Attenuation of DNA–Protein Interactions Associated with Intrinsic, Sequence-Dependent DNA Curvature†

Michal Shatzky-Schwartz, Yaffa Hiller, Ziv Reich, Rodolfo Ghirlando, Sarah Weinberger, and Abraham Minsky*
Department of Organic Chemistry, The Weizmann Institute of Science, Rehovot 76100, Israel
Received May 13, 1991; Revised Manuscript Received November 15, 1991

ABSTRACT: Inherently curved DNA segments, associated with short runs of adenines, have been identified in many gene regulatory regions, yet their physiological significance remains unknown. The observations reported in this study indicate that intrinsically bent nucleic acid fragments are characterized by substantially attenuated affinities toward DNA-binding proteins involved in structural functions, such as H1 histone and protamine, as well as toward various DNA-modifying enzymes including ligases and exo- and endonucleases. Two mechanisms might be responsible for the altered binding properties. According to the first mechanism, the attenuated binding affinities and the bending represent two independent consequences of the unique structural parameters exhibited by A-tracts. Indeed, analysis of the degradation products obtained upon exposure of the curved sequences to various chemical nucleases points toward the narrowing of the DNA minor groove, a conformational modulation known to characterize A-tracts and to run along the axially-bent motifs, as a potential determinant of the observed binding attenuation. Alternatively, the conformational constraints which result from the stable bending might act to modulate the strength of DNA–protein interactions. Although the factor directly responsible for the altered binding affinities revealed by the bent sequences cannot as yet be conclusively resolved, it is proposed that a reiteration of this specific factor, being either an A-tract or a bend, in phase with the DNA helical repeat acts to amplify the modulation of the binding. This suggestion is based on the findings obtained from ligated DNA fragments which are composed of hexadecamers as the repetitive unit, and in which the A-tracts and the bends are reiterated out-of-phase. The binding attenuation revealed by these DNA segments is found to be considerably lower than that exhibited by the globally curved sequences. On the basis of these observations, we suggest that sequence-dependent DNA bending associated with short, phased runs of adenine might be involved in gene regulation processes.

Sequence-derived nucleic acid curvature refers to the ability of specific nucleotide arrangements to confer a stable bending upon DNA segments. The largest degree of such an intrinsic bending is produced by short, phased runs of homopolymeric dA·dT base pairs (A-tracts) (Hagerman, 1985, 1986; Koo et al., 1986; Koo & Crothers, 1988; Crothers et al., 1990). Several models have been formulated in order to explain the phenomenon of intrinsic DNA bending and the conformational elements which determine its extent (Hagerman, 1988). According to the wedge model, a smooth deformation of the main DNA axis is induced by a combination of roll and tilt components between adjacent AT base pairs (Trifonov & Sussman, 1980; Ulanovsky et al., 1986; Ulanovsky & Trifonov, 1987; Trifonov & Ulanovsky, 1988), whereas the junction model points toward a rather abrupt change in the direction of the helical axis which is generated at the junction site between the non-B-DNA conformation, associated with the A-tract, and the regular B-DNA structure (Wu & Crothers, 1984; Koo et al., 1986; Levene et al., 1986; Koo & Crothers, 1987; Chuprina & Abagyan, 1988; Crothers et al., 1990). When such deformation is reiterated in-phase with the helical repeat, a long-range curvature is obtained.

Following the initial discovery of curved DNA structures in kinetoplast DNA minicircles isolated from the tropical parasite *Leishmania tarentolae* (Marini et al., 1982), such bent motifs were identified in kinetoplasts from most trypanosomatid species (Ntambi et al., 1984; Kitchin et al., 1986). Significantly, kinetoplast minicircles have been recently shown to encode guide RNA (gRNA) molecules which are involved in editing processes of various mRNA species (Strum & Simpson, 1990). Sequence-derived bends have been found in λ phage origin of replication (Zahn & Blattner, 1985) and shown to be required for DNA complexation with the initiator protein O (Zahn & Blattner, 1987). Curved motifs have been identified in SV40 replication origin (Ryder et al., 1986) and terminus (Hsieh & Griffith, 1988), in the region preceding the tRNA operon promoter of *Salmonella* (Bossi & Smith, 1984), and also in a yeast autonomously replicating sequence (ARS), a putative replication origin (Snyder et al., 1986).

The apparent ubiquity of inherently curved DNA segments in living systems, combined with the observation that such motifs are localized preferentially in gene regulatory regions, raises the intriguing possibility that bent conformations are functionally significant. On the basis of the finding that the specific location and direction of a sequence-derived curvature revealed by a particular DNA fragment are not altered upon reconstitution of that fragment into nucleosome particles, the involvement of an intrinsic bending in nucleosomal phasing has been proposed (Drew & Travers, 1985; Travers, 1987; Hayes et al., 1990). Long-range DNA curvature has also been suggested to play a role in nucleic acid packaging processes occurring in viruses (Lilley, 1986), as well as in eukaryotic heterochromatin, in which bent motifs have recently been identified (Radic et al., 1987; Balbas et al., 1990). A conclusive determination of the potential biological functions of the bent conformation must, however, await a thorough assessment of the chemical and physical features of the DNA molecules which are induced or modulated by this structural

† This work was supported by the Fund for Basic Research administered by the Israel Academy of Sciences and Humanities. Part of this study was presented at the Seventh Conversation in Biomolecular Stereodynamics, Albany, NY, 1991.
* To whom correspondence should be addressed. Incumbent of the Victor L. Erlich Career Development Chair.

0006-2960/92/0431-2339$03.00/0

motif. The most notable physical property known to date to specifically characterize the axially-curved segments is their reduced electrophoretic mobility on acrylamide gels, with a degree of retardation which correlates with the extent of the macroscopic bending (Hagerman, 1985, 1986, 1988; Koo et al., 1986; Ulanovsky et al., 1986). No obvious effects of the intrinsic curvature upon the chemical or biochemical features of the nucleic acids have been hitherto identified.

The observations presented in this study indicate that intrinsically bent structures are characterized by a distinct biochemical property, namely, a lower affinity for complexation with DNA-binding proteins, relative to the affinities exhibited by noncurved DNA sequences. The attenuated binding affinity associated with the curvature is found to be general, being observed for structural proteins such as H1 histone and protamine, as well as for ligases and exo- and endonucleases. The extent of the phenomenon is shown to correlate with the degree of the curvature. On the basis of these results, it is suggested that those structural features which are directly responsible for the induction of an intrinsic DNA curvature as well as the determination of its extent, namely, phased adenine tracts, might be involved in regulatory processes through their modulating effects upon DNA–protein interactions.

