Selective Binding to DNA Base Pair Mismatches by Proteins from Human Cells*

Claire Stephenson and Peter Karran
From the Imperial Cancer Research Fund, Clare Hall Laboratories, South Mimms, Potters Bar, Hertfordshire, EN6 3LD, United Kingdom

Using the technique of delayed oligonucleotide migration through polyacrylamide gels, we have demonstrated that cell-free extracts of the human Burkitt's lymphoma cell line Raji contain proteins which can recognize and bind to mismatched single base pairs in short fragments of DNA. One of these binding proteins resembles an activity previously reported in HeLa cells (Jiricny, J., Hughes, M., Corman, N., and Rudkin, B. B. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 8860–8864) and recognizes DNA containing G·T mismatches. Extracts of Raji cells contain an additional activity which recognizes A·C, T·C, or T·T mismatches in DNA. This second binding protein can be distinguished from the G·T binding activity by its size, substrate specificity, and its fractionation properties. In addition to Raji cells, the new mismatch binding protein is present in extracts of human lymphoblastoid cell lines from a normal individual and a xeroderma pigmentosum patient as well as the SV40-transformed human fibroblast cell line MRC5SV1. It seems likely that this novel activity is involved in a broad specificity DNA repair pathway for the correction of single base mismatches in human cells.

The detection and removal of noncomplementary DNA base pairs is one of the pathways of error avoidance in cells. Proofreading by the replication complex during DNA replication can remove transiently incorporated mispaired bases and reduce the frequency of base substitution mutations (1, 2). In addition to replicative proofreading, mismatch correction pathways remove incorrectly paired bases from DNA. In Escherichia coli, methyl-directed dam+, mutH+, mutS+, mutL+, and mutU+-dependent mismatch repair acts immediately after DNA synthesis to excise the incorrect base (3). These repair events are associated with the excision and resynthesis of relatively extensive stretches of DNA. Two additional and more specialized mismatch correction pathways have also been described in E. coli. The first of these carries out a unidirectional correction of A·G mismatches to C·G base pairs in a reaction which is independent of the products of the mutH+, mutS+, mutL+, and dam+ genes (4) and requires the product of the mutY+ gene (5). The second pathway, characterized on genetic grounds by a short excision tract (very short patch repair) (6), acts in a unidirectional fashion to convert G·T mismatches in DNA to G·C base pairs (7). This system is also independent of the mutH and mutU genes but requires functional MutS and MutL proteins (8). The very short patch system corrects the G·T mispairs that arise throughout DNA as a consequence of the deamination of 5-methylcytosine.

To characterize the pathways of mismatch correction in E. coli, defined heteroduplex DNA molecules were introduced into suitably repair-compromised bacteria by transformation. The basic properties of methyl-directed and very short patch mismatch repair were delineated by the powerful analytical techniques of the phage λ genetics system. There is no analogous method of genetic analysis in mammalian cells. Nevertheless, transfection of SV40 DNA heteroduplexes combined with restriction site analysis of progeny molecules has indicated that mismatch correction occurs in mammalian cells. Using this approach, both cytosine methylation and the presence of single-strand nicks have been shown to influence strand discrimination in mismatch repair (9), and the relative efficiencies of correction of different mispairs have been determined (10). Mismatch correction catalyzed in vitro by extracts of human cells has been reported (11), and although the individual steps in the process have not been defined, a protein which can recognize G·T mismatches in an apparently specific fashion and may initiate mismatch correction has recently been reported in human cell extracts (12).

In contrast to its role in mutation avoidance, indirect evidence suggests that mismatch repair may enhance the sensitivity of some bacterial and mammalian cells to killing by certain chemical mutagens. The modified base O6-methylguanine, which is one of the major products of the reaction of the methylating mutagens N-methyl-N'-nitro-N-nitrosoguanidine and N-methyl-N-nitrosourea with DNA, is a major contributor to the cytotoxic effect of these compounds in mammalian cells and in dam mutants of E. coli (13, 14). It has been suggested (15–19) that the cytotoxicity of N-methyl-N'-nitro-N-nitrosoguanidine and N-methyl-N-nitrosourea toward E. coli dam strains and mammalian cells may be the result of ineffective attempts at mismatch correction at O6-methylguanine-containing base pairs.

