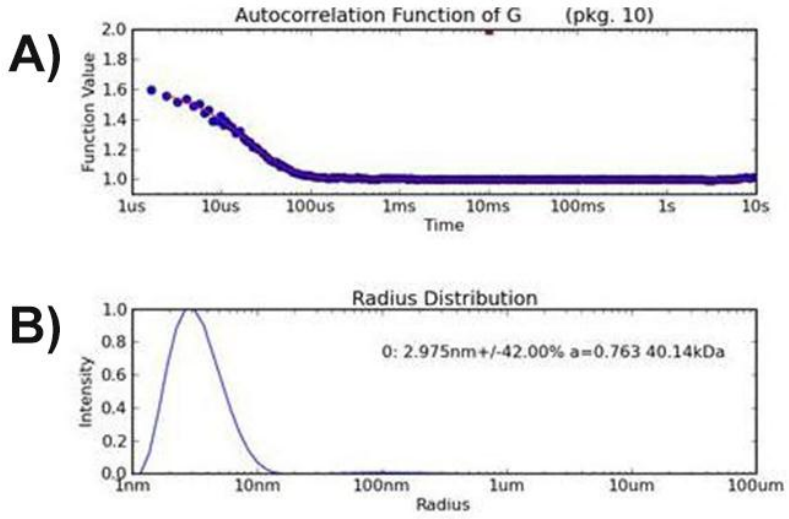

FIG. S1. The DLS profile points towards the mono-disperse state as well as the dimeric nature of the *Hp*DnaG-CTD. The molecular weight of the mono-dispersed peak obtained from hydro dynamic radius of the particle indicates ~40 KDa. The molecular weight of monomer protein is 17KDa, clearly indicating a dimeric state of the protein. A) Autocorrelation function (Y-axis) plotted as a function of time (X-axis) and (B) hydrodynamic radii (X-axis) plotted as a function of intensity distribution (Y-axis).