Experimental Procedures

Oligonucleotide Synthesis, Purification, and Ligation. The oligonucleotides specified under Results were synthesized on an automated DNA synthesizer (Applied Biosystems). Following deprotection, the oligomers were purified by gel filtration on a Sephadex G25 (Pharmacia LKB Biotechnology, Inc.) column equilibrated with 20 mM NH_4HCO_3 (pH 8.0). DNA concentrations were determined by measuring the absorbance at 260 nm, and applying the general relationship 1.0 OD = 33 μg/mL oligonucleotide. Single-strand oligomers were 5′-phosphorylated with T4 polynucleotide kinase (New England Biolabs) in a reaction mixture containing 250 μg of DNA, 80 mM Tris (pH 7.8), 15 mM $MgCl_2$, 5 mM DTT, 2 mM spermidine, and 5 mM ATP, at 19 °C for 24 h. The phosphorylated strands were annealed by heating the DNA in 50 mM Tris (pH 7.8) solution to 70 °C followed by a slow cooling to the temperature at which the ligation was conducted (specified in the legends for Figures 1 and 3). The double-strand segments were ligated with 600 units of T4 DNA ligase (New England Biolabs) in a reaction mixture containing 50 mM Tris (pH 7.8), 10 mM $MgCl_2$, and 2 mM ATP. Reaction times are specified in the figure legends.

DNA-Protein Interactions. DNA pellets, obtained by ethanol precipitation of the ligated forms, were resuspended in Tris–borate (TBE) buffer, 30 mM (pH 8.0), containing 1 mM EDTA. Histone H1 from calf thymus, protamine from salmon sperm, and bovine serum albumin (Sigma) were dissolved in the same buffer. Proteins were added to the DNA (3 μg per sample) solutions at the ratios specified in the legend of Figure 2, and the reaction volume was adjusted to 20 μL with the buffer. The solutions were gently mixed, kept for 12 h at 4 °C, and loaded on an agarose gel (1.2%, w/v). Agarose gels were run at 4 °C, with Tris–borate (30 mM) as electrophoresis buffer, at 5 V/cm, and were subsequently stained with ethidium bromide.

Enzymatic Digestion. DNA pellets were resuspended in 10 mM Tris·HCl buffer (pH 8.0) and subjected to exo- and endonucleases. BAL 31, *Taq*I, and exonuclease III were products of New England Biolabs; DNase I from bovine pancreas was purchased from Boehringer Mannheim. Specific conditions for the various digestion processes are provided in the legend of Figure 3. The digestion reactions were stopped with 50 μL of phenol, extracted with chloroform, and ethanol-precipitated.

$(OP)_2Cu^+$ Degradation. DNA (4 μg per sample) was resuspended in 20 mM Tris·HCl (pH 7.0) and treated with 100 μM 1,10-phenanthroline, 10 μM cupric sulfate, 35 mM 3-mercaptopropionic acid, and 35 mM H_2O_2 (all reagents were purchased from Aldrich). Reactions were conducted at 19 °C for the time intervals specified in the legend of Figure 4, and terminated with 35 μL of 0.1 M thiourea. The solutions were extracted with chloroform and ethanol-precipitated.

$Fe(EDTA)^{2-}$ Degradation. DNA (4 μg per sample) was dissolved in 20 mM Tris·HCl (pH 7.5) and treated with 0.4 mM $(NH_4)_2Fe(SO_4)_2 \cdot 6H_2O$, 1 mM EDTA, 0.2 M ascorbic acid, and 35 mM H_2O_2. Reactions were conducted at 19 °C for the time intervals specified in the legend of Figure 5 and terminated with 35 μL of 0.1 M thiourea. The solutions were extracted with chloroform and ethanol-precipitated.

Gels were 10% polyacrylamide (29:1 monomer/bis) and were run at 4 °C with TBE, 100 mM (pH 8.0) as running buffer. DNA molecular weight markers derived from ϕX174 RF DNA by *Hae*III digestion were purchased from New England Biolabs.

Results

The specific properties associated with intrinsic, sequence-derived DNA curvature were investigated by examining the interactions of DNA-binding proteins or chemical agents with tandemly reiterated sequences as well as by studying the effects of various DNA-modifying enzymes upon these segments. The following sequences, obtained by extensive ligation of the partially self-complementary oligomers ($N = 1$), were used:

(a) 5′-d$(CGAAAATTTT)_N$-3′

(b) 5′-d$(CGAATTAATT)_N$-3′

(c) 5′-d$(CGTTTTAAAA)_N$-3′

Notably, the repetitive units of (a–c), being decamers, are reiterated in close coherence with the DNA helical screw, thus satisfying the basic requirement for macroscopic curvature. Yet, only (a) contains adenine tracts of sufficient length and displays the appropriate polarity needed to induce such curvature (Hagerman, 1986, 1988; Koo et al., 1986; Ulanovsky et al., 1986; Koo & Crothers, 1987; Ulanovsky & Trifonov, 1987; Haran & Crothers, 1989). Indeed, only the polymers derived from (a) exhibit significant mobility anomalies on acrylamide gels, indicating that these DNA species are characterized by a stable, macroscopic curvature which originates from the in-phase reiteration of a local, sequence-dependent bending.

The ligation kinetics revealed by the decameric DNA sequences ($N = 1$) upon exposure to T4 DNA ligase were studied at two temperatures. At high temperature (30 °C), all three segments undergo ligation processes at a similar rate (Figure 1A). In clear contrast, when conducted at 4 °C, the rate of ligation of the bent sequence (a) is found to be substantially lower than the rates exhibited by the noncurved segments (b) and (c). After the shortest time interval examined at low temperature, the curved decamer undergoes less than 4-fold ligation, resulting in fragments that are too short to be detected in the 10% acrylamide gel, whereas the two other segments exhibit a significantly higher extent of polymerization. A clear difference in the length of the tandemly-repeated DNA species can still be observed after 12 h of ligation (Figure 1B). The reduced electrophoretic mobility of the bent segments (a) should be considered in order to fully

