RESEARCH ARTICLE

Protein binding sites on centromere DNA in the dimorphic yeast *Yarrowia lipolytica*

Masayoshi Matsuoka*

Affiliations:

Department of Applied Microbial Technology, Faculty of Biotechnology and Life Science, Sojo University, Ikeda 4-22-1, Kumamoto, Japan 860-0082

*** Corresponding author:** Masayoshi Matsuoka, PhD, E-mail: matsuoka@bio.sojo-u.ac.jp

Abstract

In the budding yeasts, centromere (CEN) DNA is confined within a short region of chromosome called point centromere. *Yarrowia lipolytica* forms budding-type yeast cells as well as mycelial cells depending on their growth conditions, thus imposing the interest on CEN architecture among distantly related budding yeasts. A canonical centromere-determining element (CDE) III site had been identified in *Y. lipolytica* CENs. Interestingly, CEN1 region contained two closely associated CDEIII sites, separated by 338 bp. To extend the similarity with canonical CENs in budding yeasts, we searched CDEI site(s) in CEN1 for their capacity to be bound by the purified *Y. lipolytica* transcription factor, Cbf1 protein. We found one or more possible Cbf1-binding sites at the outward region of one CDEIII site in CEN1. The strongest binding-site of Cbf1, a potential CDEI site, is involved in approximately 400 bp including the nearest CDEIII site. These results showed that the CDEI-CDEII-CDEIII architecture in budding yeast CEN is conserved in *Y. lipolytica* CEN1, but that the CDEII is longer with possible multiplication of CDEI sites.

Key words: Centromere, *Yarrowia lipolytica*, Dimorphic yeast, Transcription factor, Centromere-determining element

1. Introduction

Eukaryotic chromosomal DNAs have specialized sites called centromeres (CENs) on which kinetochores are assembled to properly segregate the replicated sister chromatids. CEN-associated proteins are evolutionally well conserved, including mammalian CENP-A, CENP-B, and CENP-C, which were identified with CEN-specific autoantibodies. In the budding yeasts, CENs are confined within short (ca. 125 bp) DNA whose sequence motifs were clarified first in

Saccharomyces cerevisiae.[1] The so-called point centromere in *S. cerevisiae* has three sequence motifs named centromere-determining element (CDE); CDEI, CDEII, and CDEIII. Each CDE is bound by specific proteinaceous factor(s). CDEI is 8 bp palindrome on which a transcription factor Cbf1 binds;[2] CDEII is ~80bp AT-rich sequence which wraps nucleosomal histones containing Cse4,[3] a centromere-specific variant of histone H3 (CENP-A orthologue); CDEIII is a partial palindromic sequence of 24 bp which is bound by CBF3 complex consisting of Ndc10, Cep3, Ctf13, and Skp1.[4,5] The CBF3 complex is a part of kinetochore and is a site of spindle microtubule attachment. Another factor, Mif2 (CENP-C orthologue) is associated with Cbf1, Cse4, and CBF3 and binds with CDEIII.[6,7] CDEIII and CDEII are essential for CEN function as measured by mitotic activity in vivo, but CDEI is negligible for this function, although CDEI is necessary for proper meiotic function of the CEN.[8] However, there is evidence indicating that transcription per se promotes CEN function in budding yeast.[9]

In contrast with point CENs in budding yeasts, other eukaryotic cells including the fission yeast *Schizosaccharomyces pombe* have longer CEN sequences (regional CENs) which recruit similar proteins involving the CENP-A homologues. Indeed, the recruitment of CENP-A is a functional marker for active CEN.[10] Thus, regional CENs are regarded to be modulated by transcriptional stalling to establish CENP-A-containing chromatin.[11]

In the present study, we analyzed CEN1 from a dimorphic yeast *Yarrowia lipolytica* which propagates by budding as well as mycelial formation. Phylogenetically, *Y lipolytica* is distantly related with *S. cerevisiae* and other budding yeasts. Six chromosomes are contained in haploid cells of *Y. lipolytica*, whose CENs had been cloned on replicative plasmids.[12, 13] For the maintenance of replicative plasmids in *Y. lipolytica*, both replication origin (ORI) and CEN are needed.[14] Plasmids containing ORI but no CEN sequence cannot be partitioned to daughter cells in *Y. lipolytica*,[15] a property contrasting with the ability of ORI (ARS)-containing plasmid to be maintained in *S. cerevisiae* albeit at low stability. By using the functional assay of replicative plasmids, *Y. lipolytica* CEN region was delimited to a short region which contained sequences homologous with CEDIII from *S. cerevisiae*.[16, 17] Thus, these CEN regions from *Y. lipolytica* were considered to retain CBF3-binding site (CDEIII) for microtubule attachment. However, CDEIII site only cannot fulfill the CEN function since longer sequence probably corresponding to CDEII site is required for replicative plasmids. In addition, two CDEIII sites closely associated on CEN1 were found, each CDEIII-containing sequence (previously designated CEN1-1 and CEN1-2) can make functional CEN when cloned on ORI-plasmid.[17] In order to clarify the CEN architecture in *Y. lipolytica*, we investigated the presence of CDEI site on CEN1 here by

assaying binding with Cbf1 protein. The revealed architecture of *Y. lipolytica* CEN would disclose the difference or similarity of CDE composition among different budding yeasts.

