acid identity between the available 20 amino acid sequence of mouse PIF and the N-terminal 20 amino acid sequence of DCD [10]. However, the identity of DCD as a human cachectic factor and the precise nature of the post-translational modifications of human PIF molecule, which seems to be tumor cell specific, have not been fully elucidated [10–12]. Here we identified DCD in human placental tissue and determined its previously uncharacterized proteolytic activity against peptide and protein substrates. Recombinant DCD induced an invasive phenotype in a human choriocarcinoma cell line JAR in vitro. The present work suggests that, by means of modulating the pericellular proteolytic cascades, DCD might participate in the regulation of trophoblast cell behaviors, and might be involved in the pathophysiology of pregnancy-related disorders such as intrauterine fetal growth retardation and preeclampsia, as well as cancer and neuronal diseases.

Materials and methods

Protein purification. Placental tissue (approximately 500 g wet weight) obtained after full term uncomplicated vaginal delivery was immediately frozen in liquid nitrogen, pulverized and boiled in four volumes of water for 5 min. Acetic acid was added at a final concentration of 0.1 N and homogenized with a Polytron homogenizer. The soluble fraction of 80% acetone was concentrated and fractionated using a μ-Bondapack C18 column (Waters, 7.8 × 300 mm, 0–50% CH_3CN gradient for 50 min and 50–100% CH_3CN gradient for 20 min in 0.1% TFA at 2 ml/min). The active fractions were separated with a TSK gel CM-2SW column (Tosoh, 4.6 × 250 mm, 10–500 mM $HCOONH_4$ gradient for 50 min and 500 mM–1.0 M $HCOONH_4$ gradient for 20 min in 10% CH_3CN at 1 ml/min) and a YMC-PACK C-8 AP column (YMC, 4.6 × 250 mm, 21–42% CH_3CN gradient for 60 min in 0.1% TFA at 1 ml/min). Protein sequence was determined by Applied Biosystems model 474A protein sequencer and HP G1005A Protein Sequencing System (Hewlett–Packard). All human tissue was collected using protocols approved by the institutional review boards.

Radioreceptor assay. Radioreceptor assay (RRA) for oxytocin was performed as described previously [13] with minor modifications. Crude membrane fraction was obtained from uterine tissue removed from the lower incision flap of a term cesarean section. Reaction mixture, consisting of 7.0 μg of membrane protein, 7000–8000 cpm of ^{125}I-ornithine vasotocin analogue (New England Nuclear) and varying concentrations of non-labeled peptides, were incubated for 60 min at room temperature. Bound and free ligands were separated by filtration over a PVDF membrane under reduced pressure (Multiscreeen 96, Millipore).

Cloning. A Blast search specified a nucleotide sequence on chromosome 12 encoding a partial amino acid sequence (KDAVEDL) obtained from purified placental tissue (bold letters in Fig. 1D). Genes were predicted computationally around the specified 21 bp nucleotide sequence on chromosome 12. FGENES predicted a gene of 333 bp encoding the amino acid sequences obtained from the purified placental tissue. This was confirmed experimentally using polymerase chain reaction with reverse transcription (RT-PCR) of the total RNA extracted from human term placenta. The specific primers used were cagccagcctttgttgactt and aggacgtctttaacgtcatgg for the first 30 cycles, and gttcccattggtccctgtc and aggacgtctttaacgtcatgg for nested amplification. PCR was performed using

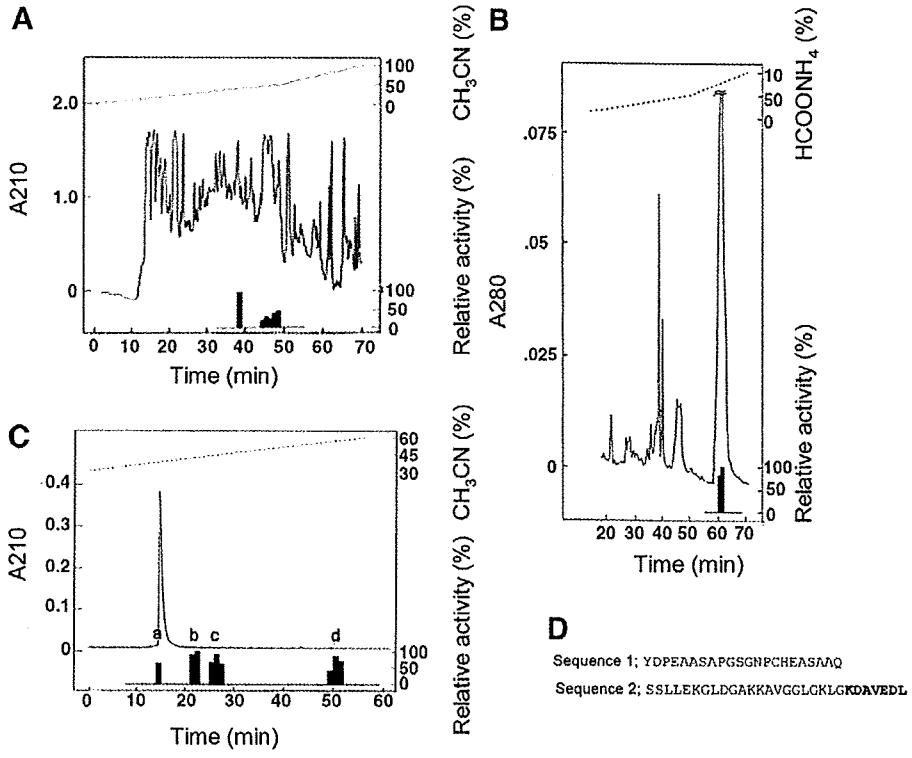


Fig. 1. Purification of DCD protein from human term placenta. (A–C) HPLC purification of DCD protein from acid extracts of human term placenta. The active fractions of preparative HPLC eluting at 44–47 min (A) were separated with ion exchange HPLC (B). The active fractions at 60–62 min (B) were separated with reverse phase HPLC (C). Filled bars represent relative activities of the corresponding HPLC fractions determined by RRA for oxytocin. (D) Edman degradation of the fraction d (C) yielded two sequences (sequence 1 and sequence 2) in equimolar amounts. Non-overlapping amino acid sequence is depicted in bold letters from which DCD gene was predicted in silico.

