Histone H1 Binds to the Putative Nuclear Factor I Recognition Sequence in the Mouse $\alpha_2(I)$ Collagen Promoter*

Jukka Ristimiemi and Jouko Oikarinen

From the Collagen Research Unit and Biocenter, University of Oulu, Department of Medical Biochemistry, Kajaanintie 52A, SF-90220 Oulu, Finland

It has previously been demonstrated that nuclear factor I (NF I) or a related protein binds to a region between -215 and -295 from the start of transcription in the mouse $\alpha_2(I)$ collagen gene promoter. In the present work we have purified this factor to homogeneity from rat liver. DNA sequence-specific proteins were isolated using heparin-agarose affinity chromatography and two successive chromatographies on a recognition site affinity matrix. Approximately 160 $\mu$g of the DNA binding proteins was obtained from 100 g of rat liver. More than 1700-fold purification over the nuclear extract and 58% recovery of the DNA binding activity was achieved.

The purified preparation contained five to six protein components ranging in molecular weight from 30,000 to 35,000, as determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. It was demonstrated using DNase I footprint analysis that the factor binds to the putative NF I binding site in the mouse $\alpha_2(I)$ collagen promoter. It has a dissociation constant of 7 nM for a short DNA fragment containing this binding site, while a constant of 0.45 nM was obtained for a similar-sized fragment containing the nuclear NF I consensus binding site.

The purified factor is identical to histone H1 in several respects. They share similar amino acid compositions and they give similar V8-protease and N-bromoessuccinimide peptides. In addition, antibodies raised to bovine histone H1 recognize the purified factor and interfere with its binding to DNA. Methylation interference and preparative gel shift assay show that histone H1 binds to the specific sequence from the preparation of the $\alpha_2(I)$ collagen promoter binding factor. It is thus evident from the present results, that histone H1 binds to the NF I recognition sequence in the mouse $\alpha_2(I)$ collagen promoter.

Type I collagen, which is one of the main secretory proteins synthesized by fibroblasts (1), confers tensile strength to various kinds of connective tissue. Its molecule is a trimer of two $\alpha_1(I)$ chains and one $\alpha_2(I)$ chain, coiled into a triple-helical structure. Interestingly, the rate of type I collagen synthesis is regulated, e.g. during development of the mouse embryo, and the protein can be detected in utero after day 8 (2-3). This induction is essential for development. Inactivation of the $\alpha_1(I)$ collagen gene, brought about by retroviral insertion of foreign DNA into the first intron of the gene in the mouse oocyte, is lethal (4-5). Major changes in the rate of synthesis of type I collagen also take place in various pathological states. A drastic decrease is observed in fibroblasts after viral transformation, for example, due to a reduced rate of transcription of the corresponding genes (6-8).

Genes for type I collagen polypeptide chains have been characterized from several species. The $\alpha_1(I)$ and $\alpha_2(I)$ collagen genes generally consist of 51 or 52 exons separated by an equal number of introns (9-10). A single copy per haplotype genome of the $\alpha_1(I)$ gene is located in the long arm of human chromosome 17 (11), and that of the $\alpha_2(I)$ gene in chromosome 7 (12). The length of the $\alpha_1(I)$ collagen gene is only one-half that of the $\alpha_2(I)$ gene, since they span 18 and 38 kilobases, respectively (13). In most situations, they are transcribed in the ratio of 2:1, which confers the same ratio on the steady-state levels of the corresponding mRNAs (14). The $\alpha_1(I)$ and $\alpha_2(I)$ collagen mRNAs are equally translated and have identical half-lives (14).

Chimeric plasmids containing part of the mouse $\alpha_1(I)$ collagen gene promoter and a bacterial chloramphenicol acetyltransferase gene have been generated (15), transient transfection experiments with these plasmids (15) and studies on transgenic mice (16) have revealed that the part of the promoter that ranges from -2000 to +54 from the start of transcription is sufficient for expression and, together with an enhancer element in the first intron (17), confers the tissue specificity necessary for the expression of the $\alpha_1(I)$ collagen gene.

Several nuclear protein factors that recognize specific DNA sequences in the mouse $\alpha_1(I)$ collagen promoter have been discovered, and two of these have been characterized in more detail. One factor binds to a region between -70 and -96 that contains a CCAAT sequence on the sense strand between -84 (18), and the other binds to a region between -295 and -315 from the start of transcription (19). It has been demonstrated by deletion analysis that the latter area is also important for efficient expression of the $\alpha_1(I)$ collagen gene (15). Less than 10% of the activity observed with the authentic promoter is obtained when a mutant having a deletion from -350 to -290 is used in transient transfection experiments. Furthermore, in the light of similarities in the recognition sequences and of competition studies, it is probable that this factor is either NF I or a related protein (20).

In order to elucidate the basic mechanisms which regulate the expression of type I collagen genes, we have investigated the nuclear protein factors which bind to specific DNA sequences in the $\alpha_1(I)$ collagen promoter. We report here on the

* This study was supported in part by grants from the Medical Research Council of the Academy of Finland. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom reprint requests should be sent.