2. Results

2.1. *Y. lipolytica* CEN1 contains two CDEIII sites with opposite orientation

Figure 1 shows a map of CEN1 region of chromosome 1 from *Y. lipolytica*. The 9063 bp CEN1 contains two regions, CEN1-1 and CEN1-2, either of which is necessary and sufficient for maintenance of ORI-containing plasmid. In addition, an ORI (ORI1001) is situated at the left side of CEN1-1. The length (bp) of each region shown in Fig. 1 is the minimum size with full functional activity. Thus, CEN1-1 is confined within 227 bp, although a shorter CEN1-1 (91 bp) could support mitotic function of recombinant plasmid with reduced transformation efficiency and mitotic stability.[16] We found previously the 21 bp consensus sequence within the CEN1-1, CEN1-2, and other chromosomal CENs (CEN3-1 and CEN5-1). The consensus sequence is responsible for binding of the same *Y. lipolytica* nuclear proteins, because point mutations as well as internal deletion of the consensus sequence abolished the protein binding, and because DNA fragment of one consensus sequence competed the protein binding with those of the other consensus sequences.[17]

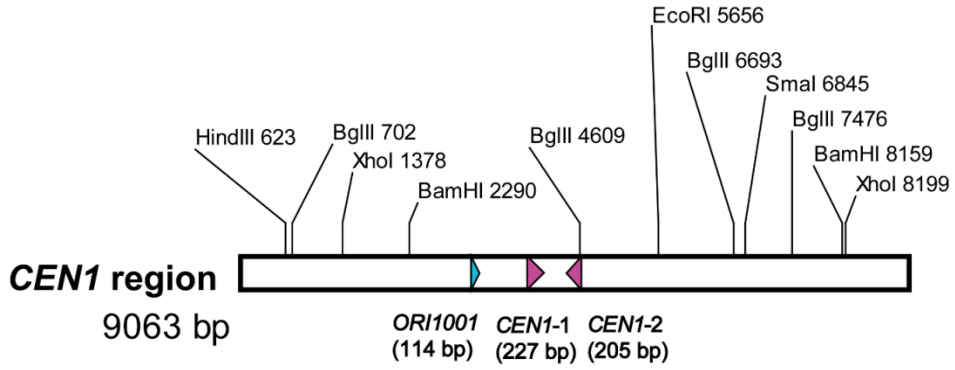

Fig.1. Restriction map for *Y. lipolytica* CEN1 region.

A 9063 bp CEN1 fragment with restriction sites is shown as an elongated square. A replication origin (ORI1001) and two regions of centromeric mitotic function (CEN1-1, CEN1-2) are indicated by triangles with their delimited sizes (in bp). CDEIII consensus sequences are found at the edge of triangles of CEN1-1 and CEN1-2.

Figure 2 shows the homology of the *Y. lipolytica* CEN consensus sequence with the canonical CDEIII of *S. cerevisiae* (*Sc*CDEIII). The consensus sequence has a partial dyad symmetry like *Sc*CDEIII, the central CC nucleotides of which are perfectly conserved in all *Y. lipolytica* CEN consensus sequences analyzed to date. Therefore, we designate here the consensus sequence as *Yl*CDEIII (CDEIII of *Y. lipolytica*) which spans approximately 27 bp. Most probably, homologous CBF3 complex is the binding proteins to *Yl*CDEIII. We observed four distinct protein-DNA complex bands in gel mobility shift assay using DNA fragments containing the *Yl*CDEIII; larger three complex bands disappeared in variously mutated *Yl*CDEIII.[16] Notably, *Yl*CDEIII sequences in CEN1-1 and CEN1-2 have opposite orientation. The 227 bp CEN1-1 and 205 bp CEN1-2 face against each *Yl*CDEIII, and the largest distance between two *Yl*CDEIII sites is 338 bp (Fig. 1). Thus, CEN1-1 extends to the left of its *Yl*CDEIII, whereas CEN1-2 does so to the right of its *Yl*CDEIII. These extended sequences might correspond to CDEII, since mitotic function of each CEN was confined in the fragment of at least 200 bp.