Initiation of mismatch correction is carried out by a protein which recognizes and binds to DNA in the region of the mismatch. As a preliminary step to characterizing the process of mismatch repair and its relation to the cytotoxicity of methylating agents in human cells, we have investigated the ability of extracts of human cells to bind to defined single base mismatches in a standard oligonucleotide sequence. In the course of our preliminary studies we observed a binding activity which differs from the previously described G·T mispair binding protein (12). Here we describe some properties of the new binding protein which suggest that its likely roles are in the correction of replication errors or the resolution of mispairs generated during recombination.

* 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.
EXPERIMENTAL PROCEDURES

Materials

Oligonucleotides were prepared by solid phase synthesis on an Applied Biosystems 380B DNA synthesizer. Biochemicals were obtained from Sigma except where stated otherwise.

Cell Culture and Preparation of Extracts

The TK- derivative of the Burkitt’s lymphoma cell line Raji (20) and the normal human lymphoblast cell line GM1953 (NIGMS, Camden, NJ) were maintained in unstirred suspension culture in RPMI 1640 medium supplemented with antibiotics and 10% fetal calf serum (GIBCO). GM2345, a lymphoblast line derived from a xeroderma pigmentosum group A patient (NIGMS) was grown in spinner culture in the same growth medium. The SV40-transformed human fibroblast cell line MRC5V1 was obtained from Dr. M. H. L. Green, MRC Cell Mutation Unit, University of Sussex. It was routinely maintained in Dulbecco’s minimum Eagle’s medium supplemented with antibiotics and 10% fetal calf serum.

Cells were harvested from suspension or from plates while in exponential growth phase. After washing, they were extracted essentially as described by Jiricny et al. (12) except that the final centrifugation was performed at 4 C for 60 min at 90,000 rpm in a Beckmann TL100 ultracentrifuge. Normally about 10⁶ cells were used per extraction which yielded approximately 2–5 mg of extract in a final volume of 500 µl. Extracts were aliquoted, snap-frozen in liquid nitrogen, and stored at -80 C. Under these conditions, binding activities were stable for at least 2 months. Protein concentrations were determined by the method of Bradford (21).

Fractionation of Binding Proteins

Ammonium Sulfate—A crude cell extract prepared from approximately 4–10⁶ Raji cells was fractionated by the sequential addition of solid ammonium sulfate to 35, 55, 75, and 100% saturation at 0 C. Precipitated material was dissolved in extraction buffer minus protease inhibitors and dialyzed overnight at 4 C against the same buffer. Insoluble material remaining after dialysis was removed by brief centrifugation before assay.

Molecular Weight Estimation—The material which precipitated between 5 and 35% ammonium sulfate saturation was applied to a column (110 x 1 cm diameter) of AcA34 (Pharmacia LKB Biotechnology Inc.) equilibrated in 25 mM Hepes-KOH, pH 8.0, 0.2 mM EDTA, 0.1 M KCl, 0.5 mM dithiothreitol, 10% glycerol. Fractions (0.8 ml) were collected and 5-µl aliquots assayed for G. T and A. C binding activity.

Oligonucleotides

Most experiments were performed using 34-mer oligonucleotides of the composition 5’-AGCTTGGCTGAGTNGAGCTGCTCCCGGGGATTC-3’ (12), where N = A, G, T, or C, annealed to a complementary strand which was either an exact match or contained a single noncomplementary base at the position corresponding to N (position 16). Annealing of end-labeled oligonucleotides with their complementary strands was carried out in 10 mM Tris-HCl, pH 8.0, 10 mM MgCl₂. The equimolar mixture was heated to 70 C for 5 min and allowed to cool slowly to room temperature. Additional experiments were performed using a 21-mer of the composition 5’-TGATCAGTACGCTAGCTAGC-3’ annealed to one of two complementary strands which contained either C or T at the position corresponding to the central G.