Advantage PCR Kits (Clontech) and 35 cycles of: 94 °C for 30 s, 60 °C for 2 min, 68 °C for 3 min. 5′UTR was amplified by nested PCR using the primers cagccagcctttgttgactt and tcttcgactgcatcttttcct for the first PCR, and cagccagcctttgttgactt and ttgctccgtctaggcctttt for nested amplification to produce a single 410 bp band. The start of the open reading frame (ORF) was confirmed by the presence of in-frame stop codons upstream of the first ATG. 3′RACE was carried out according to the manufacturer's instructions (Life Technologies). The specific sense primers were cctcttcctgacagctctgg for first amplification and aaaaggcctagacggagca for nested PCR. PCR was performed using *Taq* (Takara Shuzo) and 30 cycles of: 94 °C for 1 min, 55 °C for 1 min, 72 °C for 1 min. RT-PCR of total RNA extracted from first trimester human placental villi was performed using the specific primers gttcccattggtccctgtc and aggacgtctttaacgtcatgg for the first 30 cycles and cctcttcctgacagctctgg and aggacgtctttaacgtcatgg for nested amplification. Placental tissues were snap-frozen in liquid nitrogen in the delivery or the operating room immediately upon the delivery of the placenta and stored at −80 °C until further analysis.

Protein expression. cDNA constructs for expression in *Escherichia coli* were made by PCR using Pfu polymerase. They were sequenced completely and contained no mutations. The constructs encoding mature DCD were cloned into the LIC cloning vector to generate recombinant BL21 (DE3) cells according to the manufacturer's protocol (Stratagene). Cells were grown at 37 °C to $A_{600} = 0.6$ and induced with 1 mM isopropyl-thiogalactopyranoside for 4 h. Recombinant protein was extracted from the cells by sonication and purified through calmodulin affinity resins according to the supplier's instructions (Stratagene). To remove N-terminal calmodulin binding protein, 100 μg of recombinant DCD was incubated with 1.0 IU of recombinant enterokinase (rEK) at room temperature for 16 h according to the manufacturer's instructions (Stratagene). Cleavage was monitored by SDS–polyacrylamide gel electrophoreses (SDS–PAGE) and surface enhanced laser desorption/ionization (SELDI) mass spectrometric analysis using ProteinChip® technology (Cyphergen Inc.). Protein and peptide samples were analyzed on reverse phase (H4) chips (Cyphergen Inc.).

Proteolytic activity of recombinant DCD. Proteolytic activity of recombinant DCD was first determined with 4-methylcoumaryl-7-amide (MCA)-containing substrates. The amount of 7-amino-4-methylcoumarine (AMC) released was spectrofluorophotometrically measured using a Wallac 1420 multilabel counter (Wallac) equipped with a 355-nm excitation and a 460-nm absorption filter. 3.0 nmol of MCA-containing substrate was incubated with 1.0 μg of recombinant DCD and 0.01 IU of rEK or 0.01 IU of rEK alone at 25 °C. Proteolytic processing of synthetic peptides by recombinant DCD was monitored by ProteinChip. 2.1 μg of metastin (45–54) was incubated with 1.0 μg of recombinant DCD and 0.01 IU of rEK at 37 °C for 3 h. Proteolytic processing of pro-matrix metalloproteinases (MMPs) was analyzed by gelatin zymography. Ten nanograms of pro-MMP-9 (Sigma) was incubated with 1.0 μg of recombinant DCD and 0.01 IU of rEK at 37 °C.

In vitro invasion assay. Human JAR cells were obtained from American Type Culture Collection. JAR cells (5×10^4 cells in 500 μl RPMI 1640) were seeded in a BioCoat Matrigel Invasion Chamber® (BD Biosciences) and incubated at 37 °C for indicated periods versus a lower chamber containing 10% FBS/RPMI 1640. Matrigel and non-invading cells on the upper surface were removed. Invading cells on the lower side of the membrane were fixed, stained with Giemsa and quantified. rEK was administered with recombinant DCD (1×10^{-2} IU of rEK to 1 μg of recombinant DCD). rEK alone did not affect cell invasion within the concentrations used. 1,10-Phenanthroline was used at a concentration of 100 μg/ml.

Results

Protein purification

DCD was purified from acid extracts of term human placental tissue. Our RRA for oxytocin demonstrated cross-reactivity with human neutrophil defensin 1 and 2 (data not shown). Using this assay to monitor each purification step (Fig. 1A–C), we purified defensins from human term placental tissue to near homogeneity (the active fractions eluting at 60–62 min in Fig. 1B), as determined by Edman degradation and matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry (data not shown). Further separating this fraction by reverse phase high performance liquid chromatography (HPLC), we isolated and obtained N-terminal amino acid sequences of placental DCD from fraction d in Fig. 1C. Edman degradation of placental DCD revealed two sequences in equimolar amounts (sequences 1 and 2 in Fig. 1D): sequence 1 is an N-terminal peptide of mature DCD protein and also constitutes the N-terminal peptide fragments of Y-P 30 and PIF (Supplementary Fig. 1E); and sequence 2 is an N-terminal part of an antimicrobial peptide DCD-1 (Supplementary Fig. 1E). From each of the active fractions a–c of the reverse phase HPLC in Fig. 1C, N-terminal amino acid sequences of human neutrophil defensin 1 and defensin 2 were obtained.

Cloning of DCD and its splice variants

DCD gene expression was confirmed by RT-PCR analysis of total RNA from human term placenta (Fig. 2A, lane 2). RT-PCR analysis demonstrated two splice variants of DCD in human placenta (Fig. 2A, lanes 1 and 3). The gene encoding the partial amino acid sequences obtained from the purified placental tissue (Fig. 1D) was first predicted in silico using FGENES and then confirmed experimentally

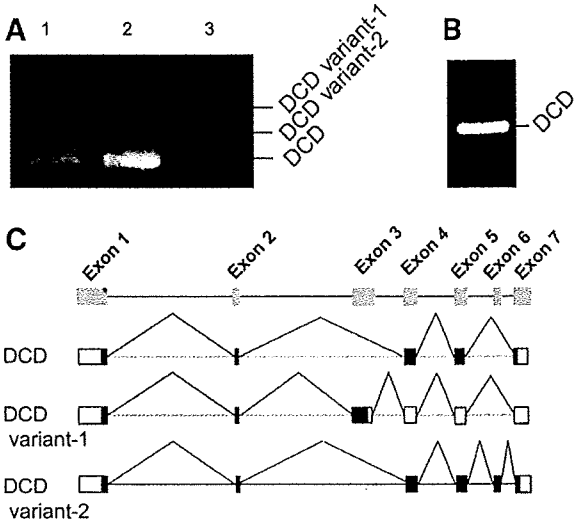


Fig. 2. Genes for DCD and its splice variants. (A) RT-PCR of the total RNA from equivalent portions of the villous parenchyma of different placentas (lanes 1 and 2) and from a reflected membrane of the placenta (lane 3). A representative result is shown. (B) RT-PCR of DCD from first trimester human placenta. (C) Schematic representation of exon/intron structures of DCDs. The protein coding regions are shown as filled bars.