CEN1-1:	TTTTGCAGATTACCTAATTTGGTAAGA
CEN1-2:	AACAATTTACTTCCTAATTTGGAAAAT
CEN3-1:	TGAAAAACACTTCCTATTTAGGAAACA
CEN5-1:	TCAATAAAATTTCCGAATAGGGCAAAA
CEN2-1&*2*-2:	TTATTAAAATTTCCGTTTACGGAAATT
CEN4-1:	ATTTTATTTTTTCCATTTATAGTAACA
consensus:	WNNWNNWNWYTWCCNWWTWNRGNAANW
CDEIII:	TGTTTWTGNTTTCCGAAANNNAAAA

Fig.2. Comparison of *Y. lipolytica* CEN consensus sequences with *S. cerevisiae* CDEIII
All the sequences of CEN1, CEN2, CEN3, CEN4, and CEN5 are found in DNA fragments cloned in replicative plasmids. A 27 bp CEN consensus is shown and aligned with *S. cerevisiae* CDEIII with a central CC in dyad symmetry.

2.2. Cbf1 protein of *Y. lipolytica* binds with CEN1 region

Since in *S. cerevisiae* and related budding yeasts, CEN architecture is a linear array of CDEI-CDEII-CDEIII, we examined whether or not *Y. lipolytica* CENs also contain a corresponding CDEI which is the binding site of a transcription factor, Cbf1 protein. There is a gene in *Y. lipolytica* genome encoding a 425 a.a. protein (GenBank: AOW01642.1) with pI 9.94, which has considerable similarity with *S. cerevisiae* Cbf1 protein. The gene (cbf1) was cloned and expressed in

Escherichia coli as described in Materials and Methods. The cbf1 gene is scattered with several arginine codons which are rarely translated in *E. coli*. Thus, six of these rare codons (AGA, AGG) were changed to CGT or CGC, to form the cbf1(Arg) gene, which is supposed to be suitable for expression in *E. coli*. In addition, a hexahistidine tag (histag) was attached to the N-terminus of Cbf1(Arg) protein for affinity purification. In contrast, we did not obtain active Cbf1(Arg) with a C-terminal histag (data not shown). Finally, the histag-cbf1(Arg) gene was expressed with a pCold vector, which preferentially produces recombinant proteins at low temperature, and the Cbf1 protein was purified by affinity column as described in Materials and Methods.

Figure 3 shows an SDS-PAGE analysis of the purified Histag-Cbf1(Arg) proteins. After binding on Ni^{2+} chelating column, recombinant proteins were eluted by a linear gradient of imidazole in a buffer. A 46 kDa prominent band appeared in fraction 13 that corresponds to elution at 200 mM imidazole. At higher imidazole concentration, proteins with lower molecular weight (<30 kDa) were eluted, that were almost inactive for binding to CEN DNAs (data not shown). The 46 kDa recombinant protein was concentrated and used for gel mobility shift assays.

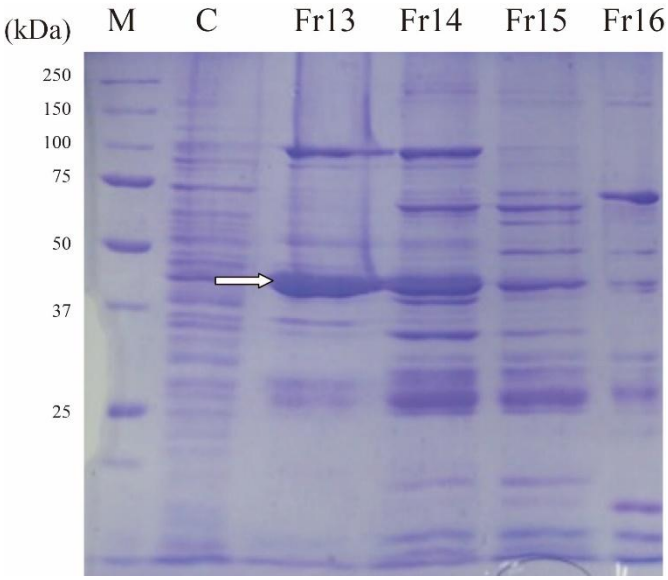

Fig.3. SDS-PAGE analysis of purified Cbf1 protein

E. coli BL21 cells harboring pColdIV-histag-cbf1(Arg) were grown at 15ºC, and disrupted in the presence of 1mM EDTA. Crude cell-free extract was passed through Sephadex G25 column, and was subjected to HisTrap HP column. After washing with a buffer, Cbf1 proteins were eluted by a linear gradient of imidazole up to 500 mM.

