HBD1 protein with a tandem repeat of two HMG-box domains is a DNA clip to organize chloroplast nucleoids in *Chlamydomonas reinhardtii*

Mari Takusagawa[a], Yusuke Kobayashi[a,b], Yoichiro Fukao[c,d], Kumi Hidaka[e], Masayuki Endo[e], Hiroshi Sugiyama[e,f], Takashi Hamaji[a], Yoshinobu Kato[a,g], Isamu Miyakawa[h], Osami Misumi[h], Toshiharu Shikanai[a], and Yoshiki Nishimura[a,1]

[a]Department of Botany, Graduate School of Science, Kyoto University, Kita-Shirakawa Oiwake-cho, Sakyo-ku, Kyoto 606-8502, Japan; [b]Graduate School of Science and Engineering, Ibaraki University, Bunkyo, Mito, Ibaraki 310-8512, Japan; [c]Plant Global Educational Project, Nara Institute of Science and Technology, Ikoma, Nara 630-0192, Japan; [d]Graduate School of Life Science, Ritsumeikan University, Kusatsu 525-8577, Japan; [e]Department of Chemistry, Graduate School of Science, Kyoto University, Kita-Shirakawa Oiwake-cho, Sakyo-ku, Kyoto 606-8502, Japan; [f]Institute for Integrated Cell-Material Sciences, Kyoto University, Sakyo-ku, Kyoto 606-8501, Japan; [g]Graduate School of Agricultural and Life Sciences, The University of Tokyo, Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan; and [h]Department of Biology, Graduate School of Sciences and Technology for Innovation, Yamaguchi University, Yoshida, Yamaguchi 753-8512, Japan

Edited by Alice Barkan, University of Oregon, Eugene, OR, and approved March 12, 2021 (received for review October 8, 2020)

Compaction of bulky DNA is a universal issue for all DNA-based life forms. Chloroplast nucleoids (chloroplast DNA–protein complexes) are critical for chloroplast DNA maintenance and transcription, thereby supporting photosynthesis, but their detailed structure remains enigmatic. Our proteomic analysis of chloroplast nucleoids of the green alga *Chlamydomonas reinhardtii* identified a protein (HBD1) with a tandem repeat of two DNA-binding high mobility group box (HMG-box) domains, which is structurally similar to major mitochondrial nucleoid proteins transcription factor A, mitochondrial (TFAM), and ARS binding factor 2 protein (Abf2p). Disruption of the *HBD1* gene by CRISPR-Cas9–mediated genome editing resulted in the scattering of chloroplast nucleoids. This phenotype was complemented when intact HBD1 was reintroduced, whereas a truncated HBD1 with a single HMG-box domain failed to complement the phenotype. Furthermore, ectopic expression of HBD1 in the mitochondria of yeast *Δabf2* mutant successfully complemented the defects, suggesting functional similarity between HBD1 and Abf2p. Furthermore, in vitro assays of HBD1, including the electrophoretic mobility shift assay and DNA origami/atomic force microscopy, showed that HBD1 is capable of introducing U-turns and cross-strand bridges, indicating that proteins with two HMG-box domains would function as DNA clips to compact DNA in both chloroplast and mitochondrial nucleoids.

chloroplast nucleoid | mitochondrial nucleoid | HMG-box domain

Mitochondria (mt) and chloroplasts (cp) (or plastids [pt]) are believed to have evolved from endosymbiosis of aerobic bacteria (α-proteobacteria) and oxygenic photosynthetic bacteria (cyanobacteria), respectively (1, 2). These organelles maintain their own DNAs (mt/cpDNAs) derived from their bacterial ancestors. mt/cpDNAs are not naked in vivo but are associated with diverse proteins to form compact nucleoprotein structures called mt/cp nucleoids.

In bacteria, a number of low-molecular-weight proteins called nucleoid-associated proteins (NAPs) have been reported (3–5). In *Escherichia coli*, more than 10 NAPs are known to maintain nucleoid structure/functions by regulating DNA compaction, supercoiling, and topology or by employing other functional components such as DNA/RNA polymerases, transcription factors, and ribosomes (5–7). Among those NAPs, heat unstable (HU) is one of the most abundant proteins in bacterial nucleoids (3). HU binds DNA without apparent sequence specificity and introduces U-turns into the DNA, thereby modulating the compaction of bacterial nucleoids in cooperation with other NAPs (8).

In mt nucleoids, the major components common to animals, fungi, and some protists are small basic proteins containing two tandemly repeated DNA-binding motifs—high mobility group box (HMG-box) domains (9, 10). They are related to eukaryotic chromosomal proteins or transcription factors, and they do not share any sequence or structural homology with bacterial NAPs. They are encoded in the nuclear genomes and designated as transcription factor A, mitochondrial (TFAM) in animals (8, 11), ARS binding factor 2 protein (Abf2p) in yeast *Saccharomyces cerevisiae* (12, 13), and agglomeration of mt chromosome (Glom) in true slime mold *Physarum polycephalum* (14). TFAM is conserved in animals that have a coelom (Coelomata) and in the nematode *Caenorhabditis elegans* (15), and the crystal structure of human TFAM revealed that it causes formation of a U-turn in mtDNA by bending DNA double helices with its two HMG-boxes (15, 16), just as HU does in bacterial nucleoids (8). TFAM and its homologs play a central role in the gene expression, maintenance,

Significance

Compaction of bulky DNA is a universal issue for all DNA-based life forms. Both chloroplasts and mitochondria maintain their own multicopy genomes organized as nucleoids, but the mechanism of DNA compaction remains obscure. Here, we discovered a chloroplast nucleoid protein (HBD1) that is highly similar to major mitochondrial nucleoid proteins transcription factor A, mitochondrial (TFAM), and ARS binding factor 2 protein (Abf2p) in terms of possessing two DNA-binding high mobility group box (HMG-box) domains. Our analyses of HBD1 based on DNA origami/atomic force microscopy showed that HBD1 is capable of compacting DNA by introducing U-turns and cross-strand bridges with the two HMG-box domains, indicating that proteins with two HMG-box domains could compact DNA in both mitochondrial and chloroplast nucleoids.

Author contributions: M.T., O.M., and Y.N. designed research; M.T., Y. Kobayashi, Y.F., K.H., M.E., H.S., T.H., I.M., O.M., and Y.N. performed research; M.T., Y. Kobayashi, Y.F., K.H., M.E., H.S., Y. Kato, I.M., O.M., T.S., and Y.N. analyzed data; and M.T., I.M., T.S., and Y.N. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: yoshiki@pmg.bot.kyoto-u.ac.jp.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2021053118/-/DCSupplemental.

Published May 11, 2021.

and organization of the mt genome and therefore are regarded as "mitochondrial histones."

In cp nucleoids, proteins with HMG-box domains have not been reported. Instead, an HU homolog that is closely related to the bacterial nucleoid protein is known to participate in the organization of cp nucleoids in red algae, some red algae endosymbionts in other algae (17–20), and the green alga *Chlamydomonas reinhardtii*—histone-like protein (HLP) (21). In land plants, however, an HU homolog has not been found. To understand the composition of cp nucleoids, a number of biochemical and proteomic studies have been performed (22–25) and have revealed numerous and diverse nuclear-encoded proteins that associate with plastid nucleoids (pt nucleoids), but their detailed functions are, in most cases, not known. The diversity of the protein composition of cp nucleoids could be explained as being a result of dynamic alteration of cp nucleoids during plant evolution, in which ancestral prokaryotic components were recurrently lost or replaced by eukaryotic components (26). However, the comprehensive picture of the molecular organization of cp nucleoids still remains a mystery, and curiously, no similarity in the molecular compositions between mt and cp nucleoids has been found so far.

In this study, we performed a mass spectrometry-based proteomic analysis of cp nucleoids isolated from *C. reinhardtii* and identified a protein that is structurally highly similar to major mt nucleoid proteins TFAM and Abf2p. The protein possessed two HMG-box domains and was designated as HBD1 (*C. reinhardtii* HMG-box domain protein 1). Investigation of the localization, molecular function, and phylogenetic background of HBD1 revealed the advantage of having two HMG-box domains for facilitating the compaction of cp and mt nucleoids.

Results

Protein Components of cp Nucleoids of *Chlamydomonas*. To identify components of cp nucleoids, we first isolated cp nucleoids that retained their morphological compactness from *Chlamydomonas* cells. Proteomic analysis of the cp nucleoid fraction revealed potential components of cp nucleoids, as shown in *SI Appendix*, Table S1. Previously reported cp nucleoid proteins such as HLP (21) and CNS (26) were found in the list, validating the reliability of our proteomic analysis. Among the numerous proteins on the list, we noticed a small basic protein (~22 kDa) with two HMG-boxes (Fig. 1*A*). Although the protein score (=77) was not high, the organization of HBD1 was highly similar to that of major mt nucleoid proteins such as TFAM, Abf2p, and Glom. It is known that the composition of structural proteins is completely different between cp and mt nucleoids, and, as far as we were aware, no homologous proteins had been found between these two organellar nucleoids (26). Therefore, the structural similarity between HBD1 and the mt nucleoid proteins TFAM, Abf2p, and Glom was striking. As the next step, we set out to investigate the precise localization and function of HBD1.