by RT-PCR. An antimicrobial peptide DCD-1 from the human sweat glands was reported thereafter [4] and our cloned gene was proven to be identical to DCD. While the DCD gene was predicted in silico, DCD splice variants were not predicted by FGENES or other gene-predicting programmes such as Genscan. In human placenta, DCD and DCD variant-1 were expressed in villous parenchyma (Fig. 2A, lanes 1 and 2) and DCD variant-2 was expressed in the reflected membrane (Fig. 2A, lane 3). DCD splice variants were not detected in the first trimester human placenta by RT-PCR (Fig. 2B) or Southern blot analysis (data not shown). Genes for DCD and its splice variants are shown in Fig. 2C and Supplementary Fig. 2.

Characterization of recombinant DCD

During the characterization of the recombinant protein, we found that recombinant DCD possessed proteolytic activities against several peptide and protein substrates. Recombinant DCD was constructed by adding N-terminally calmodulin binding protein to mature DCD (Supplementary Fig. 1A). Upon cleavage of the N-terminal calmodulin binding protein by rEK, recombinant DCD was rapidly degraded to smaller peptide fragments, analyzed by SDS–PAGE (Supplementary Fig. 1B and C). Identification of these peptide fragments by SELDI mass spectrometric analysis revealed that limited proteolysis was taking place on the carboxyl terminal peptide bond of Lys and Arg residues (Supplementary Fig. 1D and E). Calmodulin binding protein was also processed to smaller peptide fragments (Supplementary Fig. 1D and E). Degradation of recombinant DCD was not inhibited by known protease inhibitors such as leupeptin (serine and cysteine protease inhibitor), E-64 (cysteine protease inhibitor), pepstatin (aspartic protease inhibitor), bestatin (aminopeptidase inhibitor), and EDTA (data not shown).

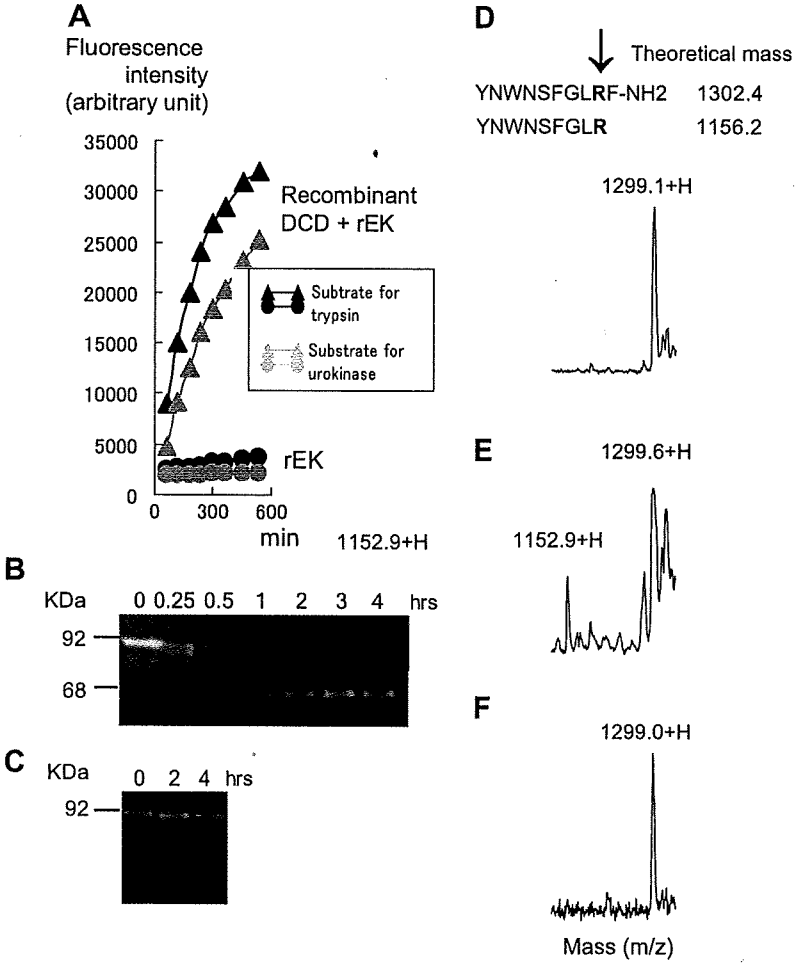


Fig. 3. Proteolytic activity of recombinant DCD. (A) MCA-containing substrates for trypsin (Boc-Gln-Ala-Arg-MCA) and urokinase (Glt-Gly-Arg-MCA) were incubated with recombinant DCD and rEK, or rEK alone. Fluorescence was measured at an excitation/emission wavelength of 355/460-nm. (B,C) Pro-MMP-9 (Sigma) was incubated with recombinant DCD and rEK (B) or rEK alone (C) and analyzed by gelatin zymography. Pro-MMP-9 was activated and, following autoprocessing, generated 68 KDa MMP-9. (D–F) Cleavage of synthetic metastin (45–54) by recombinant DCD. Amino acid sequence of metastin (45–54) is shown and cleavage site is indicated by an arrow. Synthetic metastin (45–54) (D) was incubated with either recombinant DCD and rEK (E) or rEK alone (F) and analyzed by ProteinChip. Recombinant DCD efficiently cleaved the Arg-Phe bond of metastin (45–54) to generate a 1152.9 Da mass fragment.

Proteolytic activity of recombinant DCD

To further analyze the proteolytic activity of recombinant DCD, we added various synthetic peptide and protein substrates to the above reaction mixtures. Fig. 3A shows amidolytic activity of recombinant DCD on synthetic MCA-containing substrates for trypsin and urokinase. This activity on various synthetic MCA-containing substrates indicated that the cleavage occurred only on the carboxyl-terminal side of the Arg and, to a lesser extent, Lys residues (Supplementary Table 1). We also analyzed proteolytic processing of peptide and protein substrates by recombinant DCD. Recombinant DCD effectively activated pro-MMP-9 in vitro (Fig. 3B and C). Recombinant DCD efficiently cleaved the Arg-Phe bond of metastin (45–54) (Fig. 3D–F) and the Arg-Trp bond of α-MSH (AcSYSMEHFRWGKPV-NH_2) (data not shown), but the cleavage of LHRH (<EHWSYGLRPG-NH_2) was not detected (data not shown). Proteolytic activation of pro-MMP-2 by recombinant DCD could not be detected (data not shown). Recombinant DCD did not degrade casein when analyzed by fluorescence-based EnzChek®protease assay kits (Molecular Probes) or when we analyzed the digestion of bovine milk casein by ProteinChip (data not shown). Recombinant DCD degraded neither gelatin when analyzed by gelatin zymography nor bovine serum albumin when analyzed by SDS–PAGE (data not shown).

