Characterization of a multisubunit human protein which selectively binds single stranded d(GA)$_n$ and d(GT)$_n$ sequence repeats in DNA

Anat Aharoni, Nava Baran and Haim Manor*
Department of Biology, Technion-Israel Institute of Technology, Haifa 32,000, Israel

Received June 11, 1993; Revised and Accepted September 21, 1993

ABSTRACT
A protein which selectively binds d(GA)$_n$ and d(GT)$_n$ sequence repeats in single stranded DNA has been identified in human fibroblasts. This protein, designated PGB, has been purified at least 500-fold by ammonium sulfate precipitation followed by DEAE-Sepharose column chromatography and affinity chromatography in a column of d(GA)-Sepharose. Electrophoretic mobility shift assays revealed that the PGB protein bound most avidly d(GA)$_n$ and d(GT)$_n$ tracts of n > 5. It also bound other G-rich DNA sequence repeats, including dG$_n$ tracts, with lower affinities. It did not manifest significant binding affinities to single stranded M13 DNA, or to the homopolynucleotides poly dA, poly dC and poly dT, or to various DNA sequence repeats which do not contain G residues, such as d(A-C)$_n$ and d(TC)$_n$. It did not bind double stranded d(TC)$_n$-d(GA)$_n$ tracts or other double stranded DNA sequences. In glycerol gradient centrifugation assays the d(GA)$_n$- and the d(GT)$_n$-binding activities cosedimented as a homogeneous protein species having an $S_{20,w}$ = 9.4 ± 0.7 and an estimated native molecular weight of 190,000 ± 7,000. UV crosslinking assays revealed that the protein contains 33.6 ± 2.1 kd subunits which bind d(GA)$_n$ and d(GT)$_n$. However, SDS-polyacrylamide gel electrophoresis of the purified protein followed by silver staining indicated that it may also contain other subunits that do not contact the DNA. It is proposed that binding of the PGB protein to single stranded d(GA)$_n$ or d(GT)$_n$ tracts in double stranded topologically restricted DNA may stimulate strand separation and formation of triple helices or other unusual DNA structures.

INTRODUCTION
One class of eukaryotic repetitive DNA sequences designated 'simple sequence' DNA consists of multiple repeats of short tracts of bases. One group of 'simple sequence' DNA elements comprises polypurines.polypyrimidines which have mirror symmetry and may undergo transitions into structures containing DNA triple helices (1-16; for reviews see ref. 17-19). This group includes the alternating polypurine.polypyrimidine repeats d(TC)$_n$-d(GA)$_n$ which are highly dispersed in mammalian genomes (20).

Previous studies have indicated that d(TC)$_n$-d(GA)$_n$ repeats may cause arrests of DNA replication and amplification in mammalian cells. This observation was initially made in a line of polyomavirus (Py) transformed rat cells wherein multiple rounds of replication of the chromosomally associated viral DNA can be induced by exposing the cells to carcinogens (21). The multifork replication was found to proceed into the flanking cell DNA and to be arrested at, or next to a d(TC)$_{27}$-d(GA)$_{27}$ tract (22). Other in vivo and in vitro studies have subsequently provided further support for the role of d(TC)$_n$-d(GA)$_n$ repeats as replication arrest signals (23-25). We have suggested that this biological function of d(TC)$_n$-d(GA)$_n$ repeats may be related to their ability to undergo transitions into structures including d(TC)$_n$-d(GA)$_n$-d(TC)$_n$ triple helices + single stranded d(GA)$_n$ sequences (6-10), or d(GA)$_n$-d(GA)$_n$-d(TC)$_n$ triple helices + single stranded d(TC)$_n$ sequences (13-16), and data supporting this hypothesis have been obtained (25-27). However, it appeared plausible that in living cells double stranded d(TC)$_n$-d(GA)$_n$ repeats, or the unusual structures that these sequences may form, could be associated with proteins and that such prns might play a role in regulating the arrests of DNA replication. Hence, we have undertaken a search for mammalian nuclear proteins which bind these sequence repeats in their various forms.