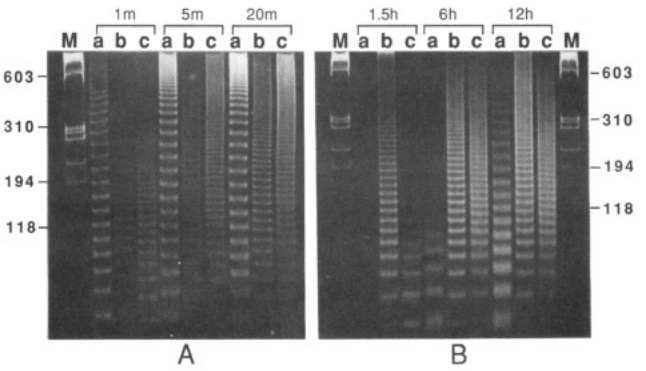

FIGURE 1: Temperature effects upon the rate of ligation of the following decamers: (a) d(CGAAAATTTT); (b) d(CGAATTAATT); (c) d(CGTTTTAAAA). Lanes a–c contain the ligation products of the corresponding decamers, obtained at (A) 30 and (B) 4 °C. Lane M contains markers derived from ϕX174 RF DNA by *Hae*III digestion. Polymers were constructed from the synthetic 5′-phosphorylated decamers by ligation with T4 DNA ligase. Ligation times were as specified in the figure.

appreciate the difference in the ligation efficiency between the curved and the nonbent sequences.

The interactions between the various ligated fragments and the DNA-binding proteins H1 histone and protamine, involved in the high-order chromatin organization processes in eukaryotic somatic nuclei and sperm cells, respectively, were investigated. The two proteins were chosen for this study as both the lysine-rich H1 and the arginine-rich protamines are characterized by high DNA-binding affinities and neither requires specific nucleic acid recognition sequences. The DNA–protein complexes, prepared at increasing ratios of protein to DNA, were analyzed by means of a gel retardation assay on agarose gels. This technique allows the determination of the relative binding affinities of a given protein toward different nucleic acid segments (Fried & Crothers, 1981), and its applicability for detailed studies of H1–DNA complexes has been demonstrated (Bernardin et al., 1986). The retardation assays were conducted on agarose gels in order to eliminate the migration anomalies associated with the intrinsically curved DNA sequences that are observed on acrylamide gels; the various fragments of the ligated DNA species are not resolved on agarose gels and appear, consequently, as a smear.

Upon increasing amounts of H1 added to the solutions containing a given amount of the DNA segments (a), (b), or (c), the electrophoretic mobility of the polymers of all three forms gradually decreases. However, the retardation effect exhibited by the nucleic acid species in the presence of the protein is found to be significantly larger for the ligated sequences (b) and (c) than that revealed by (a). Furthermore, relatively high H1 to DNA ratios result in the formation of nucleoprotein aggregates of (b) and (c)—but not of (a)—as indicated by the reduced amounts of DNA in the corresponding lanes and the concomitant appearance of fluorescent material not entering the gel (Figure 2A) (Bernardin et al., 1986). Similar results are obtained upon substituting protamine for the H1 histone (Figure 2B). In clear contrast, no mobility shifts of the various ligated species or the formation of nucleoprotein aggregates can be detected in the presence of BSA, a non-DNA-binding protein (Figure 2C). This observation indicates that the modulated electrophoretic migration is indeed associated with the formation of DNA–protein complexes and, in particular, that the different extent of such modulation is directly related to the binding affinities. The different extent of retardation exhibited by the three repetitive nucleic acid species upon interaction with H1 and protamine, as well as their different tendency to aggregate following such interactions, points to a dissimilarity in their affinities toward DNA-binding proteins; specifically, the polymers of the (a) form are found to exhibit a lower binding affinity than those characterizing the (c) and particularly the (b) repetitive sequences.

The effects of two exonucleases (BAL 31 and exonuclease III) as well as two endonucleases (DNase I and the restriction enzyme *Taq*I) on the ligated DNA fragments (a), (b), and (c) were investigated. We find that those conditions which induce exhaustive digestion of the noncurved sequences by BAL 31, exonuclease III, and DNase I fail to significantly affect the inherently bent fragments derived from (a) (Figure 3A–C). Upon ligation of the decameric sequences, a restriction site for *Taq*I (5′-TCGA-3′) is created in the polymers obtained from (a) and (b). Digestion of these two sets of fragments with *Taq*I is found to be substantially more effective for the noncurved segments (b) than for the sequences which exhibit a stable, macroscopic bending (Figure 3D). In order to prevent denaturation of the relatively short AT-rich fragments, the temperature at which the *Taq*I reaction was carried out was kept well below that required for optimal digestion with this enzyme; the incomplete degradation of the noncurved segments may result from such suboptimal conditions. It should also be noted that the digestion products resulting from

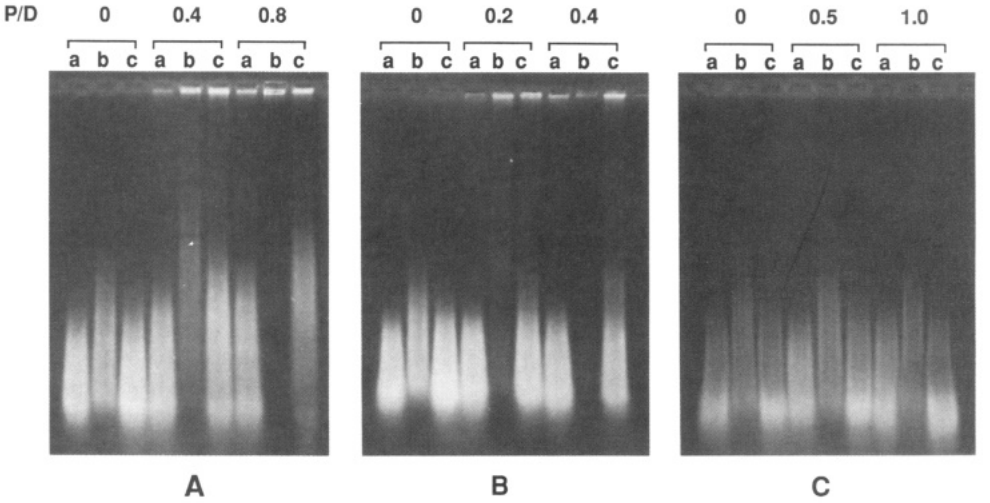

FIGURE 2: DNA–protein relative binding affinities analyzed by gel-shift assays. The polymerized DNA fragments (a), (b), and (c), specified in the legend for Figure 1, were mixed with the protein at the indicated P/D (protein to DNA) ratios (w/w), incubated at 4 °C for 12 h, and loaded on a 1.2% agarose gel. Gels were run at 4 °C. (A) H1 histone; (B) protamine; (C) bovine serum albumin (BSA).