Effects of DCD on the invasion of choriocarcinoma cell line JAR

To explore the physiological functions of DCD protein, we examined the effect of DCD on the invasiveness of choriocarcinoma cell line JAR in vitro, using a reconstituted basement membrane Matrigel. Addition of recombinant DCD resulted in a significant increase in the number of invading JAR cells through the Matrigel (Fig. 4A). 1,10-Phenanthroline significantly inhibited the effect of recombinant DCD on the invasion of JAR cells (Fig. 4B).

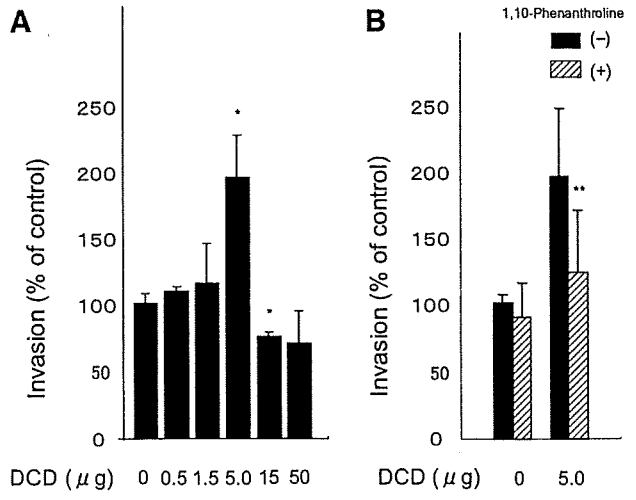


Fig. 4. Effects of DCD on invasion of JAR cells. (A) JAR cells were added to the upper chamber and incubated for 18 h, and invading cells through Matrigel were quantified. Various concentrations of recombinant DCD were added to the upper chamber. (B) Invasion of JAR cells after 18 h of incubation either with (hatched bars) or without (filled bars) 100 μg/ml of 1,10-phenanthroline. JAR cells were treated with or without recombinant DCD. $^{*}P < 0.01$ and $^{**}P < 0.05$ by Student's t-test.

Discussion

In this study, we isolated and purified DCD protein from human term placental tissue. The result of the Edman degradation of placental DCD showed two sequences in equimolar amounts (sequences 1 and 2 in Fig. 1D), suggesting the possibility of dimerization or oligomerization of placental DCD. While Y-P 30 was reported to have a tendency to form higher molecular weight aggregates [5], there has been limited data available regarding the structure or processing of mature DCD protein. Our present work offers valuable information on the probable ternary structure of DCD protein in human placenta and other tissues.

The expression pattern of DCD in humans is restricted [4,7,10] and the existence of the splice variants of the DCD gene has not been reported in human and other species investigated to date. In the present study, we identified two splice variants of DCD in human gestational tissue. The spatiotemporally restricted pattern of expression of these splice variants of DCD might suggest the involvement of these genes in the molecular mechanisms of human pregnancy.

Synthetic and recombinant DCD-1 peptide exhibits antimicrobial activities against a wide range of microbial pathogens [4,14]. Placental DCD might be processed to an antimicrobial peptide DCD-1 and, like other natural antibiotics including defensins [15], might defend against invading microorganisms that can result in chorioamnionitis, the leading cause of the premature delivery and subsequent fetal morbidity and mortality [16]. Here, by characterizing recombinant DCD, we unexpectedly identified proteolytic activity of DCD. Cleavage specificity of recombinant DCD seems to resemble that of trypsin-like serine proteinases and we could not exclude the influences of co-existing rEK. Although further characterization of the proteolytic activity of DCD is necessary, our study indicates that DCD could be involved in the regulation of the proteolytic activities on the trophoblastic cell surface.

In this study, we demonstrated that DCD might be involved in the induction of invasive phenotypes in the choriocarcinoma cell line JAR. Activation of latent metalloproteinases is implicated in the invasion of trophoblasts [17]. We showed that induction of invasiveness of JAR cells by recombinant DCD was inhibited by broad spectrum metalloproteinase inhibitor, suggesting that DCD might modulate metalloproteinase activities to induce invasive phenotypes. In xenograft experiments in rat brain, Cunningham et al. [6] reported that mouse cells that stably overexpressed DSEP/DCD were not immunosuppressed and

exhibited substantial invasion and migration from the lesion/injection site into white matter tracts and along blood vessels of the forebrain and thalamus of the host. Their experiments corroborate our present in vitro invasion experiments.

The invasion of trophoblasts is stringently controlled in normal human pregnancies, being confined spatially to the uterus and temporally to early pregnancy [18]. Reduced trophoblastic invasion is associated with serious complications of human pregnancy such as intrauterine fetal growth retardation and preeclampsia [19]. Although the role of DCD on cellular invasion in vivo remains to be determined, our present work suggests that DCD might regulate trophoblastic functions through modulation of the pericellular proteolytic activity and add another layer of refinement to the proteolytic cascades on the trophoblastic cell surface. While the detailed molecular characteristics of DCD requires further clarification, development of specific inhibitors or agonists could provide unique therapeutic approaches to pregnancy-related disorders, as well as cancer and neuronal disorders.

Acknowledgments

We thank T. Makino for comments on RRA for oxytocin. We thank S. Adachi, T. Shibata and the late M. Kobayashi for assistance with protein purifications and S. Abe for help with HPLC analysis. We thank T. Ubagai and K. Fujiwara for help with RT-PCR experiments. We thank S. Fujinoki for assistance with ProteinChip analysis and Y. Nobukuni for help with cell culture experiments.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbrc.2007.03.112.

