Identification of a 34 Amino Acid Stretch within the C-Terminus of Histone H1 As the DNA-Condensing Domain by Site-Directed Mutagenesis†

M. M. Srinivas Bharath,‡ Sneha Ramesh,‡ Nagasuma R. Chandra,§ and M. R. S. Rao*,‡

Department of Biochemistry and Bioinformatics Centre, Indian Institute of Science, Bangalore 560012, India

Received March 5, 2002; Revised Manuscript Received April 26, 2002

ABSTRACT: The C-terminus of histone H1 is necessary for the folding of polynucleosomal arrays into higher-order structure(s) and contains octapeptide repeats each having DNA binding S/TPKK motifs. These repeat motifs were earlier shown to mimic the DNA/chromatin-condensing properties of the C-terminus of histone H1 (Khadake, J. R., and Rao, M. R. S. (1995) *Biochemistry 36*, 1041−1051). In the present study, we have generated a series of C-terminal mutants of rat histone H1d and studied their DNA-condensation properties. The single proline to alanine mutation in the S/TPKK motifs either singly or in combination resulted in only a 20% decrease in the DNA-condensation property of histone H1. Deletion of all the three S/TPKK motifs resulted in a 45% decrease in DNA condensation. When the three octapeptide repeats encompassing the S/TPKK motifs were deleted, there was again a 45% decrease in DNA condensation. On the other hand, when the entire 34 amino acid stretch (residue 145−178) was deleted, there was nearly a 90% decrease in DNA condensation brought about by histone H1d. Interestingly, deletion of the 10 amino acid spacer between the octapeptide repeats (residues 161−170) also reduced the DNA condensation by 70%. Deletion of the region (residues 115−141) immediately before the 34 amino acid stretch and after the globular domain and the region (residues 184−218) immediately after the 34 amino acid stretch had only a marginal effect on DNA condensation. The importance of the 34 amino acid stretch, including the 10 amino acid spacer, was also demonstrated with the recombinant histone H1d C-terminus. We have also determined the induced α-helicity of histone H1 and its various mutants in the presence of 60% trifluoroethanol, and the experimentally determined induced helical contents agree with the theoretical predictions of secondary structural elements in the C-terminus of histone H1d. Thus, we have identified a 34 amino acid stretch in the C-terminus of histone H1d as the DNA-condensing domain.

The organization of DNA into chromatin in the eukaryotic cell nucleus is an important structural framework which influences various DNA transaction processes. The four core histones H2A, H2B, H3, and H4, in the form of an octamer, folds 146 bp of DNA around it to generate the fundamental unit of chromatin, the nucleosome core particle, whose structure has been solved recently at 2.8 Å resolution (*1*). The folding of the polynucleosomal array into chromatin fiber (popularly known as 30 nm fiber) is dependent on both the ionic strength and the fifth histone, H1 (*2*, *3*). The histone H1 protein consists of three distinct domains, a small 34 residue N-terminal fragment (nose), the central 74 residue globular domain (head) and a slightly larger 110 residue C-terminal domain (tail). In aqueous media, at physiological pH and ionic strength, the N-terminal nose is believed to have no regular structure, while the C-terminal domain, which is otherwise unfolded, folds into an ordered structure in the presence of nucleic acids. The structure of the globular domain has been well characterized both by NMR (*4*) and X-ray crystallography (*5*), whereas virtually nothing is known about the structure of the C-terminal domain. The globular domain primarily consists of a three-helical bundle, which has been described as a variant of the helix-turn-helix type of DNA-binding proteins. Biochemical experiments involving mutational analysis have demonstrated that the globular domain has two distinct DNA binding sites (*6*). The globular domain is also necessary for generating the kinetic intermediate "chromatosome particle" observed during micrococcal nuclease digestion of nucleosome, containing 166 bp DNA and histone H1 (*7*). The location of the globular domain, however, within the nucleosome has been controversial (*8*, *9*).

Although the analysis of higher-order structure of chromatin has interested researchers for decades, the field has been beset with controversies. Most studies carried out to date have utilized chromatin fragments isolated from lysed nuclei, and it has been difficult to obtain information about these structures in vivo. The models in vogue today to visualize chromatin folding portray the chromatin fiber to vary between the zigzag and the solenoid conformation.

† This work was financially supported by grants from the Council of Scientific and Industrial Research and the Department of Biotechnology, New Delhi. Use of facilities at the Interactive Graphics Based Molecular Modeling Facility and the Distributed Information Center (both supported by Department of Biotechnology) and the facilities at the Super Computer Education and Research Center are gratefully acknowledged.

* Corresponding author. Phone: 91-80-3600118, 91-80-3942547. Fax: 91-80-3600118, 91-80-3600814. E-mail: mrsrao@biochem.iisc.ernet.in.

‡ Department of Biochemistry.

§ Bioinformatics Centre.

10.1021/bi025773+ CCC: $22.00
Published on Web 05/23/2002

Recently, there has been considerable discussion as to the validity of these two models (*10*). Irrespective of the model, it is well established that histone H1 is necessary for the folding and stabilization of the linear polynucleosomal array into chromatin fiber and chromatin compaction in vivo (*11*). The structural basis of the role played by histone H1 in chromatin folding has remained a challenging problem for several years. Early studies by Allan et al. (*2*) have clearly shown that the C-terminal domain within histone H1 is necessary for the folding of chromatin into higher-order structures. Histone H1 also condenses DNA in vitro, generating what is termed as packed duplexes or liquid crystals exhibiting a characteristic ψ-type spectrum in circular dichroism (*12*). Although the molecular basis of such a phenomenon of condensation is not clear, it is generally believed that such an inherent property of the histone H1 molecule does contribute to the folding of chromatin into higher-order structures. This assay has been exploited in the literature to study the functional properties of histone H1 in vitro. It is interesting to note that this DNA-condensation property resides within the C-terminus of histone H1 (*13*). There are seven variants of histone H1 in mammals, and although the amino acid sequences of the globular domain and hence its structure are fairly conserved among all the known histone H1 subtypes known, the C-terminal domain shows a considerable divergence in its sequence. At the level of protein−DNA interaction in vitro, experiments have shown that the different subtypes of histone H1 have different DNA (*14*) and chromatin (*15*) condensation properties. It is also quite likely that the different histone H1 subtypes might generate subtly different higher-order structures in vivo. Furthermore, if we are able to study the histone H1−DNA or chromatin interactions by analyzing the regions within the C-terminus that is necessary for condensation, it could lead us to defining the structural basis of histone H1 in chromatin compaction. While studying the DNA-condensation properties of rat testis specific histone, H1t, we observed that it is a poor condenser of DNA and chromatin in vitro compared to the somatic variants H1bdec (*16*). A comparison of the amino acid sequences of the somatic and testis specific histone H1 revealed that this poor condensation property may be due to the absence of an octapeptide repeat unit, containing an S/TPKK motif that is present in histone H1d but absent in histone H1t. This octapeptide repeat which constitutes a unique structural feature of the C-termini of most of the somatic histone H1s contains a core S/TPKK sequence which was earlier shown to bind DNA (*17*). Subsequently, we were able to show that a synthetic 16-mer peptide ATPKKSTKKTPKKAKK, which corresponds to the tandem repeats of the octapeptide repeats within histone H1d C-terminus, could mimic the DNA-condensation property of full-length histone H1 (*18*). Although there are some reports in the literature explaining the role of DNA binding units such as the S/TPKK, there has been no attempt to narrow down to the region within the intact histone H1 molecule which harbors the in vitro DNA and chromatin condensing properties. Toward this objective, it is very important to understand the significance of this octapeptide repeat unit and its role in the context of full length histone H1 protein. For this purpose, we have generated various mutants of rat histone H1d in its C-terminus through site-directed mutagenesis, and all the mutants were analyzed with respect to their DNA-condensation properties.

