

Science

 AAAS

Molecular Mechanisms of HipA-Mediated Multidrug Tolerance and Its Neutralization by HipB

Maria A. Schumacher, *et al.*

Science **323**, 396 (2009);

DOI: 10.1126/science.1163806

The following resources related to this article are available online at www.sciencemag.org (this information is current as of February 4, 2009):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/cgi/content/full/323/5912/396>

Supporting Online Material can be found at:

<http://www.sciencemag.org/cgi/content/full/323/5912/396/DC1>

This article **cites 29 articles**, 13 of which can be accessed for free:

<http://www.sciencemag.org/cgi/content/full/323/5912/396#otherarticles>

This article appears in the following **subject collections**:

Microbiology

<http://www.sciencemag.org/cgi/collection/microbio>

Information about obtaining **reprints** of this article or about obtaining **permission to reproduce this article** in whole or in part can be found at:

<http://www.sciencemag.org/about/permissions.dtl>

To further investigate the detailed molecular events giving rise to LCMV/IAP hybrid DNA formation, we designed nested PCR approaches that detect recombinant cLCMV/IAP hybrid sequences (13). We were able to amplify both cNP/IAP and cGP/IAP recombination events directly from infected murine MC57G cells (Fig. 2). We also detected cGP/IAP recombination events *in vivo* in splenocytes of an acutely infected C57BL/6 mouse and in peripheral blood mononuclear cells of RAG1^{-/-} mice persistently infected with LCMV (Fig. 2B), demonstrating that recombination between IAP retrotransposons and LCMV was not an artifact of immortalized cells but also occurred spontaneously *in vivo*.

The abundance of IAP elements in the murine genome (2) makes a detailed interpretation of sequence data difficult, as the exact sequence of the IAP element involved in LCMV hybrid formation is not known. We generated a recombinant HeLa cell line (HG4) in which all the IAP elements were derived from a single molecular clone 92L23Neo^R (15) and that reliably formed cLCMV upon infection.

Three days after LCMV infection of HG4 cells we amplified several different cNP/IAP and cGP/IAP recombination events involving the 92L23Neo^R IAP element (Fig. 2). All amplified recombination products were sequenced and analyzed for homologies to 92L23Neo^R and the LCMV genome. In contrast to the sequences obtained from the subcloned infected MC57G cell lines MC57.18 and MC57.23, where IAP and LCMV sequences were directly joined to each other (Fig. 1), we observed insertions of additional nucleotides at the site of recombination in cLCMV/IAP hybrid DNA amplified from infected HG4 and MC57G cells (Fig. 2). These insertions, consisting of up to 14 nts displayed homology to the 3' end of the 76-nt-long phenylalanine tRNA (tRNA^{Phe}, GenBank accession number K02684), which has previously been shown to act as primer for the reverse transcription of IAP elements (SOM text 2) (17).

We showed that endogenous retroviral elements can recombine with exogenous nonretroviral RNA viruses presumably by copy choice during reverse transcription (SOM text 2) to yield cDNA complementary to both the endogenous retrovirus and the exogenous RNA virus. Such recombinant cDNA has the potential to integrate into the genome of the infected cell, and because IAP elements have been shown to translocate most efficiently in the thymus and in the LCMV-sanctuaries (18) of the testicles (11, 12), such recombinations theoretically could lead to germline transmission of viral genes from RNA viruses. Therefore, nonretroviral RNA viruses may have contributed more substantially and directly to the evolution of mammalian genomes than has been assumed so far (SOM text 3). However, extensive database analysis of the murine genome, which was determined with DNA from LCMV-free mice, did not yield substantial homologies

to the LCMV-genome. In humans, there is some evidence to suggest that at least one subgroup of the human endogenous retrovirus-K family may be active (19, 20). Thus, the potential risk of somatic integration into host cells during human gene therapy using RNA virus vectors should be experimentally assessed by nested PCR.

References and Notes

1. P. Klenerman, H. Hengartner, R. M. Zinkernagel, *Nature* **390**, 298 (1997).
2. E. L. Kuff, K. K. Lueders, *Adv. Cancer Res.* **51**, 183 (1988).
3. E. L. Kuff *et al.*, *Nature* **302**, 547 (1983).
4. E. Canaani *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 7118 (1983).
5. K. B. Leslie, F. Lee, J. W. Schrader, *Mol. Cell. Biol.* **11**, 5562 (1991).
6. H. Ishihara, I. Tanaka, H. Wan, K. Nojima, K. Yoshida, *J. Radiat. Res. (Tokyo)* **45**, 25 (2004).
7. X. Y. Wang, L. S. Steelman, J. A. McCubrey, *Cytokines Cell. Mol. Ther.* **3**, 3 (1997).
8. I. A. Maksakova *et al.*, *PLoS Genet.* **2**, e2 (2006).
9. E. Braun, E. Rorman, K. K. Lueders, A. Bar-Sinai, J. Hochman, *Virology* **277**, 136 (2000).
10. M. Li *et al.*, *J. Gen. Virol.* **77**, 2757 (1996).
11. A. Dupressoir, T. Heidmann, *Mol. Cell. Biol.* **16**, 4495 (1996).
12. E. L. Kuff, J. W. Fewell, *Mol. Cell. Biol.* **5**, 474 (1985).
13. Materials and methods are available as supporting material on Science Online.
14. F. Lehmann-Grube, *Virology Monographs* (Springer, New York, 1971).

15. M. Dewannieux, A. Dupressoir, F. Harper, G. Pierron, T. Heidmann, *Nat. Genet.* **36**, 534 (2004).
16. L. Naldini *et al.*, *Science* **272**, 263 (1996).
17. M. Ono, H. Ohishi, *Nucleic Acids Res.* **11**, 7169 (1983).
18. M. Recher *et al.*, *Nat. Med.* **13**, 1316 (2007).
19. N. Bannert, R. Kurth, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 14572 (2004).
20. R. Belshaw *et al.*, *J. Virol.* **79**, 12507 (2005).
21. We are very grateful to K. K. Lueders for providing plasmids and critically reading the manuscript. We thank A. Macpherson, K. McCoy, and D. Burton for their support. We would also like to thank K. McCoy for critically reading the manuscript and for helpful comments. This study was supported by the Swiss National Science Foundation and the Canton of Zurich. M.B.G. planned and performed the experiments and wrote the paper together with L.H.; J.W. performed experiments; M.D., E.G., and T.H. provided reagents, cell lines, or plasmids; H.H. and R.M.Z. were involved in the planning and discussions of experiments, helped to write the paper, and provided the infrastructure used for this study; and L.H. initiated and supervised the study, planned and conducted experiments, and wrote the paper together with M.B.G. Sequences depicted in Fig. 1 are available from GenBank under accession numbers FJ460582 (MC57.18) and FJ460583 (MC57.23).

