*Proc. Natl. Acad. Sci. USA*
Vol. 92, pp. 1901–1905, March 1995
Biochemistry

# Isolation of active genes containing CAG repeats by DNA strand invasion by a peptide nucleic acid

**(chromatin/transcription factors/microsatellite DNA)**

LIDIA C. BOFFA*, ELISABETTA M. CARPANETO*, AND VINCENT G. ALLFREY†‡

*Istituto Nazionale per la Ricerca sul Cancro IST, Genoa 16132, Italy; and †The Rockefeller University, New York, NY 10021

*Communicated by Bruce Merrifield, The Rockefeller University, New York, NY, December 7, 1994*

**ABSTRACT** **An amplification of tandem CAG trinucleotide sequences in DNA due to errors in DNA replication is involved in at least four hereditary neurodegenerative diseases. The CAG triplet repeats when translated into protein give rise to tracts of glutamine residues, which are a prominent feature of many transcription factors, including the TATA-binding protein of transcription factor TFIID. We have used a biotin-labeled, complementary peptide nucleic acid (PNA) to invade the CAG repeats in intact chromatin and then employed a method for the selective isolation of transcriptionally active chromatin restriction fragments containing the PNA·DNA hybrids. The PNA-containing chromatin fragments were captured on streptavidin–agarose magnetic beads and shown to contain all the CAG·PNA hybrids of the active chromatin fraction. DNA hybridization experiments using a DNA probe specific for unique sequences downstream of the CAG-tandem repeats confirmed that the PNA·DNA hybrids contained the transcribed gene for the TATA-binding protein. In contrast, no hybridization signal was detected with a DNA probe specific for the c-*myc* protooncogene, which is amplified and transcriptionally active in COLO 320DM cells but lacks CAG tandem repeats.**

Peptide nucleic acids (PNAs) are structural homologues of DNA and RNA in which the entire phosphate–sugar backbone has been replaced by a flexible peptide backbone consisting of 2-aminoethyl glycine units linked to the purine and pyrimidine bases (1–3). The absence of phosphate groups in the PNA molecule facilitates its invasion of negatively charged DNA duplexes containing complementary base sequences (4–6). When PNA hybridizes to a targeted DNA sequence, one strand of DNA may be rapidly displaced to form a D-loop (1, 5, 6), while the PNA binds to its complementary DNA sequence by Watson–Crick base pairing (7). PNAs show greater discrimination and form more stable associations with DNA than the corresponding DNA·DNA duplexes (7). PNA invasion of DNA strands to form stable PNA·DNA hybrids has profound implications for both positive and negative control of gene activity. For example, it has been shown that DNA loops displaced as a consequence of PNA binding act as artificial transcription promoters (8), but PNA binding to the transcribed strand of simian virus 40 DNA blocks transcript elongation beyond the site of PNA·DNA duplex formation (9).

Here, we explore the potential of a biotinylated PNA probe to invade DNA triplet repeat sequences in the chromatin of transcriptionally active genes. By using a PNA specific for multiple CAG repeats, together with a method for the separation of transcriptionally active chromatin restriction fragments from inactive chromatin restriction fragments (10), we show that the PNA probe can hybridize effectively to CAG triplet repeats in intact chromatin and that transcriptionally active chromatin fragments containing the PNA·DNA hybrids can be captured quantitatively on streptavidin-coated magnetic beads.

An amplification of CAG triplet repeats occurs in at least four inherited neurodegenerative diseases. In all cases, the severity of the disease increases with the expansion of the CAG repeat region that occurs as a consequence of errors in DNA replication. In Huntington disease, expansion of the CAG repeats in the gene results in an elongation of a glutamine-rich domain in the encoded protein from 35 to over 100 residues (11). The functional consequence of similar glutamine expansions for transcriptional control is indicated by the finding that elongation of the polyglutamine domain of the androgen receptor diminishes transcription from androgen-responsive reporter genes (12, 13). Conversely, elimination of the CAG tract in human and rat androgen receptor (AR) genes results in an elevation of transcriptional activation (13).

Polyglutamine tracts containing more than 20 residues are prominent features of at least 33 transcription factors, including the human TATA-binding protein (TBP) (14–16), the homeobox protein Abdominal B of *Drosophila* (17), and the AR from several species (18–20). Recent molecular modeling studies by Perutz *et al.* (21) point out the potential of paired polyglutamine strands to act as polar zippers, possibly joining transcription regulatory factors on separate DNA segments.

Given the evidence that polyglutamine tracts play an important role in transcriptional control, the CAG microsatellite family was selected as a target for DNA strand invasion by a complementary PNA. The aim was to isolate the transcribed genes containing the CAG repeats from a human colonic cancer cell line.