EXPERIMENTAL PROCEDURES

Cloning, Mutagenesis, Expression, and Purification of Histone H1d and Its Mutants. The plasmid 1T, containing the histone H1d genomic DNA insert, was a generous gift from Dr. W. S. Kistler. The PCR amplification of the coding region, the cloning into the pTrc 99A vector, and its expression has been described by us recently (*19*). Site-directed mutagenesis of specific residues was accomplished by a modified PCR megaprimer method, described previously (*20*). An important requirement of this method is to have two templates, in which one of the templates has an extra flanking sequence in addition to the gene or part of the gene of interest, to which one of the side primers (vector specific primer) binds while the other template has no binding site for this side primer (see Figure 2). The second template can have either different flanking sequences or sequences with mismatches. In the present study, the first template contains the rH1d coding sequence in the background of pET 22b(+). All the mutagenic primers (proline to alanine) were designed by taking care of all the parameters for proper annealing to the wild-type template. The mutagenic primers are in an orientation such that the PCR product will be formed with each mutagenic primer (m) and only one common vector specific primer (v). PCR was carried out using *Pfu* polymerase, and the sequences were confirmed on a ABI 377 Automated DNA sequencer.

For the generation of deletion mutant constructs (wherein exactly four or eight amino acids or the different regions of the C-terminus have to be deleted), we resorted to the PCR-based method since restriction enzymes could not be employed. In this method, if a particular stretch of amino acids within the coding region of histone H1d had to be deleted, specific primers flanking the region of deletion were used as shown in Figure 2. Here the histone H1d insert was amplified using the two primer sets, A/B and C/D. The PCR was carried out with *Pfu* polymerase. After confirming the sequence, we phosphorylated the products at their 5′ ends using polynucleotide kinase and ligated the blunt end. One aliquot of the ligation mixture was subjected to PCR using the extreme primers A and D. Finally, the full length PCR product obtained was subcloned into the Nco I and Bam H I sites of pTrc 99A.

The plasmid pH1d was amplified with the following primer sets to obtain the C-terminus of the protein: 5′ upstream primer, 5′ CGC CCA TAT GAA GGC GGC TTC CGG CGA A 3′, which includes the Nde I adapter; 3′ downstream primer, 5′ CGG GAT CCT TAT TGA TGG TGA TGG TGA TGT TTC TTC TTG GCT TGC T 3′, which includes the Bam H I adapter and the 6-histidine tag. For generating the C-terminus mutants, we used the mutant constructs obtained as described above as the templates. PCR amplification was done with *Pfu* polymerase, and the sequences of the products were confirmed by sequencing. After digestion with Nde I and Bam H I, they were directionally cloned into pET 22 b(+) expression vector.

All the expression clones were transformed into *Escherichia coli* BL 21 (DE3) cells and induced with 0.5 mM IPTG. Expressed proteins were purified by a combination of Ni^{2+}−agarose and heparin agarose chromatography as described by us earlier (*19*).

Circular Dichroism Spectroscopy of Histone H1d (Wild-Type and Mutants) and Its Complexes with DNA. The circular dichroism spectra of recombinant histone H1d and its various mutants were recorded in a buffer containing (1) 10 mM sodium phosphate buffer, pH 7.5, (2) the same buffer containing 1M NaCl, and (3) 60% trifluoroethanol in a JASCO J-700 spectropolarimeter. The percent helicity was calculated according to the method of Clark et al. (*21*). The circular dichroic spectra of rat oligonucleosomal DNA (0.8−2 kb in length) and its complexes obtained after adding increasing concentrations of recombinant histone H1d or its mutants in 150 mM NaCl/10 mM Tris HCl, pH 7.5/0.1 mM EDTA and incubating at room temperature for 10 min were recorded in a JASCO JA 20 spectropolarimeter. There were no changes in the spectra after 10 min, indicating that the spectra recorded represented complexes formed at equilibrium. A molar extinction coefficient of 6700 was used for calculating the concentration of DNA and a mean residue weight of 330 to calculate the mean residue ellipticity θ of DNA.

Gel Retardation Assay with C-Terminus of Wild-Type and Mutant Histone H1d. Two oligonucleotides (21-mer each) complementary to each other were used to generate a 21-mer AT-rich duplex (oligo 1, 5′ GCG CAA AAT ATT GAA AAC GCC 3′; oligo 2, 5′ GCG GCG TTT TCA ATA TTT TGC 3′; *22*). The radioactive annealed probe was used for gel retardation assay with both the wild-type and mutant C-termini of histone H1d as described by Payet et al. (*22*). Increasing concentrations of the proteins were incubated in the GRA buffer (50 mM Tris HCl, pH 7.4/10 mM $MgCl_2$/20 mM DTT/50 ug/mL BSA) with 10 ng of the probe (approximately 30 000 cpm) for 30 min at room temperature to allow the formation of DNA−protein complexes. The complexes were then analyzed on a 6% native PAGE running for 5−6 h at 50 V at 20 °C.

Analytical Methods. Protein concentrations were determined either by the turbidometric method (*23*) or by using the following formula: absorbance at 230 nm = 1.85 mg/mL. SDS−PAGE was conducted according to the method of Laemmli (*24*).

RESULTS

Figure 1A shows the amino acid sequence of rat histone H1d, wherein the three helical regions within the globular domain and the three octapeptide repeat units containing the S/TPKK motifs within the C-terminus are highlighted. Among these octapeptide repeats, the first two are in tandem and span amino acids 144−159, with the proline residues being 146 and 154. The third octapeptide unit is present 10 amino acids away from the first two and spans amino acids 170−177, wherein the proline in the S/TPKK corresponds to amino acid number 172. A multiple alignment of the C-terminal domain of various H1 subtypes from several mammalian species as shown in Figure 1B highlights the diversity in the length of the domain as well as the relative positions of the known functional motifs.

Mutational Analysis of Histone H1d with Respect to DNA Condensation. We have generated several point and deletion mutants spanning the regions corresponding to the octapeptide repeats within the C-terminus of histone H1d. All the mutant proteins were purified to homogeneity and analyzed. The condensation property was measured by circular dichroism spectroscopy (*14*, *16*, *18*). Addition of increasing concentration of histone H1 to B-type DNA progressively decreases the positive ellipticity at 270 nm to generate a ψ type of spectrum characteristic of DNA condensation. We have used rat oligonucleosomal DNA of size 0.8−2.0kb in all the circular dichroism experiments. Earlier experiments from our laboratory had indicated that at least at the peptide level, the proline-dependent β-turn structure within the S/TPKK motifs of the octapeptide repeats is essential for in vitro DNA condensation (*18*). Using this as the starting point, we mutated the three proline residues of the three SPKK motifs in the C-terminus of histone H1d to alanine either singly or in different combinations to generate single, double, and triple mutants. Figure 2 shows the strategies that we have employed for generating point and deletion mutants. Figure 3 shows the SDS−PAGE patterns of the wild type and the proline-to-alanine mutants (A, C, and E) and their effect on DNA condensation. For a better understanding of the condensation properties, we plotted the decrease in the positive ellipticity θ at 270 nm, termed as $\Delta\theta$, upon each addition of the individual histone against "*r*" values (protein/DNA ratio, mol/bp). The net condensation observed with the single mutants was decreased by 5−10% (Figure 3B). Next, combinations of two among the three proline residues were substituted by alanine to generate three double mutants. As can be seen from Figure 3D, the effect of double mutations on the ability of histone H1d to condense DNA was reduced only marginally and was equivalent to that of single mutants. Substitution of all the three prolines to alanine resulted in a 20% reduction in the ability of histone H1d to condense DNA (Figure 3F). This observation is surprising, since at the peptide level we had observed that a change in the proline to alanine resulted in a drastic reduction in the ability of the 16-mer peptide to condense DNA (*18*). The S/TPKK motifs adapt β-turn structures and are shown to snuggly fit into the minor groove of AT-rich DNA (*17*). It is possible that in the context of a smaller stretch of amino acids, the β-turn generated is greatly influenced by the proline residue per se. However, in the context of the full-length histone H1d protein, the tertiary structure generated in the C-terminal domain spanning the octapeptide units would stabilize the turn structure.