References

[1] C. Lopez-Otin, C.M. Overal, Protease degradomics: a new challenge for proteomics, Nat. Rev. Mol. Cell Biol. 3 (2002) 509–519.
[2] Z. Werb, ECM and Cell surface proteolysis: regulating cellular ecology, Cell 91 (1997) 439–442.
[3] M. Egeblad, Z. Werb, New functions for the matrix metalloproteinases in cancer progression, Nat. Rev. Cancer 2 (2002) 161–174.
[4] B. Schittek, R. Hipfel, B. Sauer, J. Bauer, H. Kalbacher, S. Stevanovic, M. Schirle, K. Schroeder, N. Blin, F. Meier, G. Rassner, C. Garbe, Dermcidin: a novel human antibiotic peptide secreted by sweat glands, Nat. Immunol. 2 (2001) 1133–1137.
[5] T.J. Cunningham, L. Hodge, D. Speicher, D. Reim, C. Tyler-Polsz, P. Levitt, K. Eagleson, S. Kennedy, Y. Wang, Identification of a survival-promoting peptide in medium conditioned by oxidatively stressed cell lines of nervous system origin, J. Neurosci. 18 (1998) 7047–7060.
[6] T.J. Cunningham, H. Jing, I. Akerblom, R. Morgan, T.S. Fisher, M. Neveu, Identification of the human cDNA for new survival/evasion peptide(DSEP): studies in vitro and in vivo of overexpression by neural cells, Exp. Neurol. 177 (2002) 32–39.
[7] D. Porter, S. Weremowicz, K. Chin, P. Seth, A. Keshaviah, J. Lahti-Domenici, Y.K. Bae, C.L. Monitto, A. Merlos-Suarez, J. Chan, C.M. Hulette, A. Richardson A, C.C. Morton, J. Marks, M. Duyao, R. Hruban, E. Gabrielson, R. Gelman, K. Polyak, A neural survival factor is a candidate oncogene in breast cancer, Proc. Natl. Acad. Sci. USA 100 (2003) 10931–10936.
[8] P. Todorov, P. Cariuk, T. McDevitt, B. Coles, K. Fearon, M. Tisdale, Characterization of a cancer cachectic factor, Nature 379 (1996) 739–742.
[9] P.T. Todorov, M. Deacon, M.J. Tisdale, Structural analysis of a tumor-produced sulfated glycoprotein capable of initiating muscle protein degradation, J. Biol. Chem. 272 (1997) 12279–12288.
[10] C.L. Monitto, S.M. Dong, J. Jen, D. Sidransky, Characterization of a human homologue of proteolysis-inducing factor and its role in cancer cachexia, Clin. Cancer Res. 10 (2004) 5862–5869.
[11] P. Cariuk, M.J. Lorite, P.T. Todorov, W.N. Field, S.J. Wigmore, M.J. Tisdale, Induction of cachexia in mice by a product isolated from the urine of cachectic cancer patients, Br. J. Cancer 76 (1997) 606–613.
[12] N. Teich, J. Kleeff, H. Lochs, J. Mossner, V. Keim, H. Friess, J. Ockenga, The presence of the proteolysis-inducing factor in urine does not predict the malignancy of a pancreatic tumour, BMC Gastroenterol. 5 (2005) 20.
[13] E. Grazzini, G. Guillon, B. Mouillac, H.H. Zingg, Inhibition of oxytocin receptor function by direct binding of progesterone, Nature 392 (1998) 509–512.
[14] Y.P. Lai, Y.F. Peng, Y. Zuo, J. Li, J. Huang, L.F. Wang, Z.R. Wu, Functional and structural characterization of recombinant dermcidin 1-L, a human antimicrobial peptide, Biochem. Biophys. Res. Commun. 3328 (2005) 243–250.
[15] D.M. Svinarich, R. Gomez, R. Romero, Detection of human defensins in the placenta, Am. J. Reprod. Immunol. 38 (1997) 252–255.
[16] R.L. Goldenberg, J.C. Hauth, W.W. Andrews, Intrauterine infection and preterm delivery, N. Eng. J. Med. 342 (2000) 1500–1507.
[17] C.L. Librach, Z. Werb, M.L. Fitzgerald, K. Chiu, N.M. Corwin, R.A. Esteves, D. Grobelny, R. Galardy, C.H. Damsky, S.J. Fisher, 92-kD type IV collagenase mediates invasion of human cytotrophoblasts, J. Cell Biol. 113 (1991) 437–449.
[18] P. Bischof, A. Meisser, A. Campana, Biochemistry and molecular biology of trophoblast invasion, Ann. N.Y. Acad. Sci. 943 (2001) 157–162.
[19] P. Kaufmann, S. Black, B. Huppertz, Endovascular trophoblast invasion: implications for the pathogenesis of intrauterine growth retardation and preeclampsia, Biol. Reprod. 69 (2003) 1–7.

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 284, NO. 38, pp. 25585–25592, September 18, 2009

Translesional DNA Synthesis through a C8-Guanyl Adduct of 2-Amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine (PhIP) *in Vitro*

REV1 INSERTS dC OPPOSITE THE LESION, AND DNA POLYMERASE κ POTENTIALLY CATALYZES EXTENSION REACTION FROM THE 3′-dC TERMINUS[*][S]

Received for publication, June 25, 2009, and in revised form, July 16, 2009 Published, JBC Papers in Press, July 23, 2009, DOI 10.1074/jbc.M109.037259

Hirokazu Fukuda[‡], Takeji Takamura-Enya[§], Yuji Masuda[¶], Takehiko Nohmi[‖], Chiho Seki[‡], Kenji Kamiya[¶], Takashi Sugimura[‡], Chikahide Masutani[**], Fumio Hanaoka[**1], and Hitoshi Nakagama[‡2]

From the [‡]Biochemistry Division, National Cancer Center Research Institute, 1-1, Tsukiji 5, Chuo-ku, Tokyo 104-0045, the [§]Department of Applied Chemistry, Faculty of Engineering, Kanagawa Institute of Technology, Ogino 1030, Atsugi, Kanagawa 243-0292, the [¶]Department of Experimental Oncology, Research Institute for Radiation Biology and Medicine, Hiroshima University, Kasumi 1-2-3, Minami-ku, Hiroshima, Hiroshima 734-8553, the [‖]Division of Genetics and Mutagenesis, National Institute of Health Sciences, Kamiyoga 1-18-1, Setagaya-ku, Tokyo 158-8501, and the [**]Cellular Biology Laboratory, Graduate School of Frontier Biosciences, Osaka University, Yamada-oka 1-3, Suita, Osaka 565-0871, Japan

