

Transcriptional activity of the homopurine – homopyrimidine repeat of the c-Ki-ras promoter is independent of its H-forming potential

Ganapathirama Raghu, Sergei Tevosian, Shrikant Anant, Kiranur N.Subramanian, Donna L.George¹ and Sergei M.Mirkin*

Department of Genetics, University of Illinois at Chicago, Chicago, IL 60612 and ¹Department of Genetics, University of Pennsylvania School of Medicine, Philadelphia, PA 19104-6145, USA

Received July 7, 1994; Accepted July 11, 1994

ABSTRACT

The mouse c-Ki-ras protooncogene promoter contains an unusual DNA element consisting of a 27 bp-long homopurine – homopyrimidine mirror repeat (H-motif) adjacent to a d(C-G)₅ repeat. We have previously shown that *in vitro* these repeats may adopt H and Z conformations, respectively, causing nuclease and chemical hypersensitivity. Here we have studied the functional role of these DNA stretches using fine deletion analysis of the promoter and a transient transcription assay *in vivo*. We found that while the H-motif is responsible for approximately half of the promoter activity in both mouse and human cell lines, the Z-forming sequence exhibits little, if any, such activity. Mutational changes introduced within the homopurine – homopyrimidine stretch showed that its sequence integrity, rather than its H-forming potential, is responsible for its effect on transcription. Electrophoretic mobility shift assays revealed that the putative H-motif tightly binds several nuclear proteins, one of which is likely to be transcription factor Sp1, as determined by competition experiments. Southwestern hybridization studies detected two major proteins specifically binding to the H-motif: a 97 kD protein which presumably corresponds to Sp1 and another protein of 60 kD in human and 64 kD in mouse cells. We conclude that the homopurine – homopyrimidine stretch is required for full transcriptional activity of the c-Ki-ras promoter and at least two distinct factors, Sp1 and an unidentified protein, potentially contribute to the positive effect on transcription.

INTRODUCTION

S1 hypersensitive sites are often associated with eukaryotic promoters within active chromatin or in superhelical plasmids. In many cases, these sites are located within homopurine – homopyrimidine stretches (reviewed in 1). It has become

increasingly clear that the formation of an unusual DNA structure, called H-DNA, by homopurine – homopyrimidine mirror repeats is responsible for nuclease hypersensitivity *in vitro* (reviewed in 2). During recent years, independent studies have shown that homopurine – homopyrimidine stretches are important for the function of several eukaryotic promoters, including human EGFR (3), *ets-2* (4), IR (5), and *c-myc* (6, 7), mouse c-Ki-ras (8) and TGF- β 3 (9), *Drosophila hsp26* (10, 11), and others. These DNA stretches also serve as binding targets for nuclear proteins, presumably functioning as transcriptional regulators (9, 10, 12–19). However, any connection between the H-forming ability of those homopurine – homopyrimidine sequences and their function as promoter elements remains to be established. An attempt to discover H-DNA *in vivo* in the *Drosophila hsp26* promoter failed to detect this structure (11).

We have studied the functional role of the S1 hypersensitive stretch within the mouse c-Ki-ras protooncogene promoter. The c-Ki-ras protooncogene plays an important role in signal transduction and cellular proliferation (reviewed in 20). The promoter of this gene is typical of housekeeping gene promoters, i.e. it is highly GC-rich, contains neither TATA nor CAAT box elements, and has multiple transcriptional start sites (8). The cloned promoter reveals strong S1 hypersensitivity associated with a 27-bp long homopurine – homopyrimidine mirror repeat flanked by a d(C-G)₅ repeat. In a previous analysis of the structure of the c-Ki-ras promoter, we found that either H-DNA, by a homopurine – homopyrimidine stretch, or Z-DNA, at a d(C-G)₅ repeat, is formed in supercoiled DNA (21). Preliminary experiments indicated that an unusual DNA element is important for c-Ki-ras transcription (8).

Here, we present a detailed structure function dissection of the c-Ki-ras promoter. Based on deletion mutagenesis coupled with transient transcription assays, we conclude that the H-motif is required for full promoter function, while the Z-forming element has little, if any, such role. Mutational analysis of the H-motif revealed that it is the sequence integrity, rather than the H-forming potential, that is important to promote transcription. Finally, we

*To whom correspondence should be addressed

have observed the H-motif in specific complexes with several nuclear proteins that likely represent sequence-specific DNA binding transcription factors.

MATERIALS AND METHODS

Oligonucleotides

Oligonucleotides were synthesized on the ABI High Throughput DNA/RNA synthesizer Model 394 as described in the User's Manual (Applied Biosystems). They were deprotected by incubation in a concentrated ammonium hydroxide for 15 hours at 55°C, followed by concentration in a SpeedVac and precipitation by 2 volumes of 2M LiClO₄ in acetone. Dried pellets were dissolved in 0.5 ml of TE buffer and additionally purified on NAP-5 columns (Pharmacia).

Plasmids

Plasmid DNA was isolated by standard alkali lysis followed by twice repeated equilibrium centrifugation in a cesium chloride-ethidium bromide gradient. All the recombinant plasmids were sequenced by the Maxam-Gilbert method to confirm their authenticity.

Chemical probing of DNA

Modification of supercoiled DNA with chloroacetaldehyde, diethyl pyrocarbonate and osmium tetroxide was carried out in 20 mM Na Acetate, pH 4.5; 1 mM EDTA; 100 mM NaCl as previously described (21). Modified DNA samples were digested by restriction enzyme *Ava*I (see Fig. 1). Top and bottom strands were ³²P-labelled using T4 polynucleotide kinase or the Klenow fragment of DNA polymerase I, respectively. A second digestion with *Bgl*II was followed by isolation of the 170 bp end-labeled fragments from a 6% native polyacrylamide gel. Samples modified with OsO₄ and DEPC were treated with 1M piperidine for 30 min. at 90°C. CAA-modified samples were treated with either formic acid, or hydrazine in high salt, followed by piperidine treatment. Piperidine-treated samples were dried, dissolved in 80% formamide, 1 mM EDTA and loaded on an 8% polyacrylamide gel with 7M urea.