To further examine the role of the S/TPKK motifs in DNA condensation, we then generated deletion mutants of these tetra peptide motifs (145−148, 153−156, and 171−174, respectively) either singly or in combination. The SDS−PAGE pattern (A, C, and E) and the effect of deletion mutants on DNA condensation are shown in Figure 4. The effect of single mutations consisting of individual deletions of each one of the three motifs Δ4I, Δ4II, and Δ 4III had only marginal effect in that there was only a 10% reduction in DNA-condensation ability as compared to the wild-type histone H1d (Figure 4B). The double mutants showed a reduction of about 25% (Figure 4D). Interestingly, when all the three S/TPKK units were deleted, there was a 45% reduction in their DNA-condensation ability (Figure 4F). We next generated deletion mutants of histone H1d wherein the three octapeptide repeats (144−151, 152−159, and 170−177 respectively) were deleted either singly or in combination and checked for their ability to condense DNA. These results are presented in Figure 5. The single, double, and triple

B. COMPARISON OF SEQUENCES OF C-TERMINUS OF HISTONE H1 IN MAMMALS

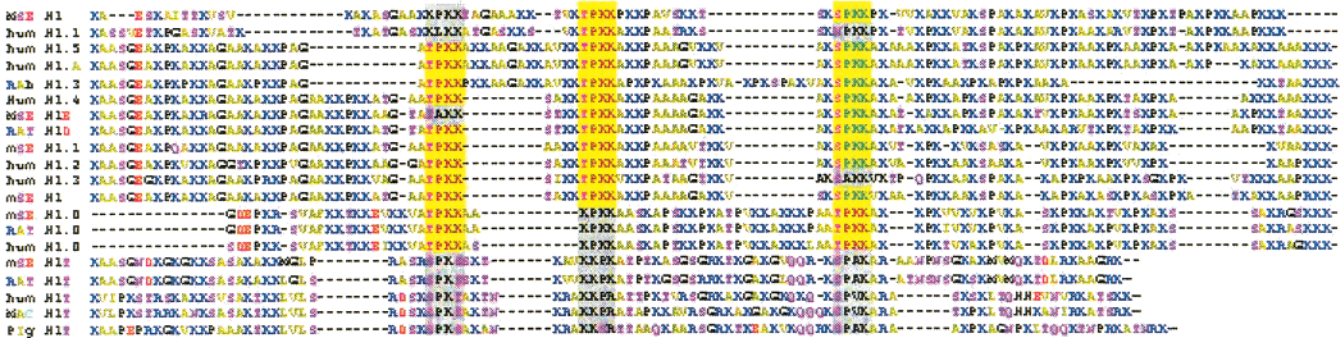


FIGURE 1: (A) Amino acid sequence of rat histone H1d showing the different structural domains. The octapeptide repeats containing the S/TPKK motifs present in the C-terminal domain are also highlighted. (B) A multiple alignment of the C-terminal domains in several mammalian linker histones. Basic residues are shown in blue, acidic residues in red, polar uncharged residues in pink, prolines and glycines in black, and nonpolar residues in yellow. The three known DNA binding motifs are highlighted in yellow. The alignment was initially obtained using CLUSTALW (*37*) and was then edited to maximize alignment of the three known functional motifs, as well as the predicted secondary structural elements. Residues in the α-helices predicted in the rat H1d sequence are enclosed in rectangles. Variants of the functional motifs found in some sequences are highlighted in gray.

deletion mutants showed a reduction of 25% (B), 40% (D), and 45% (F), respectively. We would like to add here that the number of basic amino acids (lysine + arginine) that are removed in the triple deletion mutant are 12 while there are still 32 residues present in the deleted C-terminus. This emphasizes that it is just not the presence of basic residues but a particular tertiary fold is necessary for DNA condensation.

The results described until now indicate that apart from the octapepetide repeats there may be a stretch of amino acids within the C-terminus which may also be contributing to condensation. There is a stretch of 10 amino acids (161−170) that separates the first two octapeptide repeats and the third octapeptide. We surmised that this spacer region may play an important role in facilitating DNA condensation. To test this possibility, a deletion mutant was generated wherein 34 amino acids corresponding to the three octapeptide repeats, including the intervening stretch of 10 amino acids, were deleted. This 34 amino acid deletion mutant lost 90% of the DNA-condensation ability as compared to the wild-type histone H1d (Figure 6B). This observation not only stresses the importance of the intervening sequence, it also clearly defines this 34 amino acid stretch as the "DNA-condensing unit" of histone H1d. To further evaluate the role of the spacer sequence, we created a deletion mutant of histone H1d encompassing this spacer sequence alone, keeping the repeat motifs intact. We were surprised to see that this Δ10 mutant showed only 30% of the DNA-condensation ability of the wild-type histone H1d (Figure 6 C). This result raises the possibility that the spacer may either be a part of the DNA-condensing unit or play a structural role that is important for spatial disposition of the S/TPKK motifs (see below).

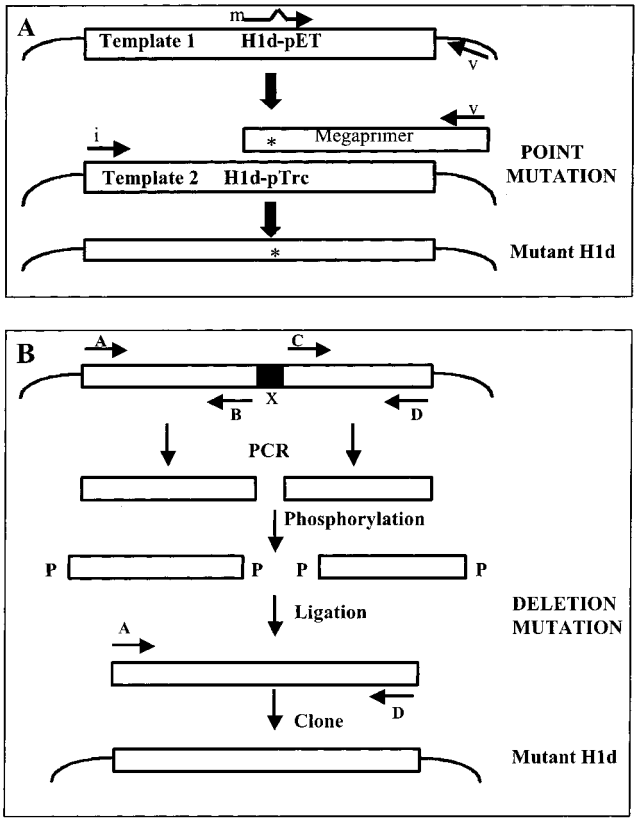


FIGURE 2: Schematic representation of the protocol employed for mutagenesis of histone H1d. Panel A depicts the modified megaprimer method of generating point mutations wherein "v" corresponds to the vector specific primer, "m" to the mutagenic primer, and "I" to the insert specific primer. Panel B shows the strategy used to generate deletion mutants wherein the region to be deleted is indicted as X. Primers used are depicted as A−D.

