Specific DNA Binding to a Major Histocompatibility Complex Enhancer Sequence by a Synthetic 57-Residue Double Zinc Finger Peptide from a Human Enhancer Binding Protein*

Kazuyasu Sakaguchi‡, Ettore Appella‡, James G. Omichinski§, G. Marius Clore§, and Angela M. Gronenborn§

From the ‡Laboratory of Cell Biology, National Cancer Institute and the §Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892

Two 57-residue peptides containing one pair of "zinc fingers" from a human enhancer binding protein were prepared by solid-phase peptide synthesis. One peptide (MBP-DF) contained the native sequence, while the second peptide ([Abu]MBP-DF) has an α-aminobutyric acid residue substituted for a nonconserved cysteine residue at position 11. The peptides were characterized by several chemical and physical methods, and their DNA binding properties were evaluated using gel retardation experiments. Spectroscopic studies demonstrated that addition of metal ions such as zinc and cobalt resulted in specific conformational changes in both peptides, indicating that cysteine-11 does not appear to be involved in metal chelation. One-dimensional 1H NMR studies indicate that a stable folded structure is formed upon addition of zinc, and the chemical shift pattern is consistent with that previously observed for one constituent single finger (Omichinski, J., Clore, G. M., Appella, E., Sakaguchi, K., and Gronenborn, A. M. (1990) Biochemistry 29, 9324–9334). Gel retardation experiments demonstrate that the peptides are capable of interacting with a 15-mer oligonucleotide comprising a portion of the major histocompatibility complex enhancer sequence and that the interaction is zinc-dependent. The dissociation constant for the [Abu]MBP-DF peptide is 1.4 × 10−7 M with maximal binding occurring at a zinc-to-peptide ratio of 2 to 1. The binding specificity observed with respect to related enhancer sequences exhibits the same relative order as noted previously for the whole protein. Studies with point mutants of the major histocompatibility complex enhancer binding sequence indicate that the last GC base pair in a four-guanine stretch plays a pivotal role in the interaction between the peptide and DNA.

Zinc finger motifs are now recognized as one of the major classes of DNA binding domains found in eukaryotic transcription factors (see Ref. 1 for a compilation of sequences and Refs. 2 and 3 for reviews). Models of the three-dimensional structure of Cys2-His2 zinc finger domains were initially proposed on the basis of amino acid sequence homologies and known x-ray structures of zinc-containing proteins (1, 4).

Subsequently, the main features of these models were confirmed by the determination of the solution structures of a number of zinc finger peptides by NMR spectroscopy (5–7). The NMR results demonstrate that these peptides exist as independently folded domains with a compact globular structure in which the zinc atom is chelated by the two invariant cysteine and two invariant histidine ligands. The three-dimensional structure consists of a short, irregular two-stranded antiparallel β-sheet, followed by an α-helix.

Several zinc finger-containing DNA-binding proteins have been identified, all of which bind to enhancer sequences within the promoter region of inducible genes. MBP-1 was identified as a transcriptional regulatory protein which recognizes a 13-base pair enhancer sequence located 165 base pairs upstream of the class I major histocompatibility complex (MHC) gene transcription initiation site (8). The DNA binding domain of this protein was found to be located within a sequence of 118 amino acids containing two zinc fingers of the Cys2-His2 family (9). Two other factors with properties similar to MBP-1 have also been identified. The first clone encodes an HIV-1 enhancer binding protein (HIV-EP1) from a human B-cell line (10), and the second clone, designated positive regulatory domain I1 of the human interferon-β promoter binding factor I (PRDII-BF1), was shown to comprise the sequences of the other clones (11). Thus, all three factors are identical and play an important role in the regulation of expression of different genes, exhibiting specific binding to the HIV-1 enhancer, and the enhancer regions of the MHC H2-Kb and the immunoglobulin κ genes (8–11).

At the present time there is no direct structural data for the detailed interaction between the DNA binding sequences and the zinc finger peptides or proteins. Methylation experiments using PRDII-BF1, as well as MBP-1, and the H2-Kb enhancer indicate that recognition involves primarily major groove contacts, as evidenced by the interference observed after methylation at N7 of several guanine bases (8, 10). Similar results were obtained for TFIIIA (12) and Sp1 (13). The length of the binding site has been proposed to contain 5–5.5 base pairs per finger (14). Not all zinc finger proteins, however, exhibit binding sites consistent with this proposal.