Delayed Migration Experiments

Reactions were carried out at 20 C in a total volume of 20 µl: 25 mM Hepes-KOH, pH 8.0, 0.5 mM EDTA, 0.1 mM ZnCl₂, 0.5 mM dithiothreitol, 10% glycerol. All incubations additionally contained poly(dC)-poly(dG) at a concentration of 50 µg/ml. In the standard assay, cell extracts (usually 10–20 µg of protein in ≤5 µl) were preincubated for 5 min with 40 fmol of standard duplex 34-mer oligonucleotide containing a G-C base pair at position 16 before addition of the radioactive substrate. Subsequent incubation with the radioactive substrate (20 fmol) was for 20 min at 20 C following which 10-µl aliquots were supplemented with bromphenol blue and loaded onto 6% polyacrylamide gels (7 x 9 x 0.75 mm thick) (Mini Protean, Bio-rad). Electrophoresis was carried out at 70 V (10 V/cm) in TBE (89 mM Tris, 89 mM boric acid, 2 mM EDTA) buffer until the dye front (and the free oligonucleotide) had just migrated out of the gels. The gels were placed in polyethylene bags without drying and exposed to x-ray film (Kodak X-Omat). In experiments in which nonradioactive competitor oligonucleotides were used, these were included in the preincubation step at the concentrations indicated in the text.

RESULTS

Specific Binding to Mismatched Substrates

End-labeled duplex 34-mer oligonucleotides which contained a single mismatched G-T or A-C base pair at position 16 were incubated in the presence of poly(dl-dC)-poly(dl-dC) with extracts of Raji cells. Analysis of the products on 6% polyacrylamide gels under conditions where free oligonucleotide migrated to the top of the gel (Fig. 1, lane 5) revealed two delayed migrating complexes in each case (Fig. 1a, lanes 3 and 7). The more slowly migrating complexes formed from each of these substrates were of slightly different mobilities; the slower migrating band formed with the G-T oligonucleotide was well resolved from the more rapidly migrating band whereas resolution between the two complexes formed from the A-C duplex was incomplete under these conditions of electrophoresis. The more rapidly migrating band formed from both substrates appeared to be of identical mobility.

In control experiments, a perfectly matched duplex of the same sequence in which either a G-C or an A-T pair replaced the mismatch at position 16 was employed as substrate. A single band, apparently identical to the faster migrating complex seen with the G-T and A-C substrates, was observed (Fig. 1a, lanes 1 and 3). The identity of these bands was confirmed by preincubating the cell extracts with a nonradioactive competitor duplex (40 fmol) containing a G-C base pair at position 16 prior to the addition of the radioactive oligonucleotide (20 fmol). As a result of this preincubation, most delayed migrating complexes were completely resolved from the perfectly matched duplex substrates (Fig. 1a, lanes 2 and 6), and with both the mismatched substrates, formation of the more rapidly migrating complex was abolished. The more slowly migrating bands formed from the mismatched substrates were unaffected by the preincubation (Fig. 1a, lanes 4 and 8). Formation of the rapidly migrating complex was also prevented by a preincubation with nonradioactive competitor oligonucleotide in which an A-T base pair replaced the G-C at position 16 (data not shown). A 10-min preincubation of the cell extract with proteinase K (1 mg/ml) abolished all complex formation (data not shown).

Thus, single purine-pyrimidine base mispairs were recognized by proteins in extracts of Raji cells. The delayed migrating complexes formed from both a G-T and A-C mismatched substrate could be resolved from an apparently non-specific complex. Formation of the non-specific complex could be avoided by the inclusion of an unlabeled perfectly matched duplex in a preincubation step. The G-T and A-C specific complexes were completely resolved from one another by more prolonged electrophoresis (see Fig. 2) and thus were probably derived from different binding proteins.