These data raise the obvious question whether the reduction in DNA condensation observed with the histone H1d mutants is because of lack of DNA binding. Although histone H1 does not have any sequence specificity, it is known to preferentially bind to AT rich sequences in the minor groove. We have carried out standard electrophoretic mobility shift assay with wild-type and mutant histone H1 using AT-rich oligonucleotide as the probe. As mentioned earlier, apart from the DNA binding and condensing capacity of the C-terminus, even the globular domain of histone H1/H5 has two DNA binding sites (*6*). Hence, if the full-length histone H1 is used for the DNA binding studies, the interpretations regarding the C-terminus binding to DNA can be misleading. To circumvent this problem, we have expressed only the C-termini of the wild type and mutants of histone H1. The C-terminus of histone H1d (residues 109−218) was PCR amplified and cloned into pET-22b expression vector, and the deletion mutants were generated. We have further confirmed the observations obtained with the full length proteins by carrying out similar condensation studies with the C-terminus of histone H1d. The SDS−PAGE profile of the C-terminus and its mutants and their DNA-condensation property are shown in panels D and F of Figure 6, respectively. Both the Δ34 and Δ10 mutants of the C-terminus showed a reduction of 80% in the DNA-condensation ability.

In a recent study, Payet et al. (*22*) have used a 21-mer AT-rich oligonucleotide for their analysis of mode of interaction of HMG-D protein with DNA. We have used the same sequence to examine the DNA binding properties of the wild type and Δ10 and Δ34 mutants by the gel retardation assay. With increasing concentrations of wild-type C-terminus, there was a retardation of the oligonucleotide probe (Figure 6F). At higher concentrations of the protein, slower mobility complexes were observed possibly due to the binding of more than one C-terminus to the same duplex. There was a decrease in the affinity of binding of the C-terminal Δ34 mutant (lanes 16, 17, and 18 as compared to lanes 4, 5, and 6), showing that this domain is an important DNA binding motif which is necessary for binding as well as subsequent DNA condensation. But the most interesting observation in Figure 6F is that there was only a marginal decrease in the DNA binding of the C-terminal Δ10 mutant although there was a considerable reduction in the DNA-condensation ability of this mutant. This probably suggests that the intervening sequence is playing a structural role in the overall folding of the C-terminus of histone H1d in a way which would facilitate optimal orientation of the DNA binding S/TPKK motifs in space for interaction with DNA. If the two mutants Δ34 and Δ10 are compared, the only missing region in the Δ34 spans the octapeptide repeats. Hence, it is obvious that the octapeptide repeats possess the condensing property and most of the DNA binding property. Since the Δ34 mutant still possesses residual DNA binding property, it is possible that there might be other regions beyond the 34-mer which can still bind DNA.

Finally, for an unequivocal assignment of the DNA-condensing function to the 34 amino acid stretch as described above, we wanted to see the effect of deletion of amino acid residues 115−141 that is immediately after the globular domain and before the 34 amino acid stretch as well as the region 184−218 that is immediately after the 34 amino acid stretch on the DNA-condensing property of histone H1d. For this purpose, we constructed the deletion mutants, and the expressed mutant proteins were purified as described above. The SDS−PAGE pattern of these histone H1d deletion mutants is shown in Figure 7A. The DNA-condensation properties of these mutant proteins in comparison with wild-type histone H1d are shown in Figure 7B. It is clear from the data that there was only a 15−20% reduction in their condensation ability, as compared to the 90% reduction observed with Δ34 mutant at equivalent protein/DNA ratios.

Alpha Helicity of Histone H1d and Its Mutants. The C-terminus of histone H1 is mostly random coil in solution, which, however, attains significant α-helicity in 60% trifluoroethanol (*21*). The low level of helicity of histone H1 in aqueous solution has been explained to be due to electrostatic repulsion between the charged lysine side chains. It is believed that the C-terminus of histone H1 upon interaction with DNA attains its optimum secondary structure, perhaps stabilized by charge neutralization. We have recorded the circular dichroism spectra of the various mutants of histone H1d that we have generated in the present study in the presence of1M NaCl as well as 60% trifluoroethanol. The helicity observed with NaCl is mostly with respect to the globular domain, while that in the presence of 60% trifluroethanol is the induced helicity observed in the C-terminus. The percent α-helicity observed with each of the proteins in the presence of trifluroethnaol is given in Table 1. Deletion of the 34 amino acid stretch (Δ 34, 145−

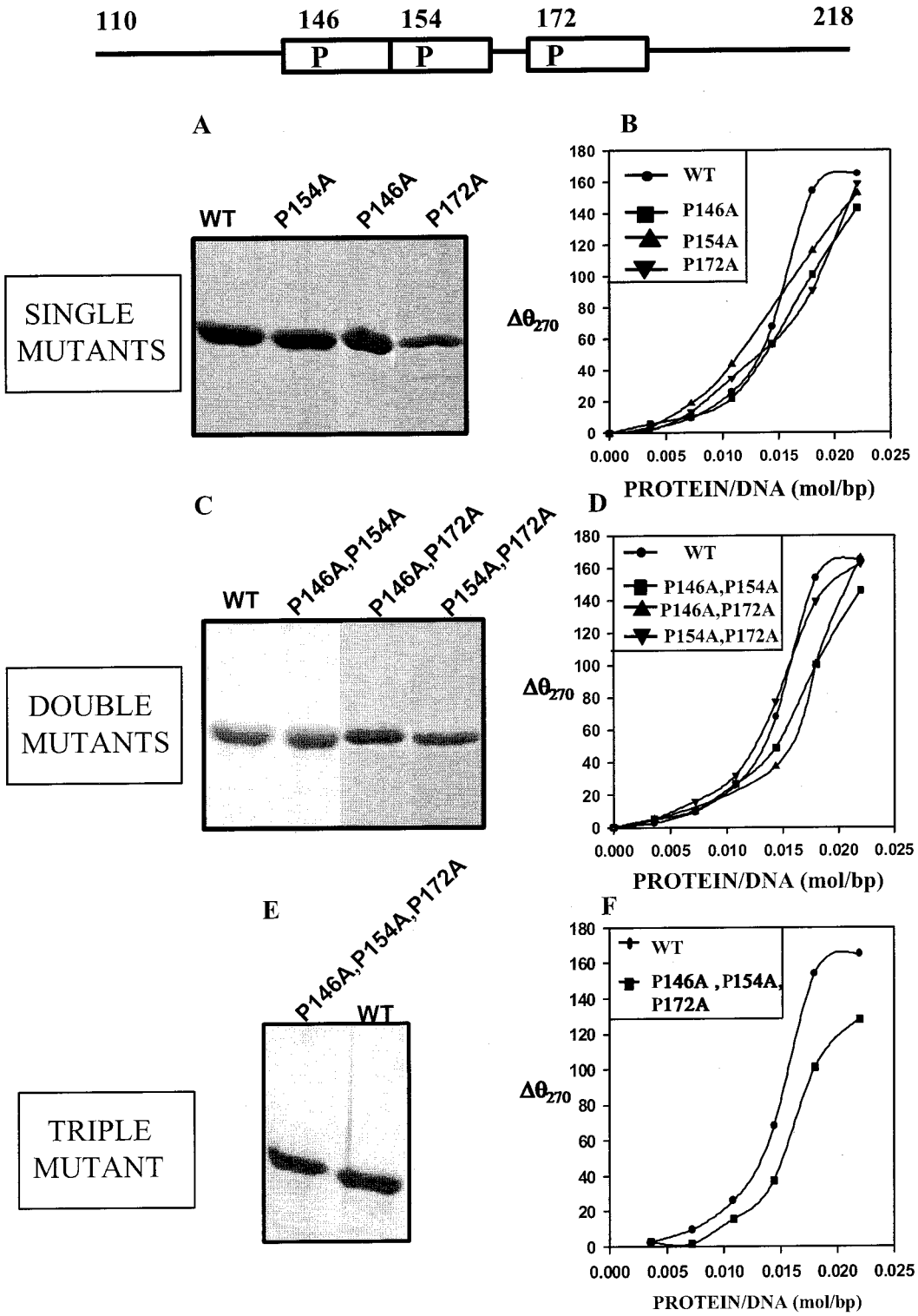


FIGURE 3: Effect of proline to alanine mutation of the S/TPKK motif in the C-terminus of histone H1d on DNA condensation. The diagram on top shows the positions of the proline residues within the octapeptide repeats that are mutated to alanine residues. The wild-type and mutant proteins are expressed in *E. coli* and purified as described in Experimental Procedures. The purity of the proteins as shown in panel A (single mutants), panel C (double mutants), and panel E (triple mutant). The net DNA condensation observed, $\Delta\theta$ at 270 nm, with oligonucleosomal DNA as a function of increasing protein/DNA ratio of the input proteins (wild-type and mutants) is shown in panel B (single mutants), panel D (double mutants), and panel F (triple mutant).

