

Purification and Characterization of the DNA-Binding Domain of BTEB, a GC Box-Binding Transcription Factor, Expressed in *Escherichia coli*¹

Yasuo Kikuchi, Kazuhiro Sogawa, Nobuaki Watanabe,² Akira Kobayashi, and Yoshiaki Fujii-Kuriyama

Department of Chemistry, Graduate School of Science, Tohoku University, Aoba-ku, Sendai 980

Received for publication, September 13, 1995

BTEB is a GC-binding protein that regulates the transcription of genes with a single GC-box or tandemly repeated GC-boxes in the promoter. The DNA-binding domain of BTEB consists of three contiguous Cys₂-His₂ zinc finger motifs and short segments adjacent to their N- and C-terminal sides [Kobayashi *et al.* (1995) *J. Biochem.* 117, 91–95]. The truncated BTEB (residues 120 to 244) containing the DNA-binding domain was expressed in *Escherichia coli* and purified to homogeneity under denaturing conditions. DNA-binding activity of the BTEB was regenerated by refolding in the presence of Zn²⁺. The efficiency in regeneration was 70±10%, and the dissociation constant (*K_d*) of the DNA-complex was 4±2 nM. Co²⁺ also regenerated the DNA-binding affinity of BTEB, albeit with less efficiency than Zn²⁺. Co-BTEB showed a slightly lower affinity to the specific DNA than Zn-BTEB. Refolding in the presence of Cd²⁺ resulted in an extremely low efficiency in regeneration of the DNA-binding activity. Zn-BTEB is in a monomer state at concentrations lower than 0.5 μM, and forms a dimer in the concentration range of about 10 to 200 μM.

Key words: DNA-binding protein, GC-box, homodimerization, transcription factor, zinc finger.

Sequence-specific DNA binding proteins are directly involved in the regulation of mRNA transcription in higher organisms, and the GC box sequence is one of the most common regulatory DNA elements in eukaryotic genes. BTE (basic transcriptional element) is a GC-rich *cis*-acting regulatory element found in the region proximal to the TATA sequence of CYP1A1 gene and is involved in the basic transcription of CYP1A1 gene (1). In the previous studies two cDNAs encoding regulatory proteins binding to the BTE sequence were obtained: Sp1, an extensively studied GC box-binding transcription factor, consisting of 788 amino acids and a novel transcription factor, BTEB (BTE-binding protein) of 244 amino acids (2). Sp1 and BTEB contain three contiguous Cys₂-His₂ zinc finger motifs at their C-terminal regions as the DNA binding domain. The amino acid sequence homology between their zinc finger motifs is 72% (2) and the two regulatory proteins are indistinguishable in DNA-binding specificity and comparable in DNA-binding affinity (3). Their transcription activation functions, however, are distinct from each other: BTEB requires tandem repeats of the GC boxes for promoter activation, while Sp1 can activate transcription from a promoter with a single GC box or tandemly repeated GC boxes (2). From functional domain analyses of BTEB in a transient expression system using Y-1 cells (mouse

adrenal cortex cell line), two novel activation domains were identified in the N-terminal regions outside the zinc finger domain (regions A and B in Fig. 1) (4). The zinc finger domain (region Z) and the regions immediately N- and C-terminal to the zinc finger structures (regions C and D, respectively) were suggested to be essential in the DNA-binding of BTEB (4). The participation of the region just N-terminal to the zinc finger motifs in the DNA-binding function was reported with *Drosophila* tramtrack protein (TTK) (5), yeast SWI5 protein (6), and ADR1 protein (7), but has not been reported with Sp1 or serum-inducible transcription activators such as Krox20 and Krox24 (also known as Zif268).

In the present study the truncated structure of BTEB containing the regions required for the DNA-binding function was expressed in *Escherichia coli* and purified to homogeneity to investigate the structure and the DNA-binding properties of these domains.