To investigate the possibility that the binding proteins were recognizing particular features of the oligonucleotide sequence, we examined the delayed migrating complexes formed with a duplex 21 mer of unrelated sequence which contained either a G-C or a G-T base pair at the central position. In this substrate, the G of the mismatched base pair was in the

1 The abbreviation used is: Hepes, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid.
FIG. 1. G·T and A·C mismatch binding proteins in Raji cell extracts. a, specific and nonspecific binding: 34 mers. Extracts of Raji cells (20 µg, 2.5 µl) were preincubated (lanes 2, 4, 6, 8) for 5 min at 20 °C in a reaction mixture (20 µl) containing 50 μg/ml poly(dl-dC)-poly(dl-dC) and nonradioactive duplex 34 mer (40 fmol) containing a G·C base pair at position 16. Following the preincubation, radioactive substrate oligonucleotide (20 fmol) was added to each sample as shown. The reactions represented in lanes 1, 3, 5, and 7 were not preincubated in the presence of the G·C-containing oligonucleotide. To these incubations, the cell extract and the substrate oligonucleotide were added simultaneously. No cell extract was added to the incubation represented in lane 9. After a further 20-min incubation at 20 °C, 2 µl of bromphenol blue (0.1%) in 50% glycerol was added to all samples, and 10 µl of each was loaded onto a 6% polyacrylamide gel. Electrophoresis was carried out at 10 V/cm until the dye front had migrated about 6 cm. The wet gel was placed in a polyethylene bag and exposed to x-ray film. The position of migration of the uncomplexed oligonucleotide is shown by the arrow.

b, 21 mers. Raji cell extract (lanes 2 and 6 = 7 pg; lanes 3 and 7 = 14 pg; lanes 1, 4, 5, and 8 = 21 pg) was preincubated with cold 21-base pair duplex oligonucleotide (40 fmol) containing a G·C base pair at the central position (lanes 2-4 and 6-8) as described above. Control samples (lanes 1 and 5) were not preincubated. Radioactive 21-mer substrates containing either a G·C (lanes 1-4) or G·T base pair (lanes 5-8) at the central position were added and incubation and sample loading carried out as above. Electrophoresis was carried out until the dye front had completely migrated through the gel.

FIG. 2. Specificity of the A·C and G·T binding proteins as determined by competition with unlabeled oligonucleotides. a, the G·T binding protein. The conditions of preincubation and incubation were as described in the legend to Fig. 1. Control samples were not preincubated (lanes 1 and 4). The remaining preincubations contained 0.5 pmol of unlabeled competitor oligonucleotide as indicated on the figure. Two protein concentrations were assayed for each substrate: 7 µg (lanes 2, 5, 7, and 9) and 14 µg (lanes 1, 3, 4, 6, 8, and 10). Radioactive G·C oligonucleotide (lanes 1-3) or G·T oligonucleotide (lanes 4-10) was added after the 5-min preincubation and incubation continued for 20 min more. Sample preparation, gel loading, and electrophoresis were carried out as in Fig. 1b. b, the A·C binding protein. Analysis was identical to that described above except that a radioactive A·T substrate replaced the G·C substrate (lanes 1-3) and a radioactive A·C substrate replaced the G·T substrate (lanes 4-10).

sequence 5'-GpG-3'. When the 21-mer oligonucleotides were incubated with extracts of Raji cells, the pattern of complex formation was similar to that observed with the 34 mer (Fig. 1b): a more rapidly migrating nonspecific band which was also formed with a perfectly matched substrate (lanes 1 and 5) and which was not formed after preincubation with competing G·C oligonucleotide (lanes 2-4) and a more slowly migrating complex which was apparently specific for the G·T mispair (lanes 5-8).

Replacing the cell extract by purified E. coli single-strand binding protein in 2-fold excess of the concentration required to completely cover a single-stranded 34-mer oligonucleotide did not generate delayed migrating bands from the mismatched 34-mer duplexes, indicating that recognition by the binding protein was probably not a result of extensive regions of single strandedness around the mismatch. Further, no bands migrating at the positions of the specific G·T or A·C complexes were generated when Raji cell extracts were incubated with single-stranded oligonucleotides nor were the specific complexes formed with the duplex substrates affected by a single-stranded 34-mer competitor of unrelated sequence (data not shown).