1 The abbreviations used are: MBP-1, major histocompatibility complex binding protein I; MBP-DF, major histocompatibility complex binding protein I double finger; HIV-EP1, human immunodeficiency virus type I enhancer binding protein; HIV 1, human immunodeficiency virus type I; PRDII, positive regulatory domain II of the human interferon-β promoter; PRDII-BF1, positive regulatory domain II of the human interferon-β promoter binding factor I; MHC, major histocompatibility complex; αEN, enhancer region of the immunoglobulin κ gene; Abu, α-aminobutyric acid; HPLC, high-performance liquid chromatography; DTT, dithiothreitol.

7306
In particular, Spl, which contains three fingers, recognizes a 10-base pair sequence (15), suggesting a binding site of only 3 base pairs or indicating the involvement of only two of the three fingers in the binding interaction. A puzzling feature of the MHC binding site recognized by MBP-1 and PRDII-BF1 is its palindromic structure which is not found for other enhancer sequences. It remains to be seen if this is of structural relevance or merely reflects the fact that different amino acids in the individual fingers are capable of interacting with identical base pairs in the two half sites in a similar fashion.

In order to gain insight into the structure and DNA binding properties of zinc finger domains we have embarked on a program to investigate peptides comprising zinc finger motifs from the proteins designated MBP-1 (8), HIV-EP1 (10), and PRDII-BF1 (11), all of which seem to be identical. Peptides corresponding to the double-finger have been chemically synthesized by solid-phase synthesis, and their DNA binding properties investigated. In this paper we report the preparation of these peptides as well as DNA binding studies involving several DNA enhancer sequences.

MATERIALS AND METHODS

Peptide Synthesis—Assembly of the polypeptide chain was carried out by stepwise solid-phase synthesis methods using an Applied Biosystems 430A automated peptide synthesizer. The following side chain protecting groups were used: tosyl (Arg), cyclohexyl (Glu, Asp), 4-chlorocarbobenzoxy (Lys), 2-bromocarbobenzoxy (Tyr), N'-benzoyloxyethyl (His), 4-methylbenzyl (Cys), and benzyl (Thr, Ser). The N'-Boc group of the final protected peptide-resin was removed by trifluoroacetic acid. Cleavage of the peptide from the resin and removal of side chain protecting groups was effected by the modified low-high pH method (16).

The crude peptide product was extracted with 5% acetic acid, 1 mM DTT applied to a Sephadex G-50 column and eluted from the column with 5% acetic acid, 1 mM DTT. Those fractions containing the peptide were further purified by reverse-phase high-performance liquid chromatography (HPLC) on a Vydac C-8 column using a water-acetonitrile solvent gradient system containing 0.05% trifluoroacetic acid. Purity of the peptide was established by HPLC (>95% as monitored at 280 nm), amino acid analysis, peptide sequencing, and mass spectrometry. Peptide concentrations were determined using an extinction coefficient of 3300 M⁻¹ cm⁻¹ for Tyr.

Circular Dichroism Studies—Circular dichroism (CD) spectra were recorded at 196 to 260 nm on a Jasco J-500 spectropolarimeter at 25 °C. The spectra were recorded in aqueous ZnCl₂ (10 mM) in the presence or absence of EDTA (15 mM) at pH 6.5. The spectra are presented as plots of mean residue ellipticity and have been corrected for background interference.

Nuclear Magnetic Resonance Spectroscopy—The sample for NMR spectroscopy contained 0.4 mM peptide in 90% H₂O, 10% D₂O, pH 6.0, in the presence of 0.8 mM ZnCl₂. One-dimensional spectra were recorded at 600 MHz on a Bruker AM-600 spectrometer at 15 °C and the water resonance was suppressed using a semiselective jump-return sequence (17). Chemical shifts are expressed relative to 4,4-dimethyl-4-silapentane-1-sulfonate.