1 The abbreviations used are: NF I, nuclear factor I; SDS, sodium dodecyl sulfate; PMSF, phenylmethylsulfonyl fluoride; DTT, dithiothreitol; EGTA, [ethylenebis(oxyethylenenitrilo)]tetraacetic acid; bp, base pair; BSA, bovine serum albumin.

2164
isolation of proteins from rat liver which bind to the putative NF I binding site.

MATERIALS AND METHODS

RESULTS

Characterization of the α2(I) Collagen Promoter Binding Proteins—Proteins which recognize the putative NF I binding site in the mouse α2(I) collagen promoter were isolated from rat liver using heparin-agarose and two consecutive DNA isolation of proteins from rat liver which bind to the putative NF I binding site.

MATERIALS AND METHODS

RESULTS

Characterization of the α2(I) Collagen Promoter Binding Proteins—Proteins which recognize the putative NF I binding site in the mouse α2(I) collagen promoter were isolated from rat liver using heparin-agarose and two consecutive DNA recognition site affinity chromatographies. The affinity purified material was subjected to electrophoresis on a SDS-polyacrylamide gel. Silver staining revealed the presence of at least five or six protein bands, ranging in molecular weight from 30,000 to 35,000 (majority of the staining at 33,000) (Fig. 1A, lane 1). Some additional bands of lower molecular weight (about 20,500, see Fig. 1B) were also observed occasionally. The preparation was more than 90% pure, as judged by polyacrylamide gel electrophoresis.

The DNA recognition sequence of the purified factor has a partial dyad of symmetry (21, 22). This implies that proteins which bind to that sequence may be dimeric in structure, like many prokaryotic DNA binding proteins. Consistent with this, the purified DNA binding factor behaves in solution mainly like a protein of molecular weight ~60,000. This is demonstrated by gel filtration of the factor on a Sephacryl S-300 matrix followed by analysis of the protein by UV spectrophotometry and DNA binding activity by the nitrocellulose filter binding assay. Most of the protein and DNA binding activity co-eluted with the bovine serum albumin standard as a distinct peak (Fig. 2). The protein nevertheless has a tendency to aggregate, and occasionally forms with an even higher molecular weight were observed.

The broad protein band ranging in molecular weight from 30,000 to 35,000 was cut out from the acrylamide gel, electroeluted from the gel and subjected to sequencing of the amino terminus and to amino acid analysis. Despite 500 pmol of protein being subjected to automated gas-phase sequencing, no sequence was obtained. It is thus probable that the amino terminus of the DNA binding factor is blocked. Its amino acid composition nevertheless turned out to be characteristic (Table 1), the most conspicuous feature being an abundance of lysine, alanine, and proline residues, while only trace amounts of aromatic amino acids and no methionine were observed.

DNA Binding Characteristics of the Purified Protein—The purified factor recognizes the putative NF I binding site in the mouse α2(I) collagen promoter, as shown by DNase I footprint analysis, where the autoradiogram pointed to an area of protection from DNase I digestion and some sites of enhanced sensitivity (Fig. 3, lanes 2 and 3). The area protected by the isolated factor actually spanned exactly the same sequence reporter to a saline concentration matrix, the factor was subjected to electrophoresis on a SDS-polyacrylamide gel (15%). Lane 1 contains 1 μg of protein and lane 3 5 μg. Lane 2 contains only the sample buffer, thus demonstrating that the bands at molecular weight about 55,000-60,000 originate from the sample buffer. Panel B demonstrates the similarity of the purified factor to bovine histone H1. The two protein preparations were subjected to electrophoresis on a glycerol-SDS-polyacrylamide gel (12%) Lane 1 contains 1 μg of protein and lane 3 5 μg. Lane 2 contains only the sample buffer, thus demonstrating that the bands at molecular weight about 55,000-60,000 originate from the sample buffer. Panel B demonstrates the similarity of the purified factor to bovine histone H1. The two protein preparations were subjected to electrophoresis on a glycerol-SDS-polyacrylamide gel (15%). Lane 2 contains 5 μg of histone H1 (from Boehringer Mannheim) and lane 3 the α2(I) collagen promoter binding factor. The bands were visualized using silver staining. The molecular weight standards (lanes 1 and 4 in panel B) from top to bottom are phosphorylase b, bovine serum albumin, ovalbumin, carbonic anhydrase, soybean trypsin inhibitor, and lysozyme. The major components of the purified factor (molecular weights 30,000-35,000 and 20,500) are marked by the arrows. BPB, bromophenol blue.

FIG. 1. SDS-polyacrylamide gel electrophoresis of the purified α2(I) collagen promoter binding factor. Panel A demonstrates the purity of the factor. After purification using heparinagarose chromatography and two successive chromatographies on the recognition site affinity matrix, the factor was subjected to electrophoresis on a SDS-polyacrylamide gel (12%). Lane 1 contains 1 μg of protein and lane 3 5 μg. Lane 2 contains only the sample buffer, thus demonstrating that the bands at molecular weight about 55,000-60,000 originate from the sample buffer. Panel B demonstrates the similarity of the purified factor to bovine histone H1. The two protein preparations were subjected to electrophoresis on a glycerol-SDS-polyacrylamide gel (15%). Lane 1 contains 1 μg of protein and lane 3 5 μg. Lane 2 contains only the sample buffer, thus demonstrating that the bands at molecular weight about 55,000-60,000 originate from the sample buffer. Panel B demonstrates the similarity of the purified factor to bovine histone H1. The two protein preparations were subjected to electrophoresis on a glycerol-SDS-polyacrylamide gel (15%). Lane 2 contains 5 μg of histone H1 (from Boehringer Mannheim) and lane 3 the α2(I) collagen promoter binding factor. The bands were visualized using silver staining. The molecular weight standards (lanes 1 and 4 in panel B) from top to bottom are phosphorylase b, bovine serum albumin, ovalbumin, carbonic anhydrase, soybean trypsin inhibitor, and lysozyme. The major components of the purified factor (molecular weights 30,000-35,000 and 20,500) are marked by the arrows. BPB, bromophenol blue.