The properties of the proteins recognizing the A·C and G·T mismatches were investigated further by comparing the
Fig. 3. Binding of the A·C binding protein to pyrimidine-purine mismatches. (a) Binding: T·T, T·C, G·A, and G·G mismatches. Preincubations (20 µg extract) were performed in the presence (lanes 2, 4, 6, 8, 10, and 12) of unlabeled G·C oligonucleotide as described in the legend to Fig. 1. Samples 1, 3, 5, 7, 9, and 11 were not preincubated. Subsequent incubations were carried out with radioactive T·T (lanes 1 and 2), T·C (lanes 3 and 4), G·A (lanes 7 and 8), or G·G (lanes 9 and 10) substrates. Control incubations with A·C substrate were also included (lanes 5 and 6, 11 and 12). Electrophoresis was carried out as in Fig. 2, b, competition. Preincubations (20 µg extract) were carried out in the presence of 0.5 pmol of unlabeled G·C (lanes 2, 6, and 10), G·T (lanes 3, 7, and 11), or A·C (lanes 4, 8, and 12) unlabeled competitor. Samples 1, 5, and 9 were not preincubated. After 5 min at 20°C, 20 fmol of radioactive substrate A·C (lanes 1–4), T·C (lanes 5–8), or T·T oligonucleotide (lanes 9–12) was added. Incubation and electrophoresis were carried out as in Fig. 2, c, binding: C·C and A·A mismatches. Cell extract was preincubated (lanes 2 and 3, 5–7, 10–12) with 40 fmol of unlabeled G·C oligonucleotide for 5 min at 20°C. Samples 1, 4, and 9 were not preincubated. Radioactive substrate (20 fmol) was added as indicated, and incubation continued for a further 20 min. Samples 2, 5, and 10 contained 8 µg of extract, samples 3, 5, 4, 6, 9, and 11, 16 µg, and samples 7 and 12, 24 µg. Electrophoresis was carried out as in Fig. 2.

relative intensities of the two bands in the presence of an excess of nonradioactive competitor mismatched duplex. In Fig. 2a, we present data derived using a G·T substrate. The nonspecific band was removed by preincubation with a G·C competitor oligonucleotide (Fig. 2a, lanes 1–6) leaving only the G·T-specific band (Fig. 2a, lanes 4–10). The intensity of the G·T band which remained was slightly reduced by the inclusion of a 25-fold excess of nonradioactive G·T oligonucleotide in the preincubation (Fig. 2a, lanes 6 and 8). Unlabeled A·C oligonucleotide was without effect (Fig. 2a, lane 10). Preincubation with a G·C competitor oligonucleotide also abolished the nonspecific band seen with both an A·T or A·C duplex (Fig. 2b, lanes 1–6). The A·C-specific band was unaffected by a G·T competitor (Fig. 2b, lanes 7 and 8) but completely removed by an A·C competitor at the same 20-fold excess (Fig. 2b, lanes 9 and 10). Thus, by the two criteria of their migration characteristics and differential sensitivity to competitor oligonucleotides, the G·T and A·C mismatch complexes are derived from different mismatch binding proteins. In addition to the A·C complex, a minor band migrating at the position of the G·T complex was generated with the A·C oligonucleotide substrate (Fig. 2b, lanes 6 and 10). This faint band, which was not formed in the presence of a competitor G·T oligonucleotide (Fig. 2b, lanes 7 and 8), was unaffected by an A·C competitor even at a concentration which resulted in a complete removal of the A·C band (Fig. 2b, lane 10). While confirming the separate nature of the G·T and A·C binding proteins, these data suggest that the G·T binding protein may have a low, but detectable, affinity for DNA containing A·C mismatches. The G·T band was relatively insensitive to competition by unlabeled G·T oligonucleotide, and only a small reduction in the intensity of this band was observed at the moderate levels of competitor used (Fig. 2a, lanes 6 and 8). In contrast, the A·C binding was effectively abolished by the same concentration of an unlabeled A·C competitor (Fig. 2b, lanes 6 and 10).