178 H1d) resulted in a decrease in the percentage helicity from 26.5% to 10.4%. The experiments described above showed that deletion of octapeptide repeats alone within the 34 amino acid stretch did result in only a 40% decrease in the DNA-condensation property of histone H1d. However, the deletion of the entire stretch which includes a 10 amino acid spacer region between the second and the third octapeptide repeats resulted in almost 90% decrease in the DNA-condensation property. We have also observed that deletion of this 10 amino acid stretch spacer (Δ 10, 161−170 H1d) resulted in a decrease in the percentage helicity to 15% (Table 1), indicating a potential helical structure in this

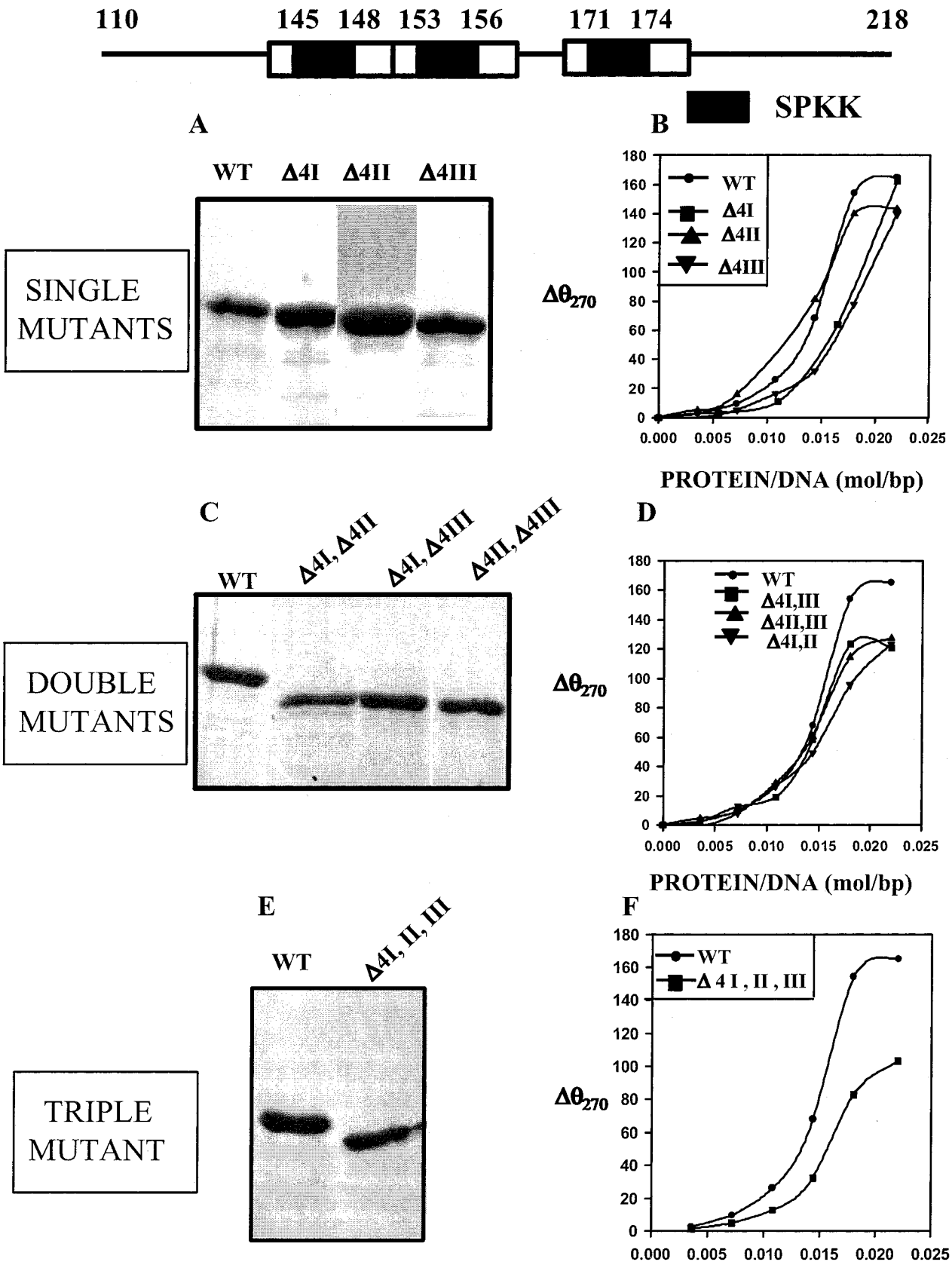


FIGURE 4: Effect of S/TPKK deletion mutation in the C-terminus of histone H1d on DNA condensation. The diagram on top shows the position of the tetrapeptide motifs present within the octapeptide repeats that are deleted as described in Figure 2B. The wild-type and deletion mutant proteins are expressed in *E. coli* and purified as described in Experimental Procedures. The SDS−PAGE patterns of the wild type and various mutants are shown in panel A for single mutants, panel C for double-deletion mutants, and panel E for triple mutants. The net DNA condensation observed, $\Delta\theta$ at 270 nm, with oligonucleosomal DNA as a function of increasing protein/DNA ratio (mol/bp) of the wild type and mutants are shown in panel B for single mutants, panel D for double mutants, and panel F for triple mutants.