Gel Mobility Shift Assays—The DNA binding reactions were carried out in 20 μl of 20 mM Tris-HCl buffer at pH 7.6 containing 50 mM NaCl, 1 mM MgCl₂, 5 mM DTT, 5% glycerol, and varying concentrations of either peptide or unlabeled DNA. Poly(dI·dC) (0.1 ug/μl) was used as a nonspecific competitor in determining the zinc dependence of binding, as well as in the direct binding of the MBP-DF and [Abu¹]MBP-DF peptides to the labeled MHC enhancer site. The reaction mixtures were incubated with 10 fmol (500 pm) of 32P-labeled oligonucleotide at room temperature for 60 min prior to electrophoresis on a 10% polyacrylamide gel (acylamide: bisacrylamide, 29:1, v/v). Electrophoresis was carried out at room temperature in Tris-borate-EDTA buffer (40 mM) at a constant voltage of 200 V. The gels were removed, dried, and autoradiographed on Kodak XAR-5 film. The data were quantified by liquid scintillation counting of gel slices.

RESULTS AND DISCUSSION

This report describes the synthesis, physical characterization, and DNA binding properties of two 57-residue peptides which contain one of two pairs of tandemly linked zinc fingers located in a human enhancer binding protein involved in transcriptional regulation of a variety of inducible eukaryotic genes. The first peptide (MBP-DF) contains the native sequence as derived from the nucleotide sequence of the corresponding gene, while the second peptide ([Abu¹]MBP-DF) comprises the identical sequence, except that an α-aminoacyl acid residue has been substituted for a cysteine residue at position 11 (see Fig. 1). This substitution was carried out with the purpose of increasing the chemical stability of the peptide, and in particular to avoid any potential problems arising from nonspecific disulfide formation via the free sulfhydryl of cysteine 11 in the folded zinc finger. Sequence comparisons with other zinc finger sequences (1) indicate that the fourth position after the second metal liganding cysteine is occupied predominately by hydrophobic amino acids. In addition, the NMR structure of the COOH-terminal zinc finger peptide in which phenylalanine is the residue in the equivalent position clearly demonstrates that the Phe side chain is part of the hydrophobic core in the folded structure (5). We thus assumed that Cys¹ in the COOH-terminal finger could be replaced by α-aminoacyl acid without perturbing the native fold of the peptide.

Peptide Synthesis and Characterization—The peptides were prepared to greater than 95% purity as judged by several criteria including HPLC, mass spectrometry, and amino acid sequencing, and the sequencing data confirmed the expected amino acid sequence. The measured molecular weights as determined by Fourier transforms mass spectrometry using electrospray ionization were 6772 for MBP-DF (calculated value, 6772.1) and 6753 for [Abu¹]MBP-DF (calculated value, 6754.1).

Spectroscopic Characterization—The CD spectrum of [Abu¹]MBP-DF in the absence of free zinc ions (+EDTA) is characteristic of a random coil conformation. The addition of zinc introduces major changes in the spectrum (Fig. 24). A significant increase in negative molar ellipticity near 222 nm is observed as well as a reduction near 205 nm. These results indicate that the peptide has an ordered structure with some α-helical characteristics in the presence of zinc, and the spectral pattern is similar to that seen with the constituent single fingers (data not shown). The CD spectrum of MBP-DF is essentially identical to that of the [Abu¹]MBP-DF peptide (data not shown).

Visible absorption spectra of both double-finger peptides in the presence of excess cobalt show the Co(II) d-d transition maximum at ~640 nm and the S-Co(II) charge-transfer bands at 310-350 nm (data not shown), indicative of tetrahedral coordination of the metal ion (18). Addition of essentially

Fig. 1. Amino acid sequence of the MBP-DF peptide, corresponding to position 1386–1442 of MBP-1 (10). The proposed cysteine coordination is indicated and the site of replacement of Cys-11 by α-aminoacyl acid is boxed.
Specific DNA Binding to an Enhancer Sequence

stoichiometric amounts of zinc to the cobalt-peptide solution results in the disappearance of the d-d transition band and the charge-transfer bands, even in the presence of a 10-fold higher concentration of cobalt versus zinc, indicating that the affinity of the peptide for zinc is considerably higher than that for cobalt.