MATERIALS AND METHODS

Construction of Expression Vector—A part of BTEB cDNA encoding amino acids 120 to 244 was amplified in a polymerase chain reaction (PCR). The PCR conditions consisted of heating at 94°C for 10 min, 25 cycles of 55°C for 1 min, 72°C for 1 min, and 94°C for 1 min, and then finally incubating at 55°C for 1 min and at 72°C for 5 min with a Quick Thermo II (Nippon Genetics). Primers used were 5'-CGCCAGCCCGCTCTCCCTC-3' and 5'-CTTCCGGG-ACCATCTCTGCC-3'. A T7 RNA polymerase-dependent expression vector, pAR2106 (8) was digested with *Bam*HI,

¹This work was supported in part by Grant-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science and Culture of Japan, and by funds from Nissan Science Foundation and Sankyo Co.

²Present address: Analytical and Metabolic Research Laboratories, Sankyo Co., Ltd., 1-2-58 Hiromachi, Shinagawa-ku, Tokyo 140.

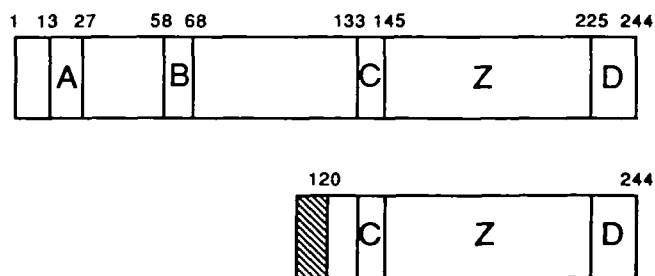


Fig. 1. Functional domains of BTEB and the truncated structure expressed in *Escherichia coli*. The functional domains were analyzed by Kobayashi *et al.* (4). A, B, activation domains; C, basic domain adjacent to N-terminus of the first zinc finger; Z, three contiguous zinc-finger domains; D, C-terminal domain. The hatched region shows the peptide of 13 amino acid residues derived from the expression vector (see "MATERIALS AND METHODS").

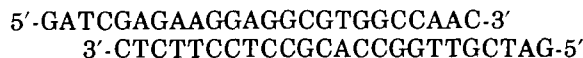
filled in with Klenow enzyme and ligated with the PCR fragment obtained above to form a fused gene coding for 13 amino acids of the plasmid (MASMTGGQMGGRD) and residues 120 to 244 of BTEB. The construction was confirmed by sequencing. The resulting plasmid was used to transform *E. coli* BL21(DE3) strain. The fusion protein with a molecular weight of 16,000 is designated BTEB(120-244).

Overexpression of BTEB(120-244)—*E. coli* cells harboring the recombinant were grown at 37°C in the LB medium containing 50 µg/ml ampicillin. Overexpression of the truncated BTEB was achieved by the addition of isopropyl-β-D-thiogalactoside to 1 mM when A_{600} of the culture reached 2.6. The cells were harvested 7.5 h later. These methods were optimum in the expression of this truncated BTEB.

Purification of BTEB(120-244)—*E. coli* pellets (24 g) obtained from 5 liters of the culture medium were sonicated in the extraction buffer solution (160 ml) containing 40 mM HEPES (pH 7.9), 8 M urea, and 0.6 M NaCl, and centrifuged at $5,000 \times g$ for 30 min at 4°C. The extract was made up to 5 mM imidazole and applied to a column (2.5 cm × 10 cm) of zinc-chelating Sepharose Fast Flow (Pharmacia) equilibrated with the extraction buffer supplemented with 5 mM imidazole. The column was washed with the same solution (150 ml) as above, and BTEB(120-244) was eluted with a linear gradient of imidazole (5 to 200 mM). The partially purified BTEB(120-244) was reduced with 160 mM DTT at 40°C for 15 h, desalted by gel filtration on a Sephadex G-25 column (2.5 cm × 60 cm) in 1 mM HCl, freeze-dried, redissolved in 0.05% trifluoroacetic acid (10 ml), and finally purified on a C18 reversed phase HPLC column (µBondapak 10 µ 25Å C18, Waters, 2.5 cm × 10 cm) in 0.05% trifluoroacetic acid with a CH₃CN gradient. The fractions containing BTEB(120-244) were pooled and freeze-dried (100 mg).