Supporting Online Material

www.sciencemag.org/cgi/content/full/323/5912/393/DC1
Materials and Methods
SOM Text
Figs. S1 to S3
Tables S1 and S2
References

20 October 2008; accepted 19 November 2008
10.1126/science.1167375

Molecular Mechanisms of HipA-Mediated Multidrug Tolerance and Its Neutralization by HipB

Maria A. Schumacher,^{1*} Kevin M. Piro,¹ Weijun Xu,¹ Sonja Hansen,² Kim Lewis,² Richard G. Brennan^{1*}

Bacterial multidrug tolerance is largely responsible for the inability of antibiotics to eradicate infections and is caused by a small population of dormant bacteria called persisters. HipA is a critical *Escherichia coli* persistence factor that is normally neutralized by HipB, a transcription repressor, which also regulates *hipBA* expression. Here, we report multiple structures of HipA and a HipA-HipB-DNA complex. HipA has a eukaryotic serine/threonine kinase-like fold and can phosphorylate the translation factor EF-Tu, suggesting a persistence mechanism via cell stasis. The HipA-HipB-DNA structure reveals the HipB-operator binding mechanism, ~70° DNA bending, and unexpected HipA-DNA contacts. Dimeric HipB interacts with two HipA molecules to inhibit its kinase activity through sequestration and conformational inactivation. Combined, these studies suggest mechanisms for HipA-mediated persistence and its neutralization by HipB.

Bacteria that are resistant or tolerant to antibiotics are an increasing threat to human health. Indeed, ~60% of infections in the developed world are caused by biofilms, which exhibit multidrug tolerance (MDT) (1, 2). MDT is caused by the presence of dormant bacterial cells called persisters, which account for only 10⁻⁶ to 10⁻⁴ cells in a growing population, making MDT difficult to study (3–5). Persisters are not mutants but phenotypic variants of wild-type cells that evade killing by somehow adopting a transient dormant state (6, 7). Dormancy provides protection because bactericidal antibiotics kill by

corrupting their active targets into producing toxic byproducts. These protected persisters can then switch back to the growth phase after the removal of antibiotics, allowing the bacterial population to survive. The first high-persistence allele, *hipA7*

¹Department of Biochemistry and Molecular Biology, University of Texas, M. D. Anderson Cancer Center, Unit 1000, Houston, TX 77030, USA. ²Department of Biology and Antimicrobial Discovery Center, Northeastern University, Boston, MA 02115, USA.

*To whom correspondence should be addressed. E-mail: rgschuma@mdanderson.org (R.G.B.); maschuma@mdanderson.org (M.A.S.)

(*high-persistence A*), was identified in *Escherichia coli* and increased the frequency of persistence by 10,000 fold (8–10). *E. coli hipA* encodes a 440-residue protein, HipA, which is cotranscribed with a smaller upstream gene, *hipB*. HipB is an 88-residue protein that represses the *hipBA* operon by binding cooperatively to four operators upstream of *hipBA* (11, 12). HipB forms a complex with HipA, and because wild-type HipA cannot be expressed in the absence of HipB because of its deleterious effects on cell growth, *hipBA* has been categorized as a toxin/antitoxin (TA) module in which HipA, the toxin, is neutralized by the antitoxin, HipB (13, 14). Toxin proteins from chromosomally encoded TA modules, of which more than 10 have been identified in *E. coli*, appear to promote cell dormancy and may play roles in the development of persistence under certain conditions (5, 7). Chromosomal TA modules can be grouped into three main superfamilies based on whether the toxin has a ribonuclease (RNase)/gyrase-like fold, RNase barnase-like structure, or a PilT N-terminus (PIN) domain (14). The corresponding antitoxins contain DNA binding domains and C termini that are largely unfolded until bound by the toxin (14). HipA and HipB show no homology to any member of these TA superfamilies. Moreover, HipA is one of the few validated biofilm tolerance factors. Indeed, it has been demonstrated that overexpression of the HipA protein leads to MDT in *E. coli* (2). However, the mechanism of HipA-mediated MDT is unknown.

To delineate the functions of HipA and HipB in MDT, we carried out biochemical and structural studies on HipA and HipA-HipB-DNA complexes. Because of wild-type HipA-mediated persistence, we used the HipA mutant Asp³⁰⁹→Gln³⁰⁹ (D309Q) (referred to as HipA^o), which can be produced in large quantities in the absence of HipB (15). The structure of HipA was solved to 1.54 Å resolution and refined to an $R_{\text{work}}/R_{\text{free}}$ of 19.5/23.2% (table S1 and fig. S1) (16–18). The HipA structure has a globular fold with 15 β strands and 15 α helices and can be divided into an N-terminal α/β domain and an all-α-helical C-terminal domain (Fig. 1A). Density is missing for residues 185 to 195, which are near the active site and probably correspond to the activation loop of other kinases. Structure-based homology searches revealed that HipA is most similar to human CDK2/cyclin A kinase (19). The structural homology between HipA and CDK2 was highest in the C-terminal region that contains CDK2 catalytic residues, suggesting that HipA functions as a protein kinase, as reported (15). Although HipA is most similar to CDK2, the proteins superimposed with a large root mean square deviation (RMSD) of 3.9 Å for 150 corresponding Cα atoms, indicating that HipA represents a previously unknown class of protein kinase (20).

HipA contains all the catalytic residues found in protein kinases, including the putative catalytic base Asp³⁰⁹ (20). The D309Q mutation abrogates persistence, strongly suggesting that kinase func-

tion is key to HipA-mediated MDT (15). Indeed, we found that HipA binds adenosine triphosphate (ATP) with a dissociation constant (K_d) of $18.0 \pm 2.0 \mu\text{M}$, which is similar to the dissociation constants obtained for ATP binding to other serine/threonine kinases (fig. S2) (20). To delineate the

ATP binding mechanism of HipA, we determined the structure of the HipA-ATP-Mg²⁺ complex to 1.66 Å resolution and refined the structure to an $R_{\text{work}}/R_{\text{free}}$ of 18.4/21.7%. Density for ATP is observed in the cleft between the HipA N and C domains (fig. S1). HipA binds ATP with high