A previously reported G·T-specific binding protein from HeLa cells (12) generated a more complex pattern of delayed migration than that from Raji cells in that three bands exhibiting delayed migration were produced by the G·T oligonucleotide. In that study, the faster migrating band was without substrate specificity and may be identical to the nonspecific band observed here. The second binding activity of HeLa cells was derived from the third apparently by a process of proteolysis in vitro. We have also observed a breakdown of the G·T binding activity during preparation of Raji cell extracts. This breakdown resulted in the formation of two G·T mismatch-protein complexes which were resolved on electrophoresis. The first of these is the one described above. The second complex was not resolved from the nonspecific band but could be distinguished from it by a relative insensitivity to perfectly matched competitor oligonucleotides. It seems probable, therefore, that the G·T binding protein in Raji cells observed here is similar to that described in HeLa cells whereas the A·C binding protein is a previously unrecognized activity.

Recognition of Other Mismatched Base Pairs

Pyrimidine-Pyrimidine Mispairs—The ability of proteins from Raji cell extracts to bind to mismatched bases other than purine-pyrimidine pairs was examined. Delayed migrating bands were observed when oligonucleotides containing pyrimidine-pyrimidine base mismatches were used as substrate (Fig. 3a). A nonspecific band apparently identical to the one previously observed with the pyrimidine-pyrimidine pairs was formed from both T·T- and T·C-containing substrates (Fig. 3a, lanes 1 and 3). The T·T mismatch substrate generated an additional band (Fig. 3a, lane 2) which migrated in a similar fashion and was of comparable intensity to the previously observed A·C band (Fig. 3a, lanes 6 and 12). A band which exhibited similar migration characteristics was also observed.
Human DNA Mismatch Binding Proteins

Fig. 4. Partial purification of the A·C binding protein on AcA34. A 5–35% ammonium sulfate fraction of a crude extract of Raji cells was separated on a column (110 × 1 cm diameter) of AcA34 precalibrated with β-amylase, alcohol dehydrogenase (D), bovine albumin (BSA), carbonic anhydrase, cytochrome c (CytC). The arrow shows the position of elution of the A·C binding activity. b, fractions (0.8 ml) were collected and 5 μl assayed for binding to labeled A·C or G·T mismatched duplexes after a preincubation with unlabeled G·C oligonucleotide under the standard conditions described under “Experimental Procedures.”

with the C·T duplex (Fig. 3a, lane 4). Incubation with a competitor A·C oligonucleotide (Fig. 3b, lanes 8 and 12) or either of the pyrimidine-pyrimidine mismatch oligonucleotides (data not shown), but not the G·T or the matched oligonucleotides (Fig. 3b, lanes 6, 7, 10, and 11), abolished these bands. When a C·C mismatched oligonucleotide was used as substrate, little or no binding was observed (Fig. 3c). Thus, the A·C mismatch binding protein additionally recognizes pyrimidine-pyrimidine mismatches with an apparent preferential affinity: T·T > C·T > C·C.

Purine-Purine Mispairs—Specific delayed migrating bands were observed with a G·G or G·A mismatched oligonucleotide (Fig. 3a, lanes 8 and 10). Their migration characteristics resembled those of the G·T- rather than the A·C-derived complex (shown in lanes 11 and 12 for comparison). Their identity was confirmed by the use of competitors; only unlabeled excess G·T (and G·G) oligonucleotides diminished the signal from these substrates, an A·C competitor being without effect (data not shown). Less binding was observed to an A·A mismatched oligonucleotide (Fig. 3c). Since the signals which were observed with these substrates did not have the migration characteristics of the A·C binding activity and were not efficiently competed by A·C mismatched oligonucleotides, we conclude that the purine-purine mismatches are not recognized by the A·C mismatch binding protein.

