Highly efficient gene expression based on chromatin engineering

クロマチン構造の人為的改変による遺伝子の高効率発現

Studies on Molecular Genetics
Major in Integrative Bioscience and Biomedical Engineering
Graduate School of Science and Engineering
Waseda University

Jun-ichi Tanase
棚瀬 潤一

2010年4月
Contents

Abbreviations ................................................................................................................. 1

Abstract ......................................................................................................................... 3

I. Introduction ................................................................................................................. 5

Chapter 1

Competence of T20 as a transcriptional activator in mouse ES cells and hepatocytes differentiated from them

II. Materials and methods ........................................................................................... 12

II-1. Plasmid construction ............................................................................................ 12

II-2. Generation of mouse ES cell lines ........................................................................ 15

II-3. Determination of transgene loci .......................................................................... 17

II-4. Cell differentiation ................................................................................................. 19

II-5. Quantitative real-time PCR analysis .................................................................... 21

II-6. DNase I footprinting assay .................................................................................. 22

II-7. ChIP assay ............................................................................................................. 23

III. Results ..................................................................................................................... 25

III-1. Effect of T20 on transcription in the genome chromatin context ....................... 25

III-2. Chromatin structure on the reporter promoter ................................................... 35
IV. Discussion .................................................................................................................................. 41

IV-1. Effect of T20 on transgene transcription before and after differentiation of mouse ES cells ......................................................................................................................... 41

IV-2. Reporter locus and T20 effect ............................................................................................... 45

Chapter 2

Activation mechanism of chromatin transcription by superhelically curved DNA segments

V. Materials and methods .............................................................................................................. 50

V-1. Plasmids and strains ............................................................................................................. 50

V-2. β-galactosidase assay ........................................................................................................... 51

V-3. MNase digestion-based analysis of chromatin structure .................................................... 52

V-4. DNase I footprinting assay .................................................................................................. 53

V-5. ChIP assay ............................................................................................................................ 54

VI. Results ..................................................................................................................................... 55

VI-1. Effect of superhelically curved DNA on transcription in yeast ........................................ 55

VI-2. Positioning of nucleosomes on and around the promoter ................................................. 58

VI-3. Confirmation of the nucleosome presence ......................................................................... 63

VI-4. Accessibility of DNA in chromatin .................................................................................... 67

VII. Discussion ................................................................................................................................ 74
VIII. Acknowledgement .............................................................................................. 78

IX. References ........................................................................................................... 79
Abbreviations

*Alb*  albumin gene
*Alox5ap*  arachidonate 5-lipoxygenase-activating protein gene
*Cdk3*  cyclin-dependent kinase 3 gene
*ChIP*  chromatin immunoprecipitation
*CK18*  cytokeratin 18
*CYC1*  iso-1-cytochrome c gene
*Cyp7a1*  cholesterol 7 alpha hydroxylase gene
*DNase I*  deoxyribonuclease I
*ES*  embryonic stem
*FGF*  fibroblast growth factor
*GAPDH*  glyceraldehyde-3-phosphate dehydrogenase
*GFP*  green fluorescent protein
*HGF*  hepatocyte growth factor
*HSV*  herpes simplex virus
*IPCR*  inverse polymerase chain reaction
*lacZ*  β-galactosidase gene
*LIF*  leukemia inhibitory factor
*MNase*  micrococcal nuclease
*Neo*<sup>r</sup>  neomycin phosphotransferase gene
*Oct4*  POU domain, class 5, transcription factor 1 gene
*PEPCK*  phosphoenolpyruvate carboxykinase 1 gene
*Rps18*  ribosomal protein S18 gene
RT-PCR  reverse transcription polymerase chain reaction
*TAT* tyrosine aminotransferase gene
TBP  TATA binding protein
*Tgfbr3* transforming growth factor β receptor III gene
*tk* thymidine kinase gene
UAS  upstream activating sequence
*URA3* orotidine-5' phosphate decarboxylase gene
Abstract

Curved DNA structures with a left-handed superhelical conformation can strongly activate transcription of transgenes in HeLa cells and COS-7 cells. T20 is an artificial 180 bp curved DNA segment of the kind that serves as a transcriptional activator in these cells. In the current study, we firstly investigated the effect of T20 on transcription in mouse embryonic stem (ES) cell lines or hepatocytes differentiated from them. I established ten sets of cell lines each harboring a single copy of the reporter construct. Each set comprised a pair of a T20-harboring cell line and a T20-less control cell line. Analyses showed that in ES cells and in hepatocytes originating from these cells, T20 both activated and repressed transcription in a manner that was dependent on the locus of reporter. The current and previous studies strongly suggest that in cells that have a strict gene regulation system, transcriptional activation by T20 occurs only when the reporter is positioned in a transcriptionally active locus in the genome.

We further studied the effect of curved DNA structures with a yeast mini-chromosome system. Using 108, 180 and 252 bp synthetic curved DNA segments, the mechanism of transcriptional activation by curved DNA structures with a left-handed superhelical conformation has been studied. Even in the presence of nucleosomes, these DNA segments activated transcription from a UAS-deleted CYC1 promoter that is silenced in the presence of nucleosomes. The fold-activations by these segments, relative to the controls that lacked such segments, were 51.4, 63.4 and 56.4,
respectively. The curved DNA structures with a left-handed superhelical conformation favored nucleosome formation. Interestingly, however, the translational positions of the nucleosomes were dynamic. The high mobility of the nucleosomes on the superhelically curved DNA structures seemed to influence the mobility of the nucleosomes formed on the promoter, and eventually enhanced the access to the center region of one TATA sequence. Functioning as a dock for the histone core and allowing nucleosome sliding seem to be the mechanisms underlying the transcriptional activation in chromatin by curved DNA structures with a left-handed superhelical conformation.

These two studies should provide important clues for designing and constructing artificial chromatin modulators, as a tool for chromatin engineering.
I. Introduction

Multifarious DNA structures are found in the genome. Their implication in DNA packaging and gene expression has long been argued. DNA with a curved trajectory of its helix axis is called bent DNA, or curved DNA. Interestingly, they frequently occur in biologically important regions such as origins of DNA replication, regions that regulate transcription, and recombination loci, and are found in a wide variety of cellular and viral genomes from bacteria to man (1-18). In 1980’s and early in 1990’s, a vast many studies have been carried out to understand their biological significance and, as the result, their essential roles seem to be largely understood (19-22).

Curved DNA structures are often implicated in transcriptional regulation in both prokaryotes and eukaryotes. In prokaryotes, they are often located from immediately upstream of the -35 hexamer to around position -100 relative to the transcription start site (+1) (23-37). They have a range of functions: facilitating RNA polymerase binding to the promoter, transition from closed to open promoter complexes, or transcription factor binding. To perform these functions, in some cases intrinsically curved DNA structures function together with DNA bends that are induced by binding of RNA polymerase, transcription factors, or nucleoid-associated proteins (23, 38-45).

In eukaryotic transcription, naturally occurring curved DNA structures function in several ways. These include acting as a structural
(conformational) signal for transcription factor binding; juxtaposing the basal transcription machinery with effector domains on upstream-bound factors; regulating transcription in association with transcription-factor-induced bending of DNA; and appropriately organizing the local chromatin structure for transcription initiation (22).

It has been revealed that the shape of the curved DNA is an important factor in transcription. In prokaryotes, right-handed superhelical curvatures activate transcription. These structures help RNA polymerase to bind to promoters and facilitate formation of the open promoter complex (21, 32, 34, 38, 45-51). On the other hand, eukaryotic transcription is sometimes activated by mimicry of negative (left-handed) supercoils; that is, by curved DNA structures with a left-handed superhelical conformation (22, 52-54). Eukaryotic genomes are first folded into nucleosomes, the fundamental unit of chromatin, where DNA forms left-handed supercoils (55, 56). The mimicry of such structures may presumably adjust the local chromatin infrastructure to make it appropriate for transcription initiation and thus can activate transcription (20, 22, 52, 57, 58). Although nucleosomes generally inhibit access or assembly of transcription factors and thus inhibit transcription (59, 60), left-handed superhelical curvatures seem to provide a mechanism to circumvent this problem.