HBD1 Localizes in cp Nucleoids. A database search (Phytozome v.12) revealed that HBD1 is encoded on chromosome 16 (Cre16.g672300), and in silico prediction programs for intracellular protein localization TargetP-2.0 (mt: 0.9963, cp: 0.0036) (27) and iPSORT (mt or cp: yes, mt: yes) (28) both predicted that this protein would be targeted to mt, while WoLF PSORT (29) predicted cp targeting (cp: 8, nuclear: 3, mt: 3). To test whether the HBD1 protein localizes to cp nucleoids, YFP-fusion protein was expressed in *Chlamydomonas* cells. Consequently, the HBD1:YFP signals were detected exclusively overlapping cp nucleoids, and no signals were detected in either mt nucleoids or the cell nucleus (Fig. 1*B*). This result showed that HBD1 is a cp nucleoid protein that possesses two HMG-box domains, similar to major mt nucleoid proteins.

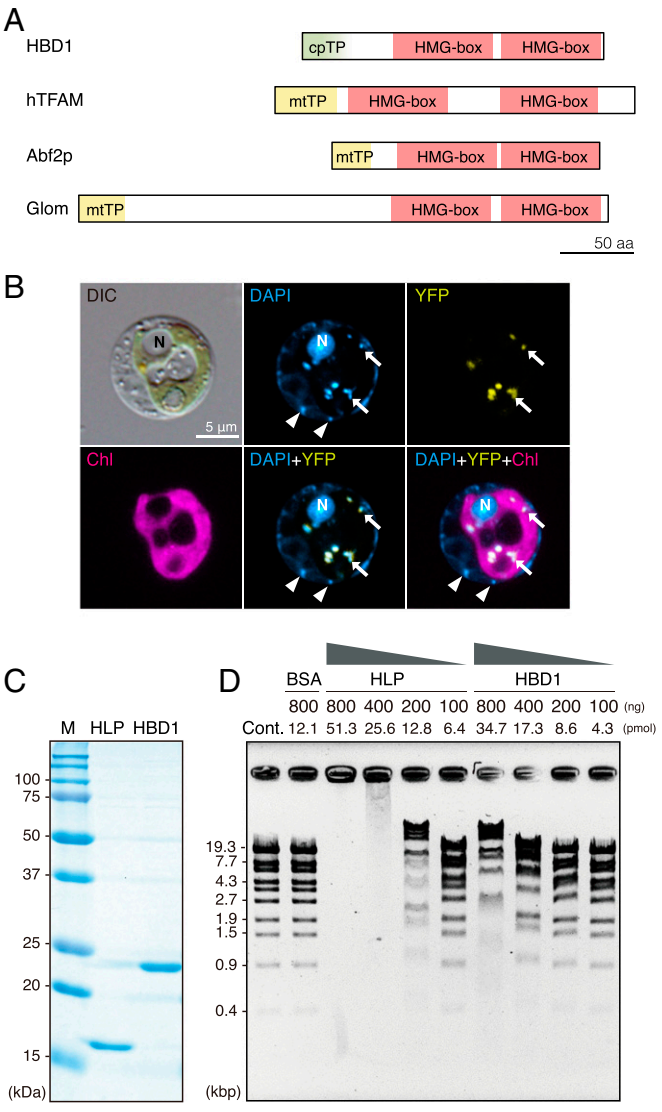

Fig. 1. HBD1 is a cp nucleoid protein that directly binds DNA. (*A*) Schematic diagrams of domains of HBD1 and major mt HMG proteins: HBD1 (*C. reinhardtii*: Uniplot ID A8J3F0), hTFAM (*Homo sapiens*: Q00059), Abf2p (*Saccharomyces cerevisiae*: Q02486), and Glom (true slime mold *Physarum polycephalum*: Q8T114). All of them have two HMG-box domains separated by a short linker region (predicted by HMMER software with Pfam database). (*B*) A cell wall–less *C. reinhardtii* cell (CW15) expressing HBD1:YFP was observed by differential interference contrast microscopy (DIC) and fluorescence microscopy (DAPI, YFP and Chl). The nucleus (N), cp nucleoids (arrows), and mt nucleoids (arrowheads) were visualized by DAPI staining (DAPI: cyan). HBD1:YFP signal (arrows) are shown in yellow (YFP). The magenta represents chlorophyll autofluorescence (Chl). DAPI+YFP+Chl is a merged image of DAPI, YFP, and Chl signals. HBD1:YFP signals exclusively overlapped with DAPI-stained cp nucleoid signals. (*C*) Coomassie Brilliant Blue-stained gel after SDS-PAGE of recombinant HLP and HBD1 proteins. (*D*) EMSA of recombinant HLP and HBD1 proteins. *Sty*I-digested λ phage DNA (400 ng) was incubated with buffer (Cont.), 800 ng BSA (BSA), or various amounts (800, 400, 200, 100 ng) of HLP or HBD1 protein.

In the nuclear genome of *C. reinhardtii*, we found two additional genes encoding proteins with two HMG-box domains: HBD2 (Cre17.g702650) and HBD3 (Cre11.g481050) (*SI Appendix*, Fig. S1). According to the TargetP-2.0 program, HBD2 was predicted to localize in mt (mt: 0.6538, cp: 0.0796), and HBD3 would be targeted to other cellular compartments (Other: 0.8167). To study their intracellular localization, we tried to express these proteins fused with YFP, but all our attempts were

unsuccessful. As an alternative approach, their localizations were investigated by the transient expression system using the leaf epidermal cells of *Nicotiana benthamiana* (*SI Appendix*, Fig. S1). HBD1:GFP appeared as a punctuate particle in the cp, consistent with the localization of HBD1 in *C. reinhardtii* (Fig. 1*B*). HBD2 was mostly targeted to mt and was also targeted to cp. The HBD3 signal was seen in the nucleus (*SI Appendix*, Fig. S1). The expression profiles of these genes during cell division were also analyzed using previously reported RNA-sequencing data (30) (*SI Appendix*, Fig. S2). The expression level of *HBD1* was up-regulated prior to cell/cp division, which was triggered by the light to dark transition. The expression level of *HBD1* reached a plateau between 0 and 1 h after the start of the dark period and then slowly decayed. The maximum expression level and decay curve were similar to those of *HLP*, which is the major component of cp nucleoids in *C. reinhardtii* (21, 26). The expression levels of *HBD2* and *HBD3* were much lower than that of *HBD1* (~15% and ~1% of *HBD1*, respectively). These analyses suggested that HBD1 is a major 2×HMG-box protein exclusively targeted to cp. HBD2 may be targeted mt or dual-targeted to mt and cp, and its expression level was much lower compared to that of HBD1. HBD3 probably functions in the cell nucleus.

HBD1 Is a DNA-Binding Protein. To test whether HBD1 is a DNA-binding protein, we expressed the protein with N-terminally fused 6×His-tag in *E. coli* (Fig. 1*C*). The HBD1 recombinant protein was purified and extensively washed using an Ni-NTA column. As a positive control, a recombinant HLP protein was prepared. The recombinant proteins were assessed for their DNA-binding activity by electrophoresis mobility shift assay (EMSA) using XbaI-digested λ phage DNA fragments as substrates (Fig. 1*D*). The addition of increasing concentrations of HBD1 resulted in slower migration of DNA-protein complexes. This analysis was repeated by using cpDNA fragments for 12 regions of 7 genes (promoter, 5′ UTR, intron or exon regions for *rbcL*, *petA*, *atpB*, *psbA*, *rpl16,* 23S rRNA, and tRNA; *trnS*). Sense and antisense single-strand DNAs (ssDNAs) were prepared and annealed to form double-strand DNA (dsDNA) fragments. GC content for each region was calculated and varied from 10 to 54%. dsDNA and ssDNA (sense) fragments were analyzed. Eventually, all fragments shifted in a similar manner, validating that HBD1 can bind both dsDNA and ssDNA without obvious sequence specificity (*SI Appendix*, Fig. S3).