and Scatchard plot analysis (Fig. 4b). A dissociation constant of 7 nM was obtained both for the PouI-PouII fragment from plasmid pBS1a and for the BglII-PstI fragment from plasmid pAZ1009. Both of these represent the mouse α2(I) collagen binding site. A dissociation constant of 0.45 nM was obtained for the PouI-PouII fragment from plasmid pNFlb, which is the same size as the corresponding fragment from plasmid pBS1a and contains the consensus sequence for NF I binding sites. No binding was observed to the PouII-PouII fragments from plasmids pUC13 and pUC19, which are the parent plasmids of pNFlb and pBS1a, respectively.

The Purified Protein Is Identical to Histone H1—The amino acid composition of the purified protein, especially its richness in lysine and alanine, suggested a possible similarity to histone H1 variants. The amino acid compositions of the two proteins were remarkably similar (Table 1). It should be noted that histone H1 complement shows less sequence homology between various species and different variants of the same
a2(I) Collagen Promoter Binding Factor

**Fig. 2.** Sephacryl S-300 chromatography of the purified factor. The purified factor (20 µg) was chromatographed on a Sephacryl S-300 matrix. The fractions were subjected to determination of DNA binding activity using the 32P-labeled PvuII-PvuII fragment from plasmid pNF1b ( ), or to spectrophotometric determination of protein ( ), the molecular weight standards were catalase (240,000), bovine serum albumin (BSA) (66,000), and pepsin (35,000).

**Fig. 3.** DNase I footprinting analysis of binding of the purified factor to the a2(I) collagen promoter binding factor and histone H1. The BglII-TaqI fragment from plasmid pAZ1009 was labeled at the BglII site using T4 polynucleotide kinase. The anti-sense strand of the collagen promoter was thus labeled. The 5’-end-labeled fragment was then incubated in the absence (lanes 1 and 4) or presence (lanes 2 and 3) of the purified factor (250 ng), treated with DNase I, and the digestion products were analyzed on a sequencing gel. An autoradiogram of the gel is shown. The Maxam-Gilbert sequencing reactions from the same fragment were run simultaneously. The sequence TCGCNCCAA can be found on the opposite strand inside the protected area. Thick arrows denote protection and narrow arrows enhancement.

**Table I**

Amino acid composition of the purified a2(I) collagen promoter binding factor and histone H1

| Amino acid residues | mmol/mol | Rat NFI | Chicken H101 | Rabbit H1.3 | Mouse H11 |
|---------------------|----------|---------|-------------|-------------|-----------|
| Asx                 | 40       | 9       | 5           | 5           | 8         |
| Thr                 | 60       | 13      | 10          | 8           | 15        |
| Ser                 | 70       | 15      | 13          | 13          | 19        |
| Glx                 | 37       | 8       | 7           | 8           | 13        |
| Pro                 | 106      | 23      | 21          | 26          | 24        |
| Gly                 | 86       | 17      | 14          | 15          | 15        |
| Ala                 | 261      | 44      | 59          | 58          | 40        |
| Cys                 | ND       | 0       | 0           | 0           | 0         |
| Val                 | 59       | 13      | 8           | 8           | 14        |
| Met                 | 0        | 0       | 0           | 0           | 0         |
| Ile                 | 21       | 5       | 2           | 2           | 2         |
| Leu                 | 70       | 15      | 10          | 10          | 9         |
| Tyr                 | Trace    | 1       | 1           | Trace       |           |
| Phe                 | Trace    | 1       | 1           | 1           | Trace     |
| His                 | 0        | 0       | 0           | 0           | 0         |
| Lys                 | 234      | 52      | 60          | 66          | 58        |
| Arg                 | 18       | 4       | 7           | 3           | 4         |
| Total               | ~220     | 218     | 224         | ~220        |

The molecular weight 30,000-35,000 protein was eluted from a SDS-polyacrylamide gel and subjected to acid hydrolysis and determination of amino acid composition. The values are means of two determinations, ND, not determined. The number of amino acid residues per chain was calculated assuming that the protein consists of 220 amino acid residues. The composition of rat NF I was then compared with those of the chicken histone H101 (27), rabbit H1.3 (42), and mouse H11 (50) variants.