Based on the knowledge described above, curved DNA structures have been thought to provide a basis for “chromatin engineering” for efficient and stable gene expression. The first possibility of chromatin
engineering was found with a synthetic left-handedly curved 36 bp DNA (subsequently named T4) in 2003. This T4 DNA increased the accessibility of the TATA box of the reporter promoter in chromatin, in a transient transfection assay system using COS-7 cells, and activated transcription by about 10-fold (52). Furthermore, when a T4-containing reporter was delivered into mouse liver by a hydrodynamics-based injection, T4 also activated transcription of the reporter gene (53). More interestingly, a recent study found that synthetic left-handedly curved DNA segments of 180 bp (named T20), 216 bp (T24), 252 bp (T28), 288 bp (T32), 324 bp (T36) and 360 bp (T40) are strong activators of transcription (61). These segments activated transcription from the herpes simplex virus (HSV) thymidine kinase (tk) promoter by about 70 to 140-fold in a transient transfection assay system in COS-7 cells. The effect of T20 on transcription in the genome chromatin context was also examined by establishment of five HeLa cell lines each harboring a T20-containing reporter in the HeLa genome and five with T20 deleted from the reporter loci. In the T20-harboring cells, transcription of the reporter gene was activated by about 2.5-, 8-, 11-, 21-, and 90-fold, respectively, compared to the control cell lines (61).

With these backgrounds of curved DNA research, the current study aimed to further develop chromatin engineering. Although it was clarified that the curved DNA segments with a left-handed superhelical conformation are definitely effective in some cells, there remain many
issues to be clarified at the very first step of chromatin engineering. Since the results obtained to date show that T20 is a strong activator of transcription, the current study firstly investigated if T20 exerts similar effects in mouse ES cells and whether these effects are maintained after cell differentiation. Secondly, this study investigated the activation mechanism of transcription by T20 using yeast mini-chromosome system (62). Heretofore, the mechanism has not been clarified, except for the case of the short segment T4 (61). This is mainly due to the experimental difficulty in using mammalian cells for fine analyses of chromatin architectures. The principal cause is the size of mammalian genomes. Mammalian genomes are very large and thus require a large amount of DNA to obtain a sufficient molarity for the analyses, which reduces the resolution. Another problem is the lack of convenient episomal systems for analyses of chromatin structure.

The experimental results showed that the chicken $\beta$-actin promoter, a well-known strong promoter, could be even activated by T20 in mouse genomic chromatin in both ES cells and the hepatocytes differentiated from them. The activation, however, occurred in a manner that was dependent on the locus of reporter. It was also shown that T20 can markedly activate transcription from a UAS-deleted CYC1 promoter in yeast chromatin. Functioning as a dock for the histone core and allowing nucleosome sliding seemed to be the mechanisms underlying the transcriptional activation by T20. The current study provides important clues for designing and
constructing artificial chromatin modulators, as a tool for chromatin engineering.
Chapter 1

Competence of T20 as a transcriptional activator in mouse ES cells and hepatocytes differentiated from them
II. Materials and Methods
II-1. Plasmid construction

Plasmid pLHC20/loxP/neo/SL constructed from pCX-GFP, pLHC20/loxP/cbact/-4 and pLHC20/loxP/TLN-6 (61) was used to establish mouse ES integrants.

Construction of pLHC20/loxP/cbact/-4

Plasmid pLNX (61) was digested with SalI and XhoI to obtain a fragment containing a loxP sequence. This fragment was inserted into the SalI site of pCX-GFP. The resulting construct was digested with SnaBI and ligated to the EcoRI-SacII fragment of pLNX, which also contains a loxP sequence. The resulting plasmid was digested with AflIII and SacI, and filled with the T20-containing KpnI-NruI fragment of pLHC20/ELN (61) to generate pLHC20/loxP/cbact.

Plasmid pLHC20/loxP/cbact/-4 is a variant of pLHC20/loxP/cbact with a deletion of 4 base pairs between T20 and the downstream loxP sequence of pLHC20/loxP/cbact. First, a variant fragment was prepared by PCR using pLHC20/loxP/cbact and the primers 5’-TAA ACA AAT AGG GGT TCC GC-3’ and 5’-ACT CGA GGG CCC ATA TGA CG-3’. The PCR product was digested with SalI and inserted between the SalI and SmaI sites of pLHC20/loxP/cbact.

Construction of pLHC20/loxP/neo/SL
Plasmid pLHC20/loxP/TLN-6 (61) was digested with \( HincII \) and the fragment containing the neomycin-resistant gene was inserted into the \( HincII \) site of pUC19 to generate pUC19neo. The \( SalI \) and \( HindIII \) fragment of pUC19neo was then inserted between the corresponding sites of pLHC20/loxP/cbact/-4 to generate pLHC20/loxP/cbact/-4/neo. We adopted the following procedure to delete the extra base pairs between T20 and each \( loxP \) site in pLHC20/loxP/cbact/-4/neo. Initially, the region spanning from the \( EcoT22I \) site to the downstream \( loxP \) sequence in the plasmid was amplified by PCR using the primers 5’-ATC TGA GCT CGT CAA TGC ATT TTT CA-3’ and 5’-CAT CGC TGC ACA AAA TAA TT-3’. The resulting fragment was digested with \( BglII \) and inserted between the \( EcoRV \) and \( BglII \) sites in the plasmid, which generated pLHC20/updel. Next, the region spanning from the \( EcoT22I \) site to the upstream \( loxP \) site in pLHC20/loxP/cbact/-4/neo was amplified by PCR using the primers 5’-GCA GGT TTA AAC AAT GCA TAT AAC-3’ and 5’-ATC TCG TCG TGA CCC ATG GC-3’. The resulting fragment was digested with \( CpoI \) and \( EcoT22I \) and ligated between the corresponding sites in pLHC20/updel. The resulting construct was digested with \( EcoT22I \), blunted with T4 DNA polymerase and subjected to self-ligation, which generated pLHC20/upcomp. This construct was digested with \( NotI \), blunted with S1 nuclease, and digested with \( SalI \). The \( SalI \) and \( PmaCl \) fragment of pLHC20/upcomp (containing T20 and the \( loxP \) sequence) was inserted between these sites to generate pLHC20/comp. Finally, to construct
multiple cloning sites upstream of the upstream $loxp$ site in pLHC20/comp, the synthetic oligonucleotides 5’-TCG ACG CGT CGC GAT CGA TTA ATT AAC TCG AGC TAG CTA CGT A-3’ and 5’-TCG ATA CGT AGC TAG CTC GAG TTA ATT AAT CGA TCG CGA CGC G-3’ were annealed and inserted into the $SalI$ site of pLHC20/comp to generate pLHC20/$loxp$/neo/SL.
II-2. Generation of mouse ES cell lines

Mouse ES cells were maintained in Glasgow minimum essential medium (G-MEM) supplemented with 10% fetal bovine serum, 0.1 mM 2-mercaptoethanol, 0.1 mM non-essential amino acids, 1 mM sodium pyruvate and 1,000 units/ml of leukemia inhibitory factor (Chemicon, Temecula, CA, USA) on gelatin-coated dish without feeder cells at 37°C in 5% CO₂. They were electroporated with 300 ng of SalI-digested pLHC20/loxP/neo/SL at 300 V/150 µF in a cuvette containing 1×10⁶ cells in 250 µl of PBS (-) (63-69). Twenty-four hours after electroporation, G418 (Sigma-Aldrich, St. Louis, MO, USA) was added at a final concentration of 200 µg/ml and the culture was continued for 10 days. Colonies of surviving cells were isolated and cultured again for 5 days in the presence of 0.4 mg/ml G418 to establish the cell lines.

Integration of the reporter construct was confirmed by PCR using the primers 5’-GAC AAT CGG CTG CTC TGA TG-3’ and 5’-TGC GAT GTT TCG CTT GGT GG-3’, which generated a 414-bp fragment when the reporter was present. Colonies harboring a single reporter were selected by Southern blot analysis. We established ten cell lines that were named MES6/T20, MES7/T20, MES25/T20, MES27/T20, MES32/T20, MES40/T20, MES62/T20, MESA2/T20, MESB2/T20 and MESn36/T20. To establish control cell lines, the cells were transfected with pBS185 (Invitrogen, Carlsbad, CA, USA), which expresses Cre recombinase. After cell cloning, T20-deleted clones were selected based on PCR analysis using
the primers described above. The established control cell lines were named
MES6, MES7, MES25, MES27, MES32, MES40, MES62, MESA2, 
MESB2 and MESn36, respectively.
II-3. Determination of transgene loci

Transgene loci were determined using IPCR (Fig.1) (70). Initially, 500 ng of genomic DNA was digested with NcoI or PstI. Each digest was purified and self-ligated in a 500 µl reaction mixture containing 500 ng of the digest, 10 units of T4 DNA ligase (TaKaRa, Otsu, Japan), 1 mM ATP and the ligation buffer at 16°C for 15 hr. After ligation, PCR was performed using the primers 5’-ACC GCT TCC TCG TGC TTT A-3’ and 5’-CCA ACG CTA TGT CCT GAT AG-3’. Nested PCR was carried out using the primers 5’-TTT ACG GTA TCG CCG CTC CC-3’ and 5’-TCC TGA TAG CGG TCC GCC A-3’. The amplified product was purified on a 0.7% agarose gel and sequenced, and the transgene locus was identified. Each locus was confirmed by checking the consistency with sequence data originating from NcoI or PstI digestion.
Fig. 1. Inverse RCR