This search has been carried out in human fibroblasts obtained from foreskins and grown on microcarrier beads. We have found that nuclear extracts prepared from these cells contained a protein that bound double stranded d(TC)$_n$-d(GA)$_n$ tracts and another protein that selectively bound the separated d(GA)$_n$ strand. Here we report the purification and initial characterization of the latter protein. This protein has been designated PGB because in addition to d(GA)$_n$ tracts, it was found to bind d(GT)$_n$ tracts and, with lower affinities, other G-rich single stranded DNA sequence repeats.

We also present in this article details of a novel method which has been developed for of the PGB protein, and which may be adopted as a general procedure for extraction of other nuclear proteins from cells attached to microcarrier beads.
MATERIALS AND METHODS

Preparation of nuclear extracts from cells attached to microcarrier beads

All steps of this procedure were performed in a 4°C cold room. In the first step, a suspension of 1,000 g of microcarrier beads carrying human fibroblasts (obtained from Interpharm Laboratories, Nes Ziona, Israel) was poured on a Whatman No. 3 filter paper (diameter 24 cm) placed in a bürchner funnel that was inserted into a suction bottle. A moderate suction was applied with a vacuum pump to wash the medium. The beads were subsequently washed twice with 1000 ml of phosphate buffered saline (PBS: 137 mM NaCl, 2.7 mM KCl, 8 mM Na2HPO4, 1.5 mM KH2PO4). The filter with the beads was placed in a plastic container and the beads were resuspended in 750 ml of buffer A and frozen at −75°C (buffer A: 10 mM N-[2-Hydroxyethyl]piperazine-N’-[2-ethanesulfonic acid] (HEPES) pH 8.0, 50 mM NaCl, 0.50 M ethylenediaminetetraacetic acid (EDTA), 0.50 mM spermidine, 0.15 mM spermine, 0.50% triton, 7.0 mM β-mercaptoethanol, 30% glycerol and 1 mM phenylmethylsulphonyl fluoride (PMSF; added shortly before use)). This suspension could be stored in the −75°C freezer for at least 18 months without loss of activity of the PGB protein.

To prepare the nuclear extracts, the frozen suspension containing 1,000 g beads carrying the attached cells (2 mg protein/g beads) was thawed at 4°C and was poured on a Whatman No. 3 filter paper. After sucking the buffer the beads carrying the cells were washed twice with 500 ml of buffer A’ (buffer A’: 10 mM HEPES pH 8.0, 50 mM NaCl, 0.50 M sucrose, 1.0 mM ethylenediaminetetraacetic acid (EDTA), 0.50 mM spermidine, 0.15 mM spermine, 0.50% triton, 7.0 mM β-mercaptoethanol, 30% glycerol and 1 mM phenylmethylsulphonyl fluoride (PMSF; added shortly before use)). This suspension could be stored in the −75°C freezer for at least 18 months without loss of activity of the PGB protein.

Purification of the PGB protein

(i) DEAE-Sepharose column chromatography. The crude nuclear extract obtained by the procedure described above contained 8–10 mg/ml of protein. Thirty ml of this protein solution was loaded on a 20 ml column of DEAE-Sepharose Fast Flow column (Sigma) (length 11.5 cm; diameter 1.5 cm) that had been previously equilibrated with 180 ml of the buffer T-0 (buffer T-0: 20 mM Tris-HCl pH 8.0, 50 mM NaCl, 15% glycerol, 1 mM EDTA, 0.50 mM spermidine, 0.15 mM spermine, 7.0 mM β-mercaptoethanol and 1 mM PMSF). The column was developed with a 140 ml linear gradient of NaCl (50 mM–400 mM) in buffer T-0. The PGB d(GA)25-binding activity was eluted as a broad peak centered at 200 mM NaCl. Fractions containing this activity were pooled and precipitated by addition of ammonium sulfate to a final concentration of 60%, as described above. The precipitated protein was dissolved in 12 ml of T-50 buffer and dialysed against 2,000 ml of TEG buffer (TEG buffer: 10 mM Tris-HCl pH 8.0, 0.10 mM EDTA, 2 mM dithiothreitol (DTT), 15% glycerol and 1 mM PMSF).