Cell culture conditions and transient expression assays

HepG2 (human liver) cells were grown in a 1:1 ratio D-MEM and F-12 nutrient mix supplemented with 10% bovine fetal serum, 10 µg/ml insulin, 100 µg/ml penicillin and 100 units/ml streptomycin. Y1 cells (from the mouse adenocortical tumor) were grown in F-10 nutrient mix supplemented with 12.5% horse serum, 2.5% bovine fetal serum, 100 µg/ml penicillin and 100 units/ml streptomycin. HeLa cells were grown in DMEM supplemented with 5% fetal bovine serum, 100 µg/ml penicillin and 100 units/ml streptomycin. Tissue culture media, nutrient mixes, sera and antibiotics were purchased from Gibco-BRL. Cells were transfected by calcium phosphate precipitation (34). Each dish received 10 µg of a test plasmid along with 2 µg of pCH110 eukaryotic β-galactosidase vector (Pharmacia) as an internal control for monitoring and normalizing transfection. 48 hr. post-transfection cell extracts were made by 3 cycles of freezing and thawing in 0.25 M Tris-HCl, pH 8.0. 5 µl of a cell extract was used for standard β-galactosidase assay (35). For CAT assays cell extracts were heated at 60°C for 10 min. and clarified by centrifugation. The volume of extract used in the assay was normalized by its β-galactosidase activity. For the derivatives of the native promoter it was equivalent to 0.6 OD

units of β-galactosidase activity, while for the minimal promoter derivatives it was 1.2 OD units. The reaction was performed in 250 mM Tris-HCl with 0.4 mg/ml of Acetyl-CoA (Pharmacia) and 1 pmol of ¹⁴C-chloramphenicol (53 mCi/mmol, Amersham) at 22°C for 15 min. (original promoter) or 30 min. (minimal promoter). The reaction products were resolved by TLC in a chloroform : methanol 95:5 mixture. The plates were dried and the amount of acetylated chloramphenicol was quantitated by scanning in a Betascope 603 analyzer.

Electrophoretic mobility shift assay (EMSA)

Nuclear extracts from Y1, HepG2 and HeLa cells in log phase were prepared according to (36). Protein concentration was estimated as described in (37). Double-stranded oligonucleotide probes were:

- GATCCCTCCCTCCCTCCTTCCCTCCCTCCCA (i)
GGAGGGAGGGAGGAAGGGAGGGAGGGTCTAG
and
GATCGGGGCGGAGAATGGGCGGAAGTGGGCGGAGTA (ii)
CCCCGCTCTTACCGCCTTGACCCGCCTCATCTAG

where (i) corresponds to the H-motif from the c-Ki-ras promoter, and (ii) contains three Sp1 consensus binding sites. In some cases the wild-type H-motif was replaced with either the corresponding DM1, DM2 or QM mutants. These probes were ³²P-labeled by end filling with Klenow enzyme.

Binding reactions were carried out as described in (27) in 20 mM Hepes-KOH pH 7.5, 70 mM KCl, 1 mM ZnSO₄, 12% (v/v) glycerol, 0.5 µg of BSA (Pharmacia), 0.05% of NP-40 (Sigma), 0.5 mM DTT. Each probe (final volume 20 µl) contained 3 µg of a nuclear extract, 10⁵ cpm of a labeled probe and 1 µg of poly d(I)·d(C) (Pharmacia). 1 µg of specific competitors was also added in selected samples. After 30 min. of incubation at room temperature, samples were resolved on a 5% native PAG in 0.5×TBE overnight at 4°C.

Southwestern blot analysis

The nuclear extract proteins were resolved on a discontinuous 10% SDS-PAGE followed by electrophoretic transfer to nitrocellulose membrane BA-S (Schleicher & Schuell). Membranes were incubated for 1–2 hr. at room temperature in a blocking solution of 20 mM Hepes-KOH pH 7.5, 70 mM KCl, 1 mM ZnCl₂, 2.5% non-fat dry milk, 0.1% NP-40 (Sigma) and 0.5 µg of poly d(I)·d(C). Then membranes were incubated overnight at 4°C in fresh blocking solution with 10⁶ cpm/ml of a labeled probe. Membranes were subsequently washed in the above solution without non-fat dry milk, dried and autoradiographed (28).

RESULTS

The H-motif is a cis-acting positive transcriptional element

As a starting construct, we used plasmid pKRS413 containing the mouse c-Ki-ras promoter in the eukaryotic expression vector pRSVA0CAT (8). This plasmid has approximately 300 bp of the mouse c-Ki-ras promoter driving the expression of the chloramphenicol acetyltransferase (*cat*) reporter gene (Fig. 1). It was previously found that this portion of the promoter is necessary and sufficient for promoter activity in mouse cells (8). This promoter region contains convenient restriction sites that we used to make various deletion constructions (Fig. 1). The original plasmid and its derivatives were transfected into human HepG2, or mouse NIH3T3, cells together with a reference

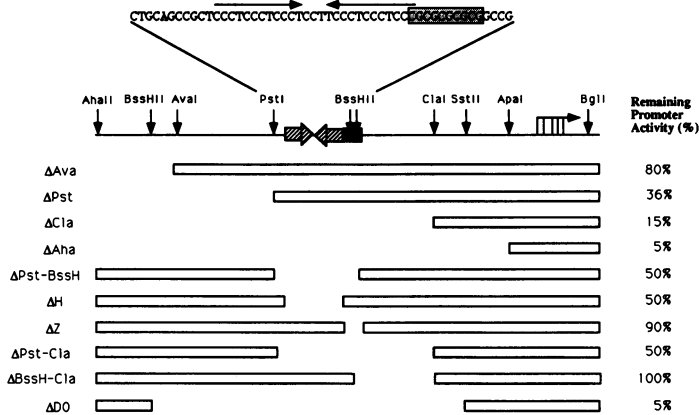


Figure 1. Deletion analysis of the c-Ki-ras promoter. Restriction sites used to obtain upstream truncations and internal deletions are shown by vertical arrows. Multiple transcription sites are indicated by vertical lines, and the direction of transcription is shown by the horizontal arrow. The homopurine-homopyrimidine mirror repeat is indicated by striated arrows, and the Z-forming stretch is shown by a black box. The primary sequence of the *PstI*-*BssHII* fragment (top strand) is expanded with horizontal arrows to mark the two halves of the mirror repeat and a shadowed box for the Z-forming sequence.