Inverse PCR is an efficient method to determine integration locus of a given transgene. (A) Genomic DNA is digested with restriction enzyme X to contain identified and unidentified sequences. (B) The product is circularized by self-ligation. (C, D) PCR is performed using primers extending outwards. The resulting product is sequenced, which clarifies the integration locus.
II-4. Cell differentiation

The integrants were differentiated into hepatocytes according to the procedure described by Teratani et al. (71), with slight modifications. In step 4 of the reported procedure, we cultured cells in Williams medium E containing insulin (5 µg/ml), dexamethasone (100 nM), non-essential amino acids (0.1 mM), l-glutamine (2 mM) and fetal bovine serum (5%) (Fig. 2). Differentiation into hepatocytes was confirmed by detecting expression of hepatocyte-specific genes such as Alb, Cyp7a1, PEPCK and TAT using RT-PCR, and CK18 with immunostaining (72). The RT-PCR was performed as follows. RNA was purified using a conventional method and cDNA was synthesized using 3 µg of total RNA, RevaTra Ace (Toyobo, Osaka, Japan) and oligo(dT). The primers used were as follows: for Alb, 5'-CAG GAT TGC AGA CAG ATA GTC-3' and 5'-GCT ACG GCA CAG TGC TTG-3'; for Cyp7a1, 5'-CCA CCT TTG ATG ACA TGG AGA AG-3' and 5'-TTC TTC TTC AGA GGC TGC TTT CAT TG-3'; for PEPCK, 5'-CAA GTG CCT GCA CTC TGT GG-3' and 5'-CCA CCA TAT CCG CTT CCA AA-3'; for TAT, 5'-ACC TTC AAT CCC ATC CGA-3' and 5'-TCC CGA CTG GAT AGG TAG-3'. For immunostaining, cultured cells were fixed in 4% paraformaldehyde at room temperature (RT) for 10 min and then washed three times with 0.1% Triton/PBS (-) at RT for 10 min. Subsequently, the cells were soaked in 5% skimmed milk/PBS (-) at RT for 1 hr, and then washed as above. Cells were incubated with goat anti-mouse CK18 (1:100) (Santa Cruz Biotechnology,
Santa Cruz, CA, USA) at 4°C for 12 hr. After washing, the cells were treated with 0.1% BSA/PBS (-) containing donkey anti-goat IgG-FITC (1:250) (Santa Cruz Biotechnology) at RT for 1 hr, washed as described above, and sealed with Vectashield Mounting Medium with DAPI (Vector, Burlingame, CA, USA).

Fig. 2. The method for inducing differentiation of mouse ES cells into hepatocytes

Cell differentiation was completed in 16 days. About 60-70% mouse ES cells differentiated into hepatocytes by this method.
II-5. Quantitative real-time PCR analysis

GFP mRNA was purified from ES cells or hepatocytes and cDNA was synthesized as described above. The product was quantified by SYBR Green with normalization against the housekeeping gene \textit{Rps18}, using a StepOne Plus real-time PCR system and StepOne software v. 2.1 (Applied Biosystems, Carlsbad, CA, USA). The following primers were used: for GFP mRNA, 5’-TTG GCG ATG GCC CTG TC-3’ and 5’-ACC TGA GGA GTG AAT TC-3’; for \textit{Rps18} mRNA, 5’-CCT GAG AAG TTC CAG CAC AT-3’ and 5’-TTC TCC AGC CCT CTT GGT G-3’.
II-6. DNase I footprinting assay

Nuclei were prepared as described previously (5, 52). Aliquots of 100 μl nuclear suspension (1×10^7 nuclei/ml) were incubated at 37°C for 1 min. Then, 0.05 U, 0.1 U or 0.3 U of DNase I was added and digestion was performed at 37°C for 2 min. Naked DNA was digested with 0.001 U, 0.003 U or 0.005 U of DNase I. After digestion, the products were purified and dissolved in water at 1 mg/ml. To detect DNase I cleavage sites, ligation-mediated PCR (73) was carried out using 3 μg DNase I digests as follows. A primer DNA of sequence 5’-TCA CCT GTG GGA GTA ACG CG-3’ was hybridized to the digests and extended to the cleavage sites with KOD-plus (Toyobo, Osaka, Japan) at 68°C for 10 min. Then, a 25-bp linker DNA was ligated to the resulting products and PCR was carried out using the primers 5’-GCG GTG ACC CGG GAG ATC TGA ATT C-3’ and 5’-AAC GC G GTC AGT CAG AGC CG-3’ under the following conditions: 98°C for 2 min and 30 cycles at 98°C for 30 sec, 58°C for 30 sec and 68°C for 2 min. Subsequently, DNase I cleavage sites were detected by PCR-based primer extension using a [5’-32P]-labeled primer, 5’-TCA GTC AGA GCC GGG GCG GG-3’, under the following conditions: 98°C for 2 min and 20 cycles at 98°C for 30 sec, 63°C for 5 sec and 74°C for 30 sec. All samples were purified and resolved in 6% polyacrylamide-7 M urea gels.
II-7. ChIP assay

The ChIP assay was performed according to Chadee et al. (74) with slight modifications. Nuclei were prepared as reported by Nishikawa et al. (52) and aliquots of 50 µl nuclear suspension (4 ×10^7 nuclei/ml) were digested with 10 U of MNase at 37°C for 15 min. Then, 50 µl of 2× lysis buffer comprising 150 mM NaCl, 25 mM Tris-HCl (pH 7.5), 5 mM EDTA, 1% Triton X-100, 0.1% SDS, and 0.5% sodium deoxycholate was added to the reaction mixture and incubated at 4°C for 10 min. Immunoprecipitation was carried out in triplicate using a OneDay ChIP Kit (Diagenode, Liège, Belgium) according to the manufacturer’s instructions, using antibody against histone H3 (ab1791; Abcam, Cambridge, UK). Using 2 µl of the DNA samples thus obtained, quantification of fragments containing a part of the chicken β-actin promoter (R1, R2 and R3) was carried out by SYBR green quantitative real-time PCR with the StepOne Plus real-time PCR system and StepOne software v. 2.1. Quantification of fragments containing a part of GAPDH was carried out simultaneously as an internal control. The primer sets were as follows: for R1, 5’-TCT CCG TAT TAG TCA TCG C-3’ and 5’-CAT CGC TGC ACA AAA TAA TT-3’; for R2, 5’-AAT TAT TTT GTG CAG CGA TGG -3’ and 5’-CGC CTC GCC ATA AAA GGA AAC T-3’; for R3, 5’-GCG CTC CGA AAG TTT CC T TTT A-3’ and 5’-TCA CCT GTG GGA GTA ACG CG-3’; for GAPDH, 5’-CCG CAT CTT CTT GTG CAG T-3’ and 5’-TCC CTA GAC CCG
TAC AGT GC-3’. The relative sample enrichment was calculated using the formula: $2^{(Ct\ target\ region-Ct\ GAPDH)}$. 
III. Results