Fractionation of Binding Proteins

The G·T and A·C binding activities differed in their biochemical properties and could be separated by conventional techniques. A crude extract of Raji cells was fractionated into four fractions representing 5–35%, 35–55%, 55–75%, and 75–100% saturation of ammonium sulfate, and after dialysis, each fraction was tested for binding activity with a G·T- or A·C-containing oligonucleotide substrate. Essentially all activity binding to A·C mispairs was recovered in the material precipitating between 5% and 35% saturation. No material binding to a G·T substrate was recovered in any individual ammonium sulfate fraction. A further fractionation of the A·C binding activity which precipitated at 35% saturation was effected by gel filtration on AcA34. The load material contained substantial A·C binding activity but very little activity toward a G·T substrate (Fig. 4b). The A·C binding activity eluted from the column in fractions 63–70, which is consistent with a monomeric protein of molecular mass around 100 kDa (Fig. 4a). A low level of binding to the G·T substrate can be seen in the fractions containing the A·C binding activity. This may represent a separate binding activity or, more likely, may reflect a low affinity recognition of a G·T mispair by the A·C binding protein analogous to the reverse situation observed in Fig. 2b. No binding activity was eluted at a position corresponding to 200 kDa. The estimated size of the A·C binding protein is therefore smaller than the 200 kDa reported for the G·T binding protein of HeLa cells and is in agreement with the more rapid migration of the A·C binding protein:oligonucleotide complex observed on polyacylamide gels.

A·C Binding Protein in Other Human Cell Lines

Both the A·C and the G·T binding activities were present in cell-free extracts of the human lymphoblastoid cell lines
GM1953 (from a normal individual) and GM2345 (from a xeroderma pigmentosum group A patient). Furthermore, extracts of the SV40-transformed human fibroblast cell line MRC5V1 prepared in the same way contained similar levels of both binding activities (data not shown).

**DISCUSSION**

The mismatch binding activity described here, which recognizes A-C, T-T, and T-C mispairs, represents a hitherto unrecognized activity which is distinguished from the previously reported G-T binding protein by its different size, substrate specificity, and biochemical properties. The range of mismatched substrates which are recognized by this A-C binding protein is only approximately correlated with the reported preferences of the mammalian mismatch correction system. The protein has an apparent preference for mispairs which contain a pyrimidine base although C-C mispairs are recognized poorly, if at all. Apart from G-T, the pyrimidine-containing mispair most frequently corrected in vivo is A-C followed by T-C, C-C, and T-T at decreasing frequencies (10). The specificity of the A-C binding protein is thus not sufficient in itself to explain the preferential correction of mismatches in mammalian cells. It is interesting, however, that the least well corrected mismatch in both *E. coli* (22) and *Pneumococcus* (23) and also the mispair for which the *E. coli* MutS binding protein has the lowest affinity (24) is the C-C base pair, which is the only pyrimidine-containing mispair (excluding G-T) not recognized well by the A-C binding protein of human cells. It is likely, however, that additional factors other than recognition by the binding proteins will determine the fate of mismatches in mammalian cells.

The experiments reported here were carried out on different single base mispairs placed within an identical DNA sequence. It is not therefore possible to assign any role of sequence context to the ability of the A-C mismatch binding protein to recognize potential substrates. However, since several of the single base mismatches are not recognized by the A-C binding protein when placed in the same sequence, it is clear that the binding observed is determined primarily by the mismatch and not by a particular sequence context. Furthermore, a cursory comparison of the interaction of the G-T binding protein with two unrelated oligonucleotides of different sizes, one of which placed the G-T mispair at the center of the sequence 5'-GTACGATC-3' and the second at the center of the sequence 5'-AGTTGACG-3', indicates that recognition by the G-T binding protein is not confined to G-T mispairs in CpG sequences. It seems unlikely that the necessary information to determine binding is provided by structural features of the mispair itself.

The presence of a mismatched base pair destabilizes the DNA helix. The G-T base pair has a relatively small destabilizing effect (25) whereas more destabilization is introduced by cytosine-containing mispairs. In particular, a C-C mispair is the least stable (26). The absence of observed binding to the C-C-containing duplex might indicate that the binding protein requires a certain minimum of secondary structural features to avoid unproductive binding to single-stranded regions in DNA. The absence of binding to either the G-T or the A-C mispairs by purified *E. coli* single-strand binding protein also supports the idea that recognition of mismatched base pairs depends on some structural features of the mispair rather than simply a localized loss of secondary structure.