## MATERIALS AND METHODS

**Preparation of Chromatin Restriction Fragments.** COLO 320DM cells, grown to a density of $4–7 \times 10^5$ cells per ml, were collected by centrifugation, and nuclei were isolated as described (10). Nuclear chromatin was cleaved *in situ* by a mixture of restriction endonucleases that do not attack CAG repeat sequences in DNA (1 unit of *Eco*RI and 1 unit of *Hin*dIII per μg of DNA) at 37°C for 1 h (10). Nuclease activity was inhibited by quick chilling samples to 0°C and adding $Na_2EDTA$ to a final concentration of 5.5 mM. The released chromatin restriction fragments, containing 14.8% ± 1.4% of the total nuclear DNA (average of nine experiments), were collected in the supernatant after centrifugation at $10,000 \times g$ for 20 min.

**Separation of Transcriptionally Active Chromatin Fragments.** The supernatant fraction was dialyzed for 2 h at 4°C against buffer A [10mM Tris·HCl, pH 7.4/25 mM NaCl/25 mM KCl/2% (wt/vol) sucrose/5 mM sodium butyrate/5 mM

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: DTNB, 5,5′-dithiobis(2-nitrobenzoic acid); DTT, dithiothreitol; EPNP, 1,2-epoxy-3-(*p*-nitrophenoxy)propane; PNA, peptide nucleic acid; TBP, TATA-binding protein; AR, androgen receptor.

‡To whom reprint requests should be addressed.

$Na_2EDTA$/0.1 mM phenylmethylsulfonyl fluoride/0.1 mM 1,2-epoxy-3-(*p*-nitrophenoxy)propane (EPNP; Kodak)], which had been bubbled with $N_2$. The supernatant was then added to a suspension of mercurated agarose magnetic beads, synthesized as described (10). The appropriate amount of beads was calculated from their SH-binding capacity (800–1000 nmol of SH groups per mg) and the reactive SH content of the chromatin digest, as determined by the 5,5′-dithiobis(2-nitrobenzoic acid (DTNB) reaction (22). A typical 15% digest contained 0.50 ± 0.05 nmol of reactive groups per μg of DNA. The SH-binding capacity of the added beads was twice the SH content of the released chromatin. The chromatin fragments were allowed to interact with the beads for at least 2 h at 4°C and the mercury-bound fragments were separated from the unbound fragments in a BioMag Separator (PerSeptive Diagnostics, Cambridge, MA). After washing the beads by repeated suspensions in buffer A and magnetic separation, the mercury-bound chromatin fragments were eluted in buffer A containing 20 mM dithiothreitol (DTT) (10).

**Detection of CAG Triplet Repeats in Chromatin Fragments.** DNA was isolated as described (23) from the nuclei, the supernatant fraction after endonuclease digestion, and the unbound and mercury-bound chromatin fragments. The DNA of each fraction was sized electrophoretically (24), its content of CAG repeat sequences was determined by slot–blot hybridization of a 21-mer oligonucleotide probe (CTGCTGCTGCTGCTGCTGCTG) to a 10 μg aliquot of DNA, and it was labeled at the 5′ end by using T4 polynucleotide kinase (New England Biolabs) and [$^{32}$P]ATP (3000 Ci/mmol; 1 Ci = 37 GBq; NEN) (23). The intensities of the hybridization signals were measured by densitometry.

**PNA Binding to Chromatin Restriction Fragments.** An antisense 21-mer PNA, CTGCTGCTGCTGCTGCTGCTG-$CONH_2$, for detecting strand invasion of CAG repeats in genomic DNA was linked at its amino terminus to biotin through two 8-amino-3,6-dioxaoctanoyl linkers to create a distance between the PNA and the biotin label (synthesized by Millipore). The purity of the biotinylated PNA was confirmed by reverse-phase chromatography and mass spectrographic analysis ($M_r$ 6200).

The biotinylated PNA probe was added to intact chromatin fractions or to DNA purified from those fractions in buffer B (5 mM Tris·HCl, pH 7.0/0.5 mM $Na_2EDTA$/0.1 mM phenylmethylsulfonyl fluoride/0.1 mM EPNP), at a concentration of 1 nmol of PNA per 400 μg of DNA (a great excess of PNA relative to the measured CAG repeat content of the DNA in all chromatin fractions). Hybridizations were carried out in buffer B for 1 h at 37°C. The biotinylated PNA·DNA complexes were separated from unbound PNA by centrifugation at 100,000 × *g* for 6 h in a sucrose density step gradient [5% over 10% (wt/vol) sucrose in buffer B].