III-1. Effect of T20 on transcription in the genome chromatin context

The reporter construct is shown in Fig. 3. The T20 sequence was placed between two loxP sequences (75-82) and the resulting sequence was placed upstream of the chicken β-actin (82) promoter linked to the GFP gene. The loxP sequences were used to establish control cell lines with a reporter gene without T20. After the construct was cleaved at the SalI site, it was introduced into mouse ES cells. We screened for cell lines with a single copy of the reporter construct using Southern blot analysis (Fig. 4). The following ten cell lines were established: MES27/T20, MES6/T20, MES62/T20, MES32/T20, MES25/T20, MESB2/T20, MES7/T20, MESA2/T20, MESA36/T20 and MES40/T20. The reporter was found integrated immediately upstream of a gene (MES32/T20, MES25/T20, MES7/T20; referred to as group A) or within a structural gene (MES62/T20, MES40/T20; group B) or in an intergenic region (MES27/T20, MES6/T20, MESB2/T20, MESA2/T20, MESA36/T20; group C) (Fig. 5). Control cell lines were established by expressing bacteriophage P1 Cre recombinase in the T20 cell lines. These were named MES27, MES6, MES62, MES32, MES25, MESB2, MES7, MESA2, MESA36 and MES40, respectively (Fig. 6). Differentiation of each cell line into hepatocytes was confirmed by the expression of genes such as Alb, Cyp7a1, PEPCK, TAT and CK18 (Fig. 7, 8) (71).
Expression of the reporter gene (GFP) was analyzed by RT-PCR in undifferentiated and differentiated cells, and GFP mRNA levels were compared between T20-harboring and T20-less cells. As shown in Fig. 8A, in which the data are aligned in the order of activation, four T20-harboring ES cell lines showed higher transcription (MES32/T20, MES25/T20, MESA2/T20, MES7/T20) and three showed lower transcription (MES27/T20, MES40/T20, MES6/T20) compared with respective control cell lines. The fold activations of transcription relative to controls in the four cell lines with higher transcription were 2.7, 1.9, 1.7 and 1.5, respectively, and three of these cell lines were in group A. Since the chicken β-actin promoter is a strong promoter (83-85), these activations are not necessarily small (discussed below). In MES32/T20, MES25/T20 and MES7/T20, the reporter construct was integrated immediately upstream of a gene: upstream of Tgfbr3 in MES32/T20; Alox5ap in MES25/T20; and Cdk3 in MES7/T20 (Fig. 5). These genes are expressed in ES cells or blastocysts (86-88). The transcription level relative to the control was unchanged in MESn36/T20 and MESB2/T20. Transcription of the reporter gene was not detected in MES62/T20 and MES62; thus, these data are not shown in Fig. 8A. The RIKEN cDNA 4930564D02 gene may not be expressed in ES cells. Except for MESA2/T20, transcriptional stimulation was not observed when the T20-harboring reporter was integrated into an intergenic region or within a structural gene.
Upon cell differentiation into hepatocytes, the effect of T20 greatly changed on the whole (Fig. 8B). For comparison, the data are shown in the same order as those in Fig. 8A. Among the differentiated MES32/T20, MES25/T20, MESA2/T20 and MES7/T20, which showed higher GFP mRNA levels than their controls before differentiation, only MES32/T20 maintained the similar fold of activation of GFP transcription relative to the control. The value slightly fell for MES7/T20 and it was 0.59 for MESA2/T20, which was the strongest transcriptional repression caused by T20. Using RT-PCR, we also clarified that the Alox5ap and Cdk3 were expressed in the hepatocytes differentiated from the ES cells (Fig. 9). The Tgfbr3 expression in hepatocytes was previously reported (89).
Fig. 3. Reporter construct and architecture of T20

(A) Reporter construct. GFP, Neo', pgk, and loxP indicate genes encoding green fluorescent protein, neomycin phosphotransferase, phosphoglycerate kinase, and loxP sequence, respectively. Nucleotide sequence of T20 is slightly different from that in the previous report (61): the sequence started with 5’-TCAGTTTTT and ended with TTTTT-3’ in the previous T20, but the corresponding sequences are TTTTT and TTTTTTCACG-3’ in the present T20. Since they have the same three-dimensional architecture, the same name was used for convenience. (B) Three-dimensional architecture of T20. The figure was drawn as reported by Sumida et al. (61). The white line indicates the superhelical axis.
**Fig. 4. Southern blot analysis of the reporter construct**

Genomic DNAs from the established cell lines were digested with restriction enzymes indicated in the figure. After separation by agarose gel electrophoresis and blotting, each digest was hybridized with a probe indicated below the autoradiograms.
The genomic region harboring the reporter construct was identified as described in the Materials and Methods. Loci of integration are indicated with arrowheads.
Fig. 6. Establishment of T20-less cell lines

Deletion of T20 from the reporter locus was confirmed by PCR. “M” indicates the size marker lane. The amplified region is shown on the right.
**Fig. 7. Differentiation of mouse ES cells into hepatocytes**

(A) RT-PCR analyses of mRNA expression of hepatocyte-specific genes *Alb*, *Cyp7a1*, *PEPCK*, and *TAT*. *Oct4* is an undifferentiated cell marker and *Rps18* is a housekeeping gene encoding the ribosomal protein S18. The latter was used as an internal control. (B) Immunofluorescent detection of expression of CK18, a marker for hepatocytes, using anti-mouse CK18 antibody (green). Nuclei were stained with DAPI (blue).
Fig. 8. Effect of T20 on the transcription of reporter gene before or after cell differentiation.

(A) Effect of T20 on transcription of reporter gene in mouse ES cells. At first, expression level of reporter mRNA in each cell line was determined using that of Rps18 mRNA for normalization. The determinations were carried out in quintuplicate and the mean values ± SD were calculated. Then, the relative GFP mRNA expression (+T20/-T20) was calculated by dividing the expression level of reporter mRNA in each T20-harboring cell line by that in the corresponding control (T20-less) cell line (only mean values were used in the calculation). The dotted, lined and striped bars indicate groups A, B and C, respectively. These groups were established based on the integration locus (see text). (B) Effect of T20 on transcription of reporter gene in hepatocytes differentiated from respective integrants.
Fig. 9. Expression of the gene adjacent to a transgene
RT-PCR was used for the analysis using undifferentiated and differentiated states of cells. All genes were found to be expressed in both states. (A) Cdk3 expression in the cell lines MES7/T20 and MES7; (B) Alox5ap expression in the cell lines MES25/T20 and MES25; (C) Tgfbr3 expression in the cell lines MES32/T20 and MES32.
III-2. Chromatin structure on the reporter promoter

The local chromatin structure on the reporter promoter was analyzed by ChIP assay in MES32/T20 and MESA2/T20 cells and their controls. The assay targeted histone H3. MES32/T20 was chosen because the reporter gene was located close to an active gene and transcription of the reporter gene was most activated by T20. MESA2/T20 was chosen because the reporter gene was integrated into an intergenic region and because T20 acted both positively (before differentiation) and negatively (after differentiation) (Fig. 8). The results of the assay are shown in Fig. 10, where the histone level is represented as fold enrichment relative to the level in a region within the GAPDH gene.

For the MESA2/T20 and MESA2 pair (Fig. 10C, D), each region of MESA2/T20 contained less histones in ES cells (Fig. 10C) and more histones in hepatocytes (Fig. 10D) compared to MESA2, which correlated well with the difference in reporter gene expression (Fig. 8). In contrast, for the MES32/T20 and MES32 pair (Fig. 10A, B), only a slight difference in histone levels was observed between the two cell lines in ES cells and hepatocytes; however, the reporter transcription level differed considerably (Fig. 8). Thus, the histone profiles for MESA2/T20 and MESA2 agreed well with the established view that a high histone level has a negative effect on transcription. However, those for MES32/T20 and MES32 could not explain the transcriptional difference between them. To understand the mechanism underlying the greater expression of the reporter gene in
MES32/T20 than in MES32, we then performed DNase I footprinting using isolated nuclei. This assay provides information on the DNA exposed to or shielded from the environment in the complex with proteins including histones. In this experiment, cleavage sites were detected by a combination of ligation-mediated PCR and primer extension (73).

No 10-bp ladder was detected in the footprinting assay (Fig. 11), indicating at least that rotationally positioned nucleosomes were not present in the region analyzed (52). In this assay, the direction of primer extension was towards the region upstream of the reporter gene. MES32/T20 and MES32 had very similar band patterns under the same state of differentiation. Before differentiation of these cells, the extended bands were generally short and remained about 15 to 30 bp upstream of the TATA box (lane 5 in each gel). Furthermore, the extent of digestion of these samples was similar to those of naked DNA samples. Thus, it was suggested that the local chromatin formed on the promoter in each cell line was relatively free from stably bound proteins before differentiation. However, considering that the bands were clearer for MES32/T20 than for MES32, MES32/T20 may have a more ordered complex with some proteins in the TATA box region compared to MES32 in the undifferentiated state, which may explain the higher transcription in the former. After differentiation of MES32/T20 and MES32, the extended fragments reached far upstream from the TATA box, near to the end of the gel (around position -110; lane 8 in each gel) and generated many discrete
bands, indicating that the major population of promoter chromatin formed an ordered structure in the hepatocytes. After differentiation, the reporter gene transcription increased dramatically in these two cell lines (Fig. 12). Thus, the ordered chromatin structure seems to have been indeed implicated in the transcriptional activation. The extended fragments in lane 8 in each gel were slightly longer and clearer for MES32/T20 than for MES32. This difference may also explain the higher transcription in the former. The emerging bands were similar in the region downstream from the transcription start site (+1) in ES cells and hepatocytes, indicating at least that downstream chromatin was not influenced by differentiation in both cell lines.
Fig. 10. Comparison of the distribution of histone H3 in T20-harboring and T20-less promoters

ChIP assays were performed using MES32/T20, MES32, MESA2/T20 and MESA2. The results for T20-less promoters are indicated with empty bars and those for T20-harboring promoters are indicated with filled bars. Values are shown as means ± SD (n = 3) of fold enrichment relative to the GAPDH gene. R1, R2 and R3 are the amplified regions and their locations are indicated at the bottom of the figure. Arrows indicate primer locations. +1 indicates the transcription start site.
Fig. 11. DNase I sensitive sites in the promoter region of the reporter locus

The nuclei of the cell lines MES32/T20 and MES32 were isolated and treated with DNase I. Cleavage sites were detected by a combination of ligation-mediated PCR and primer extension. Lanes labeled N, U-Ch and D-Ch show DNase I cleavage sites in naked DNA, chromatin DNA from ES cells, and chromatin DNA from hepatocytes, respectively. G, G ladder; +1, transcription start site.
Fig. 12. Reporter gene expression in MES32/T20 and MES32 before or after cell differentiation

(A) Expression levels of reporter mRNA in MES32/T20 before (filled bar) and after (empty bar) cell differentiation. The expression level of Rps18 mRNA was used for normalization. Values are shown as means ± SD (n = 5).