plasmid, pCH110, which contains *lacZ* gene encoding *E. coli* β -galactosidase under the control of the SV40 early promoter (22). The transfection efficiency was determined by a β -galactosidase assay and c-Ki-ras promoter activity was determined by the CAT assay. The results of the CAT assays normalized with those of the β -galactosidase assays are presented in Fig. 1. In all studied cases, we saw no significant differences when the assays were carried out in human or mouse cells. Thus, the transcriptional elements that exist in the c-Ki-ras promoter are likely to be equally functional in both human and mouse cell lines.

We studied two types of deletions: upstream truncations and internal deletions. The data obtained from examination of upstream truncation constructs indicate that three areas are important for promoter activity. Deletion of the upstream *AvaI*-*PstI* fragment causes a 3 fold decrease in promoter activity. Loss of the *PstI*-*ClaI* fragment causes an additional 2.5 fold drop. Finally, deletion to the *ApaI* site reduces promoter activity to background level. The promoter elements located upstream of the *PstI* site are yet to be elucidated, but the *PstI*-*ClaI* fragment contains the S1 hypersensitive element including H- and Z-forming DNA sequences (21). The largest deletion examined extends close to the transcriptional start sites, so the very low promoter activity observed with this construct (ΔAha) is not surprising.

We were most interested in the *PstI*-*ClaI* fragment, since both the H- and Z-forming elements are located within that region. The same fragment also contains significant flanking sequences, especially downstream from the Z-forming stretch, making it unclear, *a priori*, which sequence elements are required to promote c-Ki-ras transcription. To address this question, we benefited from the presence of a *PstI* site 11-bp upstream of the H-motif and two *BssHII* sites located within the Z-forming stretch. We obtained internal deletions of either the whole *PstI*-*BssHII* fragment (ΔPst -*BssH*) or only the putative H-motif (ΔH) (in this case oligonucleotides corresponding to sequences flanking H-DNA were inserted into the ΔPst -*BssH* plasmid). Additionally, *BssHII* digestion followed by removal of cohesive

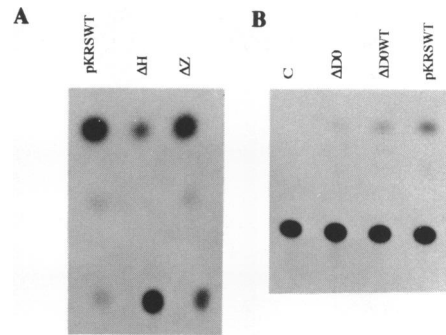


Figure 2. CAT-assays of different derivatives of the c-Ki-ras promoter. (A) pKRSWT has the *cat*-gene under the control of the whole promoter shown in Fig. 1, ΔH lacks the homopurine-homopyrimidine stretch and ΔZ lacks the d(C-G)₅ repeat. (B) pKRSWT: wild-type promoter, ΔDO has the large internal deletion diagrammed in Fig. 1, $\Delta DOWT$ was obtained from ΔDO by adding the putative H motif, (C) shows the level of CAT-activity in non-transfected cells.

termini with S1 nuclease yielded a promoter derivative with a short d(C-G)₂ repeat unable to form Z-DNA (ΔZ). We also made two additional internal deletions $\Delta BssHII$ -*ClaI* and $\Delta PstI$ -*ClaI*.

The promoter activities exhibited by these constructs are presented in Fig. 1 and 2A. While the deletion of the Z-forming sequence does not lead to a significant decrease in promoter activity, the loss of the H-forming sequence leads to a 2-3 fold decrease. The loss of only the H-motif has the same consequence for promoter activity as the larger *PstI*-*BssHII* and *PstI*-*ClaI* deletions. Conversely, the *BssHII*-*ClaI* deletion does not detectably influence promoter activity. These results clearly show that the H-motif is responsible for the transcriptional drop observed with the *PstI*-*ClaI* truncation. A previous report (8) showed that the deletion of the homopurine-homopyrimidine stretch drops transcription to almost the background level. Our results contradict this since we observed only a moderate 2-3 fold decrease in promoter activity. This contradiction can be explained because a much bigger deletion was actually used in the previous report. After sequencing the promoter derivative described in (8), we found that the construct has suffered a large internal deletion from the most upstream *BssHII* site to the *SstII* site (ΔDO , Fig. 1). This deletion removed all essential c-Ki-ras promoter elements except for the transcription start sites. The promoter activity exhibited by this derivative was only 3-5 % of the original promoter activity (Fig. 2B). Note that the results in Fig. 2A and 2B were obtained in independent transfection experiments with different extracts and have consequently different percentages of chloramphenicol conversion for pKRSWT. However, the pKRSWT conversion efficiency, normalized to β -gal activity, is similar in both sets of experiments.

The obvious disadvantage of deletion analysis is that deletions change the spatial relationships between the promoter elements as well as removing DNA sequences. One can argue that the effect of the H-motif deletion may result from promoter reorganization rather than from a direct functional role of this sequence. Though this seems unlikely due to the consistent promoter activity of constructs containing deletions and truncations differing in length and origin, this concern remained to be addressed. To this end, we cloned synthetic oligonucleotides corresponding to the H-motif into the ΔDO derivative described

homopurine–homopyrimidine stretch is vital for H-DNA formation (23), so the introduction of point substitutions in either half of such a stretch destroys mirror symmetry and consequently any H-forming ability. Combining symmetric substitutions together will restore the mirror symmetry and H-forming potential, though the resultant sequence is different from that of the original H-motif. We previously used this approach for different DNA sequences to prove the H-DNA model (23) as well as to elucidate its role in biological processes (24).