The PNA content of the PNA·DNA hybrids was measured by titration of the biotin label with $^{35}$S-labeled streptavidin (651 Ci/mmol; Amersham) using 1.5 pmol of streptavidin per 100 μg of biotinylated PNA·DNA hybrid. The $^{35}$S-labeled streptavidin adducts of the biotinylated PNA·DNA complexes were separated from unbound streptavidin by centrifugation through the two-layer sucrose gradient. The PNA·DNA hybrids in the lower layer were collected for measurement of the stoichiometrically bound $^{35}$S-labeled streptavidin by scintillation spectrometry.

**Synthesis of Streptavidin–Magnetic Beads.** Magnetic agarose beads (2.5 mg) activated with *p*-nitrophenyl chloroformate (Advanced Magnetics, Cambridge, MA) were reacted with 11 mg of streptavidin in 0.1 N $NaHCO_3$/0.5 M NaCl, pH 8.0, for 24 h at 4°C. The streptavidin–agarose beads were washed extensively in the carbonate buffer, suspended in 0.2 M glycine buffer, pH 7.1, for 5 min to block any remaining active esters, and washed repeatedly before being suspended in buffer B. Under these conditions, 50% of the added streptavidin was bound to the beads, producing a biotin-binding capacity of 23 nmol per 100 μl ($10^7$) of beads, as measured by titration with *d*-[*carbonyl*-$^{14}$C]biotin (50 mCi/mmol; Amersham).

**Isolation and Electron Microscopy of Chromatin Fragments Containing Biotinylated PNA.** Chromatin restriction fragments (190 μg of DNA) were displaced from the mercurated agarose beads in buffer A containing 20 mM DTT, and the solution was concentrated to 0.4 of its original volume. The biotinylated PNA (0.5 nmol) was added to the chromatin fragments and hybridized for 1 h at 37°C. Then, without further purification, the mixture of PNA·DNA hybrids and unhybridized fragments was added to 5 μl of streptavidin–agarose beads and gently mixed for 1 h at 25°C. The beads were collected magnetically, washed, and prepared for electron microscopy as described (10).

**DNA Sequence Analyses of Biotinylated PNA·DNA Hybrids Captured on Magnetic Streptavidin Particles.** The mercury-bound chromatin fragments were released in buffer A containing 20 mM DTT, concentrated as described above, dialyzed against buffer B, and digested with 50 μg of RNase A per ml for 1 h at 37°C, followed by 100 μg of proteinase K per ml for 2 h at 37°C. Biotinylated PNA was added in the proportion 0.25 nmol of PNA per 100 μg of DNA. After 1 h at 37°C, the mixture was centrifuged through a 5%/10% sucrose step gradient to remove unhybridized PNA. The PNA·DNA hybrids of the dense layer were shaken for 1 h at 25°C with 200 μl of BioMag–streptavidin particles with a biotin-binding capacity of 2 μg (8.2 nmol) per ml of suspension (PerSeptive Diagnostics).

Under these conditions, all of the biotinylated PNA·DNA hybrids were bound to the streptavidin particles. The particles were recovered magnetically and washed twice with buffer B. The hybridized DNA was then released from the streptavidin-bound PNA by extraction in 200 μl of 0.4 N NaOH for 5 min at 25°C, followed by 10 min at 0°C. The extract was neutralized with 0.1 vol of 4 N HCl and 0.1 vol of 0.1 M Tris·HCl, pH 7.5/10 mM $Na_2EDTA$ was added. DNA (10 μg) from each sample was blotted onto a prewetted nylon membrane (Zeta-Probe; Bio-Rad) (23). The membrane-bound DNAs were hybridized to synthetic $^{32}$P-labeled deoxyribonucleotide probes. The probe for CAG repeat sequences was the 21-mer complementary sequence CTGCTGCTGCTGCTGCTGCTG. The probe for the TBP gene was the unique sequence ACTACTAAATTGTTGGTGGGTGAGCACA corresponding to nucleotides 1060–1087 of the human TBP cDNA (15). The probe for the AR gene was CTTTTGAAGAAGACCTT corresponding to the unique sequence 1894–1910 of the human AR cDNA (25). The $^{32}$P-labeled probe for c-*myc* was a nick-translated 9-kb *Eco*RI–*Hin*dIII fragment in plasmid pHSR1, which encodes the full human c-*myc* gene. The intensity of the hybridization signals was quantitated by densitometry.

## RESULTS

**PNA-Mediated Capture of Transcribed Genes Containing CAG Triplet Repeats.** A multistep procedure was developed for the isolation of transcriptionally active chromatin fragments containing multiple tandem CAG triplets. The steps in the method are as follows: (*i*) isolated nuclei are treated with restriction endonucleases to release ≈15% of the total nuclear DNA as chromatin restriction fragments; (*ii*) the transcriptionally active chromatin fragments are separated from inactive chromatin fragments by selective binding of their thiol-reactive nucleosomes to mercurated magnetic beads; (*iii*) the active chromatin fragments are released from the mercurated beads and hybridized to a biotinylated PNA probe for CAG triplet repeats; (*iv*) the resulting PNA·DNA hybrids are captured on streptavidin-coated magnetic particles; and (*v*) DNA is released from the streptavidin-bound PNA and assayed for

its CAG triplet content and for the presence of genes known to contain multiple tandem CAG repeats.