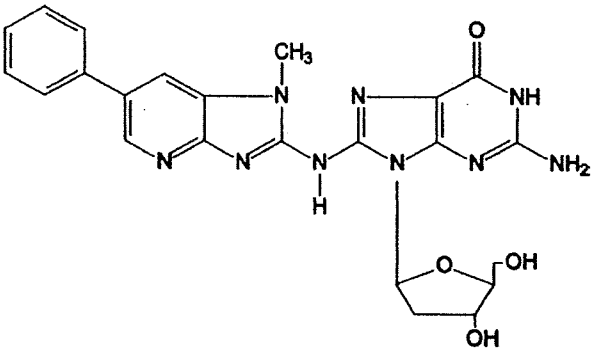
2-Amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine (PhIP) is the most abundant heterocyclic amine in cooked foods, and is both mutagenic and carcinogenic. It has been suspected that the carcinogenicity of PhIP is derived from its ability to form DNA adducts, principally dG-C8-PhIP. To shed further light on the molecular mechanisms underlying the induction of mutations by PhIP, *in vitro* DNA synthesis analyses were carried out using a dG-C8-PhIP-modified oligonucleotide template. In this template, the dG-C8-PhIP adduct was introduced into the second G of the TCC GGG AAC sequence located in the 5′ region. This represents one of the mutation hot spots in the rat *Apc* gene that is targeted by PhIP. Guanine deletions at this site in the *Apc* gene have been found to be preferentially induced by PhIP in rat colon tumors. DNA synthesis with A- or B-family DNA polymerases, such as *Escherichia coli* polymerase (pol) I and human pol δ, was completely blocked at the adducted guanine base. Translesional synthesis polymerases of the Y-family, pol η, pol ι, pol κ, and REV1, were also used for *in vitro* DNA synthesis analyses with the same templates. REV1, pol η, and pol κ were able to insert dCTP opposite dG-C8-PhIP, although the efficiencies for pol η and pol κ were low. pol κ was also able to catalyze the extension reaction from the dC opposite dG-C8-PhIP, during which it often skipped over one dG of the triple dG sequence on the template. This slippage probably leads to the single dG base deletion in colon tumors.

Heterocyclic amines (HCAs)[3] are naturally occurring genotoxic carcinogens produced from cooking meat (1). The initial carcinogenic event induced by HCAs is metabolic activation and subsequent covalent bond formation with DNA (1, 2). 2-Amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine (PhIP) is the most abundant heterocyclic amine in cooked foods, and was isolated from fried ground beef (3, 4). PhIP possesses both mutagenic and carcinogenic properties (5–8). Epidemiological studies have revealed that a positive correlation exists between PhIP exposure and mammary cancer incidence (9). PhIP induces colon and prostate cancers in male rats and breast cancer in female rats (8, 10).

The incidences of colon, prostate, and breast cancers are steadily increasing in Japan and other countries and this has been found to correlate with a more Westernized lifestyle. Elucidating the molecular mechanisms underlying PhIP-induced mutations is therefore of considerable interest. It is suspected that the carcinogenicity of PhIP is derived from the formation of DNA adducts, principally dG-C8-PhIP (11–14) (see Fig. 1). Studies of the mutation spectrum of PhIP in mammalian cultured cells and transgenic animals have revealed that G to T transversions are predominant and that guanine deletions from G stretches, especially from the 5′-GGGA-3′ sequence, are significant (15–20). Five mutations in the *Apc* gene were detected in four of eight PhIP-induced rat colon tumors, and all of these mutations involved a single base deletion of guanine from 5′-GGGA-3′ (21). These mutation spectra are thought to be influenced by various factors, including the primary structure of the target gene itself, the capacity of translesional DNA polymerases, and the activity level of repair enzymes (1). However, the molecular mechanisms underlying the formation of PhIP-induced mutations are largely unknown.

To shed further light on the molecular processes that underpin the mutations induced by PhIP, we performed *in vitro* DNA synthesis analyses using a dG-C8-PhIP-modified oligonucleotide template. We have recently reported the successful synthesis of oligonucleotides harboring a site-specific PhIP adduct

* This work was supported by Kakenhi Grant 19570144.
[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Table S1 and Figs. S1–S6.
[1] Present address: Dept. of Life Science, Faculty of Science, Gakushuin University, Mejiro 1-5-1, Toshima-ku, Tokyo 171-8588, Japan.
[2] To whom correspondence should be addressed. Tel.: 81-3-3542-2511; Fax: 81-33542-2530; E-mail: hnakagam@ncc.go.jp.
[3] The abbreviations used are: HCA, heterocyclic amines; PhIP, 2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine; TLS, translesional DNA synthesis; IQ, 2-amino-3-methylimidazo[4,5-*f*]quinoline; pol, DNA polymerase; DTT, dithiothreitol; PCNA, proliferating cell nuclear antigen; PIPES, 1,4-piperazinediethanesulfonic acid.



dG-C8-PhIP

FIGURE 1. **Structure of the dG-C8-PhIP adduct.**

(22). In our current study, we used this synthesis method to construct a 32-mer oligonucleotide template containing a 5′-TTCGGGAAC-3′ sequence with different site-specific PhIP adducts. We then utilized the resulting constructs in DNA synthesis analyses to reconstitute the PhIP-induced mutagenesis of the rat *APC* gene. DNA synthesis reactions with A- or B-family DNA polymerases, such as *Escherichia coli* pol I and human pol δ, or translesional synthesis (TLS) polymerases of the Y-family, pol η, pol ι, pol κ, and REV1, were carried out. Kinetic analyses of pol κ and REV1, for which TLS activities at the PhIP adduct were detected, were also performed.

EXPERIMENTAL PROCEDURES

Enzymes and Materials—T4 polynucleotide kinase and T4 DNA ligase were purchased from Toyobo Biochem (Osaka, Japan) and Takara Biotech (Tokyo Japan), respectively. Other materials were obtained from Sigma or Wako (Osaka, Japan).

DNA Polymerases and PCNA—Human recombinant DNA polymerases, pol δ, pol η, pol κ, and REV1, and PCNA were expressed and purified as described previously (23–27). Human DNA polymerase α and DNA polymerase ι were purchased from Chimerx. *E. coli* DNA polymerases I (Takara Biotech) and Klenow Fragment (Takara Biotech), and thermophilic bacterial DNA polymerases, r*Taq* (Toyobo Biochem) and *Tth* (Toyobo Biochem) were used.