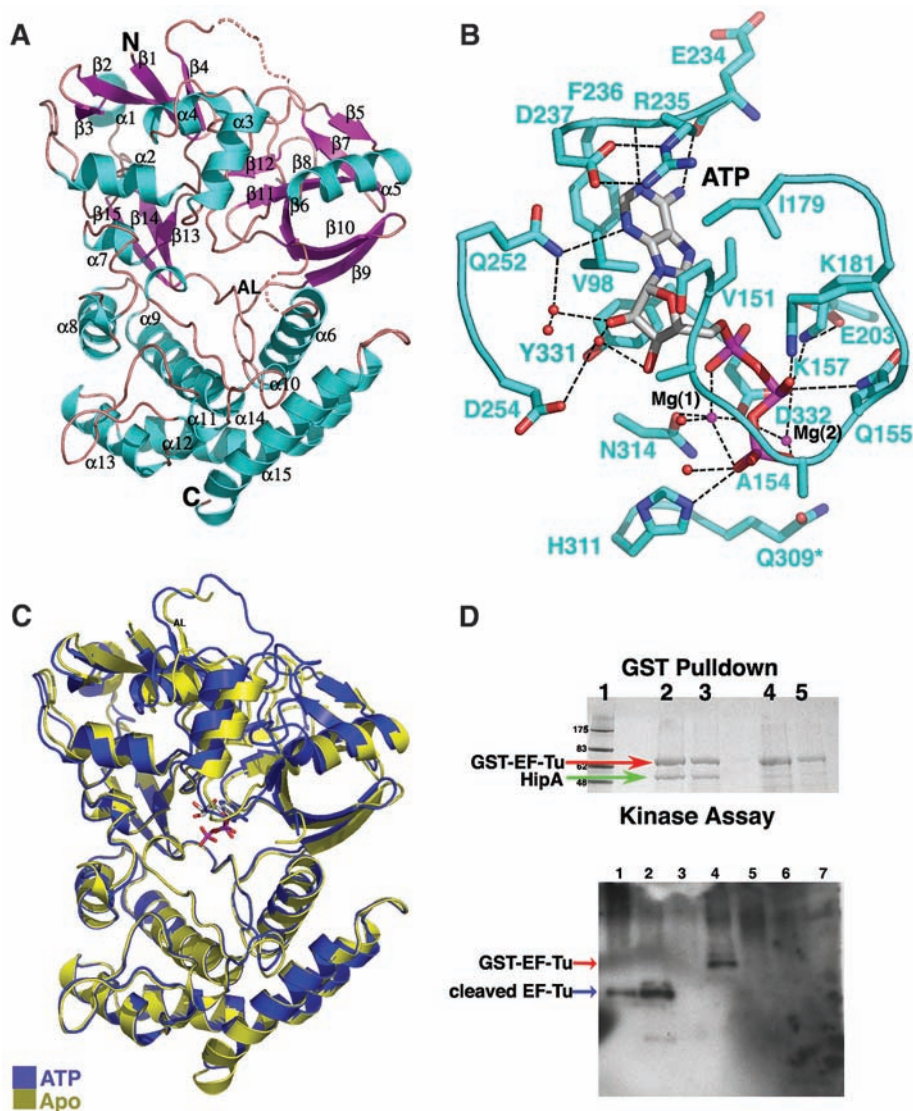


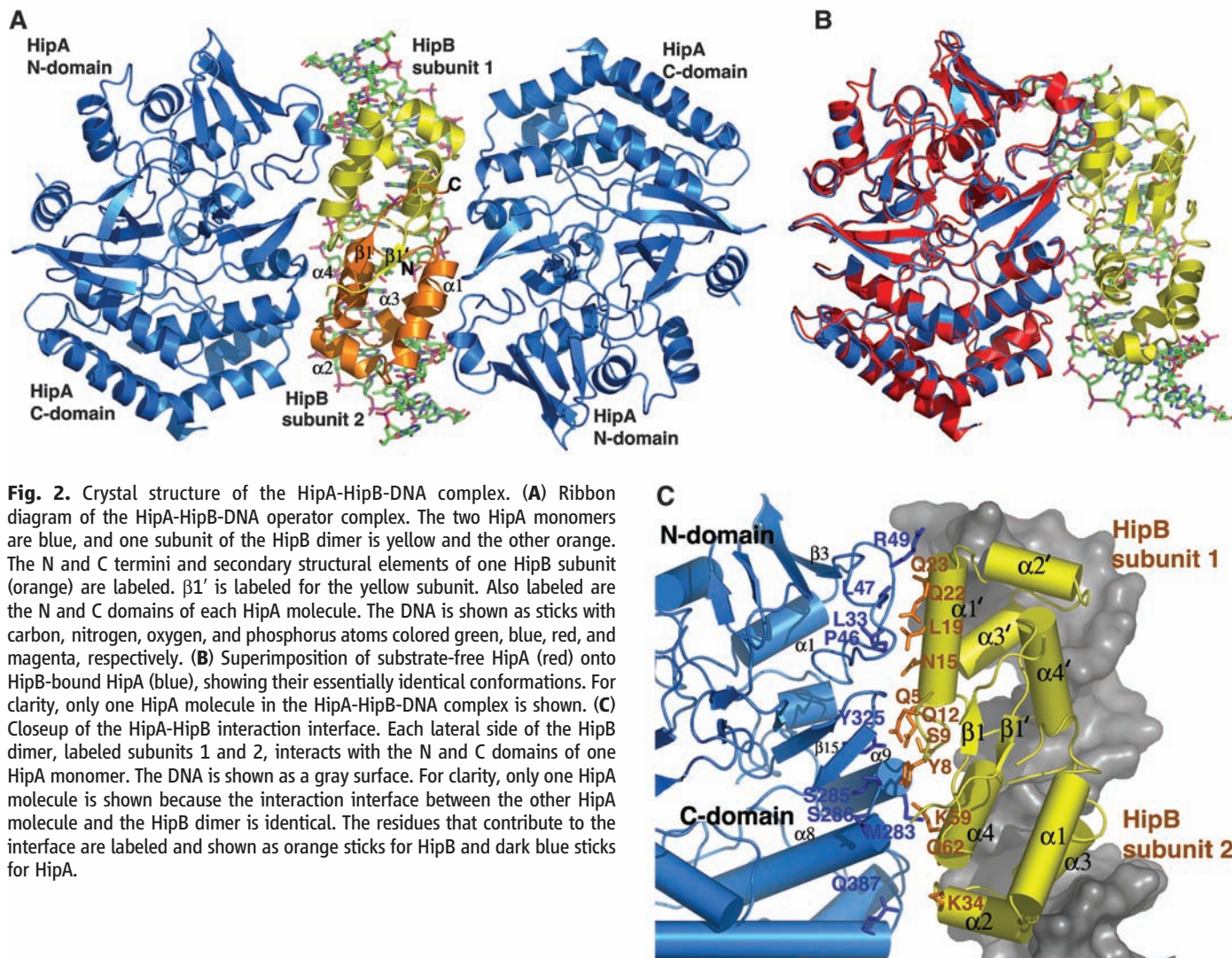
Fig. 1. HipA is a protein kinase that phosphorylates EF-Tu. (A) The *E. coli* apoHipA structure. β sheets and α helices are colored magenta and cyan, respectively. Secondary structural elements and the N and C termini are labeled. Disordered loops, including the putative activation loop (labeled AL), are indicated by dashed lines. (A) to (C) and Figs. 2, A to C; 3, B to D; and 4, A to D were made using PyMOL (31). (B) The ATP-binding pocket of HipA. Shown are ATP molecules (sticks), Mg²⁺ ions (magenta spheres), water molecules (red spheres), and hydrogen bonds (dashed black lines). (C) Superimposition of the C domains of apoHipA (yellow) and HipA-ATP-Mg²⁺ (blue) reveals only a small rotation upon ATP binding. ATP in the HipA-ATP structure is shown as sticks. (D) (Top) HipA-(GST-EF-Tu) pull-down (SDS-PAGE, stained with Coomassie brilliant blue). Lanes are as follows: 1, molecular weight ladder; 2, glutathione agarose (bead) retentate after addition of GST-EF-Tu, HipA, ATP, and GDP; 3, bead retentate after addition of GST-EF-Tu, HipA, ATP, and GDP, and washing; 4, bead retentate after addition of GST-EF-Tu, HipA, and no GDP or ATP, and washing; 5, bead retentate after addition of GST-EF-Tu, HipA, and no GDP or ATP, and washing. (Bottom) Wild-type HipA kinase assay using EF-Tu as a substrate (immunoblot). The positions of GST-EF-Tu and X₃-cleaved EF-Tu are indicated by red and blue arrows, respectively. Lanes are as follows: 1, EF-Tu (cleaved) + wild-type HipA + ATP + GTP; 2, EF-Tu (cleaved) + wild-type HipA + ATP + GDP; 3, EF-Tu (cleaved) + inactive HipA (D309Q) + ATP + GDP; 4, EF-Tu (uncleaved) + wild-type HipA + ATP + GDP; 5, EF-Tu (uncleaved) + wild-type HipA + ATP + GTP; 6, wild-type HipA + ATP; 7, native EF-Tu (cleaved) + ATP + GDP.