**Separation of Transcriptionally Active Chromatin Restriction Fragments.** The separation of active from inactive chromatin depends on the fact that nucleosomes along active genes unfold during transcription to reveal the previously shielded cysteinyl-SH groups of histone H3 molecules at the center of the nucleosome core (26). Consequently, a mixture of active and inactive nucleosomes can be separated by entrapping the SH-reactive nucleosomes on a mercurated agarose support. The mercury-bound nucleosomes (and larger chromatin fragments containing the SH-reactive nucleosomes) are recovered by displacement with DTT (27, 28). When mercury column chromatography was used to detect changes in nucleosome structure during oncogene activation, it was revealed that nucleosomes along an activated c-*fos* gene unfolded within minutes and reverted to the compact nonmercury-binding form within 15 min after transcription was suppressed (29–31).

A recent modification of this procedure permits the isolation of large chromatin fragments from active genes (10). The mixture of chromatin fragments released from cell nuclei by digestion with restriction endonucleases is fractionated by affinity chromatography on mercurated-agarose–magnetic beads. DNA hybridization experiments using $^{32}$P-labeled probes for genes known to be transcribed (or not transcribed) in the cells of origin have repeatedly shown that the mercury-bound chromatin fragments contain the transcribed sequences of the nucleus, as revealed by run-on transcription experiments (10). The technique has been shown to monitor, with accuracy, both the timing and extent of transcription of the c-*myc* oncogene and the *gas1* antioncogene in normal and malignant cells (10). Here, the same method is used to study the distribution of CAG triplet repeats in restriction fragments of a human cell line derived from a colonic adenocarcinoma.

COLO 320DM cell nuclei were isolated and digested with restriction nucleases *Eco*RI and *Hin*dIII (10). The mixture of released chromatin fragments, representing 14.8% ± 1.4% of the total nuclear DNA, was added to a suspension of mercurated magnetic beads. Chromatin fragments bound to the beads were recovered magnetically and released in 20 mM DTT. They contained 18.4% ± 1.2% of the DNA released by digestion with *Eco*RI and *Hin*dIII or 2.8% ± 0.4% of the total nuclear DNA. Protein analyses revealed a full complement of core histones (H2A, H2B, H3, and H4) in stoichiometric proportions, thus confirming the presence of intact nucleosome cores in the transcribed chromatin fraction (10). Arrays of nucleosome-sized particles were also evident in electron micrographs of chromatin fragments attached to the beads (Fig. 1 *A* and *B*).

DNA was isolated from the mercury-bound chromatin and sized electrophoretically (24). It contained DNA restriction fragments ranging in size from 4 to 23 kb, with most of the DNA present in fragments with sizes between 9 and 23 kb. A similar size distribution was observed in the unbound chromatin fraction and the total nuclear digest (Fig. 2).

**Detection of CAG Repeats in Chromatin Restriction Fragments.** DNA isolated from intact nuclei and from each of the chromatin fractions was analyzed for its content of CAG-triplet repeats by slot–blot hybridization to a $^{32}$P-labeled antisense 21-mer oligonucleotide probe (CTGCTGCTGCTGCTGCTGCTG). From the intensities of the hybridization signals (Fig. 3, lane A) as measured by densitometry and the distribution of DNA in the various fractions, it was calculated that about 2% of the nuclear CAG triplets were released during treatment with *Eco*RI and *Hin*dIII. The mercury-bound restriction fragments, which represent 2.8% of the total nuclear DNA, contained 1.12% of the total CAG repeats of the nucleus. The much larger unbound chromatin fraction, containing 12% of the total nuclear DNA, had only 0.72% of the nuclear CAG triplet repeats. It follows that the great majority of the CAG microsatellite sequences remained in the nucleus and were not released or were not accessible to this combination of restriction enzymes, possibly because of their localization in highly condensed heterochromatic regions of the nucleus.

To characterize the lengths of DNA restriction fragments containing the tandem CAG triplets, the bands in the DNA-sizing gels were transferred to nitrocellulose filters and hybridized to the $^{32}$P-labeled 21-mer CTG-oligonucleotide probe. The results show that the CAG repeat sequences are broadly distributed throughout the DNA fragments in all the chromatin fractions but most of them appear in the large restriction fragments (Fig. 2).