(ii) Affinity chromatography on d(GA)25-Sepharose columns. 150 µg of d(GA)25 were bound to 2.5 ml of packed Cyanogen Bromide-activated Sepharose 4B, as described by Kadonaga and Tjian (28). The d(GA)25-bound Sepharose was used to prepare a column (length 4.5 cm; diameter 0.85 cm). This column was preequilibrated with 50 ml of buffer AC1 (buffer AC1: 10 mM Tris–HCl pH 8.0, 10 mM NaCl, 0.10 mM EDTA, 2 mM DTT and 15% glycerol). The dialysed protein (5 mg/ml) recovered from the DEAE-Sepharose column was mixed with 25 µg/ml of E. coli DNA and the mixture was loaded on the affinity column. The solution collected from the column was passed through the column four more times. The column was subsequently washed with 10 ml of buffer AC2 (buffer AC2: 20 mM Tris–HCl pH 7.0, 125 mM NaCl, 5 mM MgCl2, 1 mM EDTA, 2 mM DTT, 15% glycerol, 0.10% NP40 and 0.25 mg/ml bovine serum albumin (BSA)). Finally, the protein was eluted with 5 ml of buffer AC3 (buffer AC3: 20 mM Tris-HCl pH 7.0, 300 mM NaCl, 5 mM MgCl2, 1 mM EDTA, 2 mM DTT, 15% glycerol, 0.10% NP40 and 0.25 mg/ml BSA). 1.0 ml PMSF was added to the eluate and the d(GA)25-binding activity was assayed as described below. Some of the affinity purified protein preparations were dialysed against TEG buffer and subjected to a second affinity chromatography in a similar column. The purified protein was frozen and stored at −75°C.
Electrophoretic mobility shift assays

Single stranded oligonucleotides were labeled at their 5'-ends using [γ-32P]-ATP and T4 polynucleotide kinase, as described (29). The labeled oligonucleotides were purified by the spun column procedure using sephadex G-50 (ibid.). Binding of the labeled oligonucleotides to the various protein fractions was carried out in 20 μl of 10 mM HEPES pH = 7.8, 100 mM NaCl, 0.10 mM EDTA, 10 mM MgCl2, 2 mM DTT and 17.5% glycerol, in the absence or in the presence of unlabeled competitor DNAs. E. coli DNA (50 μg/ml) was also added to some of the binding mixtures, as specified in the legends to the figures. The mixtures were incubated at 23°C for 20 min. and were subsequently loaded on 6% polyacrylamide gels prepared in a buffer containing 25 mM trizma base, 190 mM glycine and 1 mM EDTA. Electrophoresis was performed in the same buffer at 200 V for 1 hr at 25°C. The gels were then dried and autoradiographed.

UV crosslinking assays

The procedure used for these assays was similar to that described by Chodosh et al. (30). Briefly, oligonucleotides labeled with 32P at their 5' ends were incubated with the protein in 1.5 ml round-bottom plastic vials under the conditions described in the previous section. Subsequently, the mixtures were placed in ice and irradiated for the specified time periods with a transilluminator containing four 15 W UV lamps (UVP model C-53 emitting 254 nm UV light) that was placed at a distance of 6 cm above the samples. The mixtures were next heated and loaded on a 12% SDS-polyacrylamide gel and electrophoresed as described by Laemmli (31). The gels were subsequently dried and autoradiographed.