selectivity (Fig. 1B). Specifying contacts are provided to the adenine N6, N1, and N3 atoms by the carbonyl oxygen of Glu²³⁴, the amide nitrogen of Phe²³⁶, and the side chain Ne of Gln²⁵², respectively. The adenine ring stacks with Phe²³⁶ and Tyr³³¹, whereas Val¹⁹⁸, Val¹⁵¹, Ile¹⁷⁹, and the side-chain methylene carbons of salt-bridged residues Asp²³⁷ and Arg²³⁵ provide hydrophobic interactions. The γ phosphate hydrogen-bonds to both the side chain of His³¹¹ and the amide groups of Gly¹⁵³ and Ala¹⁵⁴, which form part of a loop analogous to the Gly loops of other protein kinases (Fig. 1B). Residues 152 to 156 of this loop are less ordered in the substrate-free HipA (apoHipA) structure, indicating that nucleotide binding is required for its stabilization. Two Mg²⁺ ions are also present in the HipA-ATP structure and probably function analogously to other protein kinases in facilitating phosphotransfer by accelerating substrate association and product dissociation (20, 21).

Comparison of the HipA-ATP and apoHipA structures revealed that binding ATP causes the N and C domains to undergo only a small rotation ($\sim 4^\circ$) relative to each other (Fig. 1C). However,

by analogy to other kinases, a more pronounced closure of the HipA domains upon binding the protein substrate is expected (20). The findings of a specialized kinase fold and high-affinity ATP binding strongly supported the hypothesis that HipA mediates persistence by phosphorylating one or more target proteins. To identify possible HipA targets, we carried out *in vitro* pulldown assays on candidate *E. coli* proteins. One protein, EF-Tu, was found to interact strongly with HipA in the presence of ATP-Mg²⁺ and guanosine diphosphate (GDP) (Fig. 1D). EF-Tu, the most abundant protein in *E. coli*, belongs to the guanosine triphosphatase superfamily and plays an essential role in translation by catalyzing aminoacyl-transfer RNA (-tRNA) binding to the ribosome (22). Upon guanosine triphosphate (GTP) hydrolysis to GDP, EF-Tu undergoes a conformational change to an open form, which cannot bind the ribosome. Previous studies showed that EF-Tu is phosphorylated on residue Thr³⁸² by an unknown kinase or kinases (23, 24). The side chain of Thr³⁸² contacts Glu¹¹⁷ to stabilize the GTP-bound closed state of EF-Tu. Phosphorylation of Thr³⁸² favors the GDP-bound open form because it would lead to

repulsion of Glu¹¹⁷ and prevent EF-Tu from adopting the GTP-bound closed conformation. Thr³⁸²-phosphorylated EF-Tu cannot bind aminoacyl-tRNA and is therefore inactive in translation (23, 24). To test whether EF-Tu is a HipA substrate, we used an *in vitro* transcription/translation system to produce the toxic wild-type HipA enzyme (fig. S3). Immunoblotting studies, using antibodies to pThr/pSer/pTyr, indicated that HipA could phosphorylate EF-Tu in a manner stimulated by GDP (Fig. 1D). Moreover, fluorescence polarization studies revealed that HipA bound the EF-Tu peptide, IREGGRTVGA (25), encompassing Thr³⁸² (shown in bold here) with a K_d of $15 \pm 5 \mu\text{M}$ (fig. S4). Subsequently, we solved a crystal structure of the HipA-(AMP-PNP)-IREGGRTVGA complex to 3.5 Å resolution. The structure revealed that the activation loop, residues 185 to 195, was now folded and density was observed for the peptide near the active site and close to the activation loop (fig. S4). These combined data suggest that HipA may phosphorylate Thr³⁸² to block aminoacyl-tRNA binding by EF-Tu. However, given that HipA affects multiple *E. coli* processes, other cellular targets are likely (9, 10).