**DNA Strand Invasion by a CAG-Specific PNA.** A biotinylated antisense 21-mer PNA (CTGCTGCTGCTGCTGCTGCTG-$CONH_2$) for detecting and measuring PNA invasion of CAG triplet repeats in genomic DNA was added in great excess over the measured CAG content of the restriction fragments. PNA interactions with restriction fragments were carried out in a low ionic strength neutral buffer for 1 h at 37°C, and the biotinylated PNA·DNA complexes were separated from free PNA by sucrose density gradient centrifugation. The PNA content of the PNA·DNA hybrids was measured by titration of the biotin label with $^{35}$S-labeled streptavidin. The high sensitivity of this isotopic technique made it possible to measure pmol quantities of the PNA in a rapid and reproducible manner.

The amount of PNA hybridized to the chromatin restriction fragments was expressed in pmol of PNA per μg of DNA and compared to the PNA content of the total nuclear digest (9.1 ± 0.2 pmol of PNA per μg of DNA). The mercury-bound chromatin contained $1.12 \pm 0.02 \times 10^{-3}$ pmol of PNA per μg of DNA, while the unbound chromatin fragments contained $0.325 \pm 0.25 \times 10^{-3}$ pmol of PNA per μg of DNA (average of five experiments). From the PNA·DNA ratios in the total nuclear digest and in the mercury-bound chromatin fraction (which represents only 2.8% of the total nuclear DNA), it was estimated that the mercury-bound chromatin fragments con-

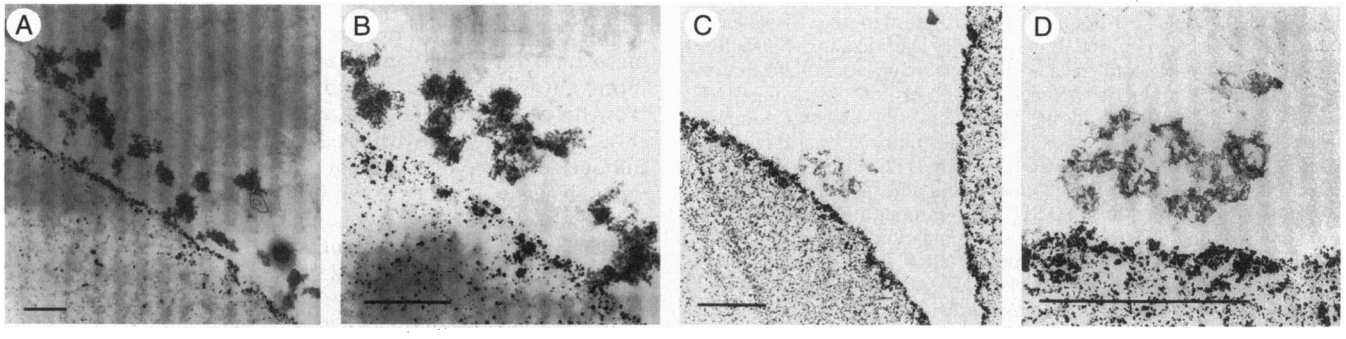


FIG. 1. Electron microscopic images of chromatin restriction fragments attached to mercurated agarose–magnetic beads (*A* and *B*) and to streptavidin-linked agarose magnetic beads (*C* and *D*). (Bar indicates 25 nm in *A*, *B*, and *C* and 45 nm in *D*.)

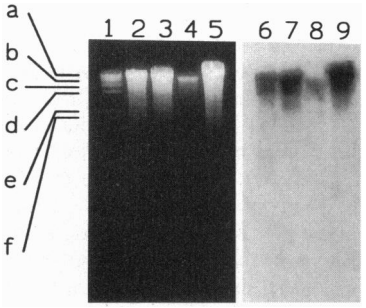


FIG. 2. DNA-sizing and detection of CAG repeat sequences in nuclear fractions. DNA was isolated from each fraction and sized electrophoretically, as described (24). Lane 1, λ phage DNA size standards; a = 23,130 bp, b = 9416 bp, c = 6557 bp, d = 4361 bp, e = 2322 bp; lane 2, DNA of mercury-bound chromatin fragments; lane 3, DNA of unbound chromatin fragments; lane 4, restriction enzyme digest supernatant; and lane 5, total nuclear restriction enzyme digest. Lanes 6–9, hybridization of DNA transferred from lanes 2 to 5 to the $^{32}$P-labeled 21-mer oligonucleotide probe for CAG repeat sequences.

tained 0.345% of the CAG repeat sequences of the nucleus. This estimate, which is a measure of PNA strand invasion of CAG repeats in transcriptionally active chromatin, is considerably lower than the 1.12% estimate derived on the basis of the hybridization of the $^{32}$P-labeled 21-mer CTG probe to purified DNA fractions (Fig. 3, lane A). The difference indicates that most of the CAG repeat sequences that were accessible in isolated DNA were not equally accessible when the DNA was associated with nucleosomal proteins. However, the important point is that PNA can invade DNA strands in intact chromatin. Moreover, the PNA·DNA hybrid content of chromatin fragments is sufficient to permit the isolation of the fragments containing multiple CAG repeats.