Each of the five or six histone H1 polypeptide chain variants encountered in all mammals has a molecular weight of about 20,500, but the protein features many posttranslational modifications such as the amino-terminal N-acetylserine (24), several poly(ADP-ribose) moieties linked to glutamates (25) and phosphates in serines, threonines, and lysines (26). These modifications are likely to increase the real molecular weight of the protein quite considerably. The NF I preparation shows amazingly similar mobility in a SDS-polyacrylamide gel to that of a bovine histone H1 preparation (Fig. 1B), to the extent that these two preparations are indistinguishable. The molecular weight 20,500 component may be a degradation product of H1. The higher molecular weight form of the a2(I) collagen promoter binding factor can be labeled by incubating nuclei in the presence of 32P-labeled NAD, thus suggesting that it contains poly(ADP-ribose) (data not shown).

In order to acquire additional support for the identity of the two proteins, the liver nuclear factor I and bovine histone H1 preparations were digested with Staphylococcus aureus V8 protease and the peptides analyzed on a glycerol-SDS-polyacrylamide gel. Both preparations gave rise to two major groups of peptides of molecular weight about 12,000 and 14,000 (Fig. 5A). These peptides apparently originate from the basic carboxyl termini, containing the part from the glutamate residue at 116 in chicken H101 variant to the
plasmid pBSla labeled in the same way. These fragments are of binding affinities. pNFlb, pBSla, and pAZ1009, respectively. approximately the same size (330 and 320 base pairs, respectively). were 760, 1420, and 640 cpm/fmol for the fragments from plasmid factor to various genes in the assay is 9 fmol. If it is assumed that the factor binds in the manner of a molecular weight 30,000-35,000 protein in the preparation the carboxyl terminus at about 220 (from Ref. 27). Other peptides obtained were more variable between the two preparations, apparently due to inter-species variations in glutamate residues in the aminoterminal part of the molecules (23).

Use was made of the ability of N-bromosuccinimide to destroy tyrosine residues in the polypeptide chain (28), in order to demonstrate a further similarity between the two proteins. The main bovine histone H1 variants have only a single tyrosine residue (the location of which corresponds to amino acid residue 71 in chicken H101, from Ref. 27), so that chemical degradation should give rise to peptides two-thirds and one-third of the size of the original polypeptide. Bovine histone H1 and the α2(I) collagen promoter binding factor from the rat gave rise to very similar peptide patterns, as demonstrated by analysis on a glycerol-SDS-polyacrylamide gel (Fig. 5B). The mobilities of the major peptides were in good agreement with cleavage of the molecules at a single site, as expected.

Immunological studies were performed to indicate that the molecular weight 30,000-35,000 protein in the preparation of the α2(I) collagen promoter binding factor is histone H1. Antibodies raised in rabbits to bovine histone H1 recognize both bovine and rat histone H1s that were isolated using the conventional methods and the molecular weight 30,000-35,000 bands in the preparation of the α2(I) collagen promoter binding factor (Fig. 6A). The antibodies apparently have equal affinities for histone H1 and the purified factor, as tested by using equal amounts of the proteins in dot blot immunoassay. From crude cell extracts the antibodies did not recognize any bands, and denaturation by boiling or an SDS-treatment does not affect the antigenicity (data not shown).
Histone H1 Binds to the NF I Recognition Sequence—In the light of the above data, it is evident that the molecular weight 33,000 protein in the preparation of α2(I) collagen promoter binding factor is identical to histone H1. Although the protein was more than 90% pure, it is not unequivocally the case that the molecular weight 33,000 protein must be

J. Ristiniemi and J. Oikarinen, unpublished results.
that which binds to the specific sequence. Methylation interference experiments performed clearly demonstrate that the proteins which cause retardation in the gel shift assay really do bind to the putative NF I recognition sequence in the \( \alpha_f(I) \) collagen promoter (Fig. 7A). In the retarded band the signals for the two adjacent G residues complementary to Cc in the GCCAA sequence were clearly missing. On the other hand, the preparative gel retardation assay combined with SDS-polyacrylamide gel electrophoresis shows that the molecular weight 33,000 protein in fact causes such retardation (Fig. 7B, lane 2).

In order to demonstrate that the protein which binds to the specific DNA sequence is recognized by the histone H1 antibodies, antisera was added during the binding incubation in the gel retardation assay. An extra band of slow mobility was observed when the factor and DNA fragments were incubated in the presence of antibodies to histone H1 and this band was not detectable when pre-immune antisera was used (Fig. 6B, lanes 4 and 6).

**DISCUSSION**

A group of proteins which bind to a putative NF I recognition sequence around -300 with respect to the start of transcription in the mouse \( \alpha_f(I) \) collagen promoter are isolated here. Nuclear extract from rat liver was subjected to heparin-agarose and two successive DNA recognition site affinity chromatographies (20). Purification of DNA binding activity was followed by two assays: a DNA gel retardation assay (19, 29–31) and a nitrocellulose filter binding assay (20). 160 µg of the affinity purified proteins was obtained from 100 g of rat liver. A 1700-fold purification over the rat liver nuclear extract was achieved, with a 58% recovery of the DNA binding activity.