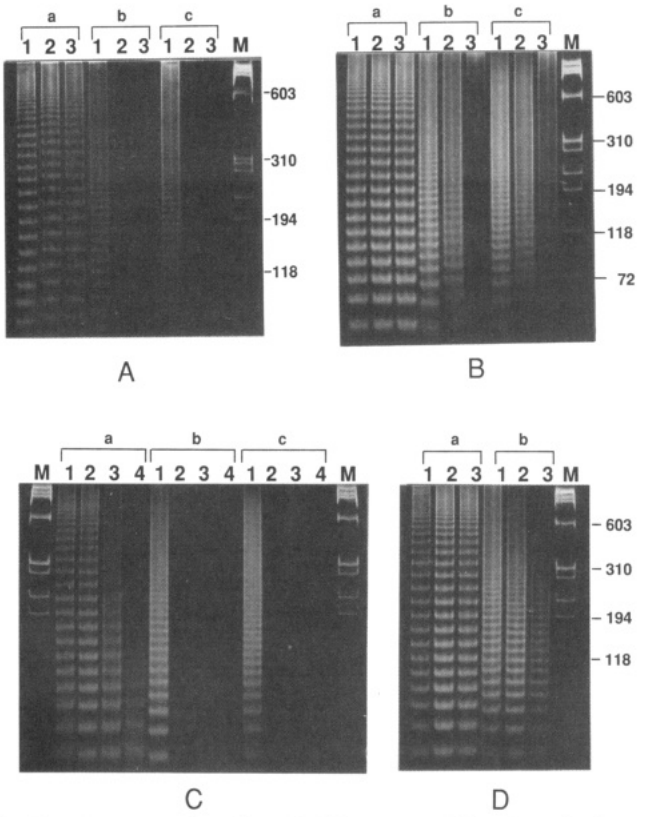

FIGURE 3: Effects of various exo- and endonucleases on the polymerized decamers. The reiterative fragments were obtained by ligation of the decamers (a), (b), and (c), which was conducted at 30 °C. (A) BAL 31, 5 units in 40 mM Tris (pH 8.0), 600 mM NaCl, 12 mM $MgCl_2$, and 12 mM $CaCl_2$. Reaction times were (lane 1) 0, (2) 8, and (3) 20 min. (B) Exonuclease III, 50 units in 25 mM Tris (pH 8.0) and 4 mM $MgCl_2$. Reaction times were (lane 1) 0, (2), and (3) 60 min. (C) DNase I, 1 unit in 50 mM Tris (pH 7.8), 0.8 mM $MnCl_2$, and 0.2 mM DTT. Reaction times were (lane 1) 0, (2) 2, (3) 4, and (4) 10 min. (D) *Taq*I, 50 units in 100 mM Tris (pH 8.4), 100 mM NaCl, and 10 mM $MgCl_2$. Reaction times were (lane 1) 0, (2) 2, and (3) 12 h. All reactions were carried out at 19 °C, except for the *Taq*I digestion that was conducted at 30 °C.

the exposure of the repetitive DNA segments to *Taq*I are multiples of 10 base pairs, in contrast with the smear which results from the activities of the other nucleases.

The properties of the various tandemly-reiterated DNA sequences were further studied by means of two chemical nucleases, namely, bis(1,10-phenanthroline)–copper [$(OP)_2Cu^+$] and iron^{II}–EDTA [$Fe(EDTA)^{2-}$]. These agents have been shown to induce nucleic acid cleavage through oxidative pathways and were extensively used for detailed structural studies as well as for footprinting experiments (Tullius et al., 1987). Exposure of the ligated DNA molecules to the nuclease activity of $(OP)_2Cu^+$ results in the digestion of the repetitive species. However, a clear difference in the rates at which the various DNA sequences undergo cleavage processes is observed. Whereas short exposure times of the (a) segments to the nuclease activity fail to affect these species, such reaction periods result in a significant degradation of the fragments (b) and (c), as indicated by the smear and reduced fluorescent material in the corresponding lanes (Figure 4). Following longer digestion periods, only limited degradation of the polymers of the (a) form can be detected, whereas the noncurved ligated species are cleaved into fragments that are too short to be observed in the gels.

Analysis of the degradation products obtained from the three DNA species (a), (b), and (c) upon exposure to $Fe(EDTA)^{2-}$ activity points toward a clear dissimilarity in the cleavage patterns exhibited by this reagent and those revealed by the $(OP)_2Cu^+$. Specifically, the enhanced relative stability that characterizes the polymers of the (a) form toward the DNase activity of $(OP)_2Cu^+$ is no longer observed when these segments are exposed to $Fe(EDTA)^{2-}$. Although (a) is found to

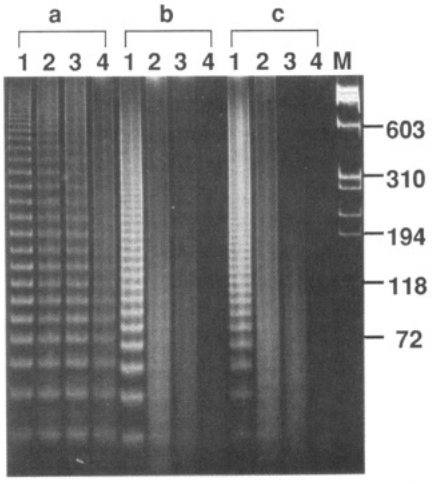

FIGURE 4: Effects of the chemical nuclease $(OP)_2Cu^+$ on the polymerized decamers (a), (b), and (c). The degradation conditions are detailed under Experimental Procedures. Reaction times were (lane 1) 0, (2), (3) 4, and (4) 10 min.

exhibit a slightly higher resistance to degradation by $Fe(EDTA)^{2-}$ than that characterizing the two other polymeric species (Figure 5; compare, in particular, lanes 3 of panels a–c), the difference in the degradation patterns is marginal.

The unique properties exhibited by curved nucleic acid sequences are further underlined by the effects exerted by distamycin, a DNA minor groove-binder drug, upon these properties. The presence of distamycin in increasing drug to DNA molar ratios is found to progressively reduce the acquired resistance of the curved segments (a) against digestion by both

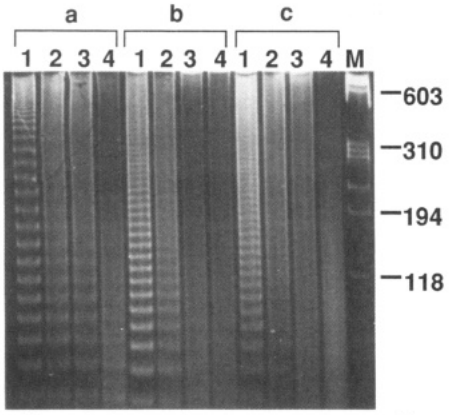

FIGURE 5: Effects of the chemical nuclease $Fe(EDTA)^{2-}$ on the ligated decamers (a), (b), and (c). Reaction times were (lane 1) 0, (2) 3, (3) 30, and (4) 90 min.