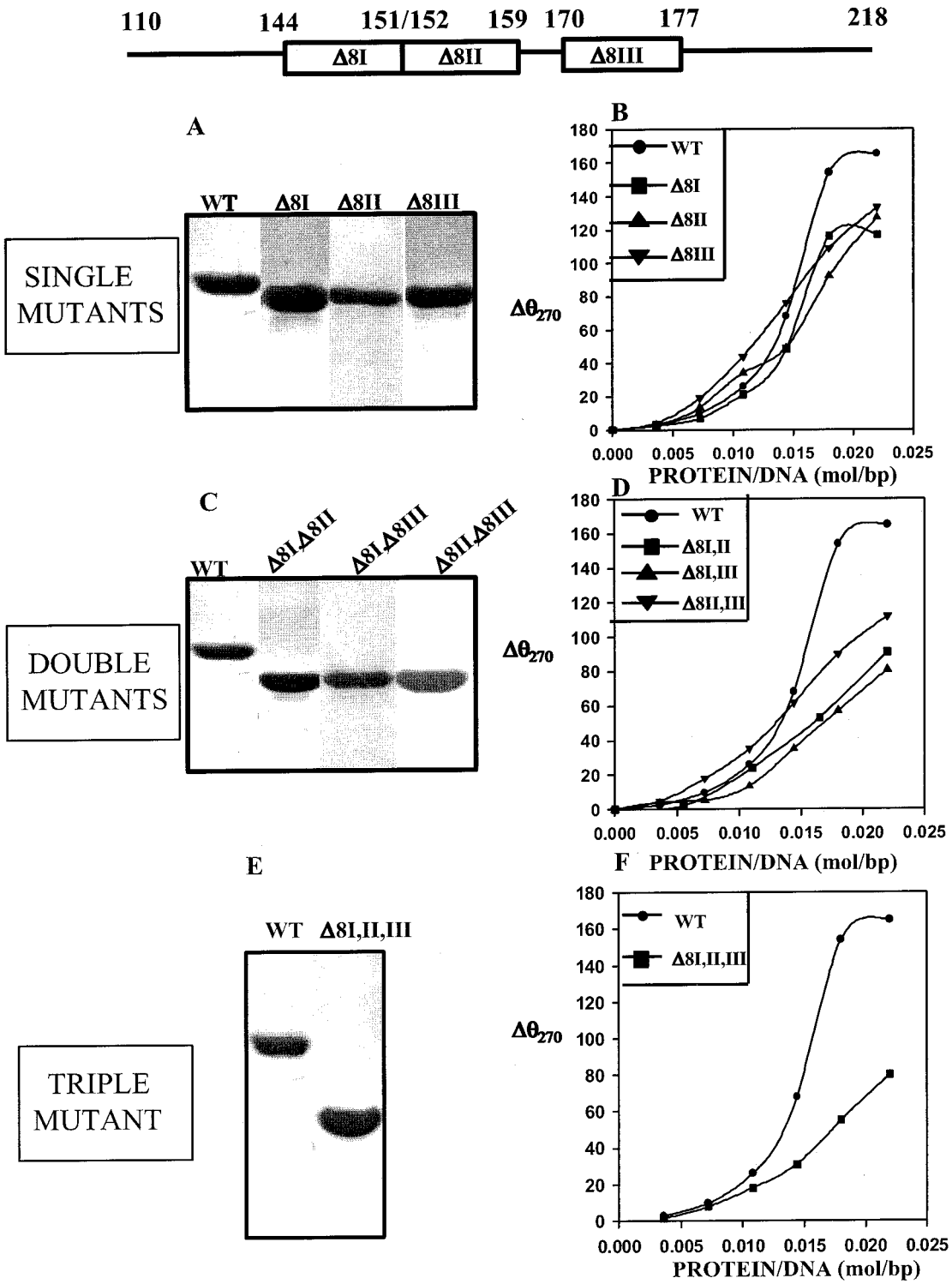


FIGURE 5: Effect of octapeptide deletion mutations in the C-terminus of histone H1d on DNA condensation. The diagram on top shows the positions of the octapeptide repeats in the C-terminus of histone H1d. The SDS−PAGE patterns of the wild type and various deletion mutants are shown in panel A for single mutants, panel C for double-deletion mutants, and panel E for triple mutants. The net DNA condensation, $\Delta\theta$ observed at 270 nm is plotted as a function of increasing protein to DNA ratio, mol/bp of the wild type and deletion mutants. Panel B, single deletion mutants; panel D, double-deletion mutants; panel F, triple-deletion mutants.

spacer region that plays an important role in spatially positioning the DNA binding S/TPKK units for effecting DNA condensation. Interestingly, deletion of the octapeptide repeats (Δ 8 I,II−H1d and Δ 8I,II,III−H1d) resulted in an increase in the induced helicity from 26.5% to 48% (Table 1), suggesting that removal of the turn regions expands the segments beyond the spacer region to attain α-helical conformation. The two other deletion mutants, namely, Δ 27,115−141H1d, and Δ 34, 184−218 H1d, also showed a decrease in the induced α-helicity of 17.2 and 10.9, respectively, as opposed to 26.5% in wild-type histone H1d, although their DNA-condensation property was not significantly reduced as compared to that of Δ 34, 145−178H1d. The experimentally observed reduction in the induced

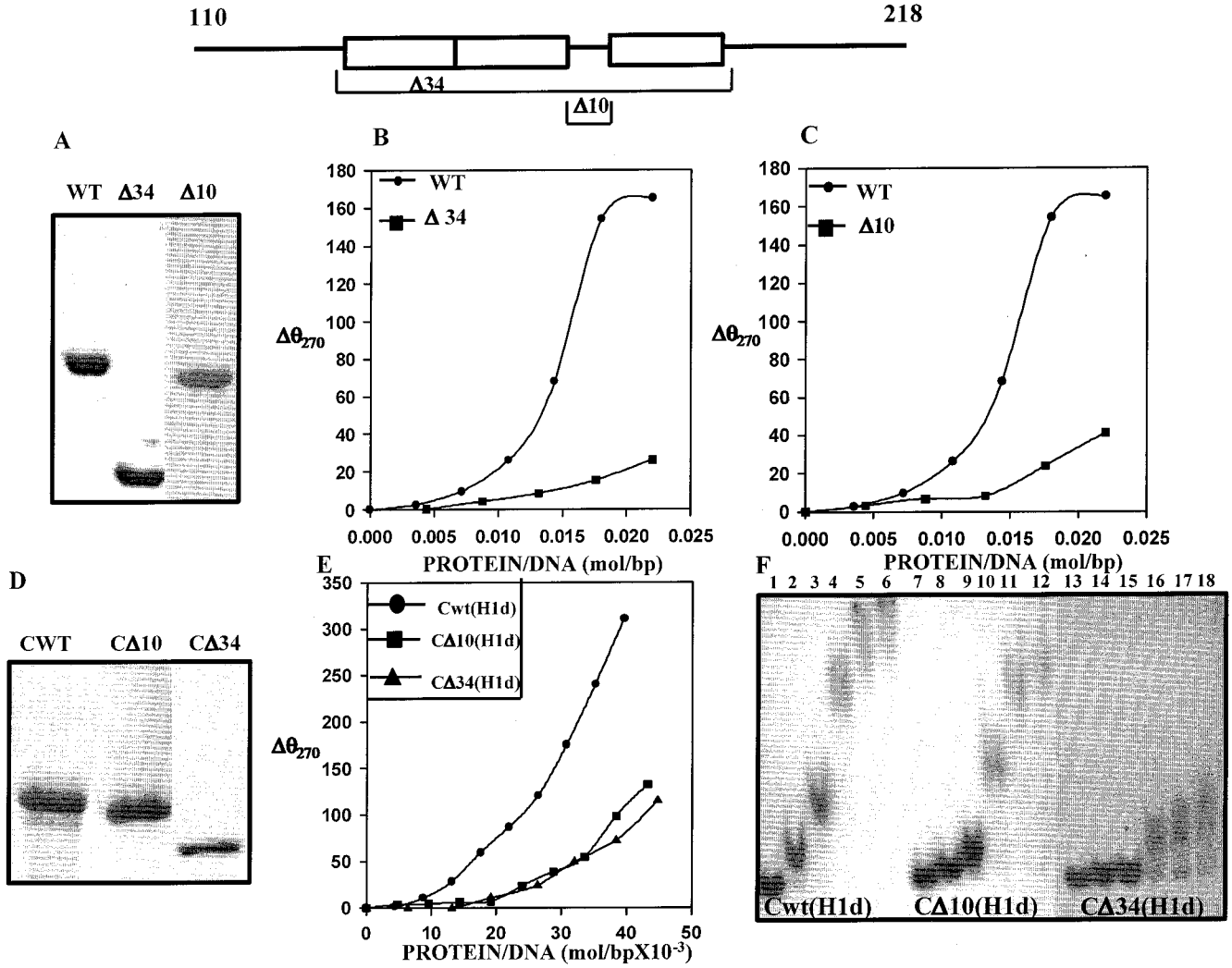


FIGURE 6: Effect of deletion mutants of histone H1d and its C-terminus on DNA condensation and DNA binding. The diagram on top shows the stretches of amino acids that are deleted in the Δ34 and Δ10 mutants in the full length and the C-terminal fragment of histone H1d. The C-terminus of histone H1d was PCR amplified, cloned, and expressed in pET 22b. The deletion mutants were generated as described in Figure 2. The wild-type and deletion mutant proteins were purified as described Experimental Procedures. The SDS−PAGE patterns of the wild type and deletion mutants are shown in panel A (full length) and panel D (C-terminal fragment). The net DNA condensation $\Delta\theta$ observed at 270 nm obtained with the Δ34 mutant is shown in panel B and that of the Δ10 mutant in panel C. The effects of the Δ34 and Δ10 mutants of the C-terminal fragments of histone H1d on DNA condensation is shown in panel E. Panel F shows the gel retardation of a 21-mer AT-rich double-stranded oligonucleotide with wild-type (lanes 1−6), Δ 34 mutant (lanes 7−12), and Δ10 mutant (lanes 13−18) fragments. Lanes 1, 7, and 13 are free probes without any protein. Lanes 2−6, 8−12, and 14−18 represent the DNA protein complexes obtained with 10.9, 21.9, 54.5, 81.75, and 109 pmoles of the wild-type, Δ10, and Δ34 C-terminal fragments, respectively.