The protein preparation obtained consisted of from five to six components, ranging in molecular weight from 30,000 to 35,000. Surprisingly, the amino acid composition of the factor was very similar to that of histone H1 (see Ref. 32). The purified rat factor and bovine histone H1 had identical mobilities in a SDS-polyacrylamide gel and gave rise to both V8 protease and N-bromosuccinimide peptide maps which very closely resembled each other. Furthermore, the antibodies raised to bovine histone H1 recognized the \( \alpha_f(I) \) collagen promoter binding factor and interfered with its binding to DNA. It was demonstrated using methylation interference (19) and preparative gel shift assay (33) that the same protein from the preparation also binds to the putative NF I recognition sequence. Thus, evidently histone H1 binds to the NF I recognition sequence in the mouse \( \alpha_f(I) \) collagen promoter.

The band pattern recognized by the histone H1 antibodies in the preparation of the promoter binding factor was slightly different from that of histone H1. However, similar pattern will be obtained after chemical acetylation of histone H1 by acetic anhydride. It is therefore evident that the molecular weight 33,000 protein in the preparation of the \( \alpha_f(I) \) collagen promoter binding factor is histone H1.

The eukaryotic genome is packed into nucleosome structures in the chromatin consisting of various kinds of histone protein. Histone H1 binds to nucleosome linker regions, probably sealing the DNA that is wrapped two turns around the core histones (34–38). From four to six variants of histone H1 exist in each mammalian species (39–45), exhibiting tissue and developmental stage-specific expressions (36, 39, 40, 46–51). The function of histone H1 has remained unelucidated. It has been hypothesized that it packs the chromatin into inactive condensed form (52–60), thus acting as a general repressor (61). This hypothesis is supported by the finding that active chromatin contains less histone H1, or else histone H1 is only loosely bound to DNA (62, 63). In addition, some results suggest that the nucleosome structure is clearly different in actively transcribed areas (64).

The factor purified here behaved in solution like a molecular weight 60,000 protein and could thus be a dimer of two identical subunits. This is consistent with previous studies which show that NF I recognizes a sequence that displays an apparent dyad of symmetry (21, 22). Interestingly, purified histone H1 also has a tendency to form aggregates (25). It is not known what is the functional complex in tissues, but it is possible to cross-link two histone H1 molecules in \textit{vivo} (28), suggesting that dimers may really exist in the nucleus.

On the other hand, no one has ever been able to show that histone H1 binds to a specific sequence in DNA. There are some hints to the effect that it could be a sequence-specific DNA binding protein, the most convincing one being the constant pattern of nucleosomes encountered in \textit{vivo} (see Ref. 65), but it seems to be very difficult to reconstitute this pattern in \textit{vitro}, probably because the factors determining nucleosome positioning are not known. Interestingly, we have recently demonstrated that the globular domain of the histone H1 molecule displays homology with the nucleotide binding domain of adenine nucleotide binding proteins (66), thus suggesting that the histone H1 molecule could possess domains that would ascribe it sequence specificity for binding. It nevertheless still remains to be explained why no one has ever noticed this characteristic of histone H1 before. One explanation might be that the conventional purification procedures used for histone H1 (67) include denaturing perchloric acid extraction and precipitation with acidified acetone.

It has previously been demonstrated that NF I will bind to several viral DNAs such as adenovirus (68), mouse mammary tumor virus-long terminal repeat (69), BK virus (69), herpes simplex virus (70) and human cytomegalovirus DNAs (71), as well as to cellular genes such as the lysozyme (72), c-myc (73) and \( \alpha_f(I) \) collagen genes (19). Several reports have come out recently on the purification of this factor. Originally, Nagata \textit{et al.} (74) reported purification of a molecular weight 47,000 protein which was though to be related to NF I activity. Later on, Kelly and Rosenfeld (20) reported purification of a factor which ranged in molecular weight from 52,000 to 66,000, and Diffley and Stillman (75) that of a factor which migrated at about 160,000 by globular standards. Very recently, purification of a protein which had NF I-like activity and a molecular weight of from 30,000 to 35,000 was reported by Rupp and Sippel (33). The factor we have purified is probably identical to the last of these. It is difficult to explain the discrepancy between the molecular weights obtained, although the factor we have purified has a tendency to aggregate. Like to those of histone H1 (25), NF I aggregates are apparently not solubilized by the SDS treatment that is usually performed prior to polyacrylamide gel electrophoresis.

It is possible that the variation in the molecular weights of the NF I preparations may at least partly be due to differences between species, since Rupp and Sippel used chicken liver, we used rat liver and the other groups cultured human HeLa cells. There is an argument against this possibility, however, namely that the molecular weight 33,000 protein can also be obtained from cultured human skin fibroblasts. It is more...
conceivable that this protein has been conserved throughout evolution in all species as is the case with histone H1 (see Ref. 22). Even so, there is still a possibility that a single species may possess multiple proteins that have NF I-related activity, and that we have purified one member of that gene family.

The dissociation constants of the present factor for binding to the NF I site in the α(1)I promoter and to the NF I consensus sequence seem to be high, thus suggesting low affinity binding. Lower values have previously been reported (20). It should be noted, however, that the dissociation constants depend on the non-specific DNA used in the assay. Therefore, it is conceivable that NF I, i.e. histone H1, may have some affinity for the [poly(dI-dC)][poly(dI-dC)] used here and that this interaction obviously interferes with the sequence-specific binding.