The H-motif mutants we constructed are presented in Fig. 3A. The original sequence (WT) is a perfect mirror repeat with a 4-bp loop at the pseudosymmetry site. Mutant DM1 has two C-to-T transitions in the left shoulder of the repeat, while DM2 has the corresponding substitutions in the right shoulder. QM combines all four substitutions to restore the mirror symmetry. Note that all four sequences still represent homopurine–homopyrimidine domains. To construct such mutants we benefited from the existence of convenient restriction sites (*Pst*I and *Bss*HI) flanking the area of interest. We cloned synthetic oligonucleotides corresponding to the whole *Pst*I–*Bss*HI stretch bearing various point substitutions within the H motif but with intact flanking sequences.

To check for H-DNA formation in all cases, we chemically probed the cloned mutant promoters. As we previously described (21), formation of H-DNA in the native *c-Ki-ras* promoter leads to hypersensitivity of a half of the purine strand and the center of the pyrimidine strand to single-strand-specific chemicals. The chemical reactivity for the H-motif mutants is presented in Fig. 3B. For QM, the reactivity is indistinguishable from that of the original (WT) promoter (as presented in 21), i.e. the center of the pyrimidine strand reacts with OsO_4 and chloroacetaldehyde, and the 5'-part of the purine strand reacts with diethyl pyrocarbonate and chloroacetaldehyde. Thus, the H-y3 isoform (Fig. 4C) is formed both by the original promoter and the quadruple mutant QM. DM2 displayed no chemical hyperreactivity (data not shown), which demonstrates its inability to form H-DNA due to point substitutions. DM1 displays weak chemical reactivity, in a pattern differing from both wild type and QM. In this case we observed reactivity towards diethyl pyrocarbonate and chloroacetaldehyde in the center of the purine strand and towards osmium tetroxide within the 5'-part of the pyrimidine strand. There is also a pronounced hyperreactivity immediately upstream of the repeat. This pattern is characteristic of another intramolecular triplex, *H-DNA, which consists of two purine and one pyrimidine strands (25). Usually it is formed by homopurine–homopyrimidine mirror repeats in the presence of bivalent cations. It was not immediately clear why *H-DNA was formed, since DM1 is asymmetrical, and we probed in the absence of bivalent cations. However, Malkov *et al.* (26) recently observed the formation of a protonated purine/purine/pyrimidine triplex composed of CG^*G , TA^*A and CG^*A^+ triads. This triplex was stable under acidic pH and in the absence of bivalent cations, i. e. under the conditions we used. The formation of the H-r3 isoform by DM1 explains its modification pattern quite well, as illustrated in Fig. 3C. The modification that was observed upstream of the repeat could be easily explained as the extension of unwinding of the 5'-part of the pyrimidine strand into adjacent region. The inability of DM2 to form a similar protonated *H-DNA is not surprising, because it must adopt an unfavorable H-r5 isoform. Fig. 3C summarizes the structural features of the

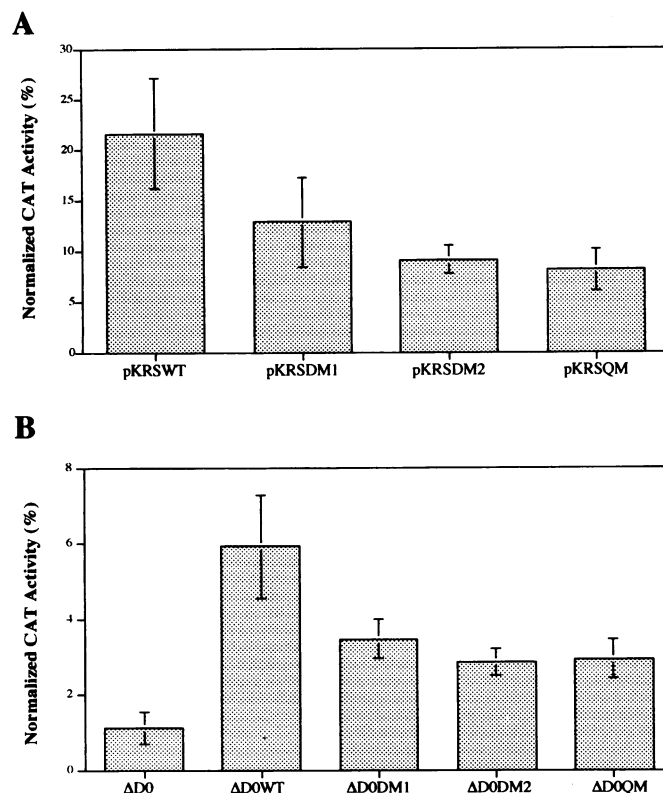


Figure 4. CAT-activity of the *c-Ki-ras* promoter carrying mutations within the H motif. (A) Derivatives of the original promoter. (B) Derivatives of the minimal promoter (ΔD0). Each graph represents the results of at least three transfections with independently isolated DNA samples, where transfection efficiency was normalized by β -galactosidase assay. For each transfection CAT-activity was analyzed in three independent extracts. CAT activity depends linearly on the amount of extract and incubation time for the range of values used here, so the results of several experiments were adjusted accordingly.

four versions of the homopurine–homopyrimidine stretch: the wild type and QM sequences form H-DNA, DM1 forms a protonated *H-DNA, and DM2 is unable to adopt any H-like structure.