Lane M: molecular weight marker (kDa sizes are indicated at left side of the gel); lane C: crude cell-free extract; Lanes Fr13 to Fr16 are fraction No. 13 to 16. An arrow shows histag-Cbf1(Arg) band.

Figure 4 shows gel mobility shift assay for CEN1 sub-fragments which were generated by digestion of 9 kb CEN1 fragment with a mixture of HpaI, AluI, and HpyCH4V. These restriction enzymes cut the CEN1 fragment to form 80 blunt-ended DNA fragments. As shown in Fig. 4A, some CEN1 sub-fragments progressively disappeared in the presence of increasing amount of Cbf1 protein (indicated by open arrows), while new DNA-protein complex bands appeared on the gel (indicated by filled arrows). We detected at least three sub-fragments from CEN1 that were bound with Cbf1 protein; i.e., the largest A and B fragments (425 and 419 bp, respectively), and one of the crowded fragments beneath the E fragment (Fig.4A). The fragment A is just flanking a BglII site at the right border of CEN1-2, whereas the fragment B is situated at further right including an EcoRI site (Fig. 1). Since the fragments A and B are closely associated as a dimer, so it was unclear which of the fragments was preferentially shifted, we tested if a fragment A or B alone can be bound with Cbf1. As shown in Fig.4B, PCR-generated A (we used 409 bp from the right side of BglII site) and B fragments were indeed shifted up by binding with Cbf1 protein, thus, indicating the presence of at least two Cbf1 binding sites at the right side of CEN1-2.

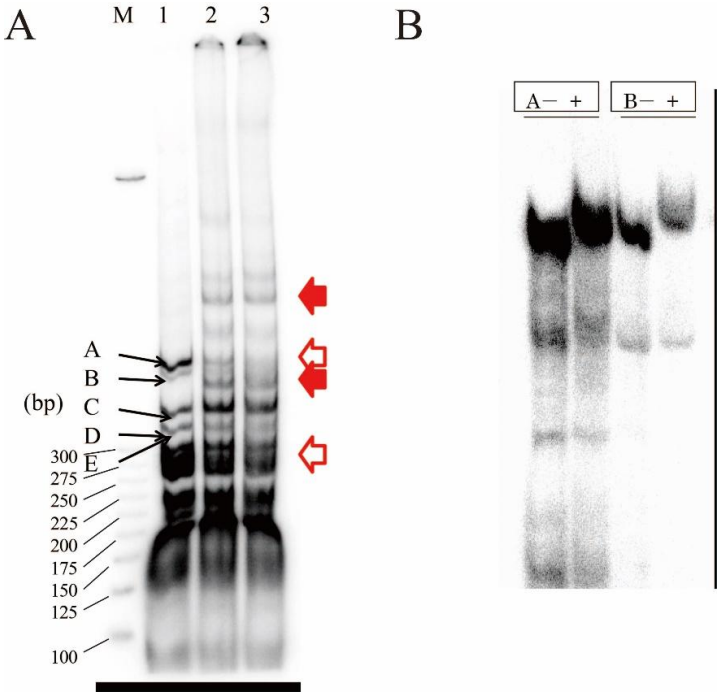

Fig.4. Search for Cbf1-binding sub-fragments of CEN1

CEN1 sub-fragments named A,B,C,D,E, according to their size were end-labeled with ^{32}P, and subjected to gel mobility shift assay with or without purified Cbf1 protein.

(A) CEN1 fragment (9063 bp) was cut with HpaI, AluI, and HpyCH4V, and used for the assay. Lane M: 25 bp step ladder; lane 1: without Cbf1; lane 2: with 1 μl Cbf1; lane 3: with 2 μl Cbf1. Free DNA bands are indicated by open arrows, while DNA-protein complex bands are shown by filled arrows.

(B) PCR-generated A fragment (409 bp) and B fragment (419 bp) were used. Lanes A-+ (left two) contained fragment A without and with Cbf1. Lanes B-+ (right two) contained fragment B without and with Cbf1.

2.3. Multiple Cbf1-binding sites are situated with increasing affinities near the CDEII-CDEIII

Figure 5 shows gel mobility shift data for delimitation of Cbf1 binding site within the fragment A and B. In order to delimit the binding site(s), the fragments A and B were divided into halves (A1 and A2, B1 and B2), each spanning ca. 200 bp. Figure 6 schematically represents the location of these fragments relative to CEN1-2. As shown in Fig. 5A, Cbf1 protein bound with A1 fragment strongly, while it did so with A2 fragment less strongly. The binding of Cbf1 with B1 and B2 fragments were much weaker than that with A1 and A2. In order to delimit the binding site(s) further, the fragment A1 was divided into two equal halves, A1-1 and A1-2, and these fragments were assayed. As shown in Fig. 5B, the unique binding of Cbf1 protein was observed in A1-2 fragment but not in A1-1 fragment. These data suggest that a high-affinity CDEI site for Cbf1 binding possibly resides on A1-2 fragment, which means that the maximum length of CDEI to CDEIII of CEN1-2 is approximately 300 to 400 bp (Fig. 6).