NF I has been assigned two different functions. Originally it was found to promote replication of adovirus (68, 76–77), but recently it has also been shown to enhance transcription of some eukaryotic genes directed both by viral and host cell promoters (70, 78). The herpes simplex virus-thymidine kinase gene promoter contains a canonical CCAAT element around −80 from the start of transcription which is required for optimal expression of the gene (70). It has been proposed that this CCAAT binding factor may be NF I, and likewise NF I has been shown to be able to enhance transcription of other genes containing the CCAAT element, such as the β-globin gene. On the other hand, it has also been demonstrated that the steroid-dependent transcriptional stimulation obtained under the control of mouse mammary tumor virus-long terminal repeat is abolished when a NF I binding site around −80 in the long terminal repeat is deleted (78). This may mean that an interaction between the two DNA sequence-specific transcription factors, i.e. steroid receptors and NF I, is needed for optimal stimulation of transcription (79).

Deletion analysis of the α(1)I collagen promoter turned out to be complicated (15). It may be that the region around −300 to which the present purified factor binds is important for expression of the gene (16), although this was difficult to demonstrate. Indirect evidence was obtained when the cis-element was cloned in front of a minimal SV40 promoter (80), wherein the element stimulated expression of a chloramphenicol acetyltransferase marker gene. On the other hand, the stimulatory effect of transforming growth factor-β has been demonstrated to be mediated via the same cis-element (80), and therefore it is obvious that the binding of a protein factor to this putative NF I binding site must have an important regulatory role.

The finding that histone H1 binds to a specific sequence in DNA, that actually has already been suggested previously (81), has interesting implications and raises a number of questions. Firstly, what is the role of histone H1 in determining nucleosome positioning, and is the NF I recognition sequence a determinant there? Secondly, how do the various effectors described above act through these linker histones? These seem to be some of the basic questions to be answered when elucidating mechanisms involved in the regulation of eukaryotic gene expression.

Recently, three reports have come out concerning the cloning of a protein that binds to the NF I recognition sequence (82–84). This protein was expressed in bacteria and it evidently stimulates transcription and replication. However, this protein is highly homologous with histone H1, and therefore we conclude that H1 and that other protein, which should be called NF I, bind to the same operator sites on the DNA.

Acknowledgments—Plasmid pAZ1009 was a gift from Benoit de Crombrugge (M. D. Anderson Hospital and Tumor Institute, Houston, TX), and placmid pNF15s from Lothar Henninghausen (National Institutes of Health, Bethesda MD). We gratefully acknowledge the expert technical assistance of Ritva Savilakos, Liisa Aijalä and Paivi Rikula. The protein sequencing was performed by Marc Baumann (University of Helsinki). We also thank Benoit de Crombrugge and Deborah Kaska for valuable comments on the manuscript.

REFERENCES

1. Burgesson, R. E., and Morris, N. P. (1987) in Connective Tissue Disease. Molecular Pathology of the Extracellular Matrix (Uitto, J., and Perejda, A. J., eds) pp. 3–28, Marcel Dekker, New York.
2. Adamson, E. D., and Ayers, S. E. (1979) Cell 16, 953–965.
3. Labro, I., Vaheri, A., Timpl, R., and Warrickova, J. (1980) Dev. Biol. 76, 100–114.
4. Schnieke, A., Harbers, K., and Jaenisch, R. (1983) Nature 304, 315–320.
5. Hartung, S., Jaenisch, R., and Breindl, M. (1986) Nature 320, 365–367.
6. Howard, B. H., Adams, S. L., Sobel, M. E., Pastan, I., and de Crombrugge, B. (1978) J. Biol. Chem. 253, 5869–5874.
7. Sandmeier, S., Gallis, B., and Bornstein, P. (1981) J. Biol. Chem. 256, 5022–5028.
8. Schmitz, A., Setoyama, C., and de Crombrugge, B. (1985) Nature 314, 286–289.
9. Boedtker, H., Fuller, F., and Tate, V. (1983) Int. Rev. Connect. Tissue Res. 10, 1–63.
10. de Crombrugge, B., Schmidt, A., Liao, G., Setoyama, C., Murdy, M., Yamada, Y., and McKeeon, K. (1986) Ann. Rev. N. Y. Acad. Sci. 480, 154–162.
11. Huerre, C., Junien, C., Weil, D., Chu, M., Morabitto, M., Foubert, M., Myers, J. C., Van Cong, N., Gross, M., Prockop, D. J., Boule, A., Kaplan, J., de la Chapelle, A., and Ramirez, F. (1982) Proc. Natl. Acad. Sci. U. S. A. 79, 6627–6630.
12. Junien, C., Weil, D., Myers, J. C., Van Cong, N., Chu, M., Foubert, M., Gross, M. S., Prockop, D. J., Kaplan, J., and Ramirez, F. (1982) Am. J. Hum. Genet. 34, 381–387.
13. Chu, M.-L., de Wet, W., Bernard, M., Ding, J.-F., Morabitto, M., Myers, J., Williams, C., and Ramirez, F. (1984) Nature 310, 357–360.
14. Hamalainen, L., Okarinen, J., and Kivirikko, K. I. (1985) J. Biol. Chem. 260, 720–725.
15. Schmidt, A., Rossi, P., and de Crombrugge, B. (1986) Mol. Cell. Biol. 6, 347–354.
16. Khillan, J. S., Schmidt, A., Overbeek, P. A., de Crombrugge, B., and Westphal, H. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 725–729.
17. Rossi, P., and de Crombrugge, B. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 5590–5594.
18. Hatamochi, A., Paterson, B., and de Crombrugge, B. (1986) J. Biol. Chem. 261, 11310–11314.
19. Okarinen, J., Hatamochi, A., and de Crombrugge, B. (1987) J. Biol. Chem. 262, 11064–11070.
20. Rosenfeld, P. J., and Kelly, T. J. (1986) J. Biol. Chem. 261, 1398–1408.
21. Gronostajski, R. M. (1986) Nucleic Acids Res. 14, 9117–9132.
22. Gronostajski, R. M. (1987) Nucleic Acids Res. 15, 5545–5559.
23. Bohn, L., and Mitchell, T. C. (1985) FEBS Lett. 185, 1–4.
24. Tsuchiya, S., Iwata, S., Kurihara, Y., and Kurihara, Y. (1984) Methods Enzymol. 106, 165–170.
25. Stone, P. R., Lorimer, W. S., III, and Kidwell, W. R. (1977) Eur. J. Biochem. 81, 9–18.
26. Fujitaki, J. M., and Smith, R. A. (1984) Methods Enzymol. 106, 23–36.
27. Coles, L. M., Robins, A. J., Madley, L. K., and Wells, J. R. E. (1987) J. Biol. Chem. 262, 9656–9663.
28. Ring, D., and Cole, R. D. (1983) J. Biol. Chem. 258, 15361–15364.
29. Garner, M. M., and Revaz, A. (1981) Nucleic Acids Res. 9, 3047–3060.
30. Fried, M., and Crothers, D. M. (1981) Nucleic Acids Res. 9, 6505–6525.
31. Singh, H., Sen, R., Baltimore, D., and Sharp, P. A. (1986) Nature 319, 154–156.
32. Shannon, M. F., and Wells, J. R. E. (1987) J. Biol. Chem. 262, 9664–9668.
33. Rupp, R. A. W., and Sippel, A. E. (1987) Nucleic Acids Res. 15, 9707–9726.
alpha(I) Collagen Promoter Binding Factor