(B) Expression levels of reporter mRNA in MES32 before (filled bar) and after (empty bar) cell differentiation. Values are shown as means ± SD (n = 5).
IV. Discussion

IV-1. Effect of T20 on transgene transcription before and after differentiation of mouse ES cells

The effect of T20 on transcription was seemingly small in mouse ES cells (maximum activation of 2.7-fold) and in hepatocytes differentiated from these cells (maximum activation of 2.5-fold) (Fig. 8). These results are in contrast to the previous data obtained in HeLa cells (maximum activation of about 90-fold) using the HSV tk promoter (Fig. 13) (61). However, the chicken β-actin promoter used in the current work is a stronger promoter (83-85), and therefore the extent of activation of this promoter is not necessarily small. As an exogenous promoter, the chicken β-actin promoter has been reported to be stronger than the simian virus 40 early promoter or Rous sarcoma virus long terminal repeat as assayed in mouse L cells (83). Furthermore, we found that the chicken β-actin promoter was 10.4-fold stronger on average \((n = 5)\) than the tk promoter in a transient transcription assay system using COS-7 cells (Fig. 14).

Down-regulation of transcription was also caused by T20 in some cell lines both before and after differentiation, which were seven cases in total (three before and four after differentiation). However, regarding group A cell lines, such phenomenon occurred only in MES25/T20 hepatocytes among six cases (three cell lines \(\times\) two states). Furthermore, transcriptional decrease was slight in this case. Therefore, although regulation of gene expression is likely to be much stricter in ES cells and hepatocytes than in
HeLa cells, T20 seems to activate transcription generally when the reporter is integrated immediately upstream of a gene. This issue is further discussed below. The effect of T20 was also very large in COS-7 cells (61) in transient transfection assays, but this result cannot be compared directly with data in stable integrants.
Fig. 13. Transgene loci in HeLa cell lines and the level of their expression

(A) Reporter loci. The loci of integration are indicated with arrowheads (Sumida et al. *FEBS J.* 2006). The corresponding normal human chromosomes are shown.

(B) Luciferase gene expression in each cell line. The relative luciferase expression (+T20/-T20) is calculated by dividing expression level of the luciferase gene (mean value) in each T20-harboring cell line by that in the corresponding control (T20-less) cell line.
Fig. 14. Strength of the chicken $\beta$-actin promoter as compared with that of the HSV tk promoter

The reporter plasmids were electroporated into COS-7 cells, and luciferase assay was carried out 21 hours after electroporation. The $\beta$-actin promoter was 10.4-fold stronger than tk promoter.
IV-2. Reporter locus and T20 effect

Integration locus seems to have restricted the function of T20 in mouse ES cells and hepatocytes differentiated from these cells. In contrast, T20 did not exert any negative effect on transcription in HeLa cells irrespective of the integration locus (61). The highest activation by T20 in HeLa cells (90-fold) occurred when T20 was located immediately upstream of an active gene (Fig. 13). Thus, when T20 is integrated into transcriptionally active regions in chromatin, it may generally activate transcription even in the cells that have much stricter gene regulation system than HeLa cells. Indeed, for MES32/T20, in which the reporter was integrated upstream of *Tgfbr3*, which is expressed in both blastocysts and liver (87, 89), activation by T20 was maintained after differentiation (Fig. 8B). For MES7/T20, activation by T20 was also maintained in the hepatocytes, although the extent of activation decreased slightly. The neighboring gene in this case was *Cdk3*, which is active in ES cells (88) and in hepatocytes differentiated from the ES cells (Fig. 9). Therefore, these results are consistent with the speculative argument made above. However, activation did not occur in MES25/T20 hepatocytes (fold activation relative to the control was 0.87). Regarding MES25/T20, the neighboring gene (*Alox5ap*) is active not only in ES cells (86) but also in hepatocytes (Fig. 9). A drastic change in chromatin structure on and around the *Alox5ap* gene that was caused by cell differentiation may have had some negative structural influence on the reporter locus in this case.
It was indicated that similar chromatin structures were formed on T20-harboring and T20-less promoters in a transcriptionally active locus under the same state of differentiation (Figs. 10A, B, 11). In such a case, a “basic transcription” (transcription without the effect of T-20) may have occurred at the same level in each cell line in the same state of differentiation since the “shielding” effect by histones was similar in two cell lines. Therefore, in T20-harboring promoters, T20 should have determined the final levels of GFP mRNA expression.

The current study also suggests that chromatin structure formed on intergenic regions may not be strictly organized in ES cells and hepatocytes differentiated from these cells (Fig. 10C, D). This hypothesis is strengthened by the data that the effect of T20 was diverse in group B cell lines (Fig. 8). Thus, in the intergenic regions of mouse ES cells and hepatocytes, the main effector of reporter gene transcription was certainly histone deposition. The different chromatin structures on T20-harboring and T20-less promoters should produce different effects on transcription. However, we do not understand why the same locus forms different chromatin structures. The *loxP* sequence may be implicated in this phenomenon, since *loxP* has a palindromic sequence that can form a cruciform under unwinding stress (77, 90). This cruciform is known to inhibit nucleosome formation (91, 92). Therefore, *loxP* may have influenced chromatin organization in different ways in some
T20-containing and T20-less promoters in intergenic regions. This effect, if any, may have been suppressed in transcriptionally active loci.

The rotational orientation of a curved DNA segment relative to the promoter and its distance from the promoter are important parameters for transcriptional activation (52, 61). However, these were not optimized in the current study, since to prepare control cell lines we had to use the Cre-\textit{loxP} system. The \textit{loxP} sequence of 34 bp between T20 and the $\beta$-actin promoter was a hindrance for phasing and positional optimization of T20. Thus, it is probable that an appropriately positioned T20 will produce greater activation of transcription.
Chapter 2

Activation mechanism of chromatin transcription
by superhelically curved DNA segments
V. Materials and methods
V-1. Plasmids and strains

A fragment containing the T20 segment was prepared by PCR from the plasmid pLHC20/loxP/TLN-6 as a template (61), using the set of primers, 5'-GCG GTA CCC TGG AGC GTC AGT CAG TTT TTC ATG-3' and 5'-GCG GTA CCC TCG AGG TTA TGA TAT CGG TGA AAA A-3'. The PCR fragment was digested with Asp718 and inserted into the KpnI (Asp718) site of pKB112, the multicopy CYC1-lacZ plasmid (93). During the plasmid screening, those harboring a T12 or T28 segment at the same site were also obtained. The plasmids were digested with XhoI to delete the UAS region from the CYC1 promoter, to construct pTM6-U11-5 (T12), pTM44-U4-5 (T20) and pTM67-U8-15 (T28). All of the plasmids thus obtained were verified by DNA sequencing. The plasmid pLGΔ312ΔSS (94), in which the UAS region was deleted from the CYC1-lacZ gene, was used as a control plasmid. These plasmids were introduced into S. cerevisiae strain AMP105 [MATa ho::LYS2 ura3 lys2 leu2::hisG].
V-2. β-galactosidase assay

The β-galactosidase assay was performed with early exponential phase cells grown in SC-Ura medium. The assay was carried out according to the standard protocol (95) with a minor modification: Y-PER (Pierce Chemical, Rockford, IL, USA) was used to prepare permeabilized yeast cells, instead of chloroform. The assay results were the average of three independent cultures of each strain.
V-3. MNase digestion-based analysis of chromatin structure

Yeast cells harboring plasmids were cultured in SC-Ura medium until OD$_{600}$ reached ~1.0. Nuclei were isolated and micrococcal nuclease (MNase) digestion was performed as described previously (96). The cleavage-sites for MNase were analyzed by indirect end-labeling, as described previously (97). Samples were digested with StuI, which cut at +442 (translation start site of $URA3$ gene is +1) of the $URA3$ gene in the plasmids. The products were separated on a 1.5% agarose gel prepared in 1×TBE buffer, transferred onto a Hybond-XL membrane (GE-Healthcare, Little Chalfont, UK), and detected by a radioactively labeled probe (201 bp) corresponding to the region from +442 to +642 in the $URA3$ gene. The nucleosome repeat assay was performed as described (97). The MNase digested samples were separated on a 1.3% agarose gel, transferred onto the Hybond-XL membrane, and detected by radioactively-labeled probes corresponding to the region of the $CYC1$ promoter, the T12, T20, or T28 segments, and the $URA3$ region. The probe for the promoter was prepared by using primers 5’-ATG GCC AGG CAA CTT TAG T-3’ and 5’-GCT ACA AAG GAC CTA ATG TAT AAG GAA-3’. The primer sets used to prepare the other probes are described in the ChIP assay section.
V-4. DNase I footprinting assay