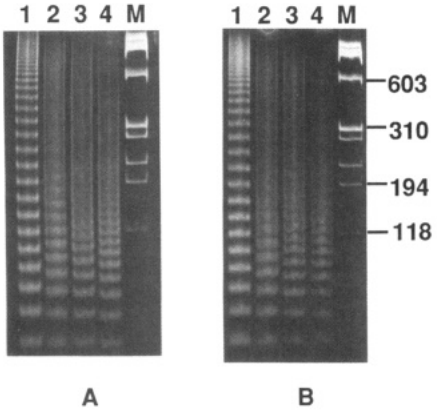

FIGURE 6: Effects of the minor-groove binder distamycin upon digestion of the ligated segments of the (a) form. The DNA species were incubated with distamycin for 6 h at 19 °C and subsequently treated with (A) exonuclease III (50 units in 25 mM Tris, pH 8.0, and 4 mM $MgCl_2$) for 1 h at 19 °C or (B) $(OP)_2Cu^+$ (conditions specified under Experimental Procedures) for 4 min at 19 °C. In both cases, the drug to base pair molar ratios were (lane 1) 0, (2) 0.2, (3) 1.0, and (4) 5.0.

exonuclease III and $(OP)_2Cu^+$ (Figure 6A,B, respectively), as well as by the exonuclease DNase I (results not shown). These observations can be interpreted in terms of the reduced degree of the sequence-derived curvature that is affected upon treatment of such a structural motif with distamycin. The curvature-neutralizing phenomenon has been detected in kinetoplast DNA fragments when treated with the drug (Wu & Crothers, 1984; Griffith et al., 1986) and is clearly exhibited by the reiterating synthetic fragments as well (Figure 6A,B, lanes 2–4; note the progressive elimination of the abnormal electrophoretic mobility observed as the drug to DNA ratios are increased). It should be noted that the digestion of the (a) fragments induced by both the enzyme and the artificial nuclease in the presence of the drug is, although significant, incomplete. This finding is due, presumably, to the protective effect exerted by the drug upon the nucleic acids following its binding to the DNA minor groove (Portugal & Waring, 1987). Indeed, it can be observed that at the highest drug to DNA ratio, the DNA molecules are more efficiently protected against digestion by both exonuclease III and $(OP)_2Cu^+$ than at the lower ratios.

Additional data concerning the properties exhibited by curved DNA sequences were obtained by a study of the effects exerted by various nucleases upon the tandemly-reiterated DNA molecules of the form:

(d) 5′-d$(CGCGAAAAATTTTTCG)_N$-3′

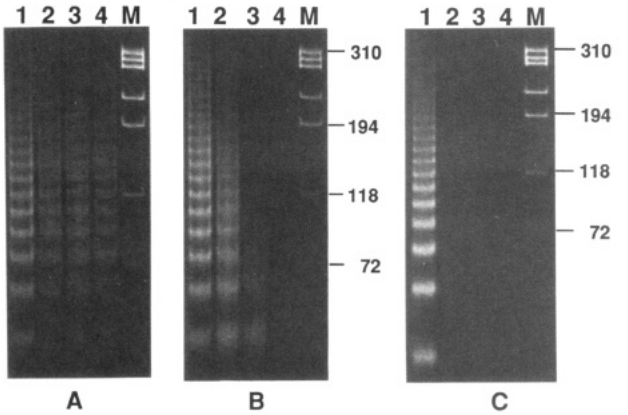

FIGURE 7: Effects of nucleases on the ligated hexadecamer d-(CGCGAAAAATTTTTCG). (A) BAL 31 (5 units); reaction times were (lane 1) 0, (2) 8, (3) 14, and (4) 20 min. (B) DNase I (1 unit); reaction times were (lane 1) 0, (2) 2, (3) 4, and (4) 10 min. (C) $(OP)_2Cu^+$; reaction times were (1) 0, (2) 2, (3) 4, and (4) 10 min. Reaction conditions are detailed in the legend for Figure 3 and under Experimental Procedures.

The polymerized species of the (d) form are composed of a monomeric unit which contains an intrinsic, sequence-inducing bent (Koo et al., 1986). Yet, these nucleic acid segments are not expected to reveal a long-range curvature as the repetitive motif, being a hexadecamer, is reiterated out of the helical phase.

Exposure of the ligated DNA fragments (d) to the exonuclease BAL 31 at those conditions which failed to affect the macroscopically bent segments (a) but induced a complete digestion of the noncurved molecules (b) and (c) (Figure 3) results in a partial degradation, as indicated by the increased smear and reduced intensity of the fragments (Figure 7A). Notably, BAL 31 is known to degrade AT-rich sequences significantly more rapidly than it degrades GC-rich regions (Sambrook et al., 1989). Moreover, it has been shown (Nelson et al., 1987) that the large propeller twist and other structural parameters which characterize A-tracts can extend to a neighboring GC base pair. These two findings combined indicate that the sequences derived from the decameric unit (a), in which only two GC base pairs separate between the A-tracts, should be more susceptible to BAL 31 digestion than the ligated (d) sequences, from both sequential and structural considerations. Thus, the observation according to which the segments of the (d) form, although containing a higher GC content, are more effectively digested by BAL 31 than the fragments of the (a) form underlines the unique properties of the macroscopically curved DNA species derived from (a). Similar results, namely, a partial digestion which is significantly more extensive than that revealed by the long-range curved molecules, but not as exhaustive as exhibited by the noncurved fragments, are observed when the ligated products derived from the 16-mer are exposed to the endonuclease DNase I (Figure 7B). Finally, exposure of the DNA molecules of the (d) form to the nuclease activity of $(OP)_2Cu^+$ results in a fast and efficient degradation of these segments, emphasizing again the dissimilarity between the polymers derived from (d) and those obtained from (a) (Figure 7C).