Analysis of CRISPR/Cas9-Mediated HBD1 Knockout Mutants. To assess the function of HBD1 in living *C. reinhardtii* cells, strains lacking the *HBD1* gene were prepared by the CRISPR/Cas9-mediated genome editing technique (*SI Appendix*, Fig. S4) (31). This was necessary because no mutants for this gene were found in the *Chlamydomonas* Library Project lines (32). Using the CRISPR/Cas9-mediated system, we obtained three independent lines in which the *HBD1* gene was disrupted by the insertion of the marker gene cassette (*SI Appendix*, Fig. S3*B*). *HBD1* mRNA levels were almost undetectable in these lines by RT-qPCR (*SI Appendix*, Fig. S4*C*). We asked whether the cp nucleoid morphology was different between wild-type (WT) and *hbd1* mutants viewed under a fluorescence microscope. In *hbd1*, the globular cp nucleoids were disintegrated and scattered throughout cp, and the sizes of cp nucleoids were smaller (~0.26 μm^2) than those in WT (~0.5 μm^2) (Fig. 2). Then we expressed HBD1 with one or two HMG-boxes (native HBD1 and truncated HBD1 retaining the N-terminal single HMG-box) in *hdb1*. When HBD1:YFP (2×HMG-box) was expressed, the size of cp nucleoids recovered to almost the WT level (~0.47 μm^2). However, the phenotype could not be complemented by expressing the truncated version of HBD1:YFP (1×HMG) (Fig. 2). The size of cp nucleoids in *hbd1* expressing 1×HMG (~0.24 μm^2) was almost the same as that in *hbd1*, suggesting that two HMG-box domains are required for

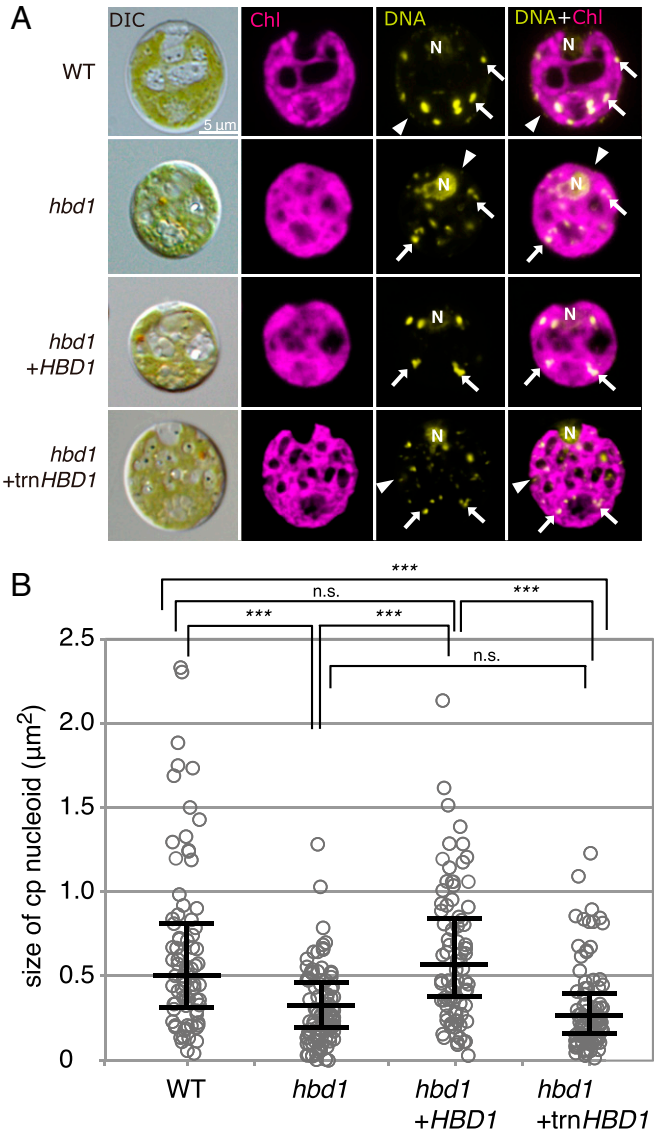

Fig. 2. Disruption of *HBD1* by CRISPR-Cas9–mediated genome editing and complementation tests. (*A*) Differential interference contrast (DIC) and fluorescence microscopic images of WT (CC-125), *hbd1* (knockout), *hbd1*/HBD1:YFP (*hbd1* cells transformed to express HBD1 [2×HMG-box] gene), and *hbd1*/truncated HBD1:YFP (*hbd1* cells expressing truncated HBD1 with 1×HMG-box [trnHBD1]) grown on TAP agar plates for 5 d. The magenta color shows chlorophyll autofluorescence (Chl). The nuclear (N), cp (arrows), and mt nucleoids (arrowheads) were visualized by SYBR Green I staining (DNA). Merged image of SYBR Green I signal and Chl (DNA+Chl). (Scale bar, 5 μm.) (*B*) Size distribution of cp nucleoids of WT, *hbd1*, *hbd1*/HBD1:YFP (2×HMG-box), and *hbd1*/truncated HBD1:YFP (1×HMG-box) analyzed by ImageJ. One-way ANOVA tests were performed with R statistical package version 3.5.0 (59) with post hoc test (Tukey's honestly significant difference), ***$P < 0.01$.

the compaction of cp nucleoids. Meanwhile, the expression level of HBD1 did not significantly affect the copy number of cpDNA (*SI Appendix*, Fig. S5), cell growth (*SI Appendix*, Fig. S6), or the accumulation levels of cp gene transcripts (*SI Appendix*, Fig. S7), unlike the effect of Abf2p, whose accumulation level is a critical determinant of mtDNA copy number in *S. cerevisiae* (33).

HBD1 Can Complement the Δabf2 Phenotype in *S. cerevisiae*. To study the function of HBD1, we examined whether HBD1 could complement a *Δabf2* mutation in *S. cerevisiae*. Abf2p is a 2×HMG-box

PLANT BIOLOGY

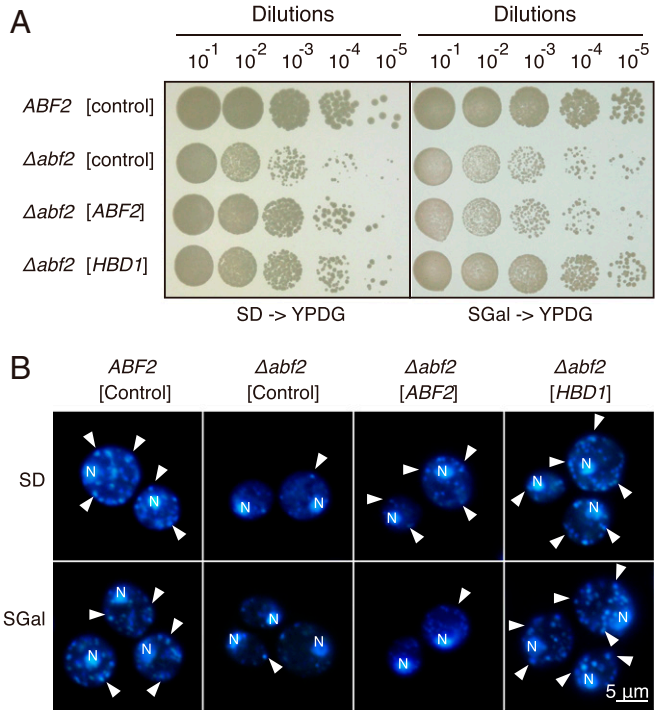

Fig. 3. HBD1 complements *Δabf2* mutation in *S. cerevisiae*. WT or *Δabf2* cells of *S. cereviciae* were transformed with the plasmid pYES2/CT (vector control), pYES2/CT-*ABF2*, or pYES2/CT-*HBD1*. (*A*) The transformants were incubated at 30 °C for 2 d in synthetic defined (SD) or SGal liquid medium, and then serial dilutions (10^{-1} to 10^{-5}) of cell suspensions (OD_{600} = 2) were spotted on YPDG plates (SD to YPDG or SGal to YPDG) and cultivated at 30 °C for 4 d. (*B*) The cells were cultivated at 30 °C for 2 d in SD or SGal liquid medium, fixed, DAPI stained, and observed by fluorescence microscopy.

protein that functions as the structural core of mt nucleoids of *S. cerevisiae*. Yeast cells lacking functional Abf2p (*Δabf2*) are able to maintain mtDNA as long as they are grown on media with nonfermentable carbon sources (such as glycerol), but they quickly lose mtDNA in medium containing fermentable carbon (such as glucose or galactose) (12). The growth of *Δabf2* cells that have lost mtDNA is severely inhibited when they are transferred to medium with nonfermentable carbon sources. We expressed HBD1 fused with the mt transit sequence of Abf2p under the control of galactose-inducible GAL1 promoter. The strains were cultured for 48 h in liquid media containing fermentable carbon sources (glucose [SD] or galactose [SGal]) and then inoculated onto YPDG agar plates containing nonfermentable carbon source medium (SD → YPDG and SGal → YPDG, respectively, in Fig. 3*A*). The GAL1 promoter drives the expression of downstream genes at high level in the presence of SGal and at low/leaky level in the presence of SD.