The one-dimensional 'H NMR spectrum of [Abu]MBP-DF in the presence of zinc is shown in Fig. 2B. The spectrum is substantially different from that obtained in the absence of zinc (not shown) which displays characteristics typical of a random coil conformation. In the presence of zinc, the chemical shift dispersion increases dramatically, and the chemical shift pattern observed with the double-finger peptide in the presence of zinc is consistent with that obtained with the constituent carboxyl-terminal single finger for which a three-dimensional solution structure has been determined (5).

The spectroscopic results indicate that the double-finger peptide is capable of forming a stable folded structure in the presence of zinc, similar to that observed for the individual single fingers. In addition, substitution of an α-amino butyric acid residue for the nonconserved cysteine residue at position 11 does not affect the metal binding properties of the peptide as judged by absorption, CD, and NMR spectroscopy, and supports the hypothesis that this residue is not involved in zinc chelation, but in all likelihood acts as a hydrophobic residue in the core of the amino-terminal finger.

Characterization of DNA Binding—The DNA-binding properties of the peptides were analyzed by a gel mobility shift assay (19, 20). First, the peptides were incubated with a synthetic 16-base pair double-stranded oligonucleotide identical to the binding region in the class I MHC enhancer site (RI; see Fig. 3) at various concentrations of zinc. Neither MBP-DF nor [Abu]MBP-DF demonstrated any ability to bind to the RI probe in the absence of zinc ions. In the presence of zinc, binding was observed with both peptides and this binding was enhanced by increasing concentration of zinc ions until a plateau was reached at a 2:1 molar ratio of zinc to peptide (Fig. 3; data is only shown for one of the two peptides). Thus, each zinc finger domain binds one zinc ion.

At very high zinc concentrations, however, the degree of binding decreased in a manner similar to that reported for a double-finger fragment of the glucocorticoid receptor (21), a protein containing Cys2-Cys2 fingers which bind zinc via four cysteine residues. This influence of a high zinc ion concentration may not be directly related to an interaction with the peptide. Rather, it is more likely that the reduced binding is caused by effects mediated through binding of the divalent zinc ions to binding sites on the base ring nitrogens in DNA (22), thereby causing steric and/or electrostatic interference in any potential peptide-base contacts.

Studies with varying concentrations of the two peptides indicated that the affinity for the RI probe appeared to be slightly higher with the α-amino butyric acid-substituted peptide (Fig. 4). The difference in apparent binding affinity,
however, is very small and most likely represents the fact that
the cysteine-containing peptide exhibits some tendency to
form disulfide-linked dimers which cannot fold correctly and,
in turn, have lost their ability to bind specifically to DNA.

Thus, the concentration of active peptide for MBF-DF is
probably somewhat lower than that measured by absorbance.
The enhanced chemical stability and, therefore, better handling
properties of the [Abu\(^{11}\)]MBP-DF, in conjunction with
basically unchanged properties, prompted us to use this pep-
tide for all further studies.

To quantify the specific binding of the double-zinc finger
peptide to the MHC enhancer sequence (RI) we performed a
nonlinear least squares fit to the binding data obtained at a
2:1 zinc-to-peptide ratio. This yielded a dissociation constant
of 1.4 \(\times 10^{-7}\) M for the [Abu\(^{11}\)]MBP-DF-RI complex (Fig. 5). The
binding constant for the entire protein may well be higher
than that measured for the peptide, assuming that the peptide
represents the minimal unit within the entire protein se-
quence that is capable of specific DNA binding. Thus, addi-
tional features of the native protein might contribute to the
overall binding interactions in the full length protein. At this
time, however, there are no quantitative data available on the
binding of the entire protein to DNA.

The specificity of the interaction of the peptide with the RI
probe was evaluated by employing unlabeled oligonucleotide
competitors in the presence of the labeled RI probe. The
competing oligonucleotides included the RI sequence as well

FIG. 4. Comparative gel shift curves for the MBP-DF pep-
tide and the modified [Abu\(^{11}\)]MBP-DF peptide. Each lane con-
tains 500 pm radiolabeled RI probe, 0.1 \(\mu\)g/\(\mu\)l poly(dI-dC), 0–100 \(\mu\)M
\(\text{ZnCl}_2\) (equivalent to 2 \(\times\) the peptide concentration), and varying
amounts of peptide.