Reconstitution of Zinc Finger Conformation—The concentration of the purified BTEB(120-244) was determined by using $A_{280} = 0.68$ for a 1 mg/ml solution. The freeze-dried preparation (2–10 mg) was dissolved in 1 mM HCl containing 1.5 molar equivalent of ZnCl₂, then the pH of the solution was slowly adjusted to 6.8 with 0.1 M NaOH in an anaerobic atmosphere and 0.1 M phosphate buffer solution (pH 6.8, 0.1 ml) was added to the solution. The reconstitution with CoCl₂ or Cd(CH₃COO)₂ was carried out similarly.

Gel Mobility Shift Assay—Synthetic oligo nucleotides



were 5'-end labeled with [γ -³²P]ATP and used as the BTE probe (2). BTEB(120-244) (3.2 ng) purified and reconstituted with Zn²⁺ was kept at 0°C for 30 min in 20 mM HEPES buffer solution (pH 7.9) containing 0.4 mg/ml poly(dI-dC), 20 mM NaCl, 5 mM MgCl₂, 5 µM ZnSO₄, 0.1 mg/ml bovine serum albumin, 20% glycerol, and 0.03% NP-40. After addition of ³²P-labeled BTE probe (16–75 nM), the mixture (10 µl) was incubated at 24°C for 30 min. The BTE-protein complex was separated from free BTE by electrophoresis on a 4.5% polyacrylamide gel containing 40 mM Tris, 306 mM glycine, and 0.1% NP-40 as described by Letovsky and Dynan (9). The radioactivities in the fractions of the BTE-protein complex and the free BTE were determined with a Fujix Bas1000 Bio-imaging analyzer. A larger amount of BTEB(120-244) reconstituted with Co²⁺ or Cd²⁺ was subjected to the assay to examine the DNA-binding activity of the preparation. The buffer solution containing 20 mM HEPES (pH 7.9), 150 mM NaCl, 0.1 mg/ml bovine serum albumin, 20% glycerol, and 0.03% NP-40 was freed from heavy metal ions with metal chelating Sepharose Fast Flow (Pharmacia).

Gel Filtration Chromatography—Zn-BTEB(120-244) (8 µg or 3 mg) in 20 mM phosphate buffer solution (pH 6.8) containing 150 mM NaCl, 5 mM MgCl₂, and 0.1 mg/ml bovine serum albumin (1 ml), was chromatographed on a column (1 cm × 95 cm) of Sephadex G-75 equilibrated with the same buffer solution at 4°C. The fractions (1 ml) were collected and assayed for specific DNA binding activity by the gel mobility shift method and for protein concentration by the method of Bradford (10).

RESULTS AND DISCUSSION

Histidine and cysteine are known to form stable coordination complexes with transition metal ions such as Zn²⁺ and Ni²⁺. BTEB(120-244) containing three contiguous Cys₂-His₂ zinc finger motifs was purified by specific affinity chromatography using a metal chelating matrix charged with Zn²⁺ ions (11). Cys₂-His₂ zinc finger is one of the motifs using zinc ions for folding protein domains [for a review see Schwabe and Klug (12)]. This purification procedure may be applied to a wide variety of zinc finger proteins which bind zinc ions through cysteinyl and/or histidyl residues. The recoveries of the protein and purification factors through the purification procedures are shown in Table I. The protein preparations were analyzed by electrophoresis on SDS-polyacrylamide gel (Fig. 2). BTEB(122-244) was expressed at a level of 20 mg/liter of culture medium, and purified to 93% homogeneity (see footnote to Table I). Since BTEB(120-244) was purified under denaturing conditions, the native conformation was reconstituted with Zn²⁺, and the DNA-binding activity of the reconstituted protein was confirmed by the gel mobility shift assay (Fig. 3A). From the Scatchard plot analysis of the results, the dissociation constant (K_d) of the BTEB(120-244)-BTE complex was calculated to be 4 ± 2 nM and the efficiency of reconstitution was $70 \pm 10\%$ (Fig. 3B, Table II). The K_d value is similar to those of the DNA-complexes of three-zinc-finger proteins such as Sp1 (0.5–3

TABLE I. Purification of BTEB(120-244).