Differential Material to

Klonal /yR Guide to the Relative Nuclear Factor X Receptor Sequence of

By John W. Taylor and John Daher

MATERIALS AND METHODS

Preparation of Nuclear Extract - About 100 g of rat liver, obtained immediately after decapitation, was homogenized using a glass-teflon homogenizer. The homogenate was centrifuged at 10,000 × g for 15 min. The pellet was resuspended in an equal volume of 10 mM Tris-Cl, pH 7.4, 0.1 mM KCl, 5 mM EDTA, 0.5 mM dithiothreitol (DTT), 0.5% Nonidet P-40, 100 µg/ml leupeptin, and 200 µg/ml aprotinin. The extract was stored at -70°C.

Preparation of Nuclear Reovirus DNA Oligonucleotides - The nuclear extract was adsorbed to chromatin affinity chromatography. The extract was diluted with nuclei to reduce the nuclei concentration to 10-15 g/ml. The mixture was incubated at 4°C for 1 h. The mixture was layered onto a 30% sucrose cushion and centrifuged at 40,000 × g for 1 h. The supernatant was collected and stored at -70°C.

Preparation of Nuclear RNA Oligonucleotides - The nuclear extract was adsorbed to chromatin affinity chromatography. The extract was diluted with nuclei to reduce the nuclei concentration to 10-15 g/ml. The mixture was incubated at 4°C for 1 h. The mixture was layered onto a 30% sucrose cushion and centrifuged at 40,000 × g for 1 h. The supernatant was collected and stored at -70°C.

Preparation of Nuclear DNA Oligonucleotides - The nuclear extract was adsorbed to chromatin affinity chromatography. The extract was diluted with nuclei to reduce the nuclei concentration to 10-15 g/ml. The mixture was incubated at 4°C for 1 h. The mixture was layered onto a 30% sucrose cushion and centrifuged at 40,000 × g for 1 h. The supernatant was collected and stored at -70°C.

Preparation of Nuclear Protein Oligonucleotides - The nuclear extract was adsorbed to chromatin affinity chromatography. The extract was diluted with nuclei to reduce the nuclei concentration to 10-15 g/ml. The mixture was incubated at 4°C for 1 h. The mixture was layered onto a 30% sucrose cushion and centrifuged at 40,000 × g for 1 h. The supernatant was collected and stored at -70°C.

Preparation of Nuclear RNA Oligonucleotides - The nuclear extract was adsorbed to chromatin affinity chromatography. The extract was diluted with nuclei to reduce the nuclei concentration to 10-15 g/ml. The mixture was incubated at 4°C for 1 h. The mixture was layered onto a 30% sucrose cushion and centrifuged at 40,000 × g for 1 h. The supernatant was collected and stored at -70°C.