Aliquots (200 µl) of the nuclear suspension were incubated at 37°C for 1 min. Then, 0.01 U, 0.1 U or 1 U of DNase I was added, and digestion was performed at 37°C for 2 min. Naked DNA was digested with 0.0001 U, 0.002 U or 0.003 U of DNase I at 37°C for 2 min. After digestion, the products were purified and dissolved in water. The DNase I cleavage sites were detected by PCR-based primer extension (52), using a [5’-32P]-labeled primer 5’-AGT GAG ACG GGC AAC AGC -3’ (to detect the cleavage sites in the promoter region) or 5’-CAC ATG CAT GCC ATA TGA T -3’ (to detect the cleavage sites in the upstream region of the promoter) under the following conditions: 95°C for 5 min and 35 cycles at 95°C for 30 sec, 63°C for 15 sec and 72°C for 2 min. All samples were purified and resolved in 6% polyacrylamide-7 M urea gels.
V-5. ChIP assay

The ChIP assay was performed according to Kuo and Allis (98) with slight modifications. To obtain the chromatin fragment, the cell lysate was digested with 20 U of MNase at 37°C for 20 min. After cell debris was removed, 60 µl of the lysate was used for immunoprecipitation and subsequent DNA purification, which was performed in triplicate with a OneDay ChIP Kit (Diagenode, Liège, Belgium) according to the manufacturer’s instructions, using antibodies against histone H3 (ab1791; Abcam, Cambridge, UK). Using 2 µl of the DNA samples, quantification of fragments containing a part of the nucleosome I, II or II’ DNA was performed by SYBR Green quantitative real-time PCR with the StepOne Plus real-time PCR system and StepOne software v. 2.1. Quantification of fragments containing part of the nucleosome III DNA was carried out simultaneously as an internal control. The primer sets were as follows: for the nucleosome I region, 5’-GTG TGC GAC GAC ACT GAT-3’ and 5’-AGA GAA AAG AAG AAA ACA AGA GTT-3’; for the nucleosome II region, 5’-GGC TGG GAA GCA TAT TTG AG-3’ and 5’-TTG AAG CTC TAA TTT GTG AGT TTA GT-3’; for the nucleosome II’ region, 5’-TGG TAC CCT GGA GAG TAG TCA G-3’ and 5’-CGG ATC TGC TCG AGG TTA TG-3’; for the nucleosome III region, 5’-AGA ACC GTG GAT GAT GTG GT-3’ and 5’-CCT TCC CTT TGC AAA TAG TCC-3’. The relative sample enrichment was calculated using the formula: $2^{-(Ct_{target\ region}-Ct_{nucleosome\ III\ region})}$.
VI. Results

VI-1. Effect of superhelically curved DNA on transcription in yeast

When the UAS region is deleted from the CYC1 promoter in a CYC1-lacZ plasmid, the remaining downstream promoters are activated by nucleosome depletion (99). This implies that the promoter activity is responsible for the nucleosome occupancy in this system. We chose it to examine mechanisms of transcriptional activation by superhelically curved DNA structures, for we assumed that the activation would be mediated by chromatin alteration.

The UAS-deleted yeast cytochrome c (CYC1) promoter is known to be a “nucleosome-responsive” promoter, which was activated 94-fold by nucleosome depletion (99). The superhelically curved DNA segments T12, T20 and T28 were introduced into the UAS-deleted CYC1 promoter (Fig. 15), and their effects on transcription were examined by measuring the β-galactosidase activity. As shown in Fig. 16, all of the curved DNA segments activated transcription, and the observed fold activations relative to the control were 51.4 (T12), 63.4 (T20) and 56.4 (T28), respectively. Thus, superhelically curved DNAs can activate transcription not only in mammalian cells but also in yeast cells. Furthermore, the extents of activation by these segments were comparable to those induced by the nucleosome-free effect described above.
Fig. 15. Reporter constructs used in this study and architecture of superhelically curved DNAs

(A) Reporter constructs. “pCYC1”, “+1” and “TATA” indicate CYC1 promoter, translation start site and TATA box, respectively. The CYC1 promoter contains five TATA sequences, and they were distinguished by the numbering according to the report by Li and Sherman (106). The nucleotide sequence of T20 is slightly different from that in the previous report (61): the sequence started with 5’-TCAGTTTTT and ended with TTTTTT-3’ in the previous T20, but the corresponding sequences are TTTTT and TTTTTTCACG-3’ in the present T20. However, since they have the same three-dimensional architecture, the same name was used for convenience. (B) Three-dimensional architectures of T12, T20 and T28. They were drawn as reported previously (61). The white line indicates the superhelical axis.
Fig. 16. Effect of superhelically curved DNA on transcription
The promoter activity was determined in a β-galactosidase assay. Values shown are means ± SD (n = 3).
VI-2. Positioning of nucleosomes on and around the promoter

To determine how T12, T20 and T28 activated transcription, at first, the potential sequence-specific localization (translational positioning) of the nucleosomes on the region containing a curved DNA and the CYC1 promoter was investigated by an indirect end-labeling analysis (Fig. 17) (97, 100-105). Plasmid-containing nuclei were isolated, and were subjected to digestion with micrococcal nuclease (MNase), and subsequently to the analysis (Fig. 18). The bands shown with white arrowheads indicate the sites that were digested by the enzyme in the naked state, but were fully or partially protected in chromatin. On the other hand, the bands shown with black arrowheads indicate the sites that were newly or highly digested in chromatin. Based on these signal profiles and the distances between the black arrowheads, the nucleosome positions were deduced.

Regarding the chromatin formed on the control construct (“control chromatin”), three nucleosomes, referred to as I, II and III, were suggested to be present as indicated in the figure: the latter two were located upstream of the TATA1 sequence and the first one was on the promoter. Nucleosomes I and III also seemed to lie on the corresponding regions in the curved DNA-harboring each construct. In the control chromatin, nucleosome I was deduced to be located between -40 and -190 relative to the translation start site, and the distance between these positions was 150 bp, while in each chromatin formed on the curved DNA-containing construct (referred to below as T12, T20 or T28 chromatin, or together as
Tn chromatin), it was presumably located between -40 and -240, and the distance between them was 200 bp. Therefore, nucleosome I seemed to cover a wider region in the Tn chromatin. The presence of a nucleosome (referred to as nucleosome II’) on the T12, T20, or T28 region was also speculated. In the locus of nucleosome II’, we detected several sites that were slightly but obviously digested in the chromatin, as indicated with black arrowheads in parentheses. They seemed to indicate the edges of nucleosomes with different translational positions. Furthermore, another nucleosome might be located between nucleosomes II’ and III, as indicated by the dashed ellipse in the figure.
Fig. 17. Procedure for mapping of nucleosome cleavage sites by indirect end-labeling

(A) Chromatin is partially digested with MNase and purified. (B) DNA is digested with restriction enzyme X. (C) The products are electrophoresed on an agarose gel and subjected to Southern blot analysis with a specific probe, and the MNase cleavage site is identified by corresponding band size and its distance from the restriction site X.
Fig. 18. Analysis of translational positions of nucleosomes on the upstream region of the reporter gene

The upstream structure of the translation start site is illustrated on the right side of each gel. The lines with a number in the CYC1 promoter (pCYC1) indicate the five TATA sequences. In each set of data, the lanes labeled Ch indicate the MNase digestion of isolated nuclei (chromatin) at three nuclease levels, and the lanes labeled DNA indicate the MNase digestion of the naked DNA at two nuclease levels, as a control. Arrowheads indicate the following sites: white, fully or partially protected in chromatin; black, highly digested or specifically digested in chromatin; parenthesized black, digested more in chromatin than in naked DNA, but the digestion itself was slight. Position numbers are relative to the translation start site (+1). The ellipses with I, II, II’ or III indicate putative nucleosomes, as deduced from the band patterns. Dashed-lined ellipses: possible additional positions.
VI-3. Confirmation of the nucleosome presence

To determine whether nucleosomes I, II, II’ and III were actually formed on the regions indicated in Fig. 18, the chromatin digests with MNase were subjected to a Southern blot analysis, using probes prepared from the relevant regions. As shown in Fig. 19, so-called “nucleosomal ladder” patterns were detected by all of the probes, which supported the presence of the putative nucleosomes indicated in Fig. 18. However, the curved DNA probes c, d and e hardly detected DNA fragments of mononucleosome size, but detected fragments with sizes between mono- and dinucleosomal DNAs, which appeared as smeared bands between these two positions. This pattern was not observed when the other probes were used. These smeared patterns could be a consequence of the difficulty in digesting the curved DNA regions, as shown in Fig. 18. In addition, it is possible that nucleosome II’ had multiple translational positionings, as described in the previous section. In such a case, the outside regions of each curved DNA could be protected from MNase digestion by the positioned histone cores. These two factors could generate DNA fragments longer than that of mononucleosomal DNA.