Discussion

The working hypothesis which initiated the current study was that sequence-derived curved DNA motifs might be physiologically significant. This hypothesis was based on the apparent ubiquity of the intrinsically bent structures in living systems, as well as their preferential localization in gene regulatory regions. In order to assess the possibility that DNA

bending is indeed involved in biological processes, the specific effects of such stable curvature upon DNA–protein interactions were investigated. Two groups of DNA-binding proteins were used in this study: proteins involved in the high-order organization of the nucleic acid molecules, namely, H1 histone, a major determinant in the packaging processes of chromatin in eukaryotic systems, and protamine, which participates in the condensation of DNA in sperm cells. The second group consists of DNA-modifying proteins, including ligases and exo- and endonucleases. By using gel retardation assays, we have found that those DNA fragments characterized by an intrinsic bending exhibit a lower affinity toward H1 and protamine than that revealed by noncurved nucleic acid structures. In addition, we have shown that inherently bent DNA structures are characterized by a significantly higher resistance against modifications induced by a broad range of DNA-modifying enzymes. On the basis of the findings obtained from the gel retardation assays, the enhanced stability associated with the curved sequences can be readily assigned to a reduced tendency of the curved motifs to interact with DNA-modifying proteins. Consequently, it may be suggested that intrinsically bent nucleic acid sequences exhibit a generic attenuation in their protein-binding affinities.

Both gel electrophoresis and electron microscopy techniques have been used to show that DNA minor groove-binding drugs such as distamycin reduce the extent of a sequence-dependent curvature following their complexation with bent nucleic acid sequences (Wu & Crothers, 1984; Griffith et al., 1986). The increased susceptibility of the curved DNA segments (a) toward nucleases that is observed in the presence of distamycin (Figure 6) might, consequently, be correlated with a partial elimination of the curvature which acts to partially restore the normal DNA-binding properties, thus allowing a more efficient digestion by the nucleases. The finding that the presence of the drug does not result in a complete digestion might be interpreted in terms of a protective effect exerted by the bound drug molecules upon the DNA species (Portugal & Waring, 1987), as well as an incomplete neutralization of the intrinsic curvature.

The notion of a general, curvature-related attenuation of the nucleic acid binding affinities is buttressed by the conspicuous temperature effect upon the ligation kinetics of the various DNA segments. It has been reported that the bending of kinetoplast DNA fragments, as well as of synthetic curved sequences, is significantly stabilized and enhanced as the temperature is decreased (Marini et al., 1984; Koo et al., 1986). The observations presented in Figure 1, according to which the reduction in the rate of the ligation reaction sustained by the bent fragments upon cooling is significantly more pronounced than the reduction of the ligation rates of the noncurved segments (b) and (c), could be assigned to a temperature-induced enhancement and stabilization of the intrinsic curvature of (a); such conformational outcomes might, in turn, result in an attenuated interaction between these DNA sequences and the ligase, leading to a relatively poorer efficiency of the ligation process. Evidently, a reduction of the extent of the DNA sequence-related bending, induced either by the groove-binding agent distamycin or by elevated temperatures, is concomitant with a reduction in the extent of the unique binding features found to characterize the bent segments.

The unusual properties of inherently curved DNA molecules are further demonstrated by the effects of the two chemical nucleases $(OP)_2Cu^+$ and $Fe(EDTA)^{2-}$ upon these structural motifs. Both species affect nucleic acid degradation by producing hydroxyl radicals which are the actual cleaving factor (Tullius et al., 1987). Yet, the strand scission activity of these two agents is accomplished through two different mechanisms. Specifically, it has been shown that the $(OP)_2Cu^+$ binds to the dsDNA, thus forming a reactive intermediate in which one phenanthroline lies within the minor groove while the second heteropolycyclic moiety extends outside the wall of the groove; the secondary structure specificity characterizing the activity of $(OP)_2Cu^+$ is indeed attributed to its different binding affinities toward various nucleic acid secondary conformations (Pope & Sigman, 1984). In contrast, the negatively charged $Fe(EDTA)^{2-}$ is not bound to the polyanionic DNA molecules, and its DNase activity lacks, consequently, any significant sequence or structure preferences (Tullius et al., 1987). These considerations indicate that the observations according to which the cleavage of the (a) polymers by $(OP)_2Cu^+$ proceeds at a lower rate than that of (b), (c), and (d) (Figures 4 and 7C), while no clear difference in the degradation rate of these species can be detected when the process is induced by $Fe(EDTA)^{2-}$ (Figure 5), can conceivably be interpreted in terms of altered binding patterns that are associated with the curved DNA sequences. Notably, the partial elimination of the resistance against enzymatic digestion, detected in the presence of the curvature-neutralizing drug distamycin, is observed in the case of $(OP)_2Cu^+$ as well (Figure 6B).

Intrinsically bent DNA sequences are, consequently, shown to exhibit attenuated binding affinities toward proteins, being either structural proteins such as histone and protamine or DNA-modifying enzymes, as well as toward minor groove binder species. Two different mechanisms might be responsible for the phenomenon. It may be suggested that the conformational constraints that are associated with a stable, sequence-dependent bending are capable of directly modulating the binding parameters of the curved DNA motifs. Alternatively, the inherent curvature, on one hand, and the altered binding properties exhibited by the bent DNA sequences, on the other, might represent two independent consequences of specific base pair arrangements. Can these two potential mechanisms be resolved?

Runs of A·T base pairs which, when reiterated in phase with the helical repeat, induce an overall bend to the DNA configuration are characterized by unusual structural parameters that differentiate such a motif from other B-type DNA sequences (Yoon et al., 1988). One of the most conspicuous features exhibited by A-tracts is a large propeller twist, which acts to enhance purine–purine stacking interactions (Nelson et al., 1987). The effect of the distinct stacking arrangement on the sugar–phosphate chains is to substantially narrow the width of the minor groove, a narrowing which increases in its extent on going from the 5′ to 3′ direction, as shown by hydroxyl radical cutting experiments (Burkhoff & Tullius, 1987). A periodic widening and narrowing of the minor groove observed in the A_NT_N sequence but not in the noncurved T_NA_N segment has been suggested as a potential curvature-inducing factor (Burkhoff & Tullius, 1988). Indeed, the distinct groove patterns and the bending are both eliminated at elevated temperatures, as well as in the presence of distamycin. An additional property of the A-tracts is concerned with the formation of an ordered spine of hydration along the minor groove which, in turn, further stabilizes the base stacking (Fratini et al., 1982; Buckin et al., 1989). Significantly, a continuous hydration network, which is removed by elevated temperatures, can be formed in A_NT_N sequences but not in T_NA_N, as it is disrupted by a TpA step and not an ApT sequence. An undisrupted base pair inclination, induced by the continuous spine of hydration characterizing the A-tracts (as

well as the A_NT_N sequences which act as an A-tract), has been suggested as a sequence-derived curvature determinant (Crothers et al., 1990).