We then evaluated the activity of these mutant promoters in transient transfection assays. If the H-forming potential of the homopurine–homopyrimidine stretch is essential for its promoter activity, the wild type and QM promoters should possess high promoter activity, while DM1 and DM2 should have low activity. In contrast, if sequence specificity is the issue, then all mutants should show some drop in transcriptional activity relative to the original promoter. The results are presented in Fig. 4A and illustrate that the promoter activities of DM2 and QM are very similar to each other, and are approximately a half of that of the wild type construct ($p \approx 0.04$). DM1 shows an intermediate 60% level of activity but this difference is not statistically significant. Thus, we find no correlation between the H-DNA-forming potential of these promoters and their transcriptional activity. Note, that the QM version of the promoter has a deletion of one GC-repeat downstream of the H-motif, making the Z-forming stretch 8 bp- rather than 10 bp-long. We don't believe, however, that this small deletion influences our interpretation of

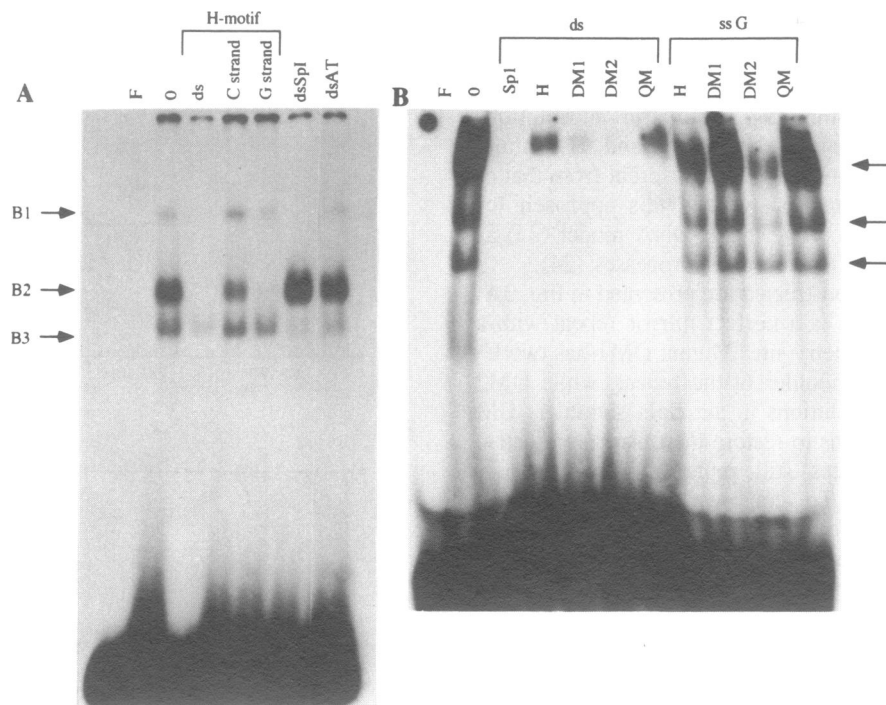


Figure 5. Electrophoretic mobility-shift assay (EMSA) for the H-motif and Sp1 probe incubated with HeLa nuclear extract. (A) H-motif EMSA. F: free probe; 0: no specific competitor; H-motif as specific competitor was used in double- (ds) or single-stranded states; dsSp1: double stranded DNA with 3 Sp1 consensus sites as a competitor; dsAT: double-stranded non-specific DNA. B1, B2 and B3 are specific H-motif–protein complexes. (B) Sp1 probe EMSA. F: free probe; 0: no specific competitor; ds: double-stranded DNA competitors; ssG: G-rich strand of H-motifs as competitors; H: wild-type H-motif; DM1, DM2 and QM: mutant H-motifs; arrows indicate specific Sp1–DNA complexes.

the transcriptional data, since the deletion of the whole Z-forming stretch does not affect promoter activity (Fig. 1).

Because the DM1 results were not statistically different, we decided to study the influence of these H-motif derivatives in the minimal version of the *c-Ki-ras* promoter ($\Delta D0$) where we expect the background to be lower. We cloned oligonucleotides corresponding to the wild type, DM1, DM2 and QM modifications of the homopurine–homopyrimidine stretch into the $\Delta D0$ construct (see Fig. 1) and analyzed their promoter activities by the CAT assay. Fig. 4B shows the results obtained with these minimal constructs, which are in accord with the original promoter data shown in Fig. 4A. DM2 and QM similarly exhibit a 2-fold decrease in promoter activity ($p < 0.075$), while DM1 is at an intermediate 63% level ($p \approx 0.1$). Note, however, that promoter activity is well above background in all cases. Thus, even the mutated homopurine–homopyrimidine stretches can meet transcriptional requirements to some extent. However, a direct correlation between the H-forming ability of these elements and their ability to promote transcription is not observed. Since all mutations decreased promoter activity to some extent, we conclude that the sequence integrity of the H-motif, rather than its structural peculiarities, is important in promoting *c-Ki-ras* transcription.

Nuclear proteins bind to the H-motif

The H-motif sequence requirement for transcription implies that this domain represents a target for the specific binding of activating protein(s). We used an electrophoretic mobility shift assay (EMSA) to look for H-motif-binding nuclear proteins. To

make a probe for these assays, we annealed synthetic oligonucleotides corresponding to the two strands of the H motif, and filled in cohesive ends with the Klenow fragment of DNA polymerase to present the H-motif in a double-helical state. We then incubated this probe in the presence of either nuclear or cytoplasmic cell extracts from human HeLa or HepG2 cells, or from mouse Y1 cells together with several specific and non-specific competitor DNA sequences. The samples were run on a low percentage PAGE as described in (27). We observed prominent band-shifts only in the presence of nuclear extracts, and there were no differences in the band shift patterns obtained using nuclear extracts from mouse versus human cell lines, in accord with the results obtained in the transient expression assays described above.