Purification step	Protein (mg) ^a	Purification (fold) ^c
1. Cell extract	2,820	1
2. Zn-affinity chromatography	192	10
3. Gel filtration on Sephadex G-25	163	13
4. HPLC reversed phase chromatography	140 (100) ^b	18

^aProtein concentration was determined by the Bradford method (10) using bovine serum albumin as standard. ^bWeight of freeze-dried preparation. ^cThe relative intensities of the silver-stained protein bands in Fig 2 were estimated by computerized densitometry

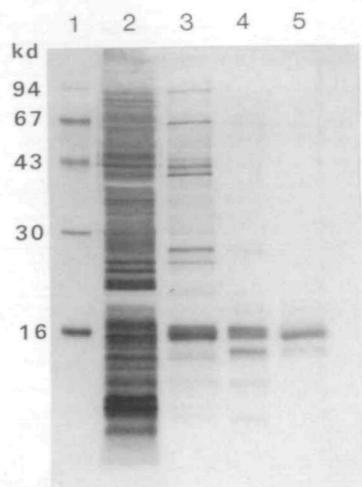


Fig. 2 SDS-polyacrylamide gel electrophoresis of the purification fractions. Proteins from each step were electrophoresed on a 16.5% SDS-polyacrylamide gel. The proteins were visualized by silver-staining. Lanes 1-5: molecular markers, cell extract, zinc chelating affinity fraction, Sephadex G-25 fraction, and reversed-phase HPLC fraction, respectively.

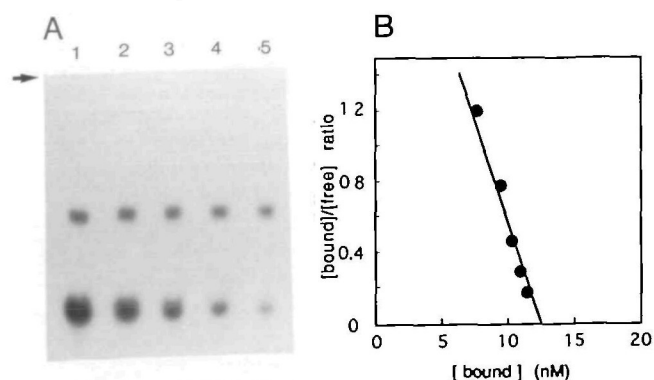


Fig. 3. DNA-binding activity of Zn-BTEB(120-244). A, a representative autoradiogram of the gel shift assay, B, Scatchard plot analysis. An arrow shows the origin of the electrophoresis. The concentration of BTEB(120-244), 20 nM. The concentrations of labeled BTE in lanes 1-5: 76.4, 50.7, 34.2, 22.0, and 14.6 nM, respectively.

nM) (9, 13) and Zif268 zinc finger peptide (6 nM) (14), and much smaller than that of the DNA complex of a two-zinc-finger protein, *Drosophila* tramtrack protein (400 nM) (5).

Co²⁺ was able to regenerate the DNA-binding activity of BTE(120-244), though less effectively than Zn²⁺, and the

TABLE II. Regeneration of DNA-binding activity.

Metal ion	Efficiency of regeneration (%)	K _d (nM)
Zn ²⁺	70 ± 10 ^a	4 ± 2 ^a
Co ²⁺	11 ± 2 ^b	16 ± 2 ^b
Cd ²⁺	0.05	30

^aData obtained from four measurements. ^bData obtained from two measurements

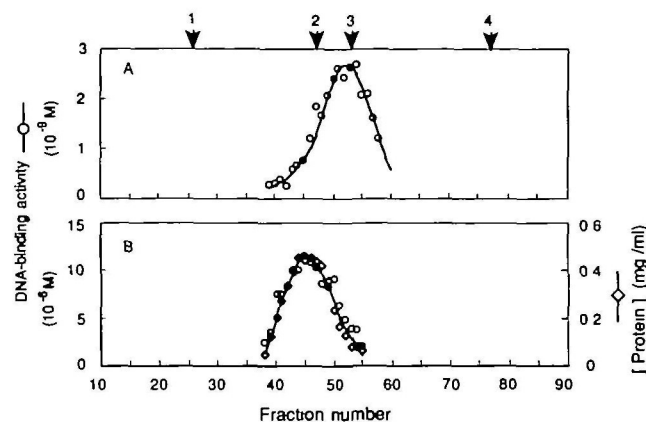


Fig. 4. Gel filtration chromatography of BTEB(120-244). (A) 8 µg or (B) 3 mg of Zn-BTEB(120-244) in 1 ml of the buffer solution was applied to a column. Arrows 1-4: elution positions of blue dextran, carbonic anhydrase (30 kDa), myoglobin (16 kDa), and glycine, respectively