In this assay, the WT cells with the empty vector were able to grow on YPDG plates, whereas *Δabf2* cells showed the petite colony phenotype (Fig. 3*A*). In the *Δabf2* cells, which formed petite colonies, mt nucleoids were small and diffusely stained by DAPI, indicating the disintegration of mt nucleoids and loss of mtDNA in both SD and SGal conditions (Fig. 3*B*). These results verify the critical importance of Abf2p in compacting and stabilizing mt nucleoids. The phenotype of *Δabf2* was complemented when intact *ABF2* was introduced (SD), but overexpression of Abf2p appeared to be deleterious because of side effects on mtDNA replication/recombination or transcription (SGal), consistent with a previous report (33). On the other hand, when HBD1 was expressed in mt of *Δabf2*, the respiratory-competent cells clearly increased in both SD → YPDG and SGal → YPDG conditions. In *Δabf2*-expressing HBD1, mt nucleoids were observed at almost the same level as in WT (Fig. 3*B*). These results suggest that HBD1 is capable of compacting and stabilizing mt nucleoids, substituting for the function of Abf2p in *S. cerevisiae*.

Possible Advantage of a Tandem Repeat of Two HMG-Box Domains for Compacting DNA. What could be the advantages of two tandemly repeated HMG-box domains in compacting DNA in both mt and cp? To address this question, recombinant proteins with one, two, or three HMG-box domains were prepared (Fig. 4*A*) and analyzed by EMSA (Fig. 4*B*). To compare the DNA-binding affinity of these proteins, a dsDNA fragment was used as a probe, and the band intensity of free DNA probe was quantified to calculate the half maximal (50%) effective concentration (EC_{50}) value (Fig. 4*C*). EC_{50} values of HMG proteins with two or three domains were lower than that with one domain, indicating that the protein with two or three HMG-box domains had much higher affinity for dsDNA compared to the one with one HMG-box domain, whereas the difference of EC_{50} between the proteins with two and three HMG-box domains was not statistically significant. These results suggest that having two HMG-box domains is advantageous for increasing DNA-binding affinity.

The crystal structures of TFAM and Abf2p revealed that these proteins bind DNA to form a U-turn with the proteins' two HMG-box domains (16, 34). Considering the structural similarity between those mt 2×HMG-box proteins and HBD1, HBD1 may also bind DNA to induce a U-turn with its two HMG-box domains. To study whether HBD1 could introduce U-turns in dsDNA, we employed DNA origami and high-speed atomic force microscopy (AFM) (35, 36). To detect dsDNA U-turns, we employed a DNA origami substrate with two parallel dsDNA strands in the center of each frame (Fig. 4*D*). When purified

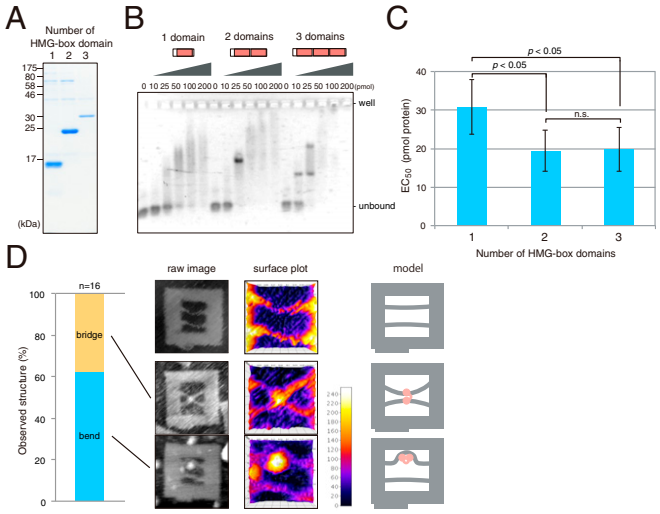

Fig. 4. Tandem repeat of two HMG-box domains is critical for increasing DNA-binding affinity and compacting DNA strands by bending and cross-strand bridges. (*A*) Coomassie Brilliant Blue-stained SDS-PAGE image showing recombinant proteins with one (1), two (2), or three (3) HMG-box domains. (*B*) The DNA-binding affinity of the recombinant proteins was analyzed by EMSA using 50 bp dsDNA as substrate. (*C*) Quantitative comparison of DNA-binding affinity between recombinant proteins with one, two, or three HMG-box domains. Band intensity of free DNA was quantified, and the EC_{50} value was calculated from *B* instead of dissociation constant (K_D). Values are the means and SDs of five independent experiments. (n = 5, Tukey–Kramer test). (*D*) Recombinant proteins with two HMG-box domains were mixed with DNA origami frames with two parallel dsDNA strands and observed by high-speed AFM. Representative AFM images showing the original DNA origami substrate (*Top*), DNA cross-strand bridge (*Middle*), and bending (*Bottom*). Raw image (*Left*), surface plot (*Middle*), and schematic model (*Right*) to show the binding of HBD1 protein (pink).

recombinant HBD1 protein was added to the DNA origami substrate, the microscopy revealed that HBD1 protein bound to the dsDNA and induced a U-turn (Fig. 4*D* and *SI Appendix*, Fig. S8). Interestingly, X-shaped structures were also observed (Fig. 4*D* and *SI Appendix*, Fig. S8), suggesting that the 2×HMG-box protein is capable of forming an interstrand bridge between two DNA strands. These results indicated that, in addition to improving the DNA-binding affinity, duplication of HMG-box domains could be important for compacting long, bulky DNA molecules into cp nucleoids by intrastrand binding (bending, U-turns) and cross-strand bridges. In other words, HBD1 may function as a DNA clip, controlling the level of DNA compaction by bending and bridging DNA strands with its two HMG-box domains (*SI Appendix*, Fig. S9).

Phylogenetic Analysis of HBD1. HBD1 is highly similar to major mt nucleoid proteins Abf2p and TFAM regarding several points including their domain structure, localization, and function. They all possess two HMG-box domains (Fig. 1*A*), are targeted to DNA-containing organelles (mt or cp) (Fig. 1*B*), and function as DNA-binding proteins (Fig. 1). HBD1 compacted cp nucleoids (Fig. 2) and was able to functionally substitute for Abf2p in mt of *S. cerevisiae* (Fig. 3). Moreover, AFM analysis showed that HBD1 is capable of introducing U-turns and cross-strand bridges to DNA strands (Fig. 4), similarly to TFAM (37). Therefore, we next analyzed the phylogenetic relationship between these HMG-box proteins.

The HMG-box domain is widely conserved among green plants, but the copy number of genes encoding HMG-box proteins varies depending on the species (38, 39). We searched for proteins with multiple HMG-box domains by hmmscan search (*SI Appendix*, Fig. S10). Proteins with two HMG domains were found in green algae including *Chlamydomonas*, *Volvox*, *Micromonas*, *Osteococcus*, etc., and it is possible that these proteins would function as DNA-compaction factors in cp or mt in a manner similar to HBD1. However, in other plants, proteins with two HMG-box domains were not detected. Instead, proteins with three HMG-box domains were identified, some of which may function in organelles.

To obtain insight into the possible origin and evolutionary scenario of 2×HMG-box proteins, phylogenetic analysis was performed, but the root of the phylogenetic tree was complicated because of the poor conservation of amino acid sequences even within the HMG-box domain. Curiously, when the two (N-terminal and C-terminal) HMG-box domains were separately analyzed (*SI Appendix*, Fig. S11), HMG-boxes of Abf2p of yeast formed independent branches consisting of its N- and C-terminal HMG-box domains. This result suggests that the extant 2×HMG-box proteins might have been acquired by independent duplications of HMG-box domains as a result of convergent evolution.

Discussion

HMG-box is a DNA-binding domain (40) that was originally discovered in nuclear chromatin–associated proteins such as transcription factors (41, 42). Later, proteins with two HMG-box domains were identified as major components of mt nucleoids in animals and fungi. Those proteins, TFAM and Abf2p, have been shown to play a central role in the organization and maintenance of the mt nucleoids and also in the transcriptional regulation of mt genes (43). On the other hand, there have been no reports suggesting the occurrence of 2×HMG-box proteins in cp (or mt) of green plants.