α-helicity of various deletion mutants agrees with the theoretical predictions of potential α-helical segments within the C-terminus of histone H1d, as shown in Table 1.

DISCUSSION

In this communication, we have addressed the DNA-condensing domain of the C-terminus of histone H1. In earlier reports, we had identified an octapeptide repeat motif in the C-terminus which could mimic all the condensing properties of histone H1. We have extended these studies in this paper in the context of the full length histone H1 by a series of site-directed mutagenesis and have identified a 34 amino acid stretch encompassing the octapeptide repeat motif that is responsible for the DNA-condensation properties of histone H1. It is interesting to note that the results presented in this paper with various deletion mutants systematically narrowed down the DNA-condensing domain to this 34 amino acid stretch within the C-terminus encompassing the octapeptide repeats containing the S/TPKK motifs. It has been shown earlier that S/TPKK motifs bind to DNA in the minor groove and in turn destabilize the groove (*25*). One of the most interesting results presented in this paper is that deletion of the 10 amino acid spacer keeping the octapeptide repeats intact resulted in an almost 80% decrease in the DNA-condensation property of histone H1d. This observation strongly suggests that this spacer probably positions the S/TPKK units optimally for interacting with the DNA minor groove to bring about DNA bending/condensation. Importantly, this region is predicted to be helical, which is confirmed experimentally by our circular dichroism studies (Table 1). This region is also rich in lysine residues. It is also interesting to note that the spacing and the length of the intervening sequence between the repeats vary among the different histone H1 subtypes (Figure 1B). It is worth-

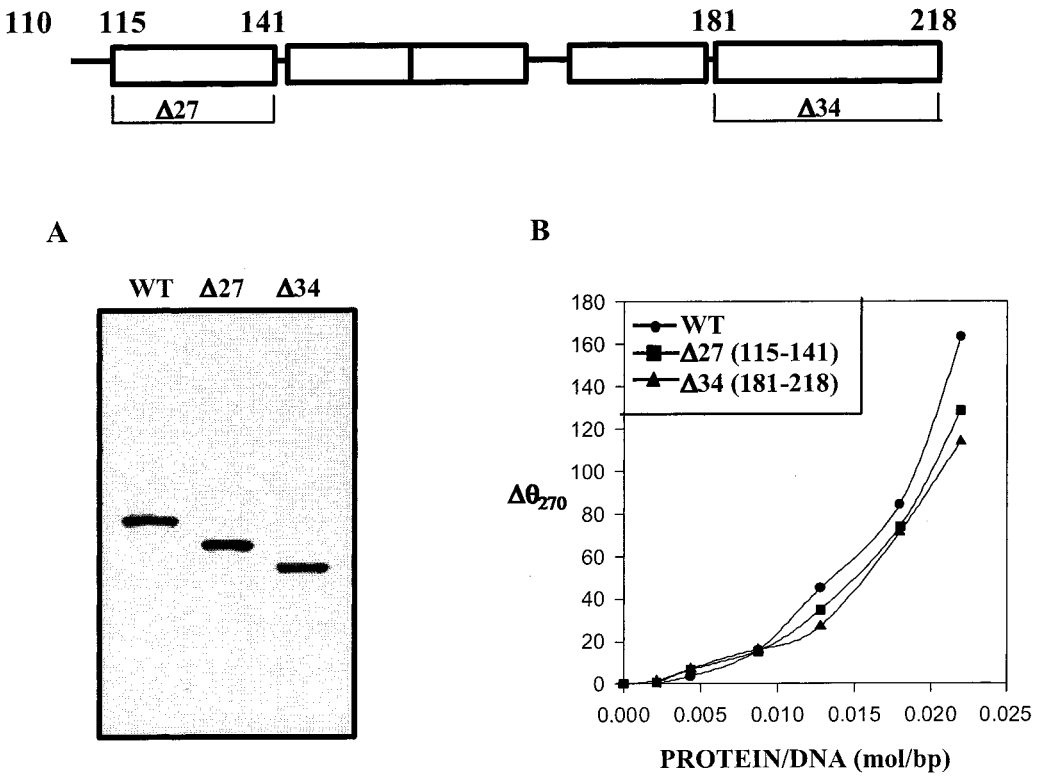
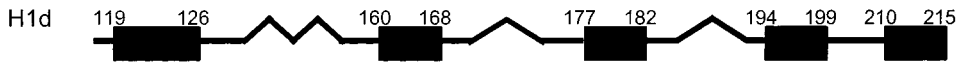


FIGURE 7: Effect of deletion mutants Δ27 (115−141) and Δ34 (184−218) of histone H1d on DNA condensation. The diagram on top shows the stretches of amino acids that are deleted in the Δ 27 (115−141) and Δ 34 (184−218) mutants. The SDS−PAGE patterns of the purified wild type and the two mutants are shown in panel A. The net DNA condensation, $\Delta\theta$ at 270 nm of the wild type and the two mutant histone H1d, is shown in panel B.

Table 1: Percentage α-Helicity Induced in Different Mutants of Histone H1d[a]

protein	θ_{220nm} (deg cm^2 dM^{-1})	ΔE	% helicity (solvent: 60% TFE)
H1d-WT	10,019	3.04	26.53
Δ8I,II-H1d	17,500	5.3	48.1
Δ8I,II,III-H1d	17,600	5.33	48.44
Δ34 (145-178)-H1d	3,800	1.15	10.4
Δ10 (161-170)-H1d	6,138	1.86	15.3
Δ27 (115-141)-H1d	6,798	2.06	17.2
Δ34 (184-218)-H1d	4,620	1.4	10.97
H1d-C(110-218)	3,142.5	0.95	6.67
H1d-C(110-218) Δ34(36-69)	1,700	0.52	2.6
H1d-C(110-218) Δ10(52-61)	1,997.2	0.61	3.4



[a] The calculations were done according to Clark et al. (*21*) using the formula $\theta = 3300\Delta E$ and percent helicity $= \Delta E - 0.25/0.105$. Those proteins indicated with C in the parentheses correspond to the mutants in the context of the C-terminus only. The rest correspond to the full length. The diagram below shows the potential secondary structure as predicted (*38*), wherein the boxes represents α-helical regions.

while pointing out here that the octapeptide repeats are fairly conserved in most of the histone H1 subtypes across species except a few, including the testis specific variant histone H1t. Our recent modeling studies also predict that the S/TPKK units of the 34 amino acid domain are optimally placed in space to facilitate interaction with DNA (*26*). We have also found in this modeling study that this 34 amino acid domain might adapt a HMG box fold similar to that found in mLEF1 and hSRY (*27*, *28*). These proteins bend the DNA very steeply upon binding to their cognate DNA sequences. At this juncture, it is tempting to speculate that the 34 amino acid domain also possibly bends the DNA in a manner analogous to the bending observed with mLEF1 and hSRY. This bending might be a major contributing factor to the circular dichroic spectral observations of a progressive decrease in the positive ellipticity at 270 nm upon interaction of DNA with histone H1. Analysis of histone H1-DNA complexes using scanning force electron microscopy (SFM) has shown that in these globular complexes, the path of DNA is not resolved, but it seems from the general appearance, the DNA is bent and has probably around a core of histone H1 molecule widening of the groove (*29*). A similar data has been obtained with the Torroids formed by the interaction of isolated C-terminal domains of histone H1 with DNA (*30*). Histone H1-induced DNA bending is physiologically relevant because the structure of the DNA obtained upon interaction with histone H1 is close to the curvature of DNA in the condensed fiber. Some experiments have also indicated that

the compaction of linker DNA in nucleosomal templates is accompanied by bending or kinking of linker DNA (*31*−*33*).