DNA-binding affinity of Co-BTEB(120-244) is slightly lower than that of Zn-BTEB(120-244) (Table II). The three-finger peptide (65 residues) of Zif268 coordinated with Zn²⁺ or Co²⁺ similarly to regenerate the DNA binding affinity, and no significant difference was detected between the X-ray crystallographic data of Zn²⁺- and Co²⁺-structures (14). When the purified BTEB(120-244) was subjected to the reconstitution procedures with Cd²⁺, the DNA-binding activity required 1,500 times as large an amount of the material as Zn-BTEB(120-244) to be detected (Table II). The results indicate that the refolding of BTEB(120-244) into a native finger conformation is practically impossible with Cd²⁺. On the other hand, Cd²⁺ can be exchanged for the Zn²⁺ in the already refolded finger protein (165 residues) of Sp1 with full retention of specific recognition (15). But the refolding of the apo-Zn-finger protein initially in the presence of Cd²⁺ was less effective than the exchange of Cd²⁺ for Zn²⁺ possibly owing to alternate liganding patterns taken up by Cd²⁺ (15).

In the gel filtration chromatography under non-denaturing conditions, Zn-BTEB(120-244) behaved as a monomer at concentrations lower than 8 µg/ml (Fig. 4A), and as a dimer in the concentration range of about 0.2 to 3 mg/ml (Fig. 4B). The gel mobility shift assay in Fig. 3 was carried out at concentrations low enough to retain Zn-BTEB(120-244) in the monomer state. Some DNA-binding proteins form a dimer when they bind DNA, i.e. mammalian transcription factors, NF-κB p50 (16, 17), estrogen receptor (18), and LFB1 (19), and yeast transcription factor GCN4 (20). The target DNA sequences of these homodimeric transcription factors are usually palindromic. In the present case the BTE sequence is not palindromic. The

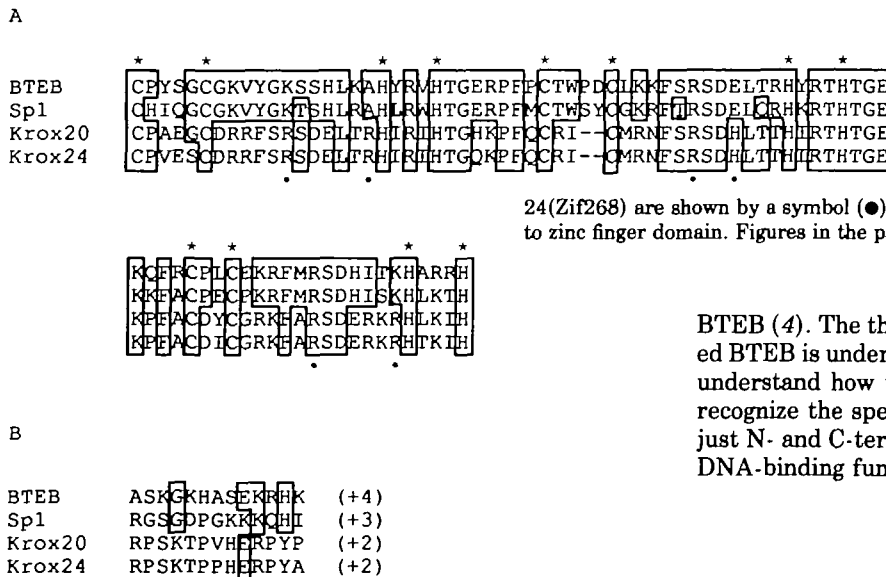


Fig. 5. Comparison of amino acid sequences. Sequences of Sp1, Krox20, and Krox24 are from Refs. 2, 22, and 23, respectively. Amino acid residues identical to those of BTEB are enclosed. A, zinc finger domain. Residues involved in zinc binding are asterisked. Residues being in contact with the bases in a DNA complex of Krox-

24(Zif268) are shown by a symbol (●) (14). B, basic amino acid sequences N-terminal to zinc finger domain. Figures in the parentheses show the net charge of the regions.