In this report, we identified and characterized the HBD1 protein with two HMG-box domains by proteomic analysis of purified cp nucleoids. The domain structure of HBD1 was highly reminiscent of major mt nucleoid proteins TFAM and Abf2p. Microscopic analysis indicated that HBD1:YFP localized exclusively in cp nucleoids. The recombinant HBD1 protein showed sequence-independent DNA-binding activity (Fig. 1). Disruption of HBD1 resulted in partial disintegration of cp nucleoids, which was abrogated by the expression of HBD1 (2×HMG), but not by the expression of truncated HBD1 with only one HMG-box domain (Fig. 2). Furthermore, ectopic expression of HBD1 in the *Δabf2* mutant of *S. cerevisiae* functionally rescued the defects (Fig. 3), suggesting that HBD1 is not only structurally but also functionally similar to Abf2p/TFAM. The results of EMSA, DNA origami, and high-speed AFM analyses suggested that two HMG-box domains would be advantageous for dsDNA compaction because of the higher dsDNA affinity and the ability to introduce bending (U-turn) and cross-strand bridges to dsDNA strands (Fig. 4). These data indicate that HBD1 with its two HMG-box domains, which is structurally and functionally similar to major mt nucleoid proteins, plays a critical role in compacting cp nucleoids.

Several mt or cp nucleoid proteins have been shown to introduce U-turns to compact DNA molecules, based on structural analyses, AFM, or electron microscopic observations (10, 15, 16, 44). Other than HMG-box proteins, bacterial HU is one of the proteins capable of introducing U-turns to dsDNA (8), so it is likely that HLP (it is structurally a HU-like protein) (21), which is one of the major components of cp nucleoids in *C. reinhardtii*, can also introduce U-turns in DNA.

As for cross-strand bridges, on the other hand, proteins with HMG-box domains could be unique. TFAM has been shown to form cross-strand bridges either by using the two HMG-box domains (37) or by dimerization (15). Our AFM analysis showed that HBD1 is also capable of forming cross-strand bridges (Fig. 4). This function would be important especially for cp nucleoids because it has been suggested that ∼10 copies of cp genomes are organized into one globular cp nucleoid in *C. reinhardtii* (45) [∼4 copies in land plants (46)], whereas mt nucleoids have been shown to contain only one copy of the mt genome (47). During cp division, the globular cp nucleoids disintegrate and become scattered into a network-like structure consisting of minute speckles, probably because of the dissociation of Holliday junctions and relaxation of DNA supercoils by MOC1 (Holliday junction resolvase), which disentangles cp nucleoids into minute speckles (42). After the completion of cp division, the minute speckles quickly merge to reform globular cp nucleoids (48). To achieve efficient reformation of globular cp nucleoids, HBD1 would serve as a core structure by forming cross-strand bridges to tether multiple copies of cp genomes. This would explain why cp nucleoids were rather scattered when *HBD1* was disrupted or truncated to lose a single HMG-box domain (Fig. 4). Based on these results, we propose a model in which HBD1 with two HMG-box domains functions as a DNA clip to form U-turns in DNA and cross-strand bridges between DNAs to form globular cp nucleoids (*SI Appendix*, Fig. S9).

HBD1, TFAM, and Abf2p have several features in common. They possess two HMG-box domains, localize in DNA-containing organelles, and function to organize nucleoids. Therefore, it is natural to ask whether they derived from the same evolutional origin or evolved convergently. This question is difficult to answer because of the rapid evolutionary rate of HMG-box domains (49). However, our phylogenetic analysis suggested that it is more likely that these genes might have occurred as a result of independent duplications of HMG-box domains (*SI Appendix*, Fig. S11). Our experiments suggested that the tandem repeat of two HMG-box domains is advantageous for enhancing cpDNA compaction (Fig. 4), and we speculate that this advantage might have been the driving force for the convergent evolution of 2×HMG-box domains as a compaction factor for both mt and cp nucleoids.

There are functional differences between HBD1 and TFAM/Abf2p, which would be reasonable if they have evolved independently. TFAM/Abf2p have multiple functions other than DNA compaction, such as regulation of transcription, replication, and homologous recombination (10). HBD1, in contrast, would not be multifunctional. Overexpression or loss of HBD1 did not change the cpDNA copy number (*SI Appendix*, Fig. S5), cell growth (*SI Appendix*, Fig. S6), or cp gene transcriptions (*SI Appendix*, Fig. S7),

unlike Abf2p and TFAM. Also, as shown in Fig. 3, both low-level and overexpression of HBD1 complemented the phenotype of yeast *Δabf2* mutant and did not inhibit the growth, in contrast to native Abf2p, which complemented the petite colony phenotype only when Abf2p was expressed at low level (in the absence of nonfermentable carbon sources). Overexpression of Abf2p inhibited the growth, consistent with the previous observation (33). It has been postulated that the growth inhibition might have been caused by an erroneous action of Abf2p in function(s) other than DNA compaction, such as homologous recombination or transcription (33). In fact, the petite colony phenotype of *Δabf2* could be rescued by the expression of other unrelated DNA-binding proteins such as bacterial HU (50). Likewise, the overexpression of HBD1 did not cause growth inhibition in yeast (Fig. 3). Taken together, it is likely that HBD1 is a protein specialized for DNA compaction and may not have additional functions in transcription or homologous recombination processes (Fig. 2*A*). However, more careful investigations would be required to investigate possible functions of HBD1, such as by analyzing the binding affinity for DNA substrates with various structures.

How common are 2×HMG box proteins as a component of cp nucleoids? Our database analysis indicated that 2×HMG-box proteins are ubiquitously found in green algae (*Chlamydomonas*, *Volvox*, *Cocoomyxa*, *Micromonas*, *Osterococcus*, etc.), and it is likely that these proteins would serve as components of cp or mt nucleoids (*SI Appendix*, Fig. S10). Otherwise, 2×HMG proteins are rather rare in green plants. Although there are many 3×HMG-box proteins, which could potentially function as cp or mt nucleoid compaction factors (*SI Appendix*, Fig. S6), in silico intracellular localization analysis predicted nuclear or cytoplasmic localization for most of the 3×HMG proteins (*SI Appendix*, Table S2). Indeed, 3×HMG proteins analyzed so far have been shown to function in the cell nucleus and, interestingly, bind to condensed mitotic chromosomes, possibly serving as condensation factors (51). It would be still interesting to pursue for HMG proteins that function as cp or mt nucleoid compaction factors in land plants.

Given that deletion of HBD1 does not affect cell-division rate, cpDNA copy number, or cp mRNA abundance, what could be the biological significance of forming globular cp nucleoids? At present, there is still no clear answer to this important question. More precise analyses such as those on the response of cp gene expression, cpDNA replication activity to environmental changes, cell cycle, or reproduction could be essential to clarify this point. One possible speculation might be that the surface of globular cp nucleoids may serve as a physical platform for numerous proteins with diverse functions to work together to achieve efficient cpDNA replication or transcription, although it might have only limited impact on the static accumulation level of cpDNA or cp mRNAs. Further analyses would shed light on still-unknown functions of cp nucleoids in plants.

Materials and Methods

Cell Culture. *C. reinhardtii* cells were usually grown on a Tris-acetate-phosphate (TAP) solid medium under continuous light at 23 °C (52). For isolation of pt nucleoids from zygotes, a mating-type plus (*mt+*) strain (CC-125) and mating-type minus (*mt−*) strain (CC-124) were grown separately on agar plates (1.2% agar in Sager and Granick medium). The mating reaction was induced by suspending the cells in nitrogen-free Tsubo mating buffer (1.2 mM Hepes [pH 6.8], 1 mM $MgSO_4$) and incubating them under strong light (200 μmol photons m^{-2} s^{-1}) for more than 3 h. Then, equal numbers of *mt+* and *mt−* gametes were mixed and allowed to mate.

Isolation of cp Nucleoids. Isolation of cp nucleoids was performed basically as described previously (supplemental figure 3 of ref. 29). In the present study, about 2 g (fresh weight) of zygotes at 60 min after mating was disrupted with an airbrush (HP-62B, Olympos) (53) at a pressure of 2.0 kg/cm^2. Cell lysates were washed with suspension buffer (150 mM sucrose, 1.2 mM Hepes-KOH [pH 6.8], 5% polyethylene glycol 6000, 0.5 mM EDTA) until the supernatant became clear. The pellet was resuspended in suspension buffer containing 30% Percoll and overlaid on top of a discontinuous Percoll gradient (45 to 80% Percoll in suspension buffer) in 15-mL tubes. After centrifugation at 2,000 × *g* at 4 °C for 40 min with a swinging-bucket rotor, the green band at the interface between the 45 to 80% layers of Percoll was collected. The cp fraction was diluted fourfold with suspension buffer and then layered onto a discontinuous sucrose density gradient (3 mL each 80, 40, and 20% sucrose in suspension buffer) in 15-mL tubes and centrifuged at 2,000 × *g* for 30 min with a swinging-bucket rotor. Green bands of cp at the 80 to 40% sucrose interface (~2 mL) were recovered and diluted to 6 mL with TAN buffer (20 mM Tris·HCl [pH 7.6], 0.5 mM EDTA, 7 mM 2-mercaptoethanol, 1.2 mM spermidine, 0.4 mM phenylmethylsulfonyl fluoride). A total of 400 μL 20% Nonidet P-40 were added to the cp suspension. After being rotated for 30 min at 4 °C, the solution was centrifuged at 100 × *g* for 5 min to remove starch grains and debris. The supernatant was filtered through a layer of nylon mesh with 5-μm pores, and the filtrate was centrifuged at 20,000 × *g* for 45 min to sediment the cp nucleoids. The sedimented cp nucleoids were resuspended in 200 μL of TAN buffer, divided into aliquots, quickly frozen, and stored at −80 °C.