**Capture and Visualization of Chromatin Restriction Fragments Containing PNA·DNA Complexes.** The transcriptionally active chromatin fragments were separated from inactive

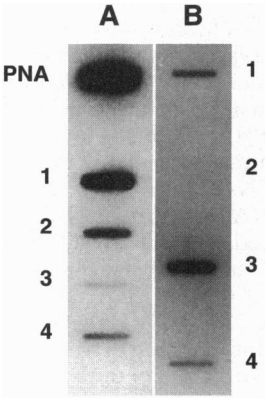


FIG. 3. Distribution of CAG repeat sequences in chromatin fractions as determined by slot–blot hybridization to the $^{32}$P-labeled antisense 21-mer CTG probe. Lane A shows hybridization signals for 10 μg of DNA from total nuclei, slot 1; mercury-bound chromatin fragments, slot 2; unbound chromatin fragments, slot 3; and digested nuclei before fractionation, slot 4. The intense slot–blot signal at the head of lane A (PNA) shows the hybridization of the $^{32}$P-labeled 21-mer CTG probe to 10 μg of the 21-mer PNA. Lane B shows detection of CAG repeats in mercury-bound and streptavidin-bound chromatin fragments. Slot 1, DNA from chromatin fragments captured on mercurated-agarose beads; slot 2, unbound DNA of the same fragments after treatment with biotinylated PNA and exposure to streptavidin-coated magnetic particles; slot 3, DNA hybridized to the biotinylated PNA and captured on streptavidin particles; and slot 4, total nuclear DNA. Note that all of the CAG repeats of the mercury-bound chromatin fragments were recovered as PNA·DNA hybrids (compare slots 2 and 3).

fragments by the mercurated agarose bead technique (10). The mercury-bound chromatin fragments (Fig. 1 *A* and *B*) were then released in 20 mM DTT and added to a suspension of magnetic agarose beads linked to streptavidin. The streptavidin–beads were captured in a magnetic field, and sectioned for electron microscopy (10).

The micrographs show that the biotinylated restriction fragments linked to the surface of the streptavidin–beads are composed of nucleosome-sized subunits (Fig. 1 *C* and *D*), but, in contrast to the large number of restriction fragments bound to the mercurated agarose beads (Fig. 1 *A* and *B*) and earlier EM studies (10), there are very few biotinylated fragments captured on the streptavidin–beads. This result is in accord with quantitative estimates (described below) that the streptavidin–bound fragments contained only 0.86% ± 0.07% of the DNA originally present in the mercury-bound chromatin fraction, and it provides visual confirmation of the specificity of the interactions between the biotinylated PNA and CAG repeat sequences in the active chromatin fraction.

**Isolation of DNA Sequences Containing CAG Triplet Repeats.** As noted above, the accessibility of the CAG repeats in chromatin is considerably lower than that observed in the purified DNA of the same chromatin fraction. Therefore, to improve the recovery of PNA·DNA complexes containing CAG repeats, DNA was isolated from the transcriptionally active chromatin fractions before adding the biotinylated PNA probe. After hybridization, the PNA·DNA hybrids were separated from free PNA by density gradient centrifugation and reacted with the streptavidin-coated magnetic particles. Under these conditions, all of the PNA·DNA hybrids were bound to the particles. The hybrids were found to contain 0.86% ± 0.07% of the DNA originally present in the mercury-bound chromatin fraction, or 0.0197% ± 0.0003% of the total nuclear DNA. The content of DNA in the streptavidin-bound complexes was also determined in control experiments by using $^{32}$P-labeled DNA (data not shown).

The DNA was released from its streptavidin–PNA complex and assayed for its content of CAG-tandem repeats by slot–blot hybridization to the $^{32}$P-labeled 21-mer CTG probe. The intensity of the hybridization signals, as assessed by densitometry (Fig. 3, lane B), revealed that the CAG repeats were at least 40 times more concentrated in the streptavidin-bound PNA·DNA hybrids than in the DNA of the mercury-bound chromatin fragments.

The CAG-enriched chromatin fragments would be expected to contain the DNA sequences of the genes encoding transcription factors with polyglutamine tracts. To confirm that the PNA had invaded those sequences, the DNA released from the streptavidin–PNA complex was hybridized to a $^{32}$P-labeled probe for a unique sequence downstream of the CAG repeats of the TBP gene. The results (Fig. 4, lane A) show an intense signal in that DNA fraction but little or no hybridization to chromatin fragments that failed to hybridize with the PNA probe.