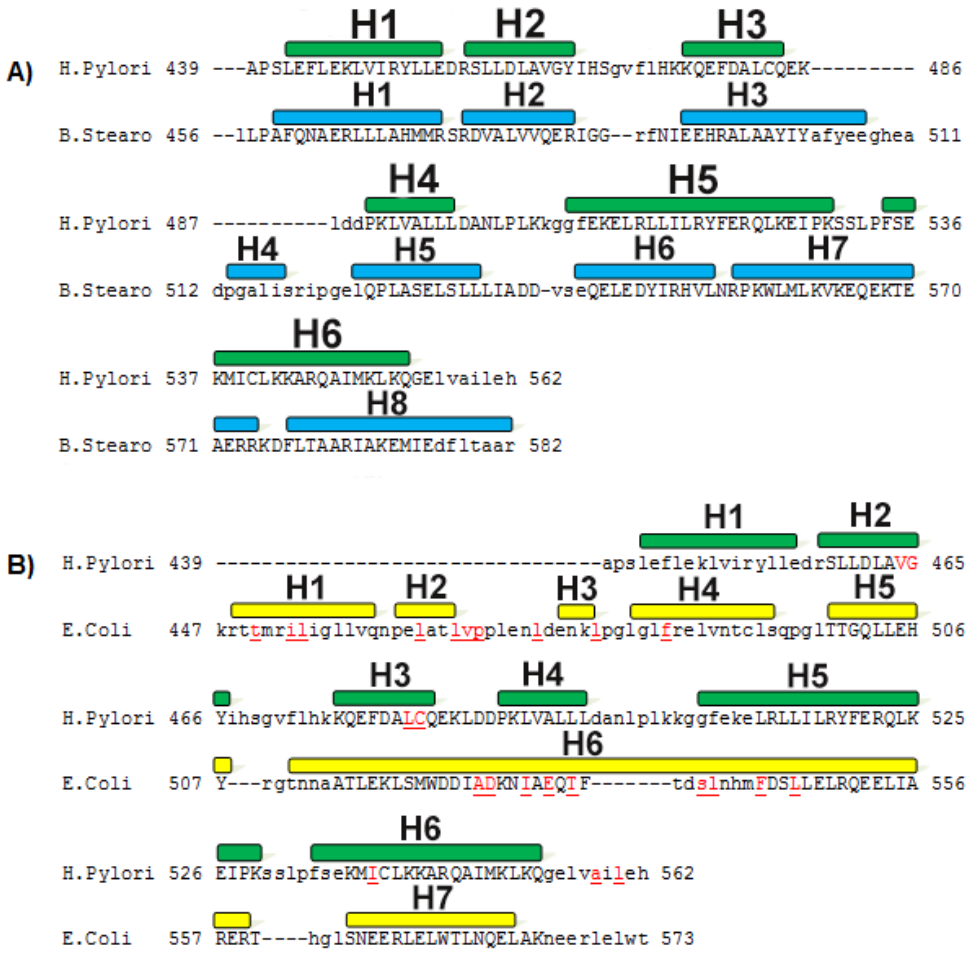

FIG. S2. Structure based sequence alignment of DnaG primase C terminal domains from *H. pylori(Hp)*, *Bacillus stearothermophilus* (*Bst*) and *Escherichia coli* (*Ec*) using RAPIDO(1). (A)Structure based sequence alignment of *Hp*DnaG-CTD with *Bst*DnaG-CTD. (B) Structure based sequence alignment of *Hp*DnaG-CTD with *Ec*DnaG-CTD.The aligned residues are indicated in capital letters while small letters are used for non-aligned residues. The underlined red letter residues are the residues seen in *Ec* and *Hp* interacting at the dimeric interface. The residues responsible for the dimerization in *Hp* and *Ec* are completely different.

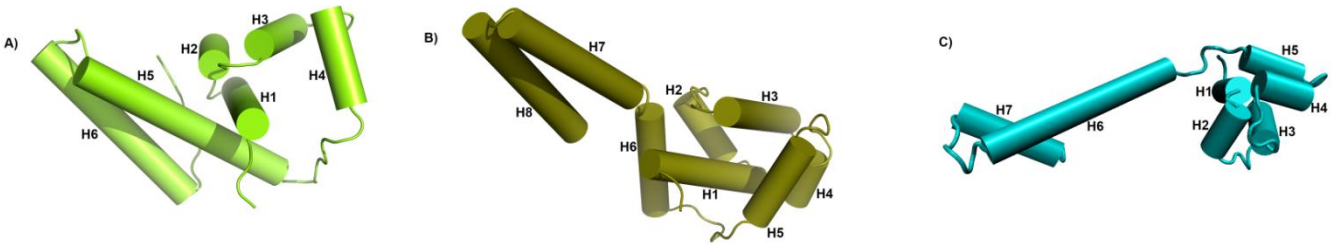

FIG. S3. Cartoons showing the helices in crystal structures of DnaG primase C- terminal domain. A) *Hp*DnaG-CTD B) *Bst*DnaG-CTD C) *Ec*DnaG-CTD. The Globe region (GR) of *Hp* has lesser number of helices (four) as compared to the other two. Also, the *Hp* globe region is

less compact. Hairpin helices are bigger in *Hp* and *Bst* than in *Ec (*H7 of hairpin is very small). Images were made with either PyMOL(2) or VMD(3).