Preparation of Nuclear DNA Oligonucleotides - The nuclear extract was adsorbed to chromatin affinity chromatography. The extract was diluted with nuclei to reduce the nuclei concentration to 10-15 g/ml. The mixture was incubated at 4°C for 1 h. The mixture was layered onto a 30% sucrose cushion and centrifuged at 40,000 × g for 1 h. The supernatant was collected and stored at -70°C.
The 24-kDa NCL, isolated from the heparan-sulphate column, was diluted with 34 g/L of NaCl at pH 7.0 to reduce chromatic concentration, and the eluant precipitation obtained was removed by lyophilization as described above. The 24-kDa NCL (50 mg) was then dissolved in 25 μL of water specific ion-free H2O. [poly(dI-dC)]20pyridyl-ethyl, was subjected to chromatography. A 15 ml column of Pharmacia DEAE-Sepharose FF (column 2) was equilibrated with 100 mM sodium acetate buffer, pH 6.0, containing 0.1 M NaCl and 0.02% NaN3. The column was washed with 80 ml of the buffer, then the protein solution was added to the column and allowed to pass through the column at a flow rate of 30 ml/h. The column was washed with 80 ml of the buffer, then the protein solution was added to the column and allowed to pass through the column at a flow rate of 30 ml/h. The column was washed with 80 ml of the buffer, then the protein solution was added to the column and allowed to pass through the column at a flow rate of 30 ml/h. The column was washed with 80 ml of the buffer, then the protein solution was added to the column and allowed to pass through the column at a flow rate of 30 ml/h. The column was washed with 80 ml of the buffer, then the protein solution was added to the column and allowed to pass through the column at a flow rate of 30 ml/h. The column was washed with 80 ml of the buffer, then the protein solution was added to the column and allowed to pass through the column at a flow rate of 30 ml/h. The column was washed with 80 ml of the buffer, then the protein solution was added to the column and allowed to pass through the column at a flow rate of 30 ml/h. The column was washed with 80 ml of the buffer, then the protein solution was added to the column and allowed to pass through the column at a flow rate of 30 ml/h.
**α_2(I) Collagen Promoter Binding Factor**

**REFERENCES**

1. Okaruru, J., Nakatsuka, H., and de Crombrugghe, B. (1982) J. Biol. Chem. 257, 10666-10670
2. Wu, C. (1984) Nature 311, 81-84
3. Aronchick, J. L., and Peilly, T. J. (1991) J. Biol. Chem. 266, 10,148
4. MacLennan, T., Frisch, E. P., and Sammons, J. (eds.) (1982) in Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York
5. Schmidt, A., Sato, S., and de Crombrugghe, B. (1983) Nature 314, 288-290
6. Mandel, M., and Hoyer, A. (1970) J. Mol. Biol. 53, 115-142
7. Kahn, R., Holter, R. D., Thomas, C., Fodor, R., Mayer, A., Benacerraf, B., and Roskoski, R. (1980) Nucleic Acids Res. 8, 268-293
8. Masi, A. M., and Gilbert, W. (1980) Nucleic Acids Res. 8, 219-220
9. Garner, M. N., and Bevan, A. (1981) Nucleic Acids Res. 9, 2147-2160
10. Fried, M., and Crossley, D. M. (1981) Nucleic Acids Res. 9, 4053-4055
11. Singh, A., Sim, K., Baltimore, D., and Sharp, P. A. (1981) Nature 291, 75-78
12. Galan, S., and Schwartz, A. (1980) Nucleic Acids Res. 8, 3753-3770
13. Lammel, C. (1970) Nature 220, 640-645
14. Murazawa, J. H. (1976) Anal. Biochem. 67, 357-362
15. Clevelander, D. W., Fisher, S. G., Roskoski, R., and Lammel, C. (1977) J. Biol. Chem. 252, 1103-1106
16. Pipp, B., and Cole, R. B. (1985) J. Biol. Chem. 260, 15517-15522
17. Schuster, K., and von Jagow, G. (1987) Anal. Biochem. 166, 108-114
18. Bradford, M. M. (1981) Anal. Biochem. 72, 248-254

**TABLE 1** (Supp.) Purification of the factor which binds to the α2(I) collagen promoter

| Fraction | Total protein (μg) | Binding capacity (μg) | Purification yield (fold) | % |
|----------|-------------------|----------------------|--------------------------|---|
| Nucleus | 495 | 165 | 3.5 | 1/100 |
| 0.6 M NaCl extract | 45 | 50 | 32 | 9/82 |
| 2.0 M NaCl extract | 0.16 | 5 | 6280 | 1782/58 |

DNA binding capacity per unit of protein was determined using the autoradiography filter binding assay. The IgG1: IgG1 fragment from punished pA17709 was labeled using T_A polymerase kinase and used to detect binding.

**Fig 1** Schematic representation of the plasmids and fragments used.

**Fig 2** Gel retardation and filter binding assays were used to detect the binding activity. Panel A shows an autoradiogram of the gel retardation assay. The IgG1: IgG1 fragment from punished pA17709 was labeled at the IgG1 site using T_A polymerase kinase. The 5'-end labeled fragment was incubated in the absence [lane 2] or presence [lane 3] of rat liver nuclear extract. Panel B demonstrates the linearity of radioactivity retention in the nitrocellulose filter binding assay when increasing amounts of the 0.6 M NaCl eluate from the phosphorylase column were added. The IgG1: IgG1 fragment from punished pA17709 was labeled at the 5' end and used to detect binding.
Fig 1 Purification scheme for the factor that recognizes the putative NF-I recognition sequence in the mouse α11(1) collagen promoter.