The successful capture of the TBP gene in the PNA·DNA hybrids of the transcriptionally active chromatin fragments confirms the expectation that a gene essential for transcription in all cells would be present in the diffuse chromatin and accessible for strand invasion by PNA. To test whether a nontranscribed gene would have been recovered as a PNA·DNA hybrid, the DNA released from the streptavidin–PNA complex was hybridized to a $^{32}$P-labeled probe for a unique sequence downstream of the CAG repeats in the AR gene. No signal was detected in that fraction (Fig. 4, lane B), but few, if any, AR gene sequences were released from COLO 320DM nuclei after treatment with *Eco*RI and *Hin*dIII, indicating that the AR gene was located in the compactly organized chromatin.

To test whether the PNA·DNA hybrids captured on streptavidin–beads include genes that are actively transcribed in COLO 320DM cells but lack CAG tandem repeats, DNA from each

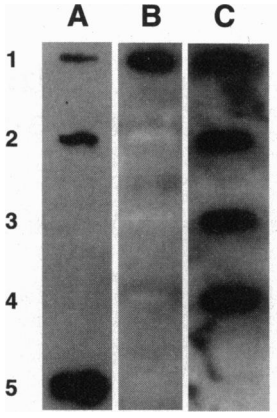


FIG. 4. Contrasting distribution of transcribed and nontranscribed genes in chromatin fractions after hybridization to the PNA probe for CAG tandem repeats. DNA (10 μg) isolated from the total nuclear restriction enzyme digest and from each chromatin fraction was hybridized to $^{32}$P-labeled probes for unique sequences of the TBP gene (lane A) and the AR gene (lane B). Hybridizations to full-length c-*myc* DNA are shown in lane C. Slot 1, total nuclear DNA; slot 2, DNA of chromatin fragments released by endonuclease digestion; slot 3, DNA of the chromatin fragments not bound to mercurated agarose beads; slot 4, DNA of mercury-bound chromatin that did not bind to streptavidin–beads after treatment with biotinylated PNA; and slot 5, DNA of mercury-bound chromatin that did bind to streptavidin–beads after hybridization to the biotinylated PNA probe.

chromatin fraction was hybridized to a $^{32}$P-labeled full-length c-*myc* DNA probe. The results (Fig. 4, lane C) show that, although c-*myc* DNA was recovered and released from the mercury-bound chromatin, no c-*myc* DNA was invaded by the biotinylated PNA probe and captured on streptavidin–beads.

## DISCUSSION

The specific targeting of CAG triplet repeats by a complementary PNA was selected as a model for the isolation of genes encoding a broad class of transcription factors containing polyglutamine-rich domains. The results have established that PNA can invade chromatin *in situ* and combine specifically with the complementary CAG repeats of a DNA sequence, as demonstrated for the TBP gene. In principle, the same techniques employed in these experiments should be applicable to any gene for which a biotinylated PNA probe can be synthesized. This would provide a direct route to the isolation and comparison of genes in their native and mutant forms in normal and malignant cells.

DNA strand invasion by a complementary PNA leads to stable PNA·DNA adducts in nucleosome-containing regions of the targeted sequence, as revealed by electron microscopy and histone content. The capture of nucleosomes bearing the biotinylated PNA·DNA complexes on streptavidin–magnetic beads provides direct access to the structural and regulatory proteins that are closely associated with a defined DNA sequence. When combined with methods such as mercury-affinity chromatography that can separate the transcriptionally active and inactive states of chromatin, PNA-mediated capture of a specific gene in its expressed or silent state may provide new insights into the microdomains of transcriptional control.

We thank Joseph Kim, Marjorie Louissaint, Maria R. Mariani, and Emmanuel Valentin for expert technical assistance, and Eleana Sphicas of the Rockefeller University Electron Microscopy Service for her perceptive advice and assistance. We thank Dr. Alexander Hoffman for the gift of a full-length cDNA probe for the human TBP gene and Dr. Shutsung Liao for the cDNA probe for the human AR gene. This work was supported by Grant CA14908 from the National Institutes of Health (to V.G.A.) and by Associazione Italiana per la Ricerca sul Cancro Grant 1993/1994 (to L.C.B.). E.M.C. is a recipient of an Associazione Italiana per la Ricerca sul Cancro fellowship.