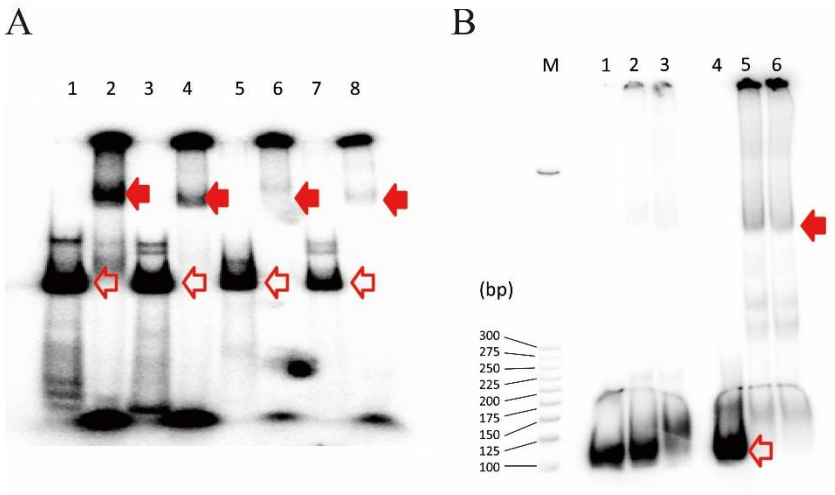

Fig. 5. Delimitation of Cbf1-binging site within the sub-fragments of CEN1-A and B

^{32}P-end labeled PCR fragments were assayed for gel mobility shift using purified Cbf1 protein. Free DNA bands are indicated by open arrows, while DNA-protein complex bands are shown by filled arrows.

(A) Sub-fragments are in lanes 1,2: A1 (205 bp); lanes 3,4: A2 (204 bp); lanes 5,6: B1 (210 bp); lanes 7,8: B2 (209 bp). Odd lanes are without Cbf1. Even lanes contained 4 µl Cbf1.

(B) Lane M: 25 bp step ladder; Sub-fragments are in lanes1 to 3: A1-1 (103 bp); lanes 4 to 6: A1-2 (102 bp). Lanes 1 and 4: without Cbf1; lanes 2 and 5: with 1µl Cbf1; lanes 3 and 6: with 2 µl Cbf1.

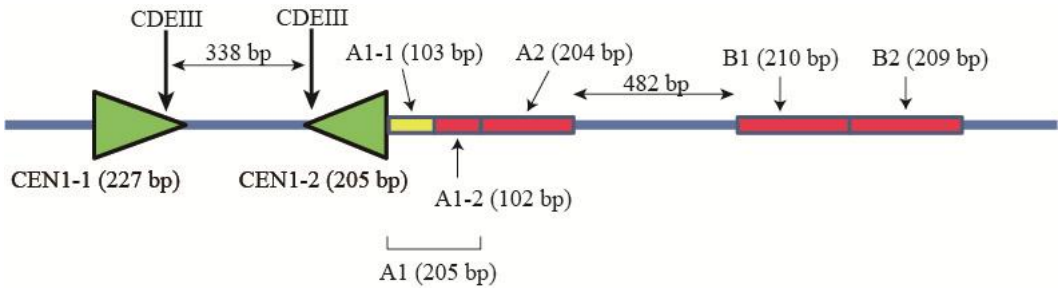

Fig. 6. Schematic representation of CEN1 with Cbf1 binding regions

Cbf1-binding sites are located at the right side of CEN1-2 within the fragments A1-2, A2, B1, and B2. These fragments are shown in red. The strongest Cbf1-binding site (potential CDEI site) is involved in A1-2 fragment, and weaker Cbf1-binding sites are possibly situated within A2, B1, and B2 fragments.