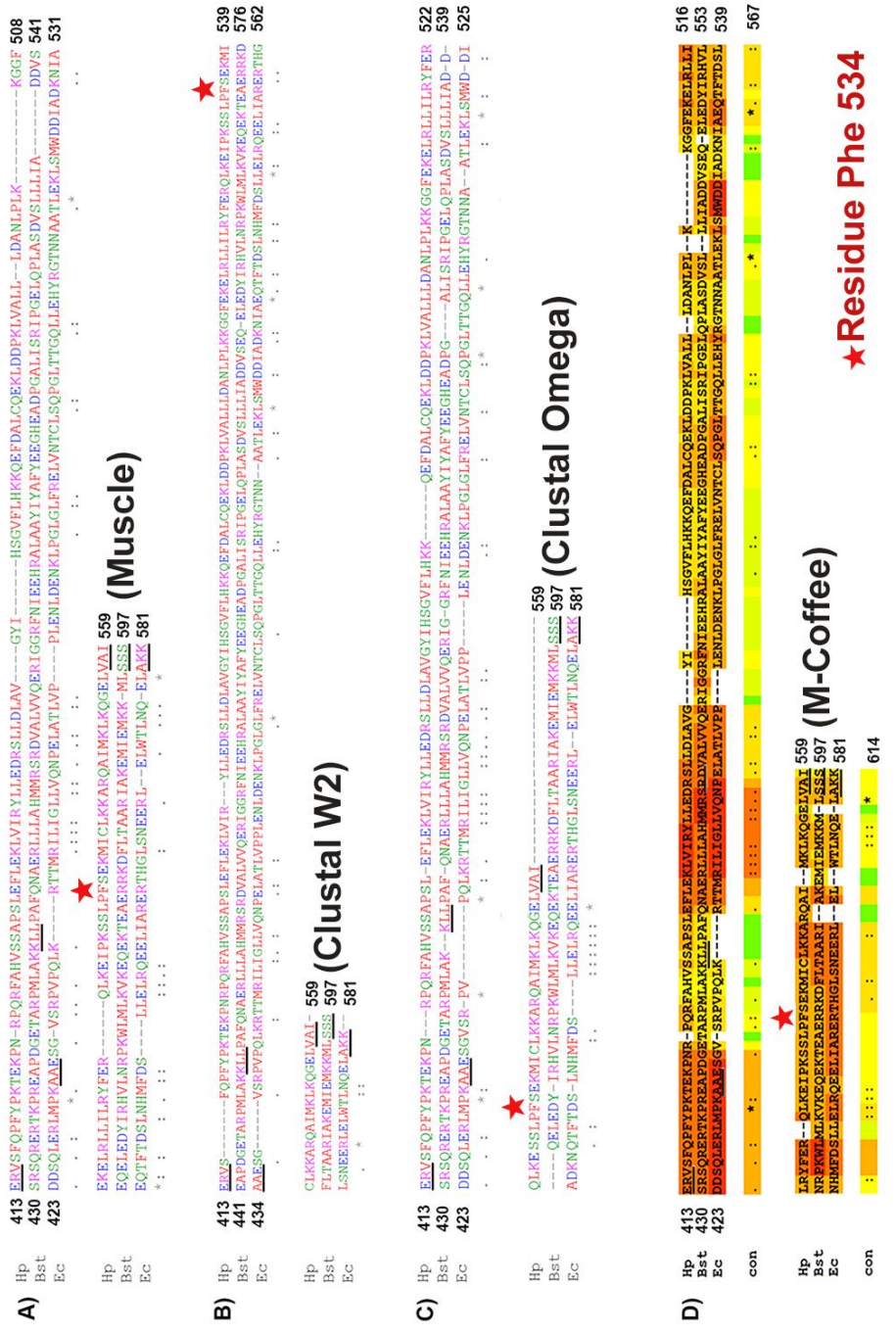

FIG. S4 Multiple sequence alignment of the DnaG primase C terminal domain (CTD). The alignment was done for the full length protein sequences and the regions corresponding to CTD were scouted out(4).The first and last three residues have been underlined according to the stretch of the gene used for the structure determination to give an idea where these sequences fall in MSA. In all, Web based servers were used. The alignment highlights the absence of

conservation. The programs used were, A) Muscle(5, 6); B) Clustal W2(7, 8); C) Clustal Omega(8, 9); D) M-Coffee(6, 10)**.**