Fig. 5A shows typical EMSA results obtained using HeLa nuclear extract. We observed three major retarded complexes in the presence of non-specific dI/dC competitor, which are designated B1, B2 and B3. This binding pattern is similar to that previously reported in (8). However, the specificity of these complexes, as demonstrated in competition assays, differ from the previous results in significant details. A nonspecific double-helical DNA segment of similar length (dsAT) showed no ability to compete for complex formation, while unlabeled double-helical H-motif DNA (dsH) competed efficiently. The formation of complexes B1 and B3 was competitively inhibited by another double-helical oligonucleotide (dsSp1, see below). However the formation of B2 was inhibited only by the unlabeled H-motif DNA. Therefore, we believe that B2 represents the specific H-motif–protein complex. In the previous study, the B2 complex

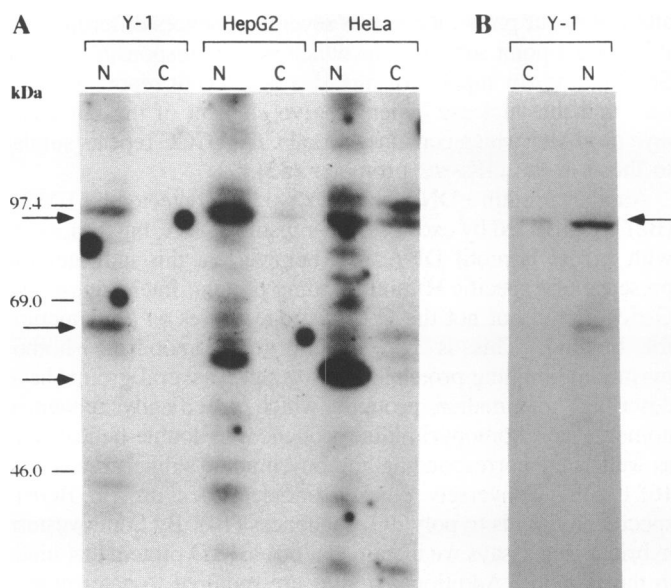


Figure 6. Detection of proteins binding to the H-motif DNA by Southwestern blot assay. (A) H-motif as probe. (B) Sp1 consensus as a probe. N: nuclear and C: cytoplasmic extracts of the Y-1, HepG2 and HeLa cells.

was considered to be non-specific, which is clearly contradicted by our results. The reasons of this contradiction are yet to be discovered.

We also used purine- and pyrimidine-rich strands of the H-motif as competitor. While the pyrimidine-rich strand was not an effective competitor, the purine-rich strand competed out B2 complex formation. This indicates that the protein(s) that recognizes the H motif in a double-helical state has specific affinity to its G-rich strand. Several other proteins known to bind homopurine-homopyrimidine stretches have been reported to prefer either their purine- or pyrimidine-rich DNA strands (13, 14, 16, 17, 19). Finally, in light of our transcription results, it was important to use double-stranded oligonucleotides corresponding to DM1, DM2 and QM mutants as competitors. It appeared, that all three DNAs compete out the three complexes (data not shown). These results are consistent with the transcriptional studies with the minimal version of the *c-Ki-ras* promoter showing that mutated stretches can promote transcription, though less efficiently than the wild-type H motif.

The nature of the B1 and B3 complexes is different. A DNA fragment containing three binding sites for transcription factor Sp1 (dsSp1) efficiently competes with the H-motif probe for complex formation, even though the consensus Sp1 binding site 5'-GGGCGG does not exist in our DNA probe. However, our probe does contain sequences that differ from this consensus by single point substitutions: (i) five copies of the sequence 5'-GGG-AGG and (ii) one copy of the sequence 5'-GGGCGC. These sequences may be responsible for Sp1 binding. In a previous study, however, it was claimed that Sp1 does not bind the H motif of the *c-Ki-ras* promoter (8).

To address this controversy in more detail, we provided EMSA experiments with HeLa nuclear extracts and double-stranded DNA containing three Sp1 consensus sites as a probe. Fig. 5B shows that three specific DNA-protein complexes are observed

(indicated by arrows). The decreasing mobility of these complexes is most likely due to binding of one, two and three Sp1 molecules, respectively, to the DNA probe. An excess of a 'cold' Sp1 DNA eliminates all three complexes. An excess of DNA samples corresponding to the wild-type H-motif and DM1, DM2 and QM mutants, compete for Sp1 binding, though somewhat less efficiently than the Sp1 probe itself (these differences are most pronounced for the slowest complex). The G-rich strand of different H-motifs shows no competition for Sp1 binding. These data are in accord with the above results showing that Sp1 DNA removes H-motif-protein complexes B1 and B3. We are convinced, therefore, that Sp1 protein can bind to the H-motif from the *c-Ki-ras* promoter, as well as to its mutated versions.

We also detected H-motif-binding proteins using a Southwestern blot analysis (28), which is based on the fact that many proteins maintain their specific DNA-binding affinities even after separation by denaturing gel-electrophoresis. Therefore, binding of labeled DNA targets with separated proteins can indicate the presence and apparent molecular weight of specific binding proteins. We separated nuclear and cytoplasmic protein extracts from mouse Y1 cells, and human HepG2 or HeLa cells on a 10% SDS-PAGE and transferred the proteins onto a nitrocellulose membrane, which we then incubated with the ³²P-labeled double-helical H-motif probe described above. Fig. 6A shows two major H-motif-binding proteins detected in these nuclear extracts. The larger protein has an apparent mass of 97 kD in both mouse and human cells, while the smaller protein has an apparent mass of 60 kD in human and 64 kD in mouse cells. The relative intensities of these bands is similar, except in the HeLa cell line where the smaller protein is more prominent. We suspect that the major 60 kD protein picked up by Southwestern blot analysis may correspond to the major gel-shifted complex B2 in HeLa nuclear extracts.

Since the mass of the Sp1 protein in HeLa cell is approximately 97 kD, it is possible that the larger protein corresponds to Sp1. In support of this conclusion, Fig. 6B shows a major binding protein of 97 kD detected when the experiment was performed by incubation of the membrane with an Sp1-specific probe.

Both the gel-retardation assays and Southwestern blot studies reveal at least two specific H-motif-protein complexes. One complex is most likely to contain Sp1 based on the competition experiments and the molecular mass of Sp1. The identity of the 60-64 kD protein is yet to be determined. The binding data fit well with the transcriptional results, in that we observe a similar protein pattern in extracts of both human and mouse cells.