BTEB (4). The three-dimensional structure of the truncated BTEB is under investigation using NMR spectroscopy to understand how the zinc finger motifs of BTEB and Sp1 recognize the specific DNA sequence and how the regions just N- and C-terminal to the zinc fingers contribute to the DNA-binding function of BTEB and Sp1.

REFERENCES

- Yanagida, A., Sogawa, K., Yasumoto, K., and Fujii-Kuriyama, Y. (1990) A novel cis-acting DNA element required for a high level of inducible expression of the rat P-450c gene. *Mol. Cell. Biol.* **10**, 1470-1475
- Imataka, H., Sogawa, K., Yasumoto, K., Kikuchi, Y., Sasano, K., Kobayashi, A., Hayami, M., and Fujii-Kuriyama, Y. (1992) Two regulatory proteins that bind to the basic transcription element (BTE), a GC box sequence in the promoter region of the rat P-4501A1 gene. *EMBO J.* **11**, 3663-3671
- Sogawa, K., Kikuchi, Y., Imataka, H., and Fujii-Kuriyama, Y. (1993) Comparison of DNA-binding properties between BTEB and Sp1. *J. Biochem.* **114**, 605-609
- Kobayashi, A., Sogawa, K., Imataka, H., and Fujii-Kuriyama, Y. (1995) Analysis of functional domains of a GC box-binding protein, BTEB. *J. Biochem.* **117**, 91-95
- Fairall, L., Harrison, S.D., Travers, A.A., and Rhodes, D. (1992) Sequence-specific DNA binding by a two zinc-finger peptide from the *Drosophila melanogaster* tramtrack protein. *J. Mol. Biol.* **226**, 349-366
- Neuhaus, D., Nakaseko, Y., Schwabe, J.W.R., and Klug, A. (1992) Solution structures of two zinc-finger domains from SWI5 obtained using two-dimensional 1-H nuclear magnetic resonance spectroscopy. A zinc-finger structure with a third strand of beta-sheet. *J. Mol. Biol.* **228**, 637-651
- Thukral, S.K., Eisen, A., and Young, E.T. (1991) Two monomers of yeast transcription factor ADR1 bind a palindromic sequence symmetrically to activate ADH2 expression. *Mol. Cell. Biol.* **11**, 1566-1577
- Studier, F.W. and Moffatt, B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**, 113-130
- Letovsky, J. and Dynan, W.S. (1989) Measurement of the binding of transcription factor Sp1 to a single GC box recognition sequence. *Nucleic Acids Res.* **17**, 2639-2653
- Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248-254
- Porath, J., Carlsson, J., Olsson, I., and Belfrage, G. (1975) Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* **258**, 598-599
- Schwabe, J.W.R. and Klug, A. (1994) Zinc mining for protein domains. *Nature Struct. Biol.* **1**, 345-349
- Harrington, M.A., Jones, P.A., Imagawa, M., and Karin, M. (1988) Cytosine methylation does not affect binding of transcription factor Sp1. *Proc. Natl. Acad. Sci. USA* **85**, 2066-2070
- Pavletich, N.P. and Pabo, C.O. (1991) Zinc finger-DNA recognition: Crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* **252**, 809-817
- Kuwahara, J. and Coleman, J.E. (1990) Role of the zinc(II) ions

concentration-dependent protein-protein interaction may contribute to the formation of multimeric complexes of BTEBs bound to the promoter region with tandemly repeated GC boxes. As suggested by Pascal and Tjian, Sp1 molecules bound to the adjacent or distal GC box sites interact to form high-order complexes responsible for the synergistic superactivation of transcription (21).