The relative populations of the nucleosome-forming constructs were analyzed by a ChIP assay. The immunoprecipitation was performed using the antibody against histone H3. The population of nucleosome III was used as a reference for comparison (Fig. 20). The relative populations of nucleosome I to nucleosome III were 61% for the control construct, 69%
for the T12-harbor ing construct (“T12 construct”), 43% for the T20 construct and 85% for the T28 construct. Except for the T20 construct, the curved DNA-flanked promoter resulted in a more nucleosome formation than the curved DNA-less (control) promoter, while the promoter activity of the former was much higher (more than 50-fold). Obviously, the transcription data in Fig. 16 could not be explained in terms of the extent of histone deposition. Furthermore, these data showed that transcription was highly activated, even from the nucleosome harboring promoter.

The promoter proximal nucleosomes were nucleosomes II and II’. The populations of nucleosome II’ gradually increased, according to the greater length of the curved DNA region. Regarding nucleosomes II and II’, the control construct and the T28 construct had comparable populations. However, T28 activated the promoter by more than 50-fold. Therefore, the presence of nucleosomes II and II’ themselves was not implicated in determining the transcription level.
Fig. 19. Southern blot analysis of DNA fragments obtained from MNase digests of nuclei

The digestion products were hybridized with probes obtained from the promoter (probe \( a \)), the curved DNA regions (probes \( c \), \( d \) and \( e \)), the \( URA3 \) region (probe \( f \)) or the region between the promoter and the \( URA3 \) region in the control construct (probe \( b \)). The positions and lengths of these probes are illustrated. Black boxes in the \( CYC1 \) promoter indicate the TATA sequences 1 to 5. Nucleosomes are shown as ellipses with Roman numerals. The MNase digestion of isolated nuclei (chromatin) was performed at three nuclease levels. ‘mono’ and ‘di’ indicate mono- and dinucleosomes, respectively.
Fig. 20. Relative populations of nucleosomes I, II, II’ and III, as determined by a ChIP assay

The assay was performed using the antibodies against histone H3. In the schematic drawing, the amplified regions and their locations are indicated with lines, with arrows indicating PCR primers. In the histogram, the fold enrichment of each nucleosome relative to nucleosome III is indicated with different bars: white, control chromatin; black, T12 chromatin; gray, T20 chromatin; striped, T28 chromatin (means ± SD of three determinations).
VI-4. Accessibility of DNA in chromatin

We subjected the same nuclear fraction to DNase I digestion (52) to investigate the accessibility of the DNA in chromatin in order to understand how curved DNAs could activate transcription. The control construct and the curved DNA-containing constructs share the same sequence, except for the curved DNA region (Fig. 15). In the region with the same sequence, a notable difference was only detected within the TATA3 sequence (Figs. 21-25). In the control construct, A-92 and A-91 were highly accessible in the naked DNA, and their accessibility was maintained even in chromatin. On the other hand, the sites with high accessibility were shifted by one nucleotide toward the center of the TATA3 sequence in the curved DNA-containing constructs in the naked states. As a result, A-91 and A-90 became highly accessible.

The accessibility of these two sites was maintained in the T20 chromatin. In addition, as in the case of the control chromatin, A-92 was also accessible in this chromatin. Although the cleavage signal at A-90 was less marked in the T12 chromatin and the T28 chromatin, as compared with that in the T20 chromatin, the signal definitely existed (see lane 9 in each panel). In conclusion, the exposed sequence was shifted towards the center of the TATA3 sequence in the curved DNA-containing constructs in the naked DNA, and this exposure was maintained even in chromatin. Since the distance between -240 and -90 is 150 bp (Fig. 18), position -90 seemed to be the downstream end of a certain population of nucleosome I.
Regarding the curved DNA region, interestingly, both the positions and intensities of the cleavage signals were almost the same between naked DNA and chromatin in each Tn chromatin. Clear “10 bp ladder” cleavage patterns with a rung spacing of 9 bp were observed, indicating that the rotational setting of the DNA was maintained even in the chromatin (Figs 22, 23, 24).
Fig. 21. Accessibility of DNA in chromatin
Nuclei harboring control construct were digested with DNase I, and the cleavage sites were analyzed as described in the Materials and Methods. ‘DNA’ and ‘Ch’ indicate DNase I cleavage sites in naked DNA and chromatin, respectively. The black arrowheads indicate the “10 bp ladder” (the rung spacing is 9 bp). The asterisk on the right of each autoradiogram indicates overlap points of the signals.
Nuclei harboring T12 construct were digested with DNase I, and the cleavage sites were analyzed as described in the Materials and Methods. ‘DNA’ and ‘Ch’ indicate DNase I cleavage sites in naked DNA and chromatin, respectively. The black arrowheads indicate the “10 bp ladder” (the rung spacing is 9 bp). The asterisk on the right of each autoradiogram indicates overlap points of the signals.

Fig. 22. Accessibility of DNA in chromatin
Fig. 23. Accessibility of DNA in chromatin

Nuclei harboring T20 construct were digested with DNase I, and the cleavage sites were analyzed as described in the Materials and Methods. ‘DNA’ and ‘Ch’ indicate DNase I cleavage sites in naked DNA and chromatin, respectively. The black arrowheads indicate the “10 bp ladder” (the rung spacing is 9 bp). The asterisk on the right of each autoradiogram indicates overlap points of the signals.
Fig. 24. Accessibility of DNA in chromatin

Nuclei harboring T28 construct were digested with DNase I, and the cleavage sites were analyzed as described in the Materials and Methods. ‘DNA’ and ‘Ch’ indicate DNase I cleavage sites in naked DNA and chromatin, respectively. The black arrowheads indicate the “10 bp ladder” (the rung spacing is 9 bp). The asterisk on the right of each autoradiogram indicates overlap points of the signals.
<table>
<thead>
<tr>
<th>TATA1</th>
<th>TATA2</th>
<th>TATA3</th>
<th>TATA4</th>
<th>TATA5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{170}$TATATATA$_{170}$</td>
<td>$T_{170}$TATATATA$_{170}$</td>
<td>$T_{26}$TATATATT$_{26}$</td>
<td>$T_{30}$TATACATT$_{30}$</td>
<td>$A_{30}$TAAATATTA$_{30}$</td>
</tr>
</tbody>
</table>

**Fig. 25. Accessibility of DNA in chromatin**

Nuclei harboring each construct were digested with DNase I, and the cleavage sites were analyzed as described in the Materials and Methods. ‘DNA’ and ‘Ch’ indicate DNase I cleavage sites in naked DNA and chromatin, respectively. The dots on the left-hand side of each gel indicate the positions of TATA-comprising nucleotides. Highly digested nucleotides within the TATA3 sequence in the chromatin are indicated with black arrowheads, and the position of $A_{90}$ is also indicated.
VII. Discussion

Using a yeast minichromosome system, the present study investigated why superhelically curved DNA segments can activate transcription in chromatin. Chromatin analyses confirmed that the major populations of the reporter constructs formed nucleosomes on the *CYC1* promoter, the curved DNA region and the *URA3* region. Regarding the promoter region, there were two clear differences between the control chromatin and the Tn chromatin. One was that nucleosome I was located between -40 and -190 in the control chromatin, while it was located between -40 and -240 in the Tn chromatin (Fig. 18). The other was that the center region of the TATA3 sequence was more exposed in the Tn chromatin than in the control chromatin (Fig. 25). Previous reports described that the TATA1 and TATA2 sequences were responsible for transcription from the *CYC1* promoter (106, 107). However, considering that the extent of exposure of these sequences in the Tn chromatin was almost the same as that in the control chromatin (Fig. 25) and that the transcription from the *CYC1* promoter is repressed in the presence of nucleosomes (99), the TATA3 sequence was presumably responsible for the transcription of the reporter gene in the present case.