Glycerol gradient centrifugation and estimation of the molecular weight of the native PGB protein

The protein was loaded on 4.0 ml linear 20%–40% (V/V) glycerol gels that also contained 20 mM Hepes pH 8.0, 100 mM NaCl, 1 mM EDTA, 2 mM DTT and 1 mM PMSF. The tubes containing the gradients were centrifuged in an SW60 rotor at 52,000 RPM at 8°C for 17 hrs in a Sorvall ultracentrifuge. 42–43 fractions were collected through the bottom of each gradient after piercing the tube. The distributions of the d(GA)2 and the d(GT)2-binding activities were determined in these fractions by electrophoretic mobility shift assays using d(GA)12 and d(GT)10 probes. An estimation of the sedimentation constant and the molecular weight of the native PGB protein was obtained by a comparison of the sedimentation velocity of these binding activities to the sedimentation velocities of known protein markers, as described by Martin and Ames (32). The protein markers used for the molecular weight determinations were: bovine liver catalase with an S20,w = 11.15 and a molecular weight of 232 kd, sweet potato β-amylase with an S20,w = 8.90 and a molecular weight of 200 kd and bovine serum albumin with an S20,w = 4.4 and a molecular weight of 66 kd.

DNA fragments and synthetic oligonucleotides

Oligonucleotides were synthesized in an Applied Biosystems model 381A DNA synthesizer. Polynucleotides were bought from Böhringer, New England Biolabs, Sigma and Pharmacia. DNA fragments were prepared by cleavage of plasmids with restriction enzymes and purified by polyacrylamide gel electrophoresis, as described (29). Single stranded bacteriophage M13 DNA was obtained from US Biochemicals and was cleaved to fragments of size ≤500 nucleotides by boiling the DNA 10 min. in 0.50N NaOH.

RESULTS

Preparation of nuclear extracts from cells attached to microcarrier beads and identification of d(GA)n-binding activity in these extracts

In order to prepare nuclear extracts from the human fibroblasts by a conventional method (33) we tried to remove and physically separate the cells from the beads in which they were grown. For this purpose we incubated the beads in buffers containing either trypsin or ethylenediaminetetraacetic acid (EDTA), two agents that are normally employed for removal of cells from plastic or glass surfaces; then, we tried to separate the cells from the beads by differential centrifugation. These attempts failed because the mass ratio beads/cells was very high and because the beads

Figure 1. Analysis of the d(GA)n-binding activity and of the polypeptide constituents recovered after various steps in the purification of the PGB protein. Nuclear extracts were prepared from human fibroblasts and the PGB protein was purified from these extracts as described in Materials and Methods. a. Electrophoretic mobility shift assays of the protein preparations obtained after the various steps of purification. These assays were performed with a 32P-labeled d(GA)12 probe. The specific radioactivity of the probe was 4×106 dpm/μg. In each assay about 50 ng/ml of the probe were incubated with a sample of the protein fraction and with 50 μg/ml of E. coli DNA as well as the other components specified in Materials and Methods, and was subsequently electrophoresed in a 6% polyacrylamide gel. The gel was dried and autoradiographed. The letters at the top of the lanes designate the protein fractions used for the assays: N — no protein. C – The (redissolved) 60% ammonium sulfate precipitate D – The protein purified by DEAE-sepharose column chromatography. A1 – The protein purified once by d(GA)25-sepharose affinity column chromatography. A2 – The protein purified a second time by d(GA)25-sepharose affinity column chromatography. b. SDS-polyacrylamide gel electrophoresis of the protein preparations obtained after the various steps of purification. Samples of the same protein preparations that were assayed in Fig. 1a were electrophoresed in an SDS-polyacrylamide gel (31). The gel was subsequently silver-stained as described (48).
swelled and did not sediment much faster than the cells. Therefore, we developed an alternative procedure for preparation of nuclear extracts from cells that were still attached to the beads. In this procedure, which is described in detail in the section on Materials and Methods, filtration rather than centrifugation was used to separate the nuclei from the cytoplasm and to prepare the nuclear extracts. Samples of these extracts were assayed for various DNA binding activities using the electrophoretic mobility shift technique (34,35). Fig. 1a (lane C) shows an assay of such a nuclear extract that was carried out with a 32P-labeled d(GA)12 probe. It can be seen that one major retarded band and one minor retarded band were generated in this assay indicating the presence of a d(GA)ₙ-binding activity in this extract. To further characterize this activity, we proceeded to purify it from the extract.