The amino acid sequence of three contiguous Cys₂-His₂ zinc finger domains of BTEB is similar to that of Sp1 but is distinct from those of Krox20 and Krox24 (also known as Zif268), serum-inducible transcription activators (22, 23) (Fig. 5A). The three contiguous zinc finger motifs of Krox20 and Krox24 recognize a GC-rich regulatory sequence similar to, but distinct from the target sequence of Sp1 (24, 25). From the crystal structure of a DNA complex of the three contiguous Cys₂-His₂ zinc finger peptide of Zif268, five arginyl residues and one histidyl residue were identified to be involved in the contact with guanine bases in one strand of double-stranded target DNA (14). These amino acid residues are all conserved in Krox20, but some of them are substituted in Sp1 and BTEB (Fig. 5A). In a DNA complex of BTEB, the guanine bases in both strands of BTEB probe were shown to be in contact with protein by footprinting methods (3). The three-dimensional structures reported more recently for the DNA-complexes of DNA-binding domains of *Drosophila* tramtrack protein (TTK) (26) and human glioblastoma protein (GLI) (27) offered versatile modes of interaction between the zinc finger and the DNA sequence. The short regions immediately N-terminal to the zinc finger domains of BTEB, Sp1, Krox20, and Krox24 are rich in basic amino acids, although they are not homologous to one another in amino acid sequence (Fig. 5B). The contribution of the basic region to the DNA-binding or the stabilization of the DNA-protein complex has been demonstrated with BTEB (4), but has not yet been reported with Sp1, Krox 20, or Krox 24 (Zif268). Conservation of the basic property in this region suggests that the basic regions play a common role in the DNA-protein interaction. The region immediately C-terminal to the zinc finger domain was also suggested to be involved in the DNA-binding of

- in the structure of the three-finger DNA binding domain of the Sp1 transcription factor. *Biochemistry* **29**, 8627-8631
16. Ghosh, G., van Duyne, G., Ghosh, S., and Sigler, P.B. (1995) Structure of NF- κ B p50 homodimer bound to a κ B site. *Nature* **373**, 303-310
 17. Müller, C.W., Rey, F.A., Sodeoka, M., Verdine, G.L., and Harrison, S.C. (1995) Structure of the NF- κ B p50 homodimer bound to DNA. *Nature* **373**, 311-317
 18. Schwabe, J.W.R., Chapman, L., Finch, J.T., and Rhodes, D. (1993) The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: How receptors discriminate between their response elements. *Cell* **75**, 567-578
 19. Ceska, T.A., Lamers, M., Monaci, P., Nicosia, A., Cortese, R., and Suck, D. (1993) The X-ray structure of an atypical homeodomain present in the rat liver transcription factor LFB1/HNF1 and implications for DNA binding. *EMBO J.* **12**, 1805-1810
 20. O'Shea, E.K., Klemm, J.D., Kim, P.S., and Alber, T. (1991) X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* **254**, 539-544
 21. Pascal, E. and Tjian, R. (1991) Different activation domains of Sp1 govern formation of multimers and mediate transcriptional synergism. *Genes Dev.* **5**, 1646-1656
 22. Chavrier, P., Zerial, M., Lemaire, P., Almendral, J., Bravo, R., and Charnay, P. (1988) A gene encoding a protein with zinc fingers is activated during G0/G1 transition in cultured cells. *EMBO J.* **7**, 29-35
 23. Lemaire, P., Revelant, O., Bravo, R., and Charnay, P. (1988) Two mouse genes encoding potential transcription factors with identical DNA-binding domains are activated by growth factors in cultured cells. *Proc. Natl. Acad. Sci. USA* **85**, 4691-4695
 24. Chavrier, P., Vesque, C., Galliot, B., Vigneron, M., Dolle, P., Duboule, D., and Charnay, P. (1990) The segment-specific gene Krox-20 encodes a transcription factor with binding sites in the promoter region of the Hox-1.4 gene. *EMBO J.* **9**, 1209-1218
 25. Lemaire, P., Vesque, C., Schmitt, J., Stunnenberg, H., Frank, R., and Charnay, P. (1990) The serum-inducible mouse gene Krox24 encodes a sequence-specific transcriptional activator. *Mol. Cell. Biol.* **10**, 3456-3467
 26. Fairall, L., Schwabe, J.W.R., Chapman, L., Finch, J.T., and Rhodes, D. (1993) The crystal structure of a two zinc-finger peptide reveals an extension to the rules for zinc-finger/DNA recognition. *Nature* **366**, 483-487
 27. Pavletich, N.P. and Pabo, C.O. (1993) Crystal structure of a five-finger GLI- DNA complex: New perspectives on zinc fingers. *Science* **261**, 1701-1707