Mass Spectrometric Analysis. The nucleoid proteins were separated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) and subjected to in-gel digestion with mass spectrometry (MS) grade modified trypsin (Promega). The peptides were extracted as described previously (26) and loaded onto a column (100-μm internal diameter, 15-cm length; L-Column, CERI) using a Paradigm MS4 HPLC pump (Michrom BioResources) and an HTC-PAL autosampler (CTC Analytics) and eluted over 26 min with a gradient of 5 to 45% (vol/vol) acetonitrile in 0.1% (vol/vol) formic acid. The eluted peptides were introduced directly into an LTQ-Orbitrap XL mass spectrometer (Thermo Fisher Scientific) with a flow rate of 500 nL-·-min^{-1} and a spray voltage of 2.0 kV. The obtained spectra were compared to those in a protein database (*Chlamydomonas* genomic information 2.4) using the in-house MASCOT server (version 2.3, Matrix Science) (54). MASCOT search parameters were set as follows: threshold of the ion score cutoff, 0.05; peptide tolerance, 10 ppm; MS/MS tolerance, 0.5 Da; and peptide charge, +1, +2, or +3. The search was also set to allow one missed cleavage by trypsin, carboxymethylation modification of Cys residues, and variable oxidation of Met residues.

Construction of YFP-Fusion Protein Expression Line. To construct YFP-fusion proteins for subcellular localization analyses, HBD1 was amplified from genomic DNA of *C. reinharditii* and cloned into NdeI/EcoRI-cleaved pNyan vector (26) using the In-fusion system (Takara Bio USA Inc.). Nuclear transformation of *C. reinhardtii* (strain CW15 *mt+*) was performed by electroporation using Gene Pulsar (Bio-Rad) (55). Briefly, the cells were collected by centrifugation at 800 × *g* for 5 min at 4 °C and resuspended in TAP medium containing 50 mM sucrose to a final density of 1 × 10^8/mL. The cell suspension (250 μL) was placed into an electroporation cuvette with a 4-mm gap (Bio-Rad), which was then chilled at 16 °C for 5 min. Then, an exponential electric pulse (2,000 V/cm) was applied to the sample with the capacitance of 25 μF. The cuvette was incubated on ice for at least 5 min, and the cell suspension was plated onto solid medium (TAP-1.2% agarose) with 10 μg/mL paromomycin. The plates were placed under 30 μmol photons m^{-2} s^{-1} at 23 °C.

Transient Expression in Leaf Epidermal Cells of *N. benthamiana*. Full-length coding sequence (CDS) of 3×HBD genes were amplified from cDNA and inserted between a cauliflower mosaic virus 35S promoter and GFP sequence into pGWB405 using Gibson assembly technology. *N. benthamiana* leaves were infiltrated with *Agrobacterium* using conventional methods. To visualize mt, *Agrobacterium* carrying mt-targeted mCherry expression vector (mCherry-mt-rk CD3-991) was coinfiltrated. After 2 d incubation, infiltrated leaves were observed under a confocal laser microscope (Leica SP5).

Construction of Knockout Line with CRISPR/Cas9 System. CRISPR/Cas9-mediated genome editing was performed following a previously reported protocol (31). *C. reinhardtii* strains CC-125 and CC-124 were grown in liquid TAP medium under fluorescent light (60 μmol photons m^{-2} s^{-1}) with rotary shaking at 110 rpm at 22 °C with a 14 h light/10 h dark cycle for at least 10 d to keep them in the exponential growth phase. Cells at a density of 1 to 2 × 10^6 cells/mL were harvested and resuspended in MAX Efficiency transformation medium (A24229; Thermo Fisher Scientific) supplemented with 40 mM sucrose to a density of 10^8 cells/mL. Then cells were heat shocked at 40 °C for 30 min. The scaffold RNA and the target sequence RNA were ordered as recommended by the manufacturer (Integrated DNA Technologies [IDT]). The two RNAs were annealed in DUPLEX buffer (100 mM potassium acetate and 30 mM Hepes [pH 7.5], IDT) to a final concentration of 10 μM by heating to 95 °C for 2 min, followed by cooling at a rate of 0.1 °C/min. Cas9 protein (Clontech Laboratories, Guide-it Recombinant Cas9, 632640) was

mixed with an equimolar amount of annealed guide RNA in 1×NEB3.1 Buffer (100 mM NaCl, 50 mM Tris·HCl [pH 7.9], 10 mM $MgCl_2$, 100 μg/mL bovine serum albumin [BSA], New England Biolabs) at a final concentration of 3 μM each and incubated for 15 min at 37 °C. Cells were mixed with 5 μL of 3 μM Cas9 ribonucleoprotein and 1.5 μL (15 pmol) of selection marker plasmid pHyg3 (APHVII) with or without 1.5 μL of Homology Directed Recombination donor dsDNA. Transformation was performed by electroporation using a NEPA21 electroporator (Nepa Gene Co.) (56). The cell suspension was plated onto solid medium (TAP-1.2% agarose) with 10 μg/mL hygromycin. The plates were placed under 30 μmol photons m^{-2} s^{-1} at 23 °C for 1 wk, and emerged colonies were cultured in 150 μL of TAP medium in the wells of a 96-well plate to use for PCR screening. For complementation, pNyan-HBD1 vector described in the section regarding construction of the YFP-fusion protein-expression line or pNyan vector with truncated (N-terminal half of) HBD1 was introduced into the knockout line using a NEPA21 electroporator.

qRT-PCR. DNA content and RNA-expression level were determined by qRT-PCR. Total DNA was extracted by the classic phenol/chloroform-based method. Total RNA extraction and reverse transcription were performed with the Maxwell 16 LEV Plant RNA Kit (Promega) and ReverTra Ace (Toyobo), respectively. qRT-qPCR was performed using FastStart SYBR Green Master (ROX, Roche Diagnostics) and the Mx3000P system following the manufacturer's instructions. The quantitative estimations were calculated with MxPro software using the ΔΔCt (cycle threshold) method (Stratagene, Agilent Technologies), using a nuclear gene *MAA7* (*SI Appendix*, Figs. S4 and S5) or *GBLP* (*SI Appendix*, Fig. S7) sequence as an internal control. The list of primers used in this study is shown in *SI Appendix*, Table S3.

Protein Expression and Purification. The pQE80L vector (Qiagen) was linearized with *Bam*HI and *Hin*dIII, and HBD1 and HLP were cloned into it using the In-fusion system. *E. coli* (Rosetta DE3, Merck Millipore) carrying pQE80L vector was grown in 200 mL Luria–Bertani (LB) medium containing 50 μg/mL carbenicillin (Nacalai Tesque) to an optical density at 600 nm (OD_{600}) of 0.4, and the genes were induced by adding isopropyl-β-D-1-thiogalactopyranoside (IPTG, Nacalai Tesque) to a final concentration of 1 mM. After 3-h incubation at 37 °C, the cells were harvested by centrifugation. The cells were frozen using liquid nitrogen and thawed on ice and suspended with lysis buffer (20 mM Hepes-KOH [pH 7.6], 500 mM NaCl, 20 mM imidazole, 0.1% lysozyme) containing EDTA-free protease inhibitor mixture (Complete, Roche) at 4 °C and disrupted by sonication. Recombinant proteins in the supernatant of the lysate were incubated with Ni-NTA agarose (Qiagen) for 2 h. After washing three times with wash buffer (20 mM Hepes-KOH [pH 7.6], 500 mM NaCl, 20 mM imidazole), proteins were collected with elution buffer (20 mM Hepes-KOH [pH 7.6], 500 mM NaCl, 250 mM imidazole). The purified protein fractions were then desalted and concentrated by ultrafiltration (Vivaspin, 3 kDa molecular weight cutoff, GE Healthcare).

To compare the proteins with one, two, or three HMG-box domains, the desired numbers of the HMG-box domains of HBD1 were artificially connected. For this purpose, the sequence equivalent to one N-terminal–side HMG-box domain, one each of N- and C-terminal domains, or two N-terminal and one C-terminal domains of HBD1, respectively, were interconnected with the respective linker amino acids and inserted into pCold-I DNA vector (Takara). *E. coli* carrying a pCold-HBD1 vector were grown in LB medium containing 50 μg/mL carbenicillin to an OD_{600} of 0.4. Protein expression was induced by 1 mM IPTG after preincubation at 16 °C for 30 min. The *E. coli* was cultured at 16 °C and harvested after 24 h. The induced proteins were purified as described above.