Under normal cellular conditions, the persistence function of HipA is somehow masked by its tight interaction with HipB (11, 12). HipB also functions as a transcriptional autoregulator of the *hipBA* operon by cooperatively binding four operators with the consensus sequence TATCCN₈GGATA (where N indicates any nucleotide), located in the *hipBA* promoter region (11, 12). HipB binds these operators with high affinity, which is enhanced by the addition of HipA to the complex (12). To delineate the mechanism of HipB-mediated inhibition of HipA, the structure of the HipA-HipB complex bound to a 21-base pair *hipB* operator (top strand ACTATCCCTTAAGGGGATAG) was solved and refined to an $R_{\text{work}}/R_{\text{free}}$ of 22.5/28.1% to 2.68 Å resolution (table S1) (Figs. 2 and 3).

HipB forms a compact dimer that specifically interacts with DNA through major groove contacts, whereas two HipA molecules sandwich the HipB-DNA complex by contacting the sides of the HipB dimer (Fig. 2). HipB binds far from the HipA active sites and, unlike other TA inhibition mechanisms, does not occlude the active site. The HipB dimer interface is extensive and buries 2700 Å² of accessible surface area (ASA), which accounts for over 36% of the total dimer ASA. HipB contains one β strand and four α helices with topology α1-α2-α3-α4-β1. Helices 2 and 3 form a canonical helix-turn-helix (HTH) motif. The first 3 and last 16 residues of each HipB subunit are disordered and located near a small β sheet that is composed of β1 and β1' (from the other subunit) and forms a “β lid” (Fig. 2A). The

HipB subunit structure showed significant homology to 434 Repressor, 434 Cro, and the restriction-modification controller protein C. AhdI from *Aeromonas hydrophila*, with RMSDs of 1.56, 1.60, and 1.51 Å for 59, 56, and 59 corresponding Cα atoms, respectively, thus placing it in the Xre-HTH family of transcriptional regulators (26). The homology between HipB and these proteins is confined to the four-helix bundle region because the β lid is found only in HipB. Despite the similarities in DNA binding domains, these proteins bind their DNA sites differently because 434 Cro does not significantly distort its DNA site, and biochemical data indicate that C. AhdI bends its DNA site by 47° (27, 28). In contrast, HipB induces a large, 70° bend in its operator (Fig. 3D). This bending may play a role in the cooperative binding of HipB to its four operator sites, which is predicted to involve DNA wrapping (11, 12). Indeed, the *hipBA* promoter also contains a binding site for the architectural protein IHF, which could further aid in DNA condensation.

HipB-induced DNA distortion aligns the recognition helices for specific binding to consecutive major grooves. Contacts from the HTH motif completely specify the nucleotides of the HipB signature motif, T₂A₃T₄C₅C₆ (Fig. 3, A and B). Ser²⁹ from α2 makes hydrophobic contacts with the Thy2 methyl group. Residues of the recognition helix provide the remaining base-specifying contacts whereby two hydrogen bonds from Gln³⁹ read Ade3, whereas two hydrophobic contacts from Ala⁴⁰ and Ser⁴³ specify Thy4. Finally, Lys³⁸ makes hydrogen bonds with the guanine O6 oxygens of base pairs 5 and 6. HipB also makes 11 phosphate contacts to each half site. Deoxyribonuclease I protection studies showed that HipA binding to the HipB-DNA complex leads to an increase in protection and binding affinity (12). This is explained by the finding that HipA provides four phosphate backbone contacts to each half site from Lys³⁷⁹ and Arg³⁸² (Fig. 3, A and C).

In the HipA-HipB-DNA complex, the HipB dimer is sandwiched on each side by one HipA molecule, and the complex is formed from non-contiguous regions of both HipA and HipB (Figs. 2 and 3C). This type of interaction contrasts sharply with structures of other TA modules in which the toxin interacts with a C-terminal region of the antitoxin that typically is structured only in the presence of toxin. Specifically, for the HipA-HipB pair, the HipA N domains interact with one HipB subunit, whereas the HipA C domains interact primarily with the other HipB subunit (Fig. 2C). This interaction interface is extensive, burying ~5000 Å² of ASA, and involves both nonpolar and polar interactions. In the HipA N-domain-HipB interface, HipB residues from the turn before α1 interact with residues on HipA β15, and residues on HipB α1 make extensive contacts with HipA residues located on a long 3₁₀-like loop between β3 and α1. The formation of the HipA C-domain-HipB interface primarily involves HipB residues from α2 and the turn between α4