While these studies were in progress, two recent reports published by Suau and co-workers (*34*, *35*) have addressed the question of secondary structure of segments of the C-terminus of histone H1°. In the first study, they have carried out NMR analysis of a peptide representing the sequence stretch from residues 91 to 121. This study showed that, although in aqueous solution the peptide is unstructured, it attains substantial α-helical structure in 60% trifluroethanol. This α-helical structure had a strong amphipathic character, with all the positively charged residues concentrated on one face of the helix and all the hydrophobic residues on the opposite face. This stretch corresponds to residues 112−131 and 144−149 in the rat histone H1d. There is an insertion of 12 amino acid residues in histone H1d. We have also observed in the present study that deletion of this stretch in histone H1d resulted in a decrease in the α-helicity in histone H1d (Table 1). In a subsequent study, they have shown by Fourier-transformed infrared spectroscopy that DNA also induces α-helical structure in the absence of trifluroethanol. Thus, although originally believed to be predominantly random coil, the C-terminus of histone H1 has a sequence with a potential to attain regular secondary structure in the presence of DNA. In conjunction with the 34 amino acid domain possessing DNA-condensation property, the structure around this condensing domain may determine the curvature and path of the linker DNA between two adjacent nucleosomes. We have also recorded the circular dichroism spectra of histone H1d and its C-terminal mutants in the presence of DNA to observe DNA induced α-helicity. However, due to large changes in the DNA spectrum, we could not come to any meaningful conclusions regarding structural alterations in the protein. The structural aspects related to the linker DNA has been a matter of controversy since the concept of higher-order structure has emerged. It is still not clear what the path of linker DNA is. There have been suggestions that the path of the linker DNA influences the higher-order structure and, if solved, will throw light on the folding pathway of polynucleosome fiber. Hamiche et al. (*36*) have recently examined mononucleosomes reconstituted with the globular domain GH5 and also the full length of histone H5 containing the C-terminus by electron microscopy. The electron micrographs reveal that the globular domain of histone H5 increases the wrapping of DNA around the nucleosome core from 1.65−1.7 to 1.8−1.9 turns. At the same time, the entering and exiting DNA duplexes are uncrossed. In the presence of full-length histone H5, a stalk is formed with the C-terminal tail bridging the entering and exiting DNA together to form a four stranded stem. The identification of a DNA-condensing domain in the C-terminus of histone H1 should stimulate our future efforts to understand the role of the C-terminus of histone H1 in folding of chromatin into higher-order structures.

REFERENCES

1. Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997) *Nature 389*, 351−360.
2. Allan, J., Cowling, G. J., Harborne, N., Cattani, P., Craigie, R., and Gould, H. (1981) *J. Cell Biol. 90*, 279−288.
3. Clark, D. J., and Kimura, T. (1990) *J. Mol. Biol. 211*, 883−896.
4. Cerf, C., Lippens, G., Muyldermans, S., Segers, A., Ramakrishnan, V., Wodak, S., Halenga, K., and Wyns, L. (1993) *Biochemistry 32*, 11345−11351.
5. Ramakrishnan, V., Finch, J. T., Graziano, V., Lee, P. L., and Sweet, R. M. (1993) *Nature 362*, 219−223.
6. Goytisolo, F. A., Gerchman, S. E., Yu, X., Rees, C., Graziano, V., Ramakrishnan, V., and Thomas, J. O. (1996) *EMBO. J. 15*, 3421−3429.
7. Allan, J., Hartman, P. G., Crane-Robinson, C., and Aviles, F. X. (1980) *Nature 288*, 675−679.
8. An, W., Leuba, S. H., van Hodle, K. E., and Zlatanova, J. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 3396−3401.
9. Zhou, Y. B., Gerchman, S. E., Ramakrishnan, V., Travers, A., and Muyldermans, S. (1998) *Nature 395*, 402−405.
10. van Holde, K. E., and Zlatanova, J. (1995) *J. Biol. Chem. 270*, 8373−8376.
11. Shen, X., Yu, H., Weir, J. W., and Gorovsky, M. A. (1995) *Cell 82*, 47−56.
12. Fasman, G. D., Schaffhaussen, B., Goldsmith, L., and Adler, A. (1970) *Biochemistry 9*, 2814−2822.
13. Moran, F., Montero, F., Azorin, J., and Suan, P. (1985) *Biophys. Chem. 22*, 125−129.
14. Liao, L. W., and Cole, R. D. (1981) *J. Biol. Chem. 256*, 6751−6755.
15. DeLucia, F., Faraone-Mannella, D., D'Erme, M. R., Quesada, P., Caiafa, P., and Farina, B. (1994) *Biochem. Biophys. Res. Commun. 198*, 32−39.
16. Khadake, J. R., and Rao, M. R. S. (1995) *Biochemistry 34*, 15792−15801.
17. Suzuki, M. (1989) *EMBO. J. 8*, 797−804.
18. Khadake, J. R., and Rao, M. R. S. (1997) *Biochemistry 36*, 1041−1051.
19. Srinivas Bharath, M. M., Khadake, J. R., and Rao, M. R. S. (1998) *Protein Exp. Purif. 12*, 38−44.
20. Meetei, A. R., and Rao, M. R. S. (1998) *Anal. Biochem. 264*, 288−291.
21. Clark, D. J., Hill, C. S., Martin, S. R., and Thomas, J. O. (1988) *EMBO J. 7*, 69−75.
22. Payet, D., Hillisch, A., Lowe, N. Diekmann, S., and Travers, A. (1999) *J. Mol. Biol. 294*, 79−91.
23. Platz, R. D., Meistrich, M. L., and Grimes, S. R., Jr. (1977) *Methods Cell Biol. 16*, 297−316.
24. Laemmli, U. K. (1970) *Nature 277*, 680−685.
25. Suzuki, M., Gerstein, M., and Johnson, T. (1993) *Protein Eng. 6*, 565−574.
26. Srinivas Bharath, M. M., Chandra, N., and Rao, M. R. S. *Proteins*, in press.
27. Werner, M. H., Herth, R., Gronenhorn, A. M., and Clore, G. M. (1995) *Cell 81*, 705−714.
28. Murphy, F. V., Sweet, R. M., and Churchill, M. E. A. (1999) *EMBO J. 18*, 6610−6618.
29. Leuba, S. H., Yang, G., Robert, C., Somori, B., van Holde, K., Zlatanova, J., and Bustamante, C. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 11621−11625.
30. Rodriguez, A. T., Perez, L., Moran, F., Montero, F., and Suau, P. (1991) *Biophys. Chem. 39*, 145−152.
31. Butler, P. J., and Thomas, J. D. (1998) *J. Mol. Biol. 281*, 401−407.
32. Van Holde, K., and Zlatanova, J. (1996a) *Proc. Natl. Acad. Sci. U.S.A. 93*, 10548−10555.
33. Yao, J., Lowary, P. T., and Widom, J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 7603−7607.
34. Vila, R., Ponte, I., Jimenez, M.a., Rico, M., and Suau, P. (2000) *Protein Sci. 9*, 627−636.
35. Vila, R., Ponte, I., Collado, M., Arrondo, J. R., and Suau, P. (2001) *J. Biol. Chem. 276*, 30898−30903.
36. Hamiche, A., Schultz, P., Ramakrishnan, V., Oudet, P., and Prunell, A. (1996) *J. Mol. Biol. 257*, 30−42.
37. Thomspon, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucl. Acids Res. 22*, 4673−4680.
38. Combet, C., Blanchet, C., Geourjon, C., and Deléage, G. (2000) *TIBS 25*, 147−150.

BI025773+