**Purification of the PGB protein**

The d(GA)ₙ-binding activity was purified by DEAE-Sepharose column chromatography followed by affinity chromatography in a column of the single stranded oligonucleotide d(GA)₂₅ bound to Sepharose. In some of the preparations the affinity chromatography step was repeated twice. Electrophoretic mobility shift assays of the fractions obtained after these purification steps were carried out with the 32P-labeled d(GA)₁₂ oligonucleotide probe and the results are shown in Fig. 1a. It can be seen that a single retarded band corresponding to the major band observed in the assays of the crude extract was obtained after the various steps of purification. Similar assays (not shown) revealed that the d(GA)ₙ-binding activity was eliminated by pretreatment of the purified preparations with pronase (100 µg/ml; 60 min at 37°C), or by heating these preparations 15 min at 60°C. These assays indicated that the d(GA)ₙ-binding activity was associated with a protein which we have designated PGB (see the Introduction). Other samples of the same preparations used for the assays shown in Fig. 1a were electrophoresed in an SDS-polyacrylamide gel and the gel was subsequently silver-stained. Fig. 1b shows the silver-stained proteins observed in this gel. Clearly, the first affinity chromatography resulted in a very substantial purification of the protein. Indeed, based on the data of Fig. 1, we estimate that the d(GA)ₙ-binding activity was purified at least 500-fold in this step. However, no further substantial purification was achieved by chromatography in the second affinity column. It can be seen that the preparations obtained after the first and the second affinity columns displayed four major bands containing 88 kd, 75 kd, 33.6 kd and 22 kd polypeptides, as well as additional minor bands. Some of the bands have apparently been enriched by the second affinity step. One of the enriched bands designated with an arrow contains the 33.6 kd polypeptide which was shown by UV crosslinking assays to bind d(GA)ₙ-tracts (see below).
Figure 4. Gel mobility shift analysis of the PGB protein: binding assays of the protein to oligonucleotides containing various G-rich sequence repeats. The oligonucleotides used for these assays were all labeled with \(^{32}\)P at specific radioactivities in the range of 2.5 - 3.7 \(\times 10^8\) dpm/\(\mu\)g. Electrophoretic mobility shift assays were performed with these oligonucleotides as described in Fig. 1, except that the concentrations of the DNA probes were 20 - 25 ng/ml. The sequences of the oligonucleotides designated \(-d(GA)_{12}-\) and \(-d(GA)_{25}-\) were 5' - AGAGAGAGATGAGAATATTCAG - 3' and 5' - AGAGAGATGAGAAATATTCAGACCT - 3', respectively. The numbers at the bottom of the gel represent the binding of the various oligonucleotides to the PGB protein relative to that of the \(d(GA)_{12}\) oligonucleotide, which was given the value 100.

DNA binding specificity of the PGB protein

Fig. 2 shows a series of electrophoretic mobility shift assays carried out with a \(^{32}\)P-labeled \(d(GA)_{25}\) oligonucleotide probe and unlabeled single stranded competitor DNAs. It can be seen that binding of the protein to this probe, like the binding to the \(d(GA)_{12}\) probe, generated one retarded band. Addition of 2 ng of unlabeled \(d(GA)_{25}\) oligonucleotide to the binding mixtures caused a substantial reduction in the intensity of the band and addition of larger amounts of this homologous competitor caused a virtual disappearance of the band. Unlabeled \(d(GA)_{12}\) competed with the radioactively labeled \(d(GA)_{25}\) probe as efficiently as the unlabeled \(d(GA)_{25}\) oligonucleotide (not shown). However, as shown in Fig. 2, addition of similar quantities of unlabeled single stranded bacteriophage M13 DNA which has been cut to fragments of \(\leq 500\) nucleotides caused only a slight reduction in the intensity of the band. Similarly, the homopolydeoxyribonucleotides poly dA, poly dT and poly dC did not effectively compete with the probe. Assays of the same type carried out with intact single stranded M13 DNA revealed that this DNA too did not compete with the probe (not shown).