If the upstream edge of nucleosome I was -240 in the Tn chromatin, then the downstream edge should have been located around -90, which is the center region of the TATA3 sequence. A certain population of chromatin seemed to adopt this positioning, since -190 was considerably
protected against the digestion (Fig. 18). However, considering that the
locus of nucleosome I (the protected region) spanned from -40 to -240,
different translational positions were also presumably adopted by the other
populations. The increased mobility of nucleosome I in the Tn chromatin
seems to have generated the wide signal for the translational positioning.
Under this condition, the TATA3-exposing population of the nucleosome I
was generated, and it was presumably implicated in the transcriptional
activation in the Tn chromatin. A quite similar result was obtained in the
system using a short, left handedly curved DNA segment, T4. When T4 (36
bp) was linked to the HSV tk promoter at a specific rotational phase and
distance, it attracted the histone core and the TATA box was thus left in the
linker DNA with its minor groove facing outwards, which led to the
activation of transcription (52). However, the activation by T4 was
approximately 10-fold. Thus, to explain the activation by more than 50-fold,
an additional parameter still seems to be necessary. It may be the sliding
ability of nucleosome II’, which would also explain the increased mobility
of nucleosome I.

Nucleosome II’ seemed to have multiple translational positions,
while, in contrast, the position of nucleosome II was almost fixed (Figs. 18
and 19). This difference was obviously caused by the underlying DNA, and
the superhelically curved DNAs were certainly implicated in the multiple
positioning of nucleosome II’. It is known that curved DNA structures
favor nucleosome formation (109-112). Therefore, the wide-ranging DNA
curvature is likely to have the ability to slide the histone core. This hypothesis is supported by the report that nucleosome sliding is a general phenomenon that is dependent on the underlying DNA sequence (113). Furthermore, nucleosome sliding is considered as an important way of regulating access to DNA sites that are near nucleosome borders (114). The data shown in Figs. 18 and 19 strongly suggest that the sliding of nucleosome II’ indeed occurred in the curved DNA region. The high mobility of the histone core seems to have increased the mobility of nucleosome II’. Furthermore, it must also be noted that TBP binding to DNA can induce nucleosome sliding (115). Therefore, we can envision a scenario in which the high mobility of nucleosome II’ enables nucleosome I to fluctuate. This fluctuation enhanced the chance of TBP binding to the TATA3 sequence, which then induced further sliding of these nucleosomes transiently, and the resulting open structure highly facilitated transcription initiation (Fig. 26). Although the dynamic features of these events could not be detected in our “static analyses”, it is quite possible that a superhelically curved DNA segment could function as an acceptor of the sliding histone core and assist the dynamic process of transcription.

The rotational orientation and the distance of a curved DNA segment relative to the promoter are important parameters for transcriptional activation (52, 61). However, these were not optimized in the present study. Therefore, further positional refinement of each segment might have produced even greater transcriptional activation.
Fig. 26. The putative mechanism of transcriptional activation in chromatin induced by superhelically curved DNA structures

The high mobility of nucleosome II’ on the curved DNA region generates a room to accept transiently and largely sliding nucleosome I that is caused by TBP binding, which exposes the promoter and activates transcription.
VIII. Acknowledgments

I greatly appreciate Prof. Takashi Ohyama (Department of Biology, School of Education, Waseda University, Tokyo, Japan) for his kind and appropriate advice and encouragement throughout this research.

I would like to express my gratitude to Prof. Toru Higashinakagawa, Prof. Hideo Namiki for the comment and discussion about this thesis. I would like to thank all collaborators, Dr. T. Mitani, Dr. M. Shimizu, Dr. J. Nishikawa, Dr. N. Morohashi Mr. K. Udagawa, Mr. H. Fujita, and Mr. S. Fujii. I would like to thank Dr. Y. Kadokawa for the plasmid pCX-GFP.

This work was supported in part by JSPS and MEXT research grants to T.O. and by a JSPS grant to M.S.
IX. References


protein-induced DNA bending and role of the nonconserved core sequence in binding site affinity, *Virology* 174, 557-575.


sp00F transcription at the initiation of sporulation in *Bacillus subtilis*, *J. Mol. Biol.* 250, 11-23.


embryonic stem cells into hepatocyte-like cells identified by cellular uptake of indocyanine green, *Stem Cells* 20, 146-154.


107. Hahn, S., Hoar, E.T., and Guarente, L. (1985) Each of three "TATA elements" specifies a subset of the transcription initiation sites at the


### 研究業績

(2010年 3月 現在)

<table>
<thead>
<tr>
<th>種 類 別</th>
<th>題名</th>
<th>発表・発行掲載誌名</th>
<th>発表・発行年月</th>
<th>連名者（申請者含む）</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4. Ohyama T., Sumida N., Tanase J. How promoters are recognized. 7th EMBL Transcription Meeting. 2006 Aug. Heidelberg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6. 棚瀨潤一、宇田川絢司、西川純一、三谷匡、大山隆．クロマチン工学による外来遺伝子の安定的発現．第32回日本分子生物学会年会 2009年 12月 横浜</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7. 伊藤美月、棚瀨潤一、山内恵里、清水光弘、大山隆．遺伝子発現に有利なクロマチン構造の人為的構築．第32回日本分子生物学会年会 2009年 12月 横浜</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8. 藤田仁、諸橋伸行、棚瀨潤一、大山隆、清水光弘．出芽酵母におけるDNA構造による遺伝子発現とクロマチン構造の制御．第32回日本分子生物学会年会 2009年 12月 横浜</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
9. 宇田川紘司、山本拓弥、棚瀬潤一、深川竜郎、大山隆. 人工ペント DNA により活性化されたトランスジーンの核内局在：マウス ES 細胞を用いた解析. 第 32 回日本分子生物学会年会 2009 年 12 月 横浜

10. 宇田川紘司、山本拓弥、棚瀬潤一、深川竜郎、大山隆. マウス ES 細胞における高発現トランスジーンの核内配置. 第 26 回染色体ワークショップ 2009 年 1 月 姫路

11. 伊藤美月、棚瀬潤一、山内恵里、清水光弘、大山隆. 遺伝子発現に有利なクロマチン構造の人为的築造. 第 31 回日本分子生物学会年会 第 81 回日本生化学会年会合同大会 2008 年 12 月 神戸

12. 藤田仁、諸橋伸行、棚瀬潤一、大山隆、清水光弘. DNA 構造によるクロマチンの改変と人为的遺伝子発現制御：出芽酵母レポータープラスマドによる解析. 第 31 回日本分子生物学会年会 第 81 回日本生化学会年会合同大会 2008 年 12 月 神戸

13. 棚瀬潤一、宇田川紘司、山本拓弥、伊藤美月、藤井俊輔、三谷匡、大山隆. クロマチン工学による外来遺伝子の安定的発現. 第 30 回日本分子生物学会年会 第 80 回日本生化学会年会合同大会 2007年12月 横浜

14. 大山隆、棚瀬潤一、伊藤美月、三谷匡、清水光弘. クロマチン工学のストラテジー. 第 30 回日本分子生物学会年会 第 80 回日本生化学会年会合同大会 2007 年 12 月 横浜

15. 三谷匡、川村織子、篠栗弘貴、横田隆徳、棚瀬潤一、大山隆、田口善智. マウス ES 細胞における非 B-型 DNA による shRNA 発現活性化に対する効果. 第 30 回日本分子生物学会年会 第 80 回日本生化学会年会合同大会 2007 年 12 月 横浜

16. 大山隆、棚瀬潤一、宇田川紘司、木村元、景山大、古屋美香、藤井俊輔、三谷匡. ゲノム DNA の特性解析と高次遺伝情報の解読. 第 7 回核ダイナミクス研究会 2007 年 9 月 北海道

17. Ohyama T., Sumida N., Fukue Y., Tanase J., Inoue S. Genetic information written in DNA conformation and properties. 第 29 回日本分子生物学会年会 第 79 回日本生化学会年会合同大会 2006 年 6 月 京都

18. 大山隆、隅田周志、棚瀬潤一. クロマチン工学の開拓. 第 6 回細胞核ダイナミクス研究会 2006 年 5 月 熊本

19. 棚瀬潤一、隅田周志、大山隆. クロマチン基盤構造の局所的改変とその遺伝子発現への影響. 第 28 回日本分子生物学会年会 2005 年 12 月 福岡

<table>
<thead>
<tr>
<th>索書</th>
<th>distinctive mechanical property common to human promoters and its transcriptional relevance. 第 28 回日本分子生物学会年会 2005 年 12 月 福岡</th>
</tr>
</thead>
<tbody>
<tr>
<td>21.</td>
<td>大山隆、 福江善朗、 井上正太郎、 隅田周志、 椎瀬潤一. DNA に印された配列以外の情報を読む. 生理研研究会 2005 年 11 月 岡崎</td>
</tr>
<tr>
<td>22.</td>
<td>隅田周志、 椎瀬潤一、 古屋貴代、 大山隆. クロマチン内 DNA の回転的ポジショニングと遺伝子発現. 第 27 回日本分子生物学会年会 2004 年 12 月 神戸</td>
</tr>
</tbody>
</table>