While it seems likely that the G-T binding activity is involved in a specialized pathway for correction of deaminated 5-methylcytosine bases (27), the A-C binding protein probably acts in a more generalized pathway of mismatch correction. Unlike G-T, A-C (or T-C, T-T) mismatches are not generated in DNA by a simple spontaneous chemical reaction such as *in situ* deamination of constituent bases. For this reason, and in view of its broader substrate range, the most likely roles of the A-C binding protein are in the selective excision of bases wrongly inserted during DNA replication or the removal of mismatched base pairs which arise during recombination.

Acknowledgments—We thank Dr. J. Jiricny for communicating a preprint of his paper and Iain Goldsmith for preparation of the oligonucleotides.

**REFERENCES**

1. Bruttig, D., and Kornberg, A. (1972) *J. Biol. Chem.* 247, 241-248
2. Mazyczka, N., Poland, R., and Bessman, M. J. (1972) *J. Biol. Chem.* 247, 7116-7122
3. Modrich, P. (1989) *J. Biol. Chem.* 264, 6597-6600
4. Lu, A. L., and Chang, D. Y. (1988) *Cell* 54, 805-812
5. Au, K. G., Cabrera, M., Miller, J. H., and Modrich, P. (1988) *Proc. Natl. Acad. Sci. U. S. A.* 85, 9163-9166
6. Lieb, M. (1986) *Mol. Gen. Genet.* 181, 118-125
7. Jones, M., Wagner, R., and Radman, M. (1987). *J. Mol. Biol.* 194, 155-159
8. Lieb, M. (1987) *J. Bacteriol.* 169, 5241-5246
9. Hare, J. T., and Taylor, J. H. (1988) *Proc. Natl. Acad. Sci. U. S. A.* 85, 7350-7354
10. Brown, T. C., and Jiricny, J. (1988) *Cell* 54, 705-711
11. Glazer, P. M., Sarkar, S. N., Chisholm, G. E., and Summers, W. C. (1987) *Mol. Cell. Biol.* 7, 218-224
12. Jiricny, J., Hughes, M., Corman, N., and Rudkin, B. B. (1988) *Proc. Natl. Acad. Sci. U. S. A.* 85, 8860-8864
13. Karran, P., and Hall, J. (1988) *Nucleic Acids Mol. Biol.* 2, 188-197
14. Karran, P., and Marinus, M. (1982) *Nature* 296, 868-869
15. Goldmacher, V., Cuzic, R. A., and Thilly, W. G. (1986) *J. Biol. Chem.* 261, 2462-2471
16. Goth-Goldstein, R. (1987) *Carcinogenesis* 6, 1449-1453
17. Ishida, R., and Takahashi, T. (1987) *Carcinogenesis* 8, 1109-1113
18. Green, M. H. L., Lowe, J. E., Petit-Frere, C., Karran, P., Hall, J., and Kataoka, H. (1988) *Carcinogenesis* 10, 855-858
19. Aquilina, G., Zijno, A., Mosenso, N., Diotti, E., and Bignami, M. (1989) *Carcinogenesis* 10, 1219-1223
20. Hampar, B., Derge, J. G., Martos, L. M., and Walker, J. L. (1972) *Proc. Natl. Acad. Sci. U. S. A.* 69, 78-82
21. Bradford, M. M. (1976) *Anal. Biochem.* 72, 249-254
22. Dohet, C., Wagner, R., and Radman, M. (1985) *Proc. Natl. Acad. Sci. U. S. A.* 82, 503-505
23. Claverey's, P. J., Mejave, V., Gasc, A. M., and Sicard, A. M. (1983) *Proc. Natl. Acad. Sci. U. S. A.* 80, 5956-5960
24. Su, S. S., Lahue, R. S., Au, K. G., and Modrich, P. (1988) *J. Biol. Chem.* 263, 6829-6835
25. Patel, D. J., Kozlowski, S. A., Ikuta, S., and Iakura, K. (1984) *Biochemistry* 23, 3218-3226
26. Abou-Elia, F., Koh, D., Tisocco, I., and Martin, P. H. (1985) *Nucleic Acids Res.* 13, 4811-4824
27. Wiebusch, K., and Jiricny, J. (1989) *Nature* 339, 234-236