DISCUSSION

Our previous studies indicated that the upstream sequences present within the mouse *c-Ki-ras* promoter may exist in either H or Z conformation *in vitro* depending on the ambient conditions (21). Here, we studied the potential role of H- and Z-forming DNA stretches in promoter function using a transient transcription assay. Our data show that while the Z-forming sequence is not essential, the H-forming stretch is responsible for approximately half of the promoter activity detected. Independent experiments with the H-motif inserted into a minimal version of the promoter demonstrated that it plays an important role in *c-Ki-ras* transcription as a positive cis-acting element. Similar conclusions have previously been made for homopurine-homopyrimidine

stretches located within different eukaryotic promoters (3–7, 9–11).

What was previously unclear is whether these sequences themselves, or their unusual structure, account for transcriptional activation function. This issue has been extensively discussed over the last 10 years. An early hypothesis by Weintraub (29) suggested that the single-stranded DNA portions of such structures could serve as an entry point for RNA polymerase. A more recent hypothesis proposes that local changes in DNA structure may regulate the interaction between promoter DNA and specific DNA-binding proteins (reviewed in 2). For example, a protein called NSEP-1 has been identified (16), which binds to the S1-hypersensitive site in the human *c-myc* gene. It also binds to non-homologous homopurine–homopyrimidine stretches from other promoters, including *c-Ki-ras*, the insulin receptor gene (IR) and the epidermal growth factor receptor gene (EGFR), implying that NSEP-1 recognizes the common structural features of these sequences, possibly their ability to form H-DNA. The partial purification of a protein that preferentially binds triple-helical DNA has also been reported (30). In a recent study (31), in which the influence of d(G)_n stretches of varying length on the activity of a downstream minimal promoter was analyzed, it was shown that while initial increases in the length of the d(G)_n stretch caused increasing activation of transcription, further increases restored the minimal promoter activity. Notably, there was a clear reverse correlation between the ability of a stretch to form *H configuration *in vitro* and its ability to activate transcription *in vivo*. The authors proposed that short d(G)_n stretches serve as binding sites for a transcriptional activator, while longer stretches adopt a triplex configuration which prevents activator binding. No direct proof of the role of H-DNA in transcription was available. However, there were indications that sequence- rather than structure-specific recognition of the homopurine–homopyrimidine stretches by transcription factors is essential for promoter activity in several cases, including the *Drosophila hsp26* (11) and human *c-myc* (32) genes.

We addressed this question by careful structure function analysis of the mutational derivatives of the murine *c-Ki-ras* promoter. Different point substitutions introduced into the homopurine–homopyrimidine stretch destroyed, or restored, its H-forming ability as shown by chemical probing. However, the results obtained in subsequent transcription studies *in vivo* showed no correlation between H-forming ability and promoter activity. It appears then, that the sequence integrity of the H-motif is crucial for transcription. Our data were obtained in a transient expression assay. As transcriptional regulation may be different in a chromosomal environment, future experiments with stable transfectants may be required to further substantiate this conclusion.

Sequence specific recognition is a property of DNA-binding proteins, and we have found evidence for at least two proteins that specifically bind to the H-motif, as assayed by gel-retardation and Southwestern hybridization assays. Note that we used linear DNA as a probe for these studies, preventing the formation of any intramolecular triplexes within the free probe. One of these proteins is probably Sp1, since an Sp1 binding sequence efficiently eliminates two of the retarded complexes (B1 and B3) formed with H-motif while H-motif sequence competes for the Sp1 binding to its consensus sequence. In addition, the molecular mass of one protein revealed by Southwestern hybridization corresponds to that of Sp1. Although there is no consensus Sp1

site within our probe, it contains several sequences differing from it by single point substitutions which may be responsible for the Sp1 binding. In support of this idea is a recent report that Sp1 can bind the nuclease-hypersensitive element of the human *c-myc* promoter which contains several CCCTCCC repeats similar to those in the *c-Ki-ras* promoter (32).

Another protein–DNA complex we have detected by EMSA (B2) is unaffected by excess Sp1 competitor DNA, but disappears with excess H-motif DNA. We believe that this indicates the presence of a specific H-motif-binding protein. Interestingly, the G-rich strand but not the C-rich strand serves as a competitor for binding. This is not unusual for homopurine–homopyrimidine-binding proteins. For example, several groups have described mammalian proteins which specifically recognize homopurine–homopyrimidine sequences in double-helical state as well as the corresponding homopyrimidine single strands (13, 16, 17, 19). Conversely, chicken zinc-dependent protein, BGP1, specifically binds to poly(dG) sequences (14). By Southwestern hybridization assays we identified a 60–64 kD protein that binds to the H-motif. Additional studies are required to determine if it is responsible for the specific retardation of the H-motif probe observed in the EMSA. Several homopurine–homopyrimidine-binding proteins have previously been described including BPG1 (14), NSEP-1 (16), MAZ (12), nm23-H2 (33), PYBP (13), Pur-1 (15) and others. Though all of these proteins differ in the molecular weight from 60–64 kD, the exact relationship between these proteins and our 60–64 kD protein remains to be studied.

Based on the data presented here we conclude that the sequence integrity of the H-motif is required for full transcriptional activity of the *c-Ki-ras* promoter. This most probably is due to the binding of specific transcriptional activator proteins binding to the H-motif DNA. This DNA motif is a 27-bp-long sequence, unusually long for a protein-binding site. It is also very redundant, containing 6 CCCT repeats and 1 CCTT stretch. It is striking, that point substitutions in only two repeats cause almost the same effect as the deletion of the whole stretch. One possibility is that the H-motif is a target for a protein with several Zn-fingers which recognize CCCT repeats, explaining why the integrity of all of the repeats is required for trans-activation.