Oligonucleotides—The method used to chemically synthesize three 9-mer oligonucleotides, 5′-TCCGGGAAC-3′, containing a PhIP adduct on either the first, second, or third G (p9B, p9C, and p9D, respectively) has been described previously (22). All other synthetic oligonucleotides were synthesized and purified using a reverse-phase cartridge (Operon Biotech Japan (Tokyo, Japan). The 23-mer oligonucleotides: p23a, 5′-TGACTCGTCGTGACTGGGAAAAC-3′, and p23b, 5′-GTCACGACGAGTCAGTTCCCGGA-3′, were used for constructing the template oligonucleotides as described below. A 32-mer oligonucleotide without the PhIP adduct, p32A, was used as a control template (see Table 1). Its 3′ complementary 29-, 28-, 27-, 26-, 22-, and 17-mer sequences (p29, p28, p27, p26, p22, and p17) were used as extension primers (see Table 1).

Construction of Template-Primer Complexes Containing the PhIP Adduct—A 32-mer template oligonucleotide p32C (see Table 1) was constructed by ligation of p9C with p23a as follows. The 5′-end of p23a was phosphorylated by T4 polynucleotide kinase and ATP. A mixture of p9C, p23a, and p23b (3 nmol each) in 250 μl of a buffer containing 5 mM Tris-HCl, 0.5 mM EDTA, 50 mM NaCl, pH 8.0, was denatured for 5 min at 95 °C, incubated for 10 min at 60 °C, and then cooled slowly to form the partial duplex structure of these three oligonucleotides (supplemental Fig. S1). The sample of the duplex oligonucleotide was mixed with 190 μl of Milli-Q water and 50 μl of ×10 ligation buffer (500 mM Tris-HCl (pH 7.5), 100 mM $MgCl_2$, 100 mM DTT, 10 mM ATP). Ligation was initiated by adding 10 μl of T4 DNA ligase (4,000 units), and the mixture was then incubated for 20 h at 16 °C. An additional incubation at 37 °C for 60 min was carried out after the addition of 1 μl of T4 DNA ligase, and the reaction was stopped by further incubation at 68 °C for 10 min. The p32C was separated by 18% PAGE containing 8 M urea, and excised and eluted as described previously (28). p32B and p32D were constructed using a similar method as for p9B and p9D, respectively (see Table 1). The purities of these oligonucleotides, p32B, p32C, and p32D, were determined by denatured PAGE after 5′-end labeling and UV absorbance at 260 and 370 nm.

Primer oligonucleotides were labeled with ^{32}P at the 5′-end as described previously (29), and then purified by MicroSpin™ G-25 or G-50 columns (GE Healthcare) as recommended by the supplier. The mixture of template and labeled primer (50 pmol each) in 400 μl of a buffer containing 8 mM Tris-HCl, 0.8 mM EDTA, 150 mM KCl (pH 8.0) was heated at 70 °C for 7 min, and then cooled slowly to room temperature. In the case of the substrates for TLS polymerases, pol η, pol ι, pol κ, and REV1, the final concentrations of template-primer and the constituents of the annealing buffers were changed to 500 nM and 10 mM Tris-HCl, 1 mM EDTA, and 50 mM NaCl (pH 8.0), respectively.

In Vitro DNA Synthesis Assay—A primer extension reaction was performed as described previously (30) with some modifications. Briefly, an aliquot of 0.75 μl of this primer-annealed template (final concentration, 12.5 nM) was mixed with 0.75 μl of ×10 Klenow buffer (100 mM Tris-HCl (pH 7.5), 70 mM $MgCl_2$, 1 mM DTT), 0.5 μl of 500 mM KCl, 0.5 μl of dNTP mixture (50 μM each), and 4.5 μl of Milli-Q water. After addition of 0.5 μl of Klenow fragment, the mixture was incubated at 37 °C for 10 min. The reaction was terminated by adding 1.5 μl of stop solution (160 mM EDTA, 0.7% SDS, 6 mg/ml proteinase K), and the samples were incubated at 37 °C for 30 min. Subsequently, 5.5 μl of the gel loading solution (30 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol, 97% formamide) was added to the samples. For pol δ, a ×10 reaction buffer containing 200 mM PIPES (pH 6.8), 20 mM $MgCl_2$, 10 mM 2-mercaptethanol, 200 μg/ml bovine serum albumin, and 50% glycerol was used instead of the buffer described above, and the reaction was carried out at 37 °C for 10 min. For other DNA polymerases, pol α, pol I, r*Taq*, and *Tth*, the constituent of each ×10 reaction buffer was altered as recommended by the suppliers.

The reaction using pol κ was performed as described above with some modifications. Briefly, an aliquot of 0.5 μl of this primer-annealed template (final 50 nM) was mixed with 0.5 μl of 10 × TLS buffer (250 mM Tris-HCl (pH 7.0), 50 mM $MgCl_2$, 50 mM DTT, 1 mg/ml bovine serum albumin), 0.5 μl of dNTP solution, and 3.0 μl of Milli-Q water. After addition of 0.5 μl of pol κ, the mixture was incubated at 30 °C for 20 min. The reac-

tion was terminated by adding 8.8 μl of the gel loading solution and a further incubation at 95 °C for 3 min. The reaction of REV1 was performed in the same manner as the reaction of pol κ with the exception that the standard reaction time was 5 min. For pol η, a ×10 reaction buffer containing 400 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 100 mM DTT, 1 mg/ml bovine serum albumin, and 450 mM KCl was used instead of the ×10 TLS buffer. The ^{32}P-labeled fragments were denatured and electrophoresed in a 9.5% polyacrylamide gel containing 8 M urea. The radioactivity of the fragments was determined using a Bio-Imaging Analyzer (BAS2500, Fuji Photo Film, Kanagawa, Japan). Kinetic parameters were determined by steady-state gel kinetic assays under similar conditions as described above. The incubation time for pol κ was changed to 10 min. K_m and k_{cat} were evaluated from the plot of the initial velocity *versus* the dCTP or dGTP concentration using a hyperbolic curve-fitting program in SigmaPlot 11 (Systat Software, Inc.). Data from two or three independent experiments were plotted together.

RESULTS

Construction of Template Oligonucleotides Containing a PhIP Adduct—We designed oligonucleotides containing a dG-C8-PhIP adduct at specific sites for use as templates in *in vitro* DNA synthesis analyses. For this purpose, we selected the 5′-TCCGGGAAC-3′ sequence as: 1) it corresponds to codon 868 – 870 of the rat *Apc* gene, one of three mutation hot spots (a single base deletion of G) in PhIP-induced colon tumors (21), and could thus be used as a model template that would reconstitute mutations of this gene; 2) two other mutation hot spots in the rat *Apc* gene and many mutated sites induced by PhIP in cultured cells and animal models contain 5′-GGGA-3′ as a core sequence (17–20). We thus speculated that the 5′-TCCGG-GAAC-3′ sequence could be used as a model sequence for these GGGA to GGA mutations to some extent; and 3) some mutagenic compounds forming dG adducts, including PhIP, are expected to react preferentially with the 5′-G of a GG dinucleotide site when compared with a single G residue (31). We thus selected a sequence containing GGG as a template for our initial analysis.