EMSA. Purified HBD1 and HLP proteins were tested for DNA binding activity by EMSA as described in ref. 14. A total of 1 μL recombinant HBD1, HLP, or BSA was added to 9 μL of reaction buffer (10 mM Tris·HCl [pH 7.5], 0.5 mM EDTA) containing 400 ng of StyI-digested λ phage DNA (Marker 6, Nippon Gene) and incubated at room temperature for 2 h. The reaction mixtures were applied to a 1.2% agarose gel in TAE buffer (40 mM Tris-acetate [pH 8.0], 2 mM EDTA). After electrophoresis, the gel was stained with ethidium bromide. In order to compare the binding strength depending on the difference in the number of HMG-box domains, a 50-bp oligonucleotide having a random sequence (GC content = 50%) was synthesized and used after annealing the complementary strands. In this experiment, a 3% agarose gel in TBE buffer (89 mM Tris·HCl [pH 8.3], 2 mM EDTA, 89 mM H_3BO_3) was used, keeping the temperature low. To compare the affinity of the tested proteins, the band intensity of free DNA was quantified, and the EC_{50} value was calculated by fitting the sigmoid.

Complementation Analysis of Δabf2 in *S. cerevisiae*. The pAMIA vector containing WT *Abf2* gene and abf2-deficient (*Δabf2*) *S. cerevisiae* mutant was a kind gift from D. A. Clayton, Howard Hughes Medical Institute, Janelia Research Campus, Ashburn, VA. The outline of complementation of *Δabf2* was described in ref. 57. Briefly, to express HBD1 in yeast mt, the sequence of the mt transit peptides of Abf2p was amplified from pAMIA-*Abf2* vector, fused to the HBD1 gene amplified from *C. reinhardtii* cDNA, and inserted into the yeast-expression vector pYES2/CT (Invitrogen) driven by the galactose-inducible GAL1 promoter. Transformation of yeast cells was performed with the lithium chloride method (58). Stability of yeast mtDNA was assayed by spotting a series of dilutions on YPDG plates (1% yeast extract, 2% peptone, 0.1% glucose, 3% glycerol, 2% agar) and incubating them for 4 d at 30 °C after growing them in liquid synthetic medium (0.17% yeast nitrogen base [Difco]) and 0.5% ammonium sulfate supplemented with adenine and appropriate amino acids) containing 3% glycerol or 2% SGal for 2 d at 30 °C.

Phylogenetic Analysis. Protein sequences of TFAM, Abf2p, and HMG proteins that have two HMG-box domains from plants and algae were collected with searches using the BLAST algorithm against public databases, and their HMG-box domains were extracted by Pfam search. The sequences were aligned using PROBCONS. RAxML was used for phylogenetic tree search and bootstrap value calculation by the maximum likelihood method. A Bayesian inference analysis was performed using MrBayes with 1,000,000 generations. In both cases, Aminosan was used for model selection.

AFM. The imaging was performed using DNA origami that consisted of a DNA frame and 70 and 80 bp dsDNA. The sequence of the DNA frame used in the experiment was described previously (35). The DNA frame was assembled in a 20 μL solution containing 10 nM M13mp18 single-stranded DNA (New England BioLabs), 100 nM staple strands (226 strands), 20 mM Tris·HCl (pH 7.6), 1 mM EDTA, and 10 mM $MgCl_2$. The mixture was annealed at a temperature gradient from 85 to 15 °C at a rate of −1.0 °C/min. The preannealed dsDNAs were incorporated into the DNA frame by annealing the mixture at a temperature gradient from 40 to 15 °C at a rate of −0.5 °C/min using a thermal cycler.

The DNA frame (at concentration 10 nM) and recombinant protein were preincubated at room temperature for 10 min in AFM buffer (20 mM Tris·HCl [pH 7.6], 10 mM $MgCl_2$, and 1 mM EDTA). After the incubation, 2 μL of the sample was deposited onto freshly cleaved, 3-aminopropyltriethoxysilane–coated mica, and then the surface was rinsed with AFM buffer. AFM imaging was performed using a high-speed atomic force microscope (Nano Live Vision, RIBM). The sample was imaged in the buffer solution at room temperature with a silicon nitride cantilever (BL-AC10EGS, Olympus Corporation). The images were obtained at a scan rate of 0.2 frames per second.

Data Availability. HBD1 has been deposited in GenBank (accession no. MW505863).

ACKNOWLEDGMENTS. We thank Dr. Kentaro Tamura (University of Shizuoka) and Dr. Hisayoshi Nozaki (University of Tokyo) for helpful advice; Dr. Kazuhiko Sugimoto, Ms. Naomi Harada, Ms. Hitomi Tanaka, and Ms. Masumi Taniguchi (Kyoto University) for technical support; and Brody Frink (Kyoto University) for checking the text. In addition, we thank Dr. D. A. Clayton (HHMI) for providing yeast strains. This work was supported by a Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Scientific Research B (18H02460), a Grant-in-Aid for Scientific Research on Innovative Areas (17H05840, 19H04861), and a Grant-in-Aid for Exploratory Challenging Research (18K19337), as well as Grants-in-Aid for Scientific Research from the Mitsubishi Foundation (201910032) and the Ohsumi Foundation (203200500032) to Y.N., a JSPS Grant-in-Aid for Early-Career Scientists (18K14733), the Sasakawa Scientific Research Grant from The Japan Science Society (29-440), and a grant from Institute for YAMATO and Kii-peninsula studies of Nara Women's University to M.T.

1. L. Margulis, *Origin of Eukaryotic Cells; Evidence and Research Implications for a Theory of the Origin and Evolution of Microbial, Plant, and Animal Cells on the Precambrian Earth* (Yale University Press, New Haven, CT, 1970).
2. M. W. Gray, The endosymbiont hypothesis revisited. *Int. Rev. Cytol.* **141**, 233–357 (1992).
3. K. Drlica, J. Rouviere-Yaniv, Histonelike proteins of bacteria. *Microbiol. Rev.* **51**, 301–319 (1987).
4. C. J. Dorman, P. Deighan, Regulation of gene expression by histone-like proteins in bacteria. *Curr. Opin. Genet. Dev.* **13**, 179–184 (2003).
5. S. C. Dillon, C. J. Dorman, Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat. Rev. Microbiol.* **8**, 185–195 (2010).
6. T. A. Azam, A. Ishihama, Twelve species of the nucleoid-associated protein from Escherichia coli. Sequence recognition specificity and DNA binding affinity. *J. Biol. Chem.* **274**, 33105–33113 (1999).