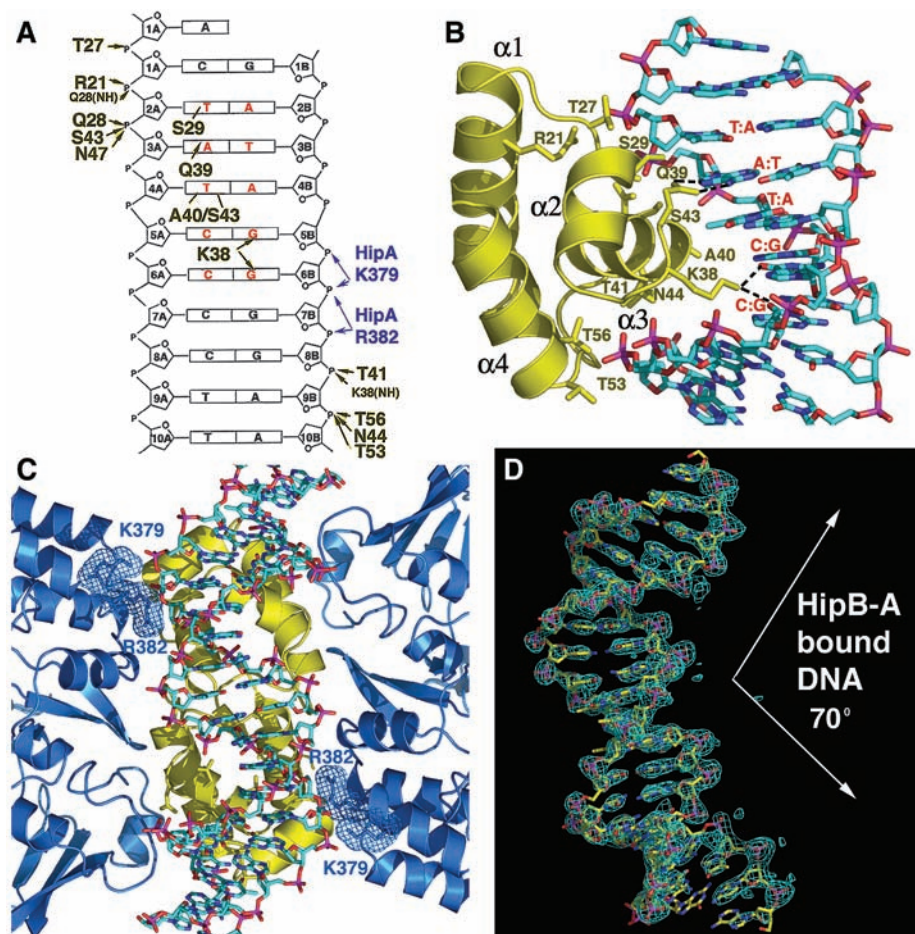


Fig. 3. HipB and HipA interactions with the *hipB* operator DNA. **(A)** Schematic representation of HipB-HipA-DNA interactions. Only one half site of the 21-oligomer duplex is shown because the identical contacts are made with each half site. The strands are labeled 1A to 10A and 1B to 10B. Bases are represented as rectangles and labeled according to sequence. The ribose groups are shown as pentagons. The operator signature motif sequence, TATCC, is red. HipB-DNA contacts are yellow. Hydrophobic contacts are indicated by lines and hydrogen bonds are indicated by arrows. Blue arrows indicate HipA-phosphate contacts. **(B)** HipB-DNA interactions. Only one HipB subunit-DNA half site is shown. The DNA and residues making side-chain contacts are shown as sticks. The signature motif sequence, TATCC, is labeled in red. For clarity, only the four-helix bundle is shown and labeled. **(C)** HipA-DNA contacts. The HipB dimer (yellow) is shown for reference. The location of the two DNA-interacting residues from HipA, K379 and R382, are shown as blue sticks and highlighted by mesh surface representations. The DNA is shown as sticks and colored as in (B). **(D)** HipA-HipB bound DNA is bent. This is an omit $F_o - F_c$ map in which the DNA was omitted from refinement to 2.68 Å resolution. The map is contoured at 2.8 σ . The DNA is shown as sticks and the bend angle is indicated.

and $\beta 1$. These residues interact primarily with HipA residues in the loop between $\alpha 8$ and $\alpha 9$ and the N terminus of $\alpha 9$. In addition, “cross-subunit” contacts are made between HipB residues Gln¹² and HipA C-domain residue Gly²⁸⁴, and between HipB residue Tyr⁸ and HipA C-domain residue Ser²⁸⁶. These cross contacts, combined with the numerous interactions of each subunit in the HipB dimer with the N and C domains of HipA, lock HipA in an open and probably inactive conformation (Fig. 2C).

To activate HipA for persistence and free it from its DNA tether, HipB must be removed or degraded. Unlike most antitoxins, HipB interacts with HipA using residues from noncontiguous, well-ordered domains and not loops. Proteases that degrade toxins typically bind and tug on disordered regions to unfold the substrate. HipB contains an exposed and flexible 16-residue C terminus attached to the small β lid that covers the hydrophobic core of the protein and would appear to be an excellent candidate for protease attack (Fig. 4A). The structure of the HipA-HipB-DNA complex also provides insight into

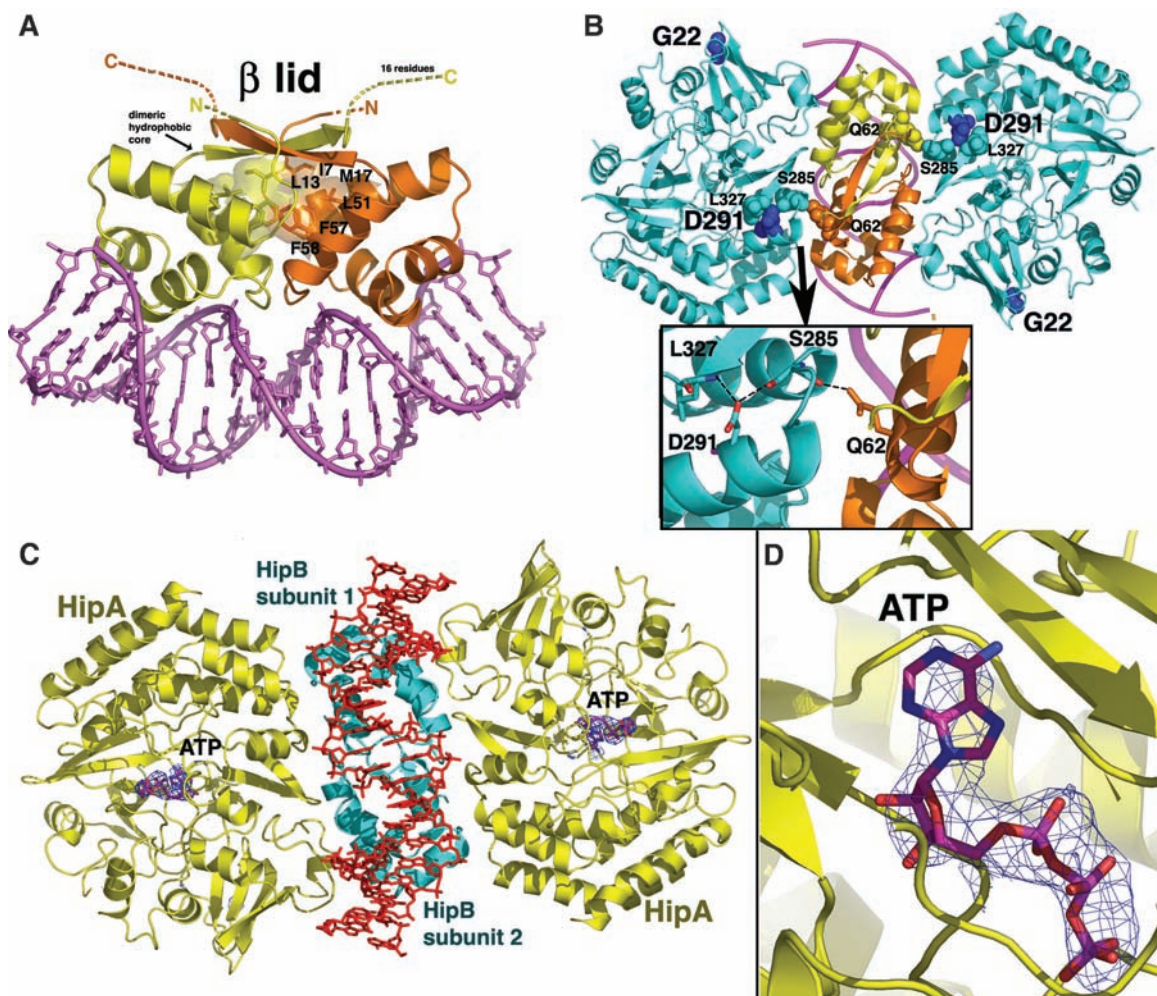
the mechanism of increased persistence of the HipA7 protein. HipA7, which contains two substitutions, G22S and D291A, confers a high-persistence phenotype on *E. coli* cells independent of HipB. Subsequent data revealed that the D291A mutation alone was sufficient for this phenotype (29). The HipA-HipB-DNA structure indicates that this phenotype probably results from a weakened HipA-HipB interaction, which unleashes HipA kinase activity. Specifically, Asp²⁹¹ makes key contacts to stabilize the HipA-HipB interface, including hydrogen bonds to the side chain of Ser²⁸⁵, which positions the Ser²⁸⁵ carbonyl oxygen to interact with HipB residue Gln⁶², and hydrogen bonds to the amide nitrogen of Leu³²⁷, which buttresses the HipA C-terminal region that interacts with HipB (Fig. 4B).