ACKNOWLEDGEMENTS

We thank Angela Tyner for her invaluable help with the tissue cultures and transfection experiments, Victor Lobanekov for his helpful discussions, and Randal Cox for his help with statistics and proofreading the manuscript. This work is supported by grants MG-25 from the American Cancer Society and R55GM46405-01A1 from the National Institutes of Health to S. M. M. and a grant from the Campus Research Board of the University of Illinois at Chicago to K. N. S.

REFERENCES

1. Wells, R.D., Collier, D.A., Hanvey, J.C., Shimizu, M. and Wohlrab, F. (1988) *FASEB J.* 2, 2939–2949.
2. Mirkin, S.M. and Frank-Kamenetskii, M.D. (1994) *Annu. Rev. Biophys. Biomol. Struct.* 23, 541–576.
3. Johnson, A.C., Jinno, Y. and Merlino, G.T. (1988) *Mol. Cell. Biol.* 8, 4174–4184.
4. Mavrothalassitis, G.L., Watson, D.K. and Papas, T.S. (1990) *Proc. Natl. Acad. Sci. USA* 87, 1047–1051.
5. Tewari, D.S., Cook, D.M. and Taub, R. (1989) *J. Biol. Chem.* 264, 16238–16245.

6. Davis, T.L., Firulli, A.B. and Kinniburgh, A. (1989) *Proc. Natl. Acad. Sci. USA* 86, 9682–9686.
7. Postel, E.H., Mango, S.E. and Flint, S.J. (1989) *Mol. Cell. Biol.* 9, 5123–5133.
8. Hoffman, E.K., Trusko, S.P., Murphy, M. and George, D.L. (1990) *Proc. Natl. Acad. Sci. USA* 87, 2705–2709.
9. Lafyatis, R., Denhez, F., Williams, T., Sporn, M. and Roberts, A. (1991) *Nucleic Acids Res.* 19, 6419–6425.
10. Gilmour, D.S., Thomas, G.H. and Elgin, S.C.R. (1989) *Science* 245, 1487–1490.
11. Glaser, R.L., Thomas, G.H., Siegfried, E., Elgin, S.C. and Lis, J.T. (1990) *J. Mol. Biol.* 211, 751–761.
12. Bossone, S.A., Asselin, C., Patel, A.J. and Marcu, K.B. (1992) *Proc. Natl. Acad. Sci. USA* 89, 7452–7456.
13. Brunel, F., Alzari, P.M., Ferrara, P. and Zakin, M.M. (1991) *Nucleic Acids Res.* 19, 5237–5245.
14. Clark, S.P., Lewis, C.D. and Felsenfeld, G. (1990) *Nucleic Acids Res.* 18, 5119–5126.
15. Kennedy, G.C. and Rattner, J.B. (1992) *Proc. Natl. Acad. Sci. USA* 89, 11498–11502.
16. Kolluri, R., Torrey, T.A. and Kinniburgh, A.J. (1992) *Nucleic Acids Res.* 20, 111–116.
17. Muraiso, T., Nomoto, S., Yamazaki, H., Mishima, Y. and Kominari, R. (1992) *Nucleic Acids Res.* 20, 6631–6635.
18. Roy, C. and Lebleu, B. (1991) *Nucleic Acids Res.* 19, 517–524.
19. Yee, H.A., Wong, A.K.C., van de Sande, J.H. and Rattner, J.B. (1991) *Nucleic Acids Res.* 19, 949–953.
20. Barbacid, M. (1987) *Annu. Rev. Biochem.* 56, 779–827.
21. Pestov, D.G., Dayn, A., Siyanova, E., George, D.L. and Mirkin, S.M. (1991) *Nucleic Acids Res.* 19, 6527–6532.
22. Hall, C.V., Jacob, P.E., Ringold, G.M. and Lee, F. (1983) *J. Mol. Appl. Genet.* 2, 101–109.
23. Mirkin, S.M., Lyamichev, V.I., Drushlyak, K.N., Dobrynin, V.N., Filippov, S.A. and Frank-Kamenetskii, M.D. (1987) *Nature* 330, 495–497.
24. Samadashwily, G.M., Dayn, A. and Mirkin, S.M. (1993) *EMBO J.* 12, 4975–4983.
25. Kohwi, Y. and Kohwi-Shigematsu, T. (1988) *Proc. Natl. Acad. Sci. USA* 85, 3781–3785.
26. Malkov, V.A., Voloshin, O.N., Veselkov, A.G., Rostapshov, V.M., Jansen, I., Soyfer, V.N. and Frank-Kamenetskii, M.D. (1993) *Nucleic Acids Res.* 21, 105–111.
27. Plumb, M.A., Lobanenkov, V.V., Nicolas, R.H., Wright, C.A., Zavou, S. and Goodwin, G.H. (1986) *Nucleic Acids Res.* 14, 7675–7693.
28. Patel, S.B. and Cook, P.R. (1983) *EMBO J.* 2, 137–142.
29. Weintraub, H. (1985) *Mol. Cell. Biol.* 5, 1538–1539.
30. Kiyama, R. and Camerini-Otero, D. (1991) *Proc. Natl. Acad. Sci. USA* 88, 10450–10454.
31. Kohwi, Y. and Kohwi-Shigematsu, T. (1991) *Genes and Devel.* 5, 2547–2554.
32. DesJardins, E. and Hay, N. (1993) *Mol. Cell. Biol.* 13, 5710–5724.
33. Postel, E.H., Berberich, S.J., Flint, S.J. and Ferrone, C.A. (1993) *Science* 261, 478–483.
34. Gorman, C.M., Moffat, L.H. and Howard, B.H. (1982) *Mol. Cell. Biol.* 2, 1044–1051.
35. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: a Laboratory Manual* - 2nd ed. Cold Spring Harbor Laboratory Press. New York.
36. Dignam, J.D., Lebovitz, R.M. and Roeder, R.G. (1983) *Nucleic Acids Res.* 11, 1475–1489.
37. Bradford, M.M. (1976) *Anal. Biochem.* 72, 248–254.