We conclude that the PGB protein is not a general single stranded DNA binding protein (SSB).

Fig. 3 shows gel shift competition assays in which the single stranded \(^{32}\)P-labeled \(d(GA)_{12}\) oligonucleotide was used as a probe. It can be seen that the unlabeled \(d(GA)_{25}\) competed efficiently with the radioactively labeled \(d(GA)_{12}\) probe. On the other hand, equivalent concentrations of a double stranded DNA fragment designated TER, which consists of 231 base pairs including a \(d(TC)_{27}-d(GA)_{27}\) tract, did not compete with the probe at all. It can also be seen that double stranded bacteriophage \(\lambda\) DNA did not compete with the probe. Hence, the PGB protein does not bind double stranded DNA in general, and double stranded \(d(TC)_{27}, d(GA)_{27}\) tracts in particular.

Fig. 4 shows electrophoretic mobility shift assays designed to further determine the binding specificity of the PGB protein. These assays were carried out with the various oligonucleotide probes indicated at the top of the gel, all of which were labeled with \(^{32}\)P at approximately the same specific radioactivities; the DNA and the protein concentrations were also the same in all these assays. The numbers at the bottom of the gel designate the relative extents of binding of the oligonucleotides to the PGB protein. It can be seen that (a) the \(d(GT)_{10}\) oligonucleotide bound the PGB protein 1.5 times more avidly than the \(d(GA)_{12}\) oligonucleotide. (b) The 24 base oligonucleotides including a \(d(GA)_{15}\), or a \(d(GA)_{3}\) tract, bound the PGB protein two times...
b. protein, samples of the affinity purified protein were centrifuged in glycerol gradients. Fractions collected from such gradients To obtain an estimate of the molecular weight of the native PGB protein, oligonucleotides did not bind the PGB protein. Similar assays (not shown) revealed that d(A-) oligonucleotide. It can be seen that, other oligonucleotides, but in the range at which the binding concentration of DNA used in the experiment shown in Fig. 4 was within the linear range of the binding curves for the other oligonucleotides, but in the range at which the binding curves for the d(GA)12 and the d(GT)10 oligonucleotides began to level off.

Fig. 5 shows gel shift competition assays in which the probe was the 32P-labeled d(GA)12 oligonucleotide. It can be seen that, in accordance with the data presented in Fig. 4, the d(GT)10 oligonucleotide very effectively competed with the probe for binding the PGB protein, whereas a dG18 oligonucleotide was a rather poor competitor. Based on these data, it was estimated that the affinity of the PGB protein to dGn oligonucleotides was less than 5-fold the affinity in which it bound the d(GT)10 oligonucleotide. Similar assays (not shown) revealed that d(A-C)10 and d(TC)12 oligonucleotides did not bind the PGB protein.

Chemical physical properties of the PGB protein
To obtain an estimate of the molecular weight of the native PGB protein, samples of the affinity purified protein were centrifuged in glycerol gradients. Fractions collected from such gradients were analysed by electrophoretic mobility shift assays using the radioactively labeled d(GA)12 and d(GT)10 probes. Fig. 6 shows the sedimentation profiles of the two binding activities determined in aliquots withdrawn from fractions of a single gradient. It can be seen that these binding activities cosedimented in the gradient, in support of the notion that the same protein binds the two DNA sequences. This and additional glycerol gradient centrifugation assays yielded an estimate of a sedimentation coefficient of $S_{w,0} = 9.4 \pm 0.7$ and an estimated molecular weight of 190,000 ± 7,000 for the native PGB protein (see Materials and Methods).