FIG. 5. Titration curve for the [Abu\(^{11}\)]MBP-DF peptide with the
MHC wild-type enhancer probe (RI). A, gel shift assay; the migration
positions of free (\(F\)) and bound oligo-
nucleotide (\(B\)) are marked by arrows. B,
binding curve derived from A by plotting
the percentage of radioactivity bound versus the peptide concentration.
as enhancer sequences from the \( \kappa \) immunoglobulin (\( \kappa \)EN), HIV-1 (HIV), the MHC class I region II (RII) promoter regions and a double-point mutant of the RI sequence (mhc). The sequences for these nucleotide sequences are summarized in Fig. 6. The unlabeled RI sequence was the most potent inhibitor of complex formation, while the \( \kappa \)EN, the HIV enhancer, and the double-point mutant (mhc) sequences were able to inhibit complex formation at slightly higher concentrations (Fig. 6). The RII probe was essentially unable to efficiently inhibit the peptide-RI complex formation. At 100 nM of competitor oligonucleotide, the relative strength for effective competition was RI > \( \kappa \)EN ≥ HIV ≥ mhc >> RII. This is the same relative order as was observed with the truncated MBP-1 protein, expressed in vitro, which exhibited high affinity binding to the class I MHC region I and \( \beta \)-microglobulin sites and a lower affinity for the \( \kappa \) enhancer and HIV sites (9, 10). All those sequences that exhibit appreciable binding contain the last two guanines in the stretch of four guanines in the left half of the symmetrical sequence of RI, whereas the RII oligonucleotide does not share the same symmetrical sequence pattern. RII, however, does contain two stretches of four guanines, the presence of which alone seems to be insufficient to promote significant binding specificity.

To further probe which nucleotides are responsible for the specific binding interactions between the peptide and the RI probe, five point mutants (M1-M5) of the RI probe were synthesized in which each guanine was replaced in turn by cytosine and one adenine was replaced by thymine (Fig. 7). Competitive binding studies employing the mutant oligonucleotides (Fig. 7) demonstrate that the M1 (G-4 to C-4) and M2 (G-5 to C-5) variants competed for the peptide at nearly identical concentrations as that observed with RI itself. The M3 (G-6 to C-6) and the M5 (A-8 to T-8) variants were somewhat less effective, while the M4 (G-7 to C-7) variant failed to compete with the RI probe even at extremely high concentrations. These results indicate that the guanine at position 7 is crucial for the interaction of the peptide with the RI probe, and that the guanine at position 6 and the adenine at position 8 are also involved.

Concluding Remarks—In this paper we have demonstrated that a synthetic 57-residue double-zinc finger peptide from the human enhancer binding protein MBP-1/PRDII-BFl/HIV-EP1 is capable of specific DNA binding in its own right. So far the demonstration of specific DNA binding of Cys2-His, zinc finger containing proteins have, to our knowledge, always involved larger polypeptide domains, the smallest unit used comprising an 89-residue polypeptide containing the three zinc finger domains of SW15 (23). Thus, the present double-finger appears to be the smallest unit so far capable of directing specific DNA binding. The DNA binding experiments with a range of synthetic oligonucleotides reported here are in agreement with the methylation interference studies which implicated the involvement of the four-guanine stretch in the binding with the full length MBP-1 protein (8), and extend this observation in pointing to the GGA trinucleotide stretch as being the most important region for sequence-specific binding. Several factors have been reported to bind to the symmetrical MHC enhancer site and with different affinity to the \( \beta \)-microglobulin, \( \kappa \) chain, SV40, \( \beta \)-interferon, and HIV sites (9). The relationship between these factors and the MBP-1/PRDII-BFl/HIV-EP1 protein (9, 11) remains unknown. Recently, the nuclear factor NF-\( \kappa \)B, which was first described as a protein binding to the immunoglobulin \( \kappa \) enhancer (24), has been cloned (25, 26). Analysis of its nucleotide sequence reveals that no zinc finger domain is present. This indicates
that proteins containing different types of DNA binding motifs can bind to the same DNA sites and that DNA binding represents only one of several interactions that are important for the cellular function of these factors.

Acknowledgments—We thank Dr. D. Hunt, University of Virginia, Charlottesville, for the mass spectrometry and Dr. J. Romano, NCI for help in labeling the oligonucleotides.

REFERENCES