3. Discussion

In an attempt to find a possible CDEI site elsewhere in the CEN1 fragment from *Y. lipolytica*, we utilized a search for Cbf1-binding fragments covering the entire CEN1 region. Thus, restriction digestion of CEN1 DNA into 80 fragments enabled us to find out at least three DNA fragments with possible Cbf1 binding (Fig. 4A). Two sub-fragments A and B of CEN1 were indeed bound with purified histag-Cbf1(Arg) protein, justifying our approach (Fig.4B). Interestingly, possible Cbf1-binding sites are located within the A1, A2, B1, B2 fragments, although the binding affinities of Cbf1 on fragments distal to the CEN1-2 were progressively weakened (Fig.5A). A strongest Cbf1 binding was observed within A1-2 (102 bp) fragment which is located from 103 to 205 bp at the right side of CEN1-2 (Fig.5B, Fig.6). Since CDEIII site is located at the left end of CEN1-2, the maximum length spanning CDEIII-CDEII- CDEI should be less than 400 bp, or the minimal length is greater than 300 bp (Fig. 6). These data indicate that the canonical CDEI-CDEII-CDEIII architecture of budding yeast CENs is also conserved in *Y. lipolytica*, but the size spanning the CDEI to CDEIII is about twice as long as that in *S. cerevisiae*. In addition, multiple CDEI sites could be positioned to the CDEI-distal side. The precise localization of CDEI site in the A1-2 fragment is now in progress.

The Cbf1 protein is well conserved within budding yeasts.[18] In *S. cerevisiae*, it is a transcription factor acting on CEN and methionine biosynthesis gene.[19, 20, 21] The mammalian homologue of Cbf1 is CENP-B, which has weak but significant homology with microbial Cbf1. Mammalian CENP-B structure is well conserved even in a fission yeast and a holocentric insect.[22, 23] Both CENP-B and Cbf1 contain helix-loop-helix DNA binding domain, and act as a dimer on special palindromic sequences.[24, 25] Both CENP-B and Cbf1 introduce kink or bend of

DNA when bound to the recognition sequences.[26, 27] These properties of Cbf1 and CENP-B are consistent with its role on the formation of CEN-specific chromatin structure. To this end, *Y. lipolytica* Cbf1 protein would be one of the key components for kinetochore assembly.

Among the budding yeasts, *Y. lipolytica* is unique to have closely associated duplicated CDEIII sites on CEN1 and possibly on other CENs. Whether or not the duplicated CEN architecture might correlate with a specific mitotic function such as dimorphism in *Y. lipolytica* is an interesting question. In a dimorphic yeast, *Candida maltosa*, the presence of CEN sequence on replicating plasmid induced the pseudohyphal growth.[28] Although *Y. lipolytica* CEN sequences on plasmids did not seem to induce pseudohyphal growth, the specific architecture of CEN in this species would be worth investigating the relationship between CEN organization and control on budding and mycelial reproduction.

4. Materials and Methods

4.1. Gel mobility shift assay for CEN1 sub-fragments

CEN1 sequence is annotated in GenBank: EU340887.1. Polymerase chain reaction (PCR) was performed using PrimeStar DNA polymerase (TaKaRa) as described by the manufacturer. The PCR-generated sub-fragments of CEN1 are follows (nucleotide numbers are shown in parenthesis): A1(4612-4816), A2(4817-5020), B1(5503- 5712), B2(5713-5921), A1-1(4612-4714), and A1-2(4715-4816). We used pSL20 as a template for PCR.[12] Restriction digested DNAs were dephosphorylated with calf intestine alkaline phosphatase, treated with phenol/ chloroform prior to 5'-phosphrylation with T4 polynucleotide kinase and [γ-^{32}P] ATP. PCR-generated DNA fragments were purified and 5'-end labeled as above. Labeled fragments were purified with Micro BioSpin-30 (BioRad).

A typical binding assay mixture contained in 30 μl, 20mM HEPES (pH7.6), 1mM EDTA, 1mM DTT, 0.05M NaCl, 5%(w/v) glycerol, 0.5μg poly(dI-dC), ^{32}P-labeled DNA, and purified Cbf1 protein. The mixture was incubated at 25°C for 30 min, and was subjected to electrophoresis on polyacrylamide gel. After electrophoresis, gel was dried, and autoradiogram was obtained using FLUOR- image analyzer (PharosFX Plus, BioRad and FLA 9500, GE Healthcare).

4.2. Cloning and expression of cbf1 gene from *Y. lipolytica*

The cbf1 gene ORF with hexahistidine tag at the C terminal position was initially cloned between SgfI and PmeI sites of pF1A T7 vector (Promega). The cbf1 gene contains 6 Arg codons which are rarely translatable in *E. coli*. The C-terminal Arg codon (AGG) was simultaneously changed to CGT within the PCR primer. Other 5 Arg codons, AGA and

AGG, were changed to CGT and CGC, respectively, using QuikChange site-directed mutagenesis kit (Agilent Technologies). The mutated cbf1(Arg) gene with an N-terminal hexahistidine tag was then subcloned into pColdIV vector (TaKaRa). The PCR primers used were as follows; CBF1 Cold-1: 5'-GGTACCCTCGAGCACCACCATCACCACCACGACCTCAAATC-3'; CBF1 Cold-2: 5'-GTTTCTAGATTAACGCTCACTGTTCATGGGAGGCAGCTTATCG-3'.