HipA undergoes only a small conformational change upon binding ATP, suggesting that HipA could bind ATP when in complex with HipB-DNA. Indeed, HipA(ATP)-HipB-DNA crystals isomorphous to HipA-HipB-DNA crystals could be grown de novo. Alternatively, ATP could be soaked into preformed HipA-HipB-DNA crys-

als. In both cases, difference $F_o - F_c$ electron density maps revealed clear density for ATP in the HipA active site (Fig. 4, C and D). In addition, isothermal titration calorimetry studies revealed a K_d of $15.0 \pm 1.0 \mu\text{M}$ for ATP-Mg²⁺ binding to HipA in the HipA-HipB-DNA complex, which is essentially identical to that obtained for ATP-Mg²⁺ binding to HipA alone (fig. S2). If the HipA active site is not blocked for ATP binding, then how does HipB binding neutralize HipA? Data from other protein kinase structures indicate that although ATP binding causes only small domain movements such as those we observed in HipA, the binding of protein substrates causes substantial domain closure (20, 21). This large-scale movement brings the two substrates into proximity for catalysis and precludes bulk solvent from the active site. HipB binding would appear to prevent this conformational change in HipA by locking the enzyme into an inactive open conformation by its extensive interactions with the HipA N and C domains. Finally, the recent finding that *E. coli* EF-Tu is localized primarily to the cytosolic and membrane fractions, which is far from the

Fig. 4. HipB is a vulnerable antitoxin that neutralizes HipA. (A)

The HipB dimer is shown as ribbons with one yellow and one orange subunit. The DNA is shown as pink sticks. Hydrophobic core residues are shown as sticks and transparent surfaces. The β lid is composed of a short two-stranded β sheet. Disordered C-terminal residues extend from the lid and are depicted as colored dashes. (B) Mutation sites (G22S and D291A) in the HipA7 protein mapped onto HipA proteins of the HipA-HipB-DNA complex. The HipA molecules are cyan, the HipB subunits are yellow and orange, and the DNA is a pink ribbon. Residues G22 and D291 are shown as blue Corey, Pauling, Koltan structures (CPKs). HipB residue Q62 and HipA residues S285 and L327 are also shown as CPKs. A closeup of the interactions involving HipA residue D291 is shown (inset). Hydrogen bonds are indicated by dashed lines. (C) Structure of the HipA(ATP)-HipB-DNA complex. HipA molecules are yellow, the HipB dimer is cyan, and the DNA and ATP are shown as sticks. The blue mesh represents a 2.98 Å resolution $F_o - F_c$ map, contoured at 4.0 σ , in



which the ATP was omitted. (D) Closeup view of the ATP-binding pocket and omit map. Carbon, nitrogen, oxygen, and phosphorus atoms are colored magenta, blue, red, and purple, respectively.

nucleoid where the HipA-HipB-DNA complex would reside, suggests that HipB-DNA binding may also inactivate HipA by its sequestration (30).

These studies have provided important insight into the mechanisms by which HipA mediates persistence and HipB neutralizes HipA. The high conservation of HipA among Gram-negative bacteria indicates its central role in the development of persistence. Thus, inhibitors that specifically target the substrate-binding sites of HipA, may prove effective against persistence and MDT.

References and Notes

1. K. Lewis, *Nat. Rev. Microbiol.* **5**, 48 (2007).
2. K. Lewis, *Curr. Top. Microbiol. Immunol.* **322**, 107 (2008).
3. N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, S. Leibler, *Science* **305**, 1622 (2004).
4. J. W. Bigger, *Lancet* **244**, 497 (1944).
5. I. Keren, D. Shah, A. Spoering, N. Kaldalu, K. Lewis, *J. Bacteriol.* **186**, 8172 (2004).
6. I. Keren, N. Kaldalu, A. Spoering, Y. Wang, K. Lewis, *FEBS Microbiol. Lett.* **230**, 13 (2004).
7. D. Shah *et al.*, *Biomed. Cent. Microbiol.* **12**, 6 (2006).
8. R. Scherrer, H. S. Moyed, *J. Bacteriol.* **170**, 3321 (1988).
9. H. S. Moyed, S. H. Broderick, *J. Bacteriol.* **166**, 399 (1986).
10. H. S. Moyed, K. P. Bertrand, *J. Bacteriol.* **155**, 768 (1983).
11. D. S. Black, B. Irwin, H. S. Moyed, *J. Bacteriol.* **176**, 4081 (1994).
12. D. S. Black, A. J. Kelly, M. J. Mardis, H. S. Moyed, *J. Bacteriol.* **173**, 5732 (1991).
13. K. Gerdes, S. K. Christensen, A. Löbner-Olesen, *Nat. Rev. Microbiol.* **3**, 371 (2005).
14. L. Buts, J. Lah, M.-H. Dao-Thi, L. Wyns, R. Loris, *Trends Biochem. Sci.* **30**, 672 (2005).
15. F. F. Correia *et al.*, *J. Bacteriol.* **188**, 8360 (2006).
16. T. C. Terwilliger, J. Berendzen, *Acta Crystallogr. D Biol. Crystallogr.* **55**, 849 (1999).
17. A. T. Brünger *et al.*, *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905 (1998).
18. T. A. Jones, J. Y. Zou, S. W. Cowan, M. Kjeldgaard, *Acta Crystallogr. A* **47**, 110 (1991).
19. R. Honda *et al.*, *EMBO J.* **24**, 452 (2005).
20. M. Huse, J. Kuriyan, *Cell* **109**, 275 (2002).
21. D. R. Knighton *et al.*, *Science* **253**, 407 (1991).
22. J. Nyborg, A. Liljas, *FEBS Lett.* **430**, 95 (1998).
23. C. Lippmann *et al.*, *J. Biol. Chem.* **268**, 601 (1993).
24. C. Alexander *et al.*, *J. Biol. Chem.* **270**, 14541 (1995).
25. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
26. R. Wintjens, M. Rooman, *J. Mol. Biol.* **262**, 294 (1996).
27. C. Wolberger, Y. C. Dong, M. Ptashne, S. C. Harrison, *Nature* **335**, 789 (1988).
28. I. Papapanagiotou, S. D. Streeter, P. D. Cary, G. G. Kneale, *Nucleic Acids Res.* **35**, 2643 (2007).
29. S. B. Korch, T. M. Hill, *J. Bacteriol.* **188**, 3826 (2006).
30. C. Archambaud, E. Gouin, J. Pizarro-Cerda, P. Cossart, O. Dussurget, *Mol. Microbiol.* **56**, 383 (2005).
31. W. L. Delano, the PyMOL Molecular Graphics System (DeLano Scientific, San Carlos, CA, 2002).
32. We thank C. R. Knudsen for her generous gift of the GST-EFTu expression construct. Coordinates and structure factor amplitudes for the apoHipA, HipA-ATP, HipA-HipB-DNA, and HipA(ATP)-HipB-DNA complexes have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (www.rcsb.org/) under accession codes 3DNU, 3DNT, 3DNV, and 3DNW, respectively. We acknowledge support from Burroughs Wellcome Career Development Award 992863 and NIH grant GM074815 (to M.A.S.), NIH grant GM061162 (to K.L.), and the Robert A. Welch Foundation (grant G0040) and NIH grant AI048593 (to R.G.B.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/323/5912/396/DC1
Materials and Methods
Figs. S1 to S4
Table S1