The above presented observations and considerations underline the basic, causal dependence of a sequence-derived curvature upon the distinct structural features characterizing the A-tracts. This dependence renders the elucidation of the factor that is directly responsible for the altered binding affinities exhibited by the bent DNA sequences intrinsically ambiguous. It can be suggested that the attenuated binding affinities result from the unique patterns of the minor groove that are exhibited by the A-tracts and accompany the curvature. Significantly, it has been shown that mutations within the Pribnow box of the lac operon, which result in stronger promoters, are accompanied by an enhanced sensitivity of these sequences to $(OP)_2Cu^+$, indicating an increased accessibility of the promoter region (Sigman et al., 1985; Spassky & Sigman, 1985). The effect might originate from the widening of the minor groove which is the primary target of this nuclease, a widening which may, in turn, be responsible for the increased affinity of the mutated Pribnow box for RNA polymerase. According to these considerations, DNA–protein interactions might be attenuated by a curvature-associated narrowing of the minor groove and enhanced by its widening. Alternatively, the altered binding properties of the curved DNA species can be ascribed to the unique base stacking and base inclination that are induced and stabilized in A-tracts by the strong minor groove hydration. These conformational features, suggested as a main determinant of the curvature, act to reduce the flexibility of the DNA segments and as such could affect protein binding affinities. Finally, the modulated nucleic acid complexation parameters might be directly attributed to an intrinsic curvature. DNA bendability, namely, the ability of DNA segments to undergo protein-induced bending, is recognized as an important factor in determining the strength of many regulatory protein–DNA complexes such as CAP–DNA and Cro–DNA species (Gartenberg & Crothers, 1988; Brennan et al., 1990). It might be suggested that whereas an induced bending facilitates and stabilizes nucleic acid–protein interactions, a large, stable intrinsic curvature can interfere with such processes.

Evidently, the factor directly responsible for the altered binding properties found to be associated with the curved DNA sequences cannot be unambiguously resolved as yet. We propose, however, that the phased reiteration of such a factor, being either A-tracts or a sequence-derived bending, acts to enhance the extent of the binding modulations. This suggestion is based upon the results obtained from the experiments conducted on the repetitive DNA segments of the (d) form. These polymers contain an A-tract and, conceivably, a local bending which are, however, reiterated out of the helical phase. The observations according to which the (d) segments are more efficiently digested by enzymatic and chemical nucleases than the globally curved (a) sequences (Figure 7), although they should, from mere base composition considerations, be less susceptible to cleavage by BAL 31, indicate that a phased reiteration of the binding–modulating factor could amplify the extent of these modulations. Specifically, it may be envisaged that phased A-tracts, or long-range curvature which results from appropriately phased, repetitive bends, may act as a single cooperative unit, thus cooperatively enhancing their effects upon the binding parameters, whereas nonphased A-tracts or bends behave independently.

The extent of DNA curvature can be tuned by intrinsic factors, namely the length, number of reiterations, and precise position of the A-tracts, as well as by extrinsic factors which include the presence of ions such as Cu^{2+} and Mg^{2+} (Laundon & Griffith, 1987) and polyamines. This observation, the occurrence of phased A-tracts and, hence, globally curved DNA motifs in regulatory regions, combined with our results which point toward a correlation between A-tracts, intrinsically bent sequences, and substantial modulations of the DNA-binding properties raises the possibility that adenine tracts or a bending associated with such tracts may act as a hitherto unconsidered determinant in gene regulation processes.

References

Balbas, A. M., Campos, A. R., Ramirez, M. G. R., Carrera, J. S. P., Aymami, J., & Azorin, F. (1990) *Biochemistry 29*, 2342–2348.

Bernardin, W., Losa, R., & Koller, T. (1986) *J. Mol. Biol. 189*, 503–517.

Bossi, L., & Smith, D. M. (1984) *Cell 39*, 643–652.

Brennan, R. G., Roderick, S. L., Takeda, Y., & Matthews, B. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8165–8169.

Buckin, V. A., Kankiya, B. I., Bulichov, N. V., Lebedev, A. V., Gukovsky, I., Chuprina, V. P., & Williams, A. R. (1989) *Nature 340*, 321–322.

Burkhoff, A. M., & Tullius, T. D. (1987) *Cell 48*, 935–943.

Burkhoff, A. M., & Tullius, T. D. (1988) *Nature 331*, 455–457.

Chuprina, V. P., & Abagyan, R. A. (1988) *J. Biomol. Struct. Dyn. 6*, 121–138.

Crothers, D. M., Haran, T. E., & Nadeau, J. G. (1990) *J. Biol. Chem. 265*, 7093–7096.

Drew, H. R., & Travers, A. A. (1985) *J. Mol. Biol. 186*, 773–790.

Fratini, A. V., Kopka, M. L., Drew, H. R., & Dickerson, R. E. (1982) *J. Biol. Chem. 257*, 14686–14707.

Fried, M., & Crothers, D. M. (1981) *Nucleic Acids Res. 9*, 6506–6525.

Gartenberg, M. R., & Crothers, D. M. (1988) *Nature 333*, 824–829.

Griffith, J. D., Bleyman, M., Rauch, C. A., Kitchin, P. A., & Englund, P. T. (1986) *Cell 46*, 717–724.

Hagerman, P. J. (1985) *Biochemistry 24*, 7033–7037.

Hagerman, P. J. (1986) *Nature 321*, 449–450.

Hagerman, P. J. (1988) in *Unusual DNA Structures* (Wells, R. D., & Harvey, S. C., Eds.) pp 225–236, Springer-Verlag, New York.

Haran, T. E., & Crothers, D. M. (1989) *Biochemistry 28*, 2763–2767.

Hayes, J. J., Tullius, T. D., & Wolfe, A. P. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 7405–7409.

Hsieh, C.-H., & Griffith, J. D. (1988) *Cell 52*, 535–544.

Kitchin, P. A., Klein, V. A., Rian, K. A., Gann, K. L., Rauch, C. A., Kang, D. S., Wells, R. D., & Englund, P. T (1986) *J. Biol. Chem. 261*, 11302–11309.

Koo, H.-S., Wu, H. M., & Crothers, D. M. (1986) *Nature 320*, 501–506.