An independent estimation of a molecular weight in the range of 160,000-180,000 was obtained by gel filtration of the PGB protein in a column of Sephacryl S-200 HR (not shown).

Electrophoretic mobility shift assays of the same gradient fractions performed with a radioactively labeled double stranded d(T-C)12-d(GA)12 probe, or with a single stranded d(TC)12 probe, confirmed that the PGB protein did not contain any detectable d(TC)n-d(GA)2n-, or d(TC)n-binding activities (not shown).

We also performed UV crosslinking assays using as a probe the 32P-labeled d(GA)12 oligonucleotide. In these assays, mixtures of the affinity purified protein and the probe were exposed to UV light for various time periods. Subsequently, the UV irradiated mixtures were electrophoresed in an SDS-polyacrylamide gel and the gel was autoradiographed. Fig. 7a shows that one polypeptide was crosslinked to the probe in a UV dose-dependent reaction. The band containing this polypeptide apparently corresponds to the band that has been enriched by
the second affinity chromatography step shown in Fig. 1b. Five independent UV crosslinking assays of this type have yielded an estimate of 33.6 ± 2.1 kd for the molecular weight of the crosslinked polypeptide. Fig. 7b shows that the radioactively labeled d(GA)_{12} probe could be competed from the complex formed between the probe and the 33.6 kd polypeptide by unlabeled d(GA)_{12} and d(GT)_{10} oligonucleotides, respectively. A minor band which is also seen in Fig. 7b, is not clearly discernible in Fig. 1b and hence it may not be specifically associated with the PGB protein. Similar assays carried out with a radioactively labeled d(GT)_{10} oligonucleotide confirmed that this oligonucleotide too binds the 33.6 kd polypeptide (not shown). Taken together these results have indicated that PGB is a multisubunit protein containing at least one 33.6 kd subunit which contacts the DNA (see Discussion below).

**DISCUSSION**

We described in this article a novel method for preparation of nuclear extracts from cells attached to microcarrier beads. The development of this method was essential for extraction of the PGB protein since we were unable to separate intact cells from the beads. This method should be also suitable for extraction of other nuclear proteins from cells attached to beads; furthermore, it can be scaled up and adopted for industrial production of such proteins.

We also described a rather simple procedure for purification of the PGB protein from the nuclear extracts. Using this procedure we have purified the d(GA)$_n$-binding activity at least 500-fold. The electrophoretic mobility shift assays of the purified protein revealed that in addition to d(GA)$_n$-tracts, it also binds d(GT)$_n$-tracts with a high affinity, but has lower affinities to other single stranded G-rich repeats, including d$_G$-tracts. It does not bind double stranded d(TC)$_n$-d(GA)$_n$-tracts or other double stranded DNA molecules. It does not appear to be a general single stranded DNA binding protein (SSB), because it does not manifest significant binding affinities to single stranded M13 DNA and various polynucleotides or oligonucleotides such as poly dA, poly dT, poly dC, d(AC)$_n$ and d(TC)$_n$. Our data do not provide detailed information on the features in the d(GA)$_n$ and the d(GT)$_n$-tracts that the PGB protein recognizes. The fact that the affinity of the PGB protein to the members of the oligonucleotide series d(GA)$_n$ decreases as $i$ increases (Fig. 4), indicates that the protein recognizes features related to the repetitive G residues within these sequences. However, it is not clear as yet whether the protein recognizes chemical groups in the alternating G bases, or other structural features of the DNA repeats including these bases, or both. Since the protein binds both the d(GA)$_n$ and the d(GT)$_n$ repeats, it appears unlikely that it recognizes the special structures that d(GA)$_n$-tracts may assume, namely the parallel stranded duplexes (36) or the single stranded helix stabilized by ionic bonds (37).