The amplified DNA fragment was cut with XhoI and XbaI, and was cloned into the same restriction sites of pColdIV. The resulting plasmid, pCold-histag-cbf1(Arg) was used to transform *E. coli* BL21, and one of the ampicillin-resistant transformants was retained as enzyme source. The BL21/pCold-histag-cbf1(Arg) cells were grown in LB medium containing 100μg/ml ampicillin, and the culture was transferred to 15°C when OD_{610} of the culture became equal to 0.5. The culture was then added with 0.5mM IPTG, and incubated at 15°C for 24 hours. The cells were harvested by centrifugation, and were suspended in a binding buffer (20mM Na_2HPO_4, 0.5M NaCl, 20mM imidazole, 10%(w/v) glycerol, pH7.4). To the cell suspension, 1mM PMSF, 2.5μg/ml pepstatin A, 2.5μg/ml leupeptin, 5mM DTT, and 1mM EDTA were added, and the ice-cold cell suspension were treated with ultrasonic processor (XL2020, Astrason). The disrupted cell suspension was centrifuged at 10,000 g for 20min at 4°C, and the supernatant was retained as cell-free extract. The cell-free extract was stored at -80°C.

The purification of histag-Cbf1(Arg) protein was performed on BioCAD SPRINT station (Applied Biosystems) equipped with 1ml HisTrap HP column (GE Healthcare). The column was injected with cell-free extract (three times of 5 ml) and washed by the binding buffer. The bound Cbf1 proteins were eluted with a linear gradient of imidazole in the buffer up to 500 mM imidazole. The eluted protein was concentrated with centrifugal filter units (10,000 NMWL, Amicon Ultra, Merck Millipore) and stored at -80°C. Protein concentration was measured by DC Protein assay (BioRad). Proteins were analyzed by SDS-PAGE on 12% polyacryl-amide gel and stained with CBB.

5. Acknowledgement

Excellent experimental support of Takayoshi Yamane, Jun Kaseda, Yoshiaki Fujikawa, Ai Miyagawa, and Misaki Hiroshige is greatly appreciated. This study was supported by research grant from Sojo University.

6. References

1. Fitzgerald-Hayes M, Clarke L, Carbon J. Nucleotide sequence comparisons and functional analysis of yeast centromere DNAs. Cell 1982; 29(1):235-244.

2. Niedenthal R, Stoll R, Hegemann JH. In vivo characterization of the *Saccharomyces cerevisiae* centromere DNA element I, a binding site for the helix-loop-helix protein CPF1. Mol. Cell. Biol. 1991;

11(7):3545-3553.

3. Krassovsky K, Henikoff J, Henikoff S. Tripartite organization of centromeric chromatin in budding yeast. Proc. Natl. Acad. Sci. USA 2012; 109(1):243-248.

4. Lechner J, Carbon J. A 240 kd multisubunit protein complex, CBF3, is a major component of the budding yeast centromere. Cell 1991; 64(4):717-725.

5. Stemmann O, Lechner J. The *Saccharomyces cerevisiae* kinetochore contains a cyclin-CDK complexing homologue, as identified by *in vitro* reconstitution. EMBO J. 1996; 15(14):3611-3620.

6. Meluh PB, Koshland D. Evidence that the *MIF2* gene of *Saccharomyces cerevisiae* encodes a centromere protein with homology to the mammalian centromere protein CENP-C. Mol. Biol. Cell 1995; 6(7):793-807.

7. Cohen RL, Espelin CW, De Wulf P, Sorger PK, Harrison SC, Simons KT. Structural and functional dissection of Mif2p, a conserved DNA-binding kinetochore protein. Mol. Biol. Cell 2008; 19(10):4480-4491.

8. Cumberledge S, Carbon J. Mutational analysis of meiotic and mitotic centromere function in *Saccharomyces cerevisiae*. Genet. 1987; 117(2):203-212.

9. Ohkuni K, Kitagawa K. Endogenous transcription at the centromere facilitates centromere activity in budding yeast. Curr. Biol. 2011; 21(20):1695-1703.

10. Furuyama S, Biggins S. Centromere identity is specified by a single centromeric nucleosome in budding yeast. Pro. Natl. Acad. Sci. USA 2007; 104(37):14706-14711.

11. Catania S, Pidoux AL, Allshire RC. Sequence features and transcriptional stalling within centromere DNA promote establishment of CENP-A chromatin. PLoS Genet. 2015; 11(3):e1004986.