Koo, H.-S., & Crothers, D. M. (1987) *Biochemistry 26*, 3745–3748.

Koo, H.-S., & Crothers, D. M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1763–1767.

Laundon, C. H., & Griffith, J. D. (1987) *Biochemistry 26*, 3759–3762.

Levene, S. D., Wu, H.-M., & Crothers, D. M. (1986) *Biochemistry 25*, 3988–3995.

Lilley, D. (1986) *Nature 320*, 487–488.

Marini, J. C., Levene, S. D., Crothers, D. M., & Englund, P. T. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 7664–7668.

Nelson, H. C. M., Finch, J. T., Luisi, B. F., & Klug, A. (1987) *Nature 330*, 221-226.
Ntambi, J. M., Marini, J. C., Bangs, J. D., Hajduk, S. L., Jimenez, H. E., Kitchin, P. A., Klein, V, K., & Englund, P. T. (1984) *Mol. Biochem. Parasitol. 12*, 273-286.
Pope, L. E., & Sigman, D. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 3-7.
Portugal, J., & Waring, M. J. (1987) *FEBS Lett. 225*, 195-200.
Radic, M. Z., Lundgren, K., & Hamkalo, B. A. (1987) *Cell 50*, 1101-1108.
Ryder, K., Silver, S., Delucia, A. L., Fanning, E., & Tegtmeyer, P. (1986) *Cell 44*, 719-725.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) in *Molecular Cloning*, p 5.75, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sigman, D. S., Spassky, A., Rimsky, S., & Buc, H. (1985) *Biopolymers 24*, 183-197.
Snyder, M., Buchman, A. R., & Davis, R. W. (1986) *Nature 324*, 87-89.
Spassky, A., & Sigman, D. S. (1985) *Biochemistry 24*, 8050-8056.
Sturm, N. R., & Simpson, L. (1990) *Cell 61*, 879-884.
Travers, A. A. (1987) *Trends Biochem. Sci. (Pers. Ed.) 12*, 108-112.
Trifonov, E. N., & Sussman, J. L. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 3816-3820.
Trifonov, E. N., & Ulanovsky, L. (1988) in *Unusual DNA Structures* (Wells, R. D., & Harvey, S. C., Eds.) pp 173-187, Springer-Verlag, New York.
Tullius, T. D., Dombroski, B. A., Cherchill, M. E. A., & Kam, L. (1987) *Methods Enzymol. 155*, 537-558.
Ulanovsky, L., & Trifonov, E. N. (1987) *Nature 326*, 720-722.
Ulanovsky, L., Bonder, M., Trifonov, E. N., & Choder, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 862-866.
Wu, H. M., & Crothers, D. M. (1984) *Nature 308*, 509-513.
Yoon, C., Prive, G. G., Goodsell, D. S., & Dickerson, R. E. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6332-6336.
Zahn, K., & Blattner, F. R. (1985) *Nature 317*, 451-453.
Zahn, K., & Blattner, F. R. (1987) *Science 236*, 416-422.

cDNA Cloning and Sequencing of Rat α_1-Macroglobulin[†,‡]

Birgitta Wärmegård, Nancy Martin, and Staffan Johansson*
Department of Medical and Physiological Chemistry, The Biomedical Center, Box 575, S-751 23 Uppsala, Sweden
Received September 6, 1991; Revised Manuscript Received November 20, 1991

ABSTRACT: cDNA clones coding for the plasma protease inhibitor α_1-macroglobulin were isolated from a rat liver library. The obtained cDNA sequence contained 4701 nucleotides and had an open reading frame coding for a 1500 amino acid long protein, including a 24 amino acid signal peptide. The identity of the deduced protein sequence as α_1-macroglobulin was established by comparison with published peptide sequences of the protein. The mature protein shares 53% and 57% overall amino acid identity with the two other identified members of the rat α-macroglobulin family, α_1-inhibitor 3 and α_2-macroglobulin. A sequence typical for an internal thiol ester was identified. Of the 24 cysteines, 23 are conserved with α_2-macroglobulin. However, instead of the two most C-terminal cysteines in α_2-macroglobulin, which forms a disulfide bridge in the receptor binding domain, α_1-macroglobulin contains phenylalanine. One mRNA species hybridizing with the α_1-macroglobulin probe was observed in rat and mouse liver RNA (~6.2 kb), whereas no corresponding transcript was detected in RNA from human liver.

α-Macroglobulins (αMs)[1] are large plasma proteins which act as inhibitors of proteases of all subclasses [for reviews, see Sottrup-Jensen (1987, 1989)]. Upon cleaving of a particular peptide stretch in the αM, the "bait" region, the protease becomes trapped due to a conformational change of the inhibitor. In addition, an internal thiol ester in the αM is activated and may form a covalent linkage with the protease. A recently identified cell-surface receptor (Moestrup & Gliemann, 1989; Strickland et al., 1990) recognizes activated αMs and mediates internalization of the protease–αM complex by endocytosis (Van Leuven et al., 1979). Interestingly, also the complement factors C3 and C4, which are structurally related to the αMs (Sottrup-Jensen, 1987), are activated by limited proteolysis, bind their target proteins by use of a labile thiol ester, and are recognized by specific receptors after complex formation.

In rat, three members of the αM protein family have been described (Lonberg-Holm et al., 1987). α_2M and α_1M are tetramers of 180-kDa subunits whereas α_1I_3 is monomeric. α_1M differs from the other two proteins in that each monomer is composed of a heavy chain (~140 kDa) and a light chain (~40 kDa) which are held together by disulfide bonds (Lonberg-Holm et al., 1987). While the primary structures of α_2M and α_1I_3 are known (Gehring et al., 1987; Braciak et al., 1988), only limited peptide and cDNA data of α_1M have

[†] This work was supported by grants from the Swedish Medical Research Council (7147), the King Gustaf V's 80th Anniversary Fund, and the Medical Faculty at Uppsala University.
[‡] The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number JO5359.

[1] Abbreviations: αMs, α-macroglobulins; α_1M, α_1-macroglobulin; α_2M, α_2-macroglobulin; α_1I_3,α_1-inhibitor 3; PZP, pregnancy zone protein; bp, base pair(s); IPTG, isopropyl β-D-thiogalactopyranoside; PAGE, polyacrylamide gel electrophoresis; SDS, sodium dodecyl sulfate; SSC, 0.15 M NaCl and 15 mM sodium citrate.

0006-2960/92/0431-2346$03.00/0