1. Nielsen, P. E., Egholm, M., Berg, R. H. & Buchardt, O. (1991) *Science* **254,** 1497–1500.
2. Egholm, M., Buchardt, O., Nielsen, P. E. & Berg, R. H. (1992) *J. Am. Chem. Soc.* **114,** 1895–1897.
3. Egholm, M., Nielsen, P. E., Buchardt, O. & Berg, R. H. (1992) *J. Am. Chem. Soc.* **114,** 9677–9678.
4. Nielsen, P. E., Egholm, M., Berg, R. H. & Buchardt, O. (1993) *Nucleic Acids Res.* **21,** 197–200.
5. Peffer, N. J., Hanvey, J. C., Bisi, J. E., Thomson, S. A., Hassman, C. F., Noble, S. A. & Babiss, L. E. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 10648–10652.
6. Cherny, D. Y., Belotserkovskii, B. P., Frank-Kamenetskii, M. D., Egholm, M., Buchardt, O., Berg, R. H. & Nielsen, P. E. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 1667–1670.
7. Egholm, M., Buchardt, O., Christensen, L., Behrens, C., Freier, S. M., Driver, D. A., Berg, R. H., Kim, S. K., Norden, B. & Nielsen, P. E. (1993) *Nature (London)* **365,** 566–568.
8. Mollegaard, N., Buchardt, O., Egholm, M. & Nielsen, P. E. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 3892–3895.
9. Hanvey, J. C., Peffer, N. J., Bisi, J. E., Thomson, S. A., Cadilla, R., Josey, J. A., Ricca, D. J., Hassman, C. F., Bonham, M. A., Au, K. G., Carter, S. G., Bruckenstein, D. A., Boyd, A. L., Noble, S. A. & Babiss, L. E. (1992) *Science* **258,** 1481–1485.
10. Chen-Cleland, T. A., Boffa, L. C., Carpaneto, E. M., Mariani, M. R., Valentin, E., Mendez, E. & Allfrey, V. G. (1993) *J. Biol. Chem.* **268,** 23409–23416.
11. The Huntington's Disease Collaborative Research Group (1993) *Cell* **72,** 971–983.
12. Mahtre, A. N., Trifiro, M. A., Kaufman, M., Kazemi-Esfarani, P., Figlewicz, D., Rouleau, G. & Pinsky, L. (1993) *Nat. Genet.* **5,** 184–188.
13. Chamberlain, N. L., Driver, E. D. & Miesfeld, R. L. (1994) *Nucleic Acids Res.* **22,** 3182–3186.
14. Hoffman, A., Sinn, E., Yamamoto, T., Wang, J., Roy, A., Horikoshi, M. & Roeder, R. G. (1990) *Nature (London)* **346,** 387–390.
15. Peterson, M. G., Tanese, N., Pugh, B. F. & Tjian, R. (1990) *Science* **248,** 1625–1630.
16. Zhou, Q., Boyer, T. G. & Berk, J. (1993) *Genes Dev.* **7,** 180–187.
17. Ali, N. & Bienz, M. (1991) *Mech. Dev.* **35,** 55–64.
18. Chang, C., Kokontis, J. & Liao, S. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 7211–7215.
19. Tan, J. A., Joseph, D. R., Quarmby, V. E., Lubahn, D. B., Sar, M., French, F. S. & Wilson, E. M. (1988) *Mol. Endocrinol.* **2,** 1276–1285.
20. Gaspar, M. L., Meo, T. & Tosi, M. (1990) *Mol. Endocrinol.* **4,** 1600–1610.
21. Perutz, M. F., Johnson, T., Suzuki, M. & Finch, J. T. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 5355–5358.
22. Soper, T. S., Jones, W. M. & Manning, J. M. (1979) *J. Biol. Chem.* **254,** 10901–10905.
23. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
24. Maniatis, T., Jeffrey, A. & Van Desande, H. (1975) *Biochemistry* **14,** 3787–3794.
25. Tilley, W. D., Marcelli, M., Wilson, J. D. & McPhaul, M. J. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 327–331.
26. Prior, C. P., Cantor, C. R., Johnson, E. M., Littau, V. C. & Allfrey, V. G. (1983) *Cell* **34,** 1033–1042.
27. Walker, J., Chen, T. A., Sterner, R., Berger, M., Winston, F. & Allfrey, V. G. (1990) *J. Biol. Chem.* **265,** 5736–5746.
28. Allfrey, V. G. & Chen, T. A. (1991) *Methods Cell Biol.* **35,** 315–335.
29. Chen, T. A. & Allfrey, V. G. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 5252–5256.
30. Chen, T. A., Sterner, R., Cozzolino, A. & Allfrey, V. G. (1990) *J. Mol. Biol.* **212,** 481–493.
31. Boffa, L. C., Walker, J., Chen, T. A., Sterner, R., Mariani, M. R. & Allfrey, V. G. (1990) *Eur. J. Biochem.* **194,** 811–823.