Based on glycerol gradient centrifugation, the molecular weight of the native PGB protein has been estimated to be 190,000 ± 7,000. The UV crosslinking assays have indicated that the same protein includes a 33.6 kd polypeptide which directly binds the DNA. In addition, it may contain other polypeptides which do not directly contact the DNA (see Fig. 1). However, our data are not conclusive as to whether these other polypeptides are integral components of the native PGB protein. In view of the specificity that this protein exhibits towards the repeat sequences d(GA)$_n$ or d(GT)$_n$, it appears plausible that each of the protein molecules may contain several 33.6 kd subunits, which simultaneously bind these repeats.

A human protein designated Pur Factor has been shown to specifically bind the purine-rich separated strands of sequence elements found near putative mammalian replication origins and gene flanking regions (38). UV crosslinking assays of this protein revealed that it contains a 28 kd subunit. Proteins that bind double stranded polypurine.polypyrimidine repeats have been also reported in Drosophila melanogaster, chicken, mouse and human cells (39—44). These proteins may be involved in regulation of transcription from promoters containing such sequence elements (ibid.). As mentioned in the Introduction, the human nuclear extracts from which the PGB protein has been purified also contained a protein that specifically bound double stranded d(TC)$_n$.d(GA)$_n$-tracts (Aharoni, Baran and Manor, unpublished results). It would be interesting to find out whether the genes encoding these various proteins are evolutionarily related.

The selective binding of the PGB protein to d(GA)$_n$-tracts suggests possible biological functions for the protein. As already mentioned in the Introduction, d(TC)$_n$.d(GA)$_n$-tracts inserted in topologically restricted DNA may undergo transitions into structures including d(TC)$_n$.d(GA)$_n$.d(TC)$_n$-triplexes + single stranded d(GA)$_n$-sequences (3—19). If such transitions occur in living cells, then the PGB protein may bind the single stranded d(GA)$_n$-tracts in these unusual structures. Furthermore, the binding of the protein to these single stranded DNA sequences might enhance the transitions. Since, as pointed out in the Introduction, transitions of this type may cause arrests of DNA replication and amplification, the PGB protein could regulate the arrests by enhancement of the transitions. Thus, the concentration of this protein in the cell nuclei may determine whether these arrest signals would be turned off or on. It is conceivable, for example, that interaction of the PGB protein with d(T-C)$_n$.d(GA)$_n$-tracts that map at two putative mammalian origins of DNA replication (45,46), may switch the replication initiated at these origins from a bidirectional to a unidirectional mode. Also, since d(TC)$_n$.d(GA)$_n$-tracts are widely dispersed in mammalian genomes and could be present in every replicon (20), the PGB protein may be involved in regulating DNA replication at most or all replicons. Another role for the PGB protein could be to regulate arrests of multiple forks which are presumably generated rarely at some origins (47). Thus, binding of the PGB protein to d(TC)$_n$.d(GA)$_n$-tracts found next to those origins could stop the amplification of genes located beyond these sequences, thereby preventing the potentially harmful amplification of extensive regions of DNA. Experiments designed to test some of these ideas are currently in progress in our laboratory.

We do not have any specific suggestions as to the biological significance of the rather strong affinity of the PGB protein to single stranded d(GT)$_n$-tracts. It appears possible that such binding should stimulate strand separation of d(GT)$_n$.d(AC)$_n$-tracts, which are also prevalent in mammalian genomes, and perhaps promote generation of unusual DNA structures involving these separated strands.

**ACKNOWLEDGEMENTS**

We thank Dr R. Knippers, our collaborator in the research supported by the G.I.F, for helpful discussions and suggestions, and Dr T. Haran for helpful discussions. We thank Interpharm for giving us some of the cells used for preparing the PGB...
protein. This research was supported by a Grant from the G.I.F., the German—Israeli Foundation for Scientific Research and Development.

REFERENCES