25 July 2008; accepted 18 November 2008
10.1126/science.1163806

Chromatin-Associated Periodicity in Genetic Variation Downstream of Transcriptional Start Sites

Shin Sasaki,^{1*} Cecilia C. Mello,² Atsuko Shimada,³ Yoichiro Nakatani,¹ Shin-ichi Hashimoto,⁴ Masako Ogawa,⁴ Kouji Matsushima,⁴ Sam Guoping Gu,² Masahiro Kasahara,¹ Budrul Ahsan,¹ Atsushi Sasaki,¹ Taro Saito,¹ Yutaka Suzuki,⁵ Sumio Sugano,⁵ Yuji Kohara,⁶ Hiroyuki Takeda,³ Andrew Fire,^{2†} Shinichi Morishita^{1,7†}

Might DNA sequence variation reflect germline genetic activity and underlying chromatin structure? We investigated this question using medaka (Japanese killifish, *Oryzias latipes*), by comparing the genomic sequences of two strains (Hd-rR and HNI) and by mapping ~37.3 million nucleosome cores from Hd-rR blastulae and 11,654 representative transcription start sites from six embryonic stages. We observed a distinctive ~200–base pair (bp) periodic pattern of genetic variation downstream of transcription start sites; the rate of insertions and deletions longer than 1 bp peaked at positions of approximately +200, +400, and +600 bp, whereas the point mutation rate showed corresponding valleys. This ~200-bp periodicity was correlated with the chromatin structure, with nucleosome occupancy minimized at positions 0, +200, +400, and +600 bp. These data exemplify the potential for genetic activity (transcription) and chromatin structure to contribute to molding the DNA sequence on an evolutionary time scale.

Mutation and repair characteristics of DNA sequence in experimental systems have been shown in a number of cases to reflect structures in chromatin. For one well-studied experimental system, ultraviolet-irradiated yeast (*Saccharomyces cerevisiae*), repair rates for a set of DNA nucleosome core regions are lower than in the surrounding linker regions (1–4). Correlations between chromatin structure and mutation rates have also been suggested in analysis of human and yeast genomes (5–7). The draft genome sequences of two inbred medaka strains, Hd-rR and HNI (8), provide a remarkable opportunity for extensive comparison

between genomic variation and structural features in the genome. The two strains are cross-fertile, yet their genomes are substantially different [~3.42% single-nucleotide polymorphism (SNP)] (8). For analysis of chromatin and transcriptional effects on genetic variation, tissue samples including totipotent (germline tissue) would be most relevant, as mutational events in the germ line would uniquely contribute to shaping the genome over evolutionary time (9–11).

To characterize transcriptional activity patterns from the medaka genome at embryonic stages, we collected 25-nucleotide (nt) 5'-end mRNA tags for 1-, 2-, 3-, 5-, 10-, and 14-day Hd-rR medaka

embryos (12). Among a total of ~38.5 million 5'-end tags collected, ~26.2 million (68.14%) were successfully aligned to unique positions in the medaka genome (fig. S1). Starting with a rough assumption that one cell contains ~300,000 mRNA molecules (13), single-copy-per-cell RNAs would be represented by ~100 of the ~26.2 million tags. To define a set of active transcription start sites (TSSs), we used a clustering algorithm yielding 11,654 ≥100-tag clusters. More than 98.4% of neighboring clusters were separated by >100 base pairs (bp) from their nearest neighbor (fig. S3B). A reference TSS for each cluster was defined as the position with the most 5'-end tags.

The substitution and indel rates within 1000 bp of the reference TSSs in the 11,654 TSS clusters tend to reach a valley at the TSSs (Fig. 1A), suggesting relative selective constraint within promoters. This is consistent with reports of high conservation around TSS regions in mammals (14). Our analysis in medaka uncovers an additional pattern: The substitution rate (blue line) showed peaks at +100 and +300 bp and

¹Department of Computational Biology, Graduate School of Frontier Sciences, the University of Tokyo, Kashiwa, 277-0882, Japan. ²Departments of Pathology and Genetics, School of Medicine, Stanford University, Stanford, CA 94305–5324, USA. ³Department of Biological Sciences, Graduate School of Science, the University of Tokyo, Tokyo, 113-0033, Japan. ⁴Department of Molecular Preventive Medicine, School of Medicine, the University of Tokyo, Tokyo, 113-0033, Japan. ⁵Department of Medical Genome Sciences, Graduate School of Frontier Sciences, the University of Tokyo, Tokyo, 108-8639, Japan. ⁶Center for Genetic Resource Information, National Institute of Genetics, Mishima, 411-8540, Japan. ⁷Bioinformatics Research and Development (BIRD), Japan Science and Technology Agency (JST), Tokyo, 102-8666, Japan.

*Present address: Mitsubishi Research Institute, Inc., Tokyo, 100-8141, Japan.

†To whom correspondence should be addressed. E-mail: afire@stanford.edu (A.F.); moris@cb.k.u-tokyo.ac.jp (S.M.)