7. R. L. Ohniwa, Y. Ushijima, S. Saito, K. Morikawa, Proteomic analyses of nucleoid-associated proteins in Escherichia coli, Pseudomonas aeruginosa, Bacillus subtilis, and Staphylococcus aureus. *PLoS One* **6**, e19172 (2011).
8. K. K. Swinger, K. M. Lemberg, Y. Zhang, P. A. Rice, Flexible DNA bending in HU-DNA cocrystal structures. *EMBO J.* **22**, 3749–3760 (2003).
9. F. Malka, A. Lombès, M. Rojo, Organization, dynamics and transmission of mitochondrial DNA: Focus on vertebrate nucleoids. *Biochim. Biophys. Acta* **1763**, 463–472 (2006).
10. C. Kukat, N. G. Larsson, mtDNA makes a U-turn for the mitochondrial nucleoid. *Trends Cell Biol.* **23**, 457–463 (2013).
11. B. A. Kaufman *et al*., The mitochondrial transcription factor TFAM coordinates the assembly of multiple DNA molecules into nucleoid-like structures. *Mol. Biol. Cell* **18**, 3225–3236 (2007).
12. J. F. Diffley, B. Stillman, A close relative of the nuclear, chromosomal high-mobility group protein HMG1 in yeast mitochondria. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 7864–7868 (1991).
13. I. Miyakawa, N. Sando, S. Kawano, S. Nakamura, T. Kuroiwa, Isolation of morphologically intact mitochondrial nucleoids from the yeast, Saccharomyces cerevisiae. *J. Cell Sci.* **88**, 431–439 (1987).
14. N. Sasaki *et al*., Glom is a novel mitochondrial DNA packaging protein in Physarum polycephalum and causes intense chromatin condensation without suppressing DNA functions. *Mol. Biol. Cell* **14**, 4758–4769 (2003).
15. A. Rubio-Cosials *et al*., Human mitochondrial transcription factor A induces a U-turn structure in the light strand promoter. *Nat. Struct. Mol. Biol.* **18**, 1281–1289 (2011).
16. H. B. Ngo, G. A. Lovely, R. Phillips, D. C. Chan, Distinct structural features of TFAM drive mitochondrial DNA packaging versus transcriptional activation. *Nat. Commun.* **5**, 3077 (2014).
17. K. D. Grasser *et al*., The recombinant product of the Chryptomonas phi plastid gene hlpA is an architectural HU-like protein that promotes the assembly of complex nucleoprotein structures. *Eur. J. Biochem.* **249**, 70–76 (1997).
18. T. Kobayashi *et al*., Detection and localization of a chloroplast-encoded HU-like protein that organizes chloroplast nucleoids. *Plant Cell* **14**, 1579–1589 (2002).
19. E. V. S. R. Ram, R. Naik, M. Ganguli, S. Habib, DNA organization by the apicoplast-targeted bacterial histone-like protein of Plasmodium falciparum. *Nucleic Acids Res.* **36**, 5061–5073 (2008).
20. S. L. Wang, X. Q. Liu, The plastid genome of Cryptomonas phi encodes an hsp70-like protein, a histone-like protein, and an acyl carrier protein. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 10783–10787 (1991).
21. D. Karcher, D. Köster, A. Schadach, A. Klevesath, R. Bock, The Chlamydomonas chloroplast HLP protein is required for nucleoid organization and genome maintenance. *Mol. Plant* **2**, 1223–1232 (2009).
22. M. Powikrowska *et al*., SVR4 (suppressor of variegation 4) and SVR4-like: Two proteins with a role in proper organization of the chloroplast genetic machinery. *Physiol. Plant.* **150**, 477–492 (2014).
23. W. Majeran *et al*., Nucleoid-enriched proteomes in developing plastids and chloroplasts from maize leaves: A new conceptual framework for nucleoid functions. *Plant Physiol.* **158**, 156–189 (2012).
24. Y. Yagi, T. Shiina, Recent advances in the study of chloroplast gene expression and its evolution. *Front. Plant Sci.* **5**, 61 (2014).
25. J. Pfalz, T. Pfannschmidt, Plastid nucleoids: Evolutionary reconstruction of a DNA/protein structure with prokaryotic ancestry. *Front. Plant Sci.* **6**, 220 (2015).
26. Y. Kobayashi *et al*., Eukaryotic components remodeled chloroplast nucleoid organization during the green plant evolution. *Genome Biol. Evol.* **8**, 1–16 (2015).
27. O. Emanuelsson, S. Brunak, G. von Heijne, H. Nielsen, Locating proteins in the cell using TargetP, SignalP and related tools. *Nat. Protoc.* **2**, 953–971 (2007).
28. H. Bannai, Y. Tamada, O. Maruyama, K. Nakai, S. Miyano, Extensive feature detection of N-terminal protein sorting signals. *Bioinformatics* **18**, 298–305 (2002).
29. P. Horton *et al*., WoLF PSORT: Protein localization predictor. *Nucleic Acids Res.* **35**, W585–W587 (2007).
30. J. M. Zones, I. K. Blaby, S. S. Merchant, J. G. Umen, High-resolution profiling of a synchronized diurnal transcriptome from Chlamydomonas reinhardtii reveals continuous cell and metabolic differentiation. *Plant Cell* **27**, 2743–2769 (2015).
31. A. Greiner *et al*., Targeting of photoreceptor genes in Chlamydomonas reinhardtii via zinc-finger nucleases and CRISPR/Cas9. *Plant Cell* **29**, 2498–2518 (2017).
32. X. Li *et al*., An indexed, mapped mutant library enables reverse genetics studies of biological processes in Chlamydomonas reinhardtii. *Plant Cell* **28**, 367–387 (2016).
33. O. Zelenaya-Troitskaya, S. M. Newman, K. Okamoto, P. S. Perlman, R. A. Butow, Functions of the high mobility group protein, Abf2p, in mitochondrial DNA segregation, recombination and copy number in Saccharomyces cerevisiae. *Genetics* **148**, 1763–1776 (1998).
34. A. Chakraborty *et al*., DNA structure directs positioning of the mitochondrial genome packaging protein Abf2p. *Nucleic Acids Res.* **45**, 951–967 (2017).
35. Y. Kobayashi *et al*., Holliday junction resolvases mediate chloroplast nucleoid segregation. *Science* **356**, 631–634 (2017).
36. M. Endo, Y. Katsuda, K. Hidaka, H. Sugiyama, Regulation of DNA methylation using different tensions of double strands constructed in a defined DNA nanostructure. *J. Am. Chem. Soc.* **132**, 1592–1597 (2010).
37. C. Kukat *et al*., Cross-strand binding of TFAM to a single mtDNA molecule forms the mitochondrial nucleoid. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 11288–11293 (2015).
38. K. D. Grasser, Chromatin-associated HMGA and HMGB proteins: Versatile co-regulators of DNA-dependent processes. *Plant Mol. Biol.* **53**, 281–295 (2003).
39. K. D. Grasser, D. Launholt, M. Grasser, High mobility group proteins of the plant HMGB family: Dynamic chromatin modulators. *Biochim. Biophys. Acta* **1769**, 346–357 (2007).
40. J. O. Thomas, A. A. Travers, HMG1 and 2, and related 'architectural' DNA-binding proteins. *Trends Biochem. Sci.* **26**, 167–174 (2001).
41. G. H. Goodwin, C. Sanders, E. W. Johns, A new group of chromatin-associated proteins with a high content of acidic and basic amino acids. *Eur. J. Biochem.* **38**, 14–19 (1973).
42. B. Kiilerich *et al*., Chromosomal high mobility group (HMG) proteins of the HMGB-type occurring in the moss Physcomitrella patens. *Gene* **407**, 86–97 (2008).
43. F. Legros, F. Malka, P. Frachon, A. Lombès, M. Rojo, Organization and dynamics of human mitochondrial DNA. *J. Cell Sci.* **117**, 2653–2662 (2004).
44. B. M. Hallberg, N. G. Larsson, TFAM forces mtDNA to make a U-turn. *Nat. Struct. Mol. Biol.* **18**, 1179–1181 (2011).
45. Y. Nishimura, T. Higashiyama, L. Suzuki, O. Misumi, T. Kuroiwa, The biparental transmission of the mitochondrial genome in Chlamydomonas reinhardtii visualized in living cells. *Eur. J. Cell Biol.* **77**, 124–133 (1998).
46. S. Greiner *et al*., Chloroplast nucleoids are highly dynamic in ploidy, number, and structure during angiosperm leaf development. *Plant J.* **102**, 730–746 (2020).
47. C. Kukat *et al*., Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 13534–13539 (2011).
48. Y. Kamimura, H. Tanaka, Y. Kobayashi, T. Shikanai, Y. Nishimura, Chloroplast nucleoids as a transformable network revealed by live imaging with a micro fluidic device. *Commun. Biol.* **1**, 47 (2018).
49. J. Nosek, L. Tomaska, M. Bolotin-Fukuhara, I. Miyakawa, Mitochondrial chromosome structure: An insight from analysis of complete yeast genomes. *FEMS Yeast Res.* **6**, 356–370 (2006).
50. T. L. Megraw, C. B. Chae, Functional complementarity between the HMG1-like yeast mitochondrial histone HM and the bacterial histone-like protein HU. *J. Biol. Chem.* **268**, 12758–12763 (1993).
51. M. Antosch, S. A. Mortensen, K. D. Grasser, Plant proteins containing high mobility group box DNA-binding domains modulate different nuclear processes. *Plant Physiol.* **159**, 875–883 (2012).
52. E. H. Harris, *The Chlamydomonas Sourcebook* (Academic Press, ed. 2, 2009).
53. Y. Nishimura *et al*., An mt(+) gamete-specific nuclease that targets mt(-) chloroplasts during sexual reproduction in C. reinhardtii. *Genes Dev.* **16**, 1116–1128 (2002).
54. D. N. Perkins, D. J. Pappin, D. M. Creasy, J. S. Cottrell, Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567 (1999).
55. K. Shimogawara, S. Fujiwara, A. Grossman, H. Usuda, High-efficiency transformation of Chlamydomonas reinhardtii by electroporation. *Genetics* **148**, 1821–1828 (1998).
56. T. Yamano, H. Iguchi, H. Fukuzawa, Rapid transformation of Chlamydomonas reinhardtii without cell-wall removal. *J. Biosci. Bioeng.* **115**, 691–694 (2013).
57. I. Miyakawa *et al*., Mitochondrial nucleoids from the yeast Candida parapsilosis: Expansion of the repertoire of proteins associated with mitochondrial DNA. *Microbiology (Reading)* **155**, 1558–1568 (2009).
58. H. Ito, Y. Fukuda, K. Murata, A. Kimura, Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.* **153**, 163–168 (1983).
59. R Core Team, Joy in playing. R statistical package version 3.5.0 (R Core Team, 2018).