12. Matsuoka M, Matsubara M, Daidoh H, Imanaka T, Uchida K, Aiba S. Analysis of regions essential for the function of chromosomal replicator sequences from *Yarrowia lipolytica*. Mol. Gen. Genet. 1993; 237(3):327-333.

13. Fournier P, Abbas A, Chasles M, Kudla B, Ogrydziak DM, Yaver D, Xuan J-W, Peito A, Ribet A-M, Feynerol C, He F, Gaillardin C. Colocalization of centromeric and replicative functions on autonomously replicating sequences isolated from the yeast *Yarrowia lipolytica*. Proc. Natl. Acad. Sci. USA 1993; 90(11):4912-4916.

14. Vernis L, Abbas A, Chasles M, Gaillardin CM, Brun C, Huberman JA, Fournier P. An origin of replication and a centromere are both needed to establish a replicative plasmid in the yeast *Yarrowia lipolytica*. Mol. Cell. Biol. 1997; 17(4):1995-2004.

15. Vernis L, Poljak L, Chasles M, Uchida K, Casaregola S, Kas E, Matsuoka M, Gaillardin C, Fournier P. Only centromeres can supply the partition system required for *ARS* function in the yeast *Yarrowia lipolytica*. J. Mol. Biol. 2001; 305(2):203-217.

16. Yamane T, Sakai H, Nagahama K, Ogawa

T, Matsuoka M. Dissection of centromeric DNA from yeast *Yarrowia lipolytica* and identification of protein-binding site required for plasmid transmission. J. Biosci. Bioeng. 2008; 105(6):571-578.

17. Yamane T, Ogawa T, Matsuoka M. Derivation of consensus sequence for protein binding site in *Yarrowia lipolytica* centromere. J. Biosci. Bioeng. 2008; 105(6):671-674.

18. Stoyan T, Gloeckner G, Diekmann S, Carbon J. Multifunctional centromere binding factor 1 is essential for chromosome segregation in the human pathogenic yeast *Candida glabrata*. Mol. Cell. Biol. 2001; 21(15):4875-4888.

19. Cai M and Davis R. Yeast centromere binding protein CBF1, of the helix-loop-helix protein family, is required for chromosome stability and methionine prototrophy. Cell 1990; 61(3):437-446.

20. Baker RE, Masison DC. Isolation of the gene encoding the *Sacchromyces cerevisiae* centromere-binding protein CP1. Mol. Cell. Biol. 1990; 10(6):2458-2467.

21. Mellor J, Jiang W, Funk M, Rathjen J, Barnes CA, Hinz T, Hegemann JH, Philippsen P. CPF1, a yeast protein which functions in centromeres and promoters. The EMBO J. 1990; 9(12):4017-4026.

22. Halverson D, Baum M, Stryker J, Carbon J, Clarke L. A centromere DNA-binding protein from fission yeast affects chromosome segregation and has homology to human CENP-B. J. Cell Biol. 1997; 136(3):487-500.

23. d'Alencon E, Negre N, Stanojcic S, Alassoeur B, Gimenez S, Leger A, Abd-Alla A, Juliant S, Fournier P. Characterization of a CENP-B hololog in the holocentric Lepidoptera *Spodoptera frugiperda*. Gene 2011; 485(2):91-101.

24. Dowell SJ, Tsang SH, Mellor. The centromere and promoter factor 1 of yeast contains a dimerization domain located carboxy-terminal to the bHLH domain. Nuc. Acids Res. 1992; 20(16):4229-4236.

25. Yoda K, Kitagawa K, Matsumoto H, Muro Y, Okazaki T. A human centromere protein, CENP-B, has a DNA binding domain containing four potential α helices at the NH_2 terminus, which is separable from dimerizing activity. J. Cell Biol. 1992; 119(6):1413-1427.

26. Niedenthal RK, Sen-Gupta M, Wilmen A, Hegemann JH. Cpf1 protein induced bending of yeast centromere DNA element I. Nuc. Acids Res. 1993; 21(20):4726-4733.

27. Tanaka Y, Nureki O, Kurumizaka H, Fukai S, Kawaguchi S, Ikuta M, Iwahara J, Okazaki T, Yokoyama S. Crystal structure of the CENP-B protein-DNA complex: the DNA-binding domains of CENP-B induce kinks in the CENP-B box DNA. The EMBO J. 2001; 20(23):6612-6618.

28. Nakazawa T, Motoyama T, Horiuchi H, Ohta A, Takagi M. Evidence that part of a centromeric DNA region induces pseudohyphal growth in a dimorphic yeast, *Candida maltose*. J. Bacteriol. 1997; 179(16):5030-5036.