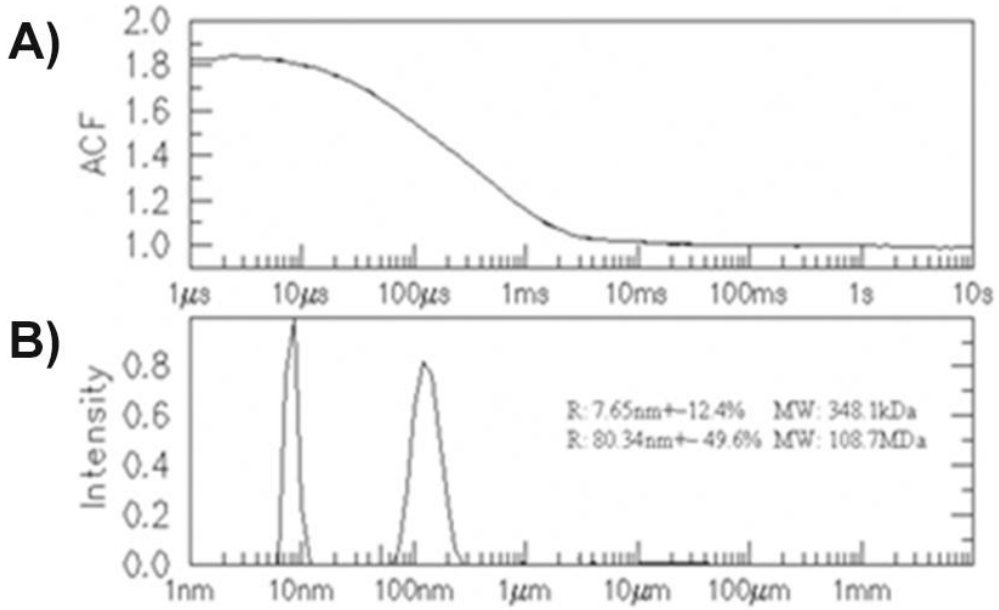

FIG.S5. Dynamic Light Scattering profile of purified wtHpDnaB helicase at pH 8.0 (a) Autocorrelation function (Y-axis) plotted as a function of time (X-axis) and (b) hydrodynamic radii (X-axis) plotted as a function of intensity distribution (Y-axis). From the hydrodynamic radius of the first peak, the molecular weight of the particle is estimated about ~350kDa, which is close to the hexameric state (~330kDa) of the protein. The high molecular weight second peak may be dust or higher aggregates of the protein. {Adapted from Ph.D thesis of Tara Kashav, 2009 (unpublished data from the lab)}.

References

1. **Mosca R, Schneider TR.** 2008. RAPIDO: a web server for the alignment of protein structures in the presence of conformational changes. Nucleic Acids Res **36:**W42-46.
2. **DeLano WL.** 2002. The PyMOL User's Manual. DeLano Scientific, San Carlos, CA, USA.
3. **Humphrey W, Dalke A, Schulten K.** 1996. VMD: visual molecular dynamics. Journal of molecular graphics **14:**33-38, 27-38.
4. **Jacob FF, Brenner S, Cuzin F.** 1963. On the regulation of DNA replication in bacteria. Cold Spring Harbor symposia on Quantitative Biology **28:**329-348.
5. **Edgar RC.** 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC bioinformatics **5:**113.
6. **Moretti S, Armougom F, Wallace IM, Higgins DG, Jongeneel CV, Notredame C.** 2007. The M-Coffee web server: a meta-method for computing multiple sequence alignments by combining alternative alignment methods. Nucleic Acids Res **35:**W645-648.

7. **Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG.** 2007. Clustal W and Clustal X version 2.0. Bioinformatics **23:**2947-2948.
8. **Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, Lopez R.** 2010. A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res **38:**W695-699.
9. **Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG.** 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular systems biology **7:**539.
10. **Wallace IM, O'Sullivan O, Higgins DG, Notredame C.** 2006. M-Coffee: combining multiple sequence alignment methods with T-Coffee. Nucleic Acids Res **34:**1692-1699.