We have recently synthesized three 9-mer oligonucleotides separately harboring a PhIP adduct on each G within the sequence 5′-TCC GGG AAC-3′ (22). Three 32-mer template oligonucleotides, p32B, p32C, and p32D, were constructed in our present study by ligation of these 9-mer oligonucleotides containing the dG-PhIP adduct with a 23-mer oligonucleotide, p23a, (Table 1 and supplemental Fig. S1). The purities of these oligonucleotides were tested after resolution by electrophoresis. In our present study, we principally describe the results of our *in vitro* DNA synthesis analysis using p32C as the template to avoid complexity.

In Vitro DNA Synthesis by A- and B-family DNA Polymerase—Many of the chemical compounds that can form DNA adducts *in vivo* and that show mutagenicity have been reported to impede the progress of DNA synthesis to different extents. The molecular size of PhIP is greater than most other mutagenic chemicals that form adducts. Hence, dG-PhIP was expected to block DNA synthesis to a considerable extent. To examine the effects of the dG-C8-PhIP adduct upon DNA synthesis, primer

TABLE 1
Oligonucleotide templates and primers

Oligonucleotide	Sequence[a]
p32A	5′-TCC GGG AAC TGACTCGTC GTGACTGGG AAAAC-3′
p32B	5′-TCC GGG AAC TGACTCGTC GTGACTGGG AAAAC-3′
p32C	5′-TCC GGG AAC TGACTCGTC GTGACTGGG AAAAC-3′
p32D	5′-TCC GGG AAC TGACTCGTC GTGACTGGG AAAAC-3′
p29	5′-GTT TTC CCA GTCACGACG AGTCAGTTC CC-3′
p28	5′-GTT TTC CCA GTCACGACG AGTCAGTTC C-3′
p27	5′-GTT TTC CCA GTCACGACG AGTCAGTTC-3′
p26	5′-GTT TTC CCA GTCACGACG AGTCAGTT-3′
p22	5′-GTT TTC CCA GTCACGACG AGTC-3′
p17	5′-GTT TTC CCA GTCACGAC-3′

[a] The bold G indicates the site of the PhIP-C8-dG adduct. Underlined sequences correspond to codon 868 – 870 at nucleotides 2602–2610 of the rat APC gene.

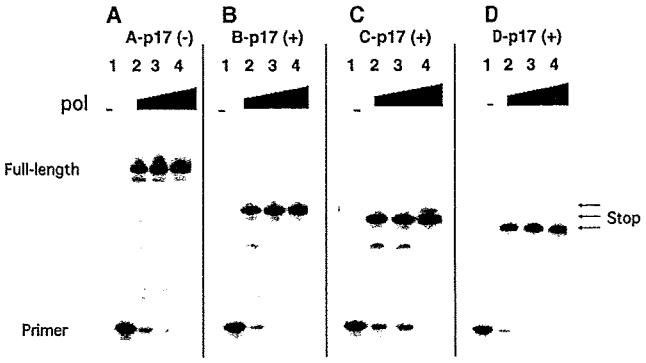


FIGURE 2. ***In vitro* DNA synthesis using Klenow fragment.** Gel electrophoresis indicating the primer extensions obtained using the 32-mer oligonucleotide templates, p32A (*A*), p32B (*B*), p32C (*C*), and p32D (*D*), which have no PhIP adduct, and a PhIP adduct on the first, second, and third G within the triple G sequence, respectively. The 3′ complementary 17-mer sequence, p17, was used as the extension primer. The final concentration of each template-primer complex was 12.5 nM. Concentrations of Klenow fragment were 0 (*lane 1*), 7.8 (*lane 2*), 23 (*lane 3*), and 78 units/ml (*lane 4*).

extension experiments using p32B, p32C, and p32D as templates were carried out (see Table 1). The length of each produced fragment was precisely determined using ladders of oligonucleotide fragments as markers (data not shown). The Klenow fragment of *E. coli* DNA polymerase I, a member of the A-family DNA polymerases, was first used in this analysis. The production of a 28-, 27-, and 26-mer from these primer extension reactions using B-p17, C-p17, and D-p17, respectively, using a template-primer complex, and lack of longer fragments indicated that the Klenow fragment stalled just before the dG-C8-PhIP adduct (Fig. 2). On the other hand, control experiments using p32A without the adduct as a template produced a 32-mer fragment (Fig. 2*A*). Similar results were obtained with *E. coli* DNA polymerase I (exo$^+$) and B-family DNA polymerases, such as the thermophilic bacterial DNA polymerases, r*Taq* and *Tth*, and human DNA polymerase α (data not shown) (supplemental Fig. S2), suggesting that stalling at the dG-C8-PhIP adduct occurs for all replicative DNA polymerases. Stalling of r*Taq* and *Tth* at the PhIP adduct was observed at 65 °C, as well as at 37 °C, indicating that this is the result of a physical hindrance of the adduct itself and not from secondary DNA structures. Moreover, there was no difference found between the stalling of *E. coli* DNA polymerase I (exo$^+$) and that of the Klenow fragment (exo$^-$). This indicates that the physical blocking of DNA polymerases at the dG-C8-PhIP adduct does not depend upon their proofreading function.

Finally, DNA synthesis analyses with human DNA polymerase δ (pol δ), a member of the B-family DNA polymerases and a truly replicative polymerase, were carried out. In the case of using p32C and p17 (C-p17) as a template-primer complex, the production of 27-mer fragments indicated the stalling of pol δ just before the PhIP adduct (Fig. 3, *lane 11*). From a control reaction using A-p17, a template-primer complex without the PhIP adduct, a full-length product of 32-mer was generated (Fig. 3, *lane 8*). In addition to these major products, minor products extended one nucleotide further (28- and 33-mer) and ladders of bands indicating degradation of primer (<17-mer) were observed (Fig. 3), corresponding with previous results reporting terminal dA transferase and exonuclease activities of pol δ (32). PCNA, an accessory protein acting as a sliding clamp for pol δ, was previously reported to promote DNA synthesis by pol δ past several template lesions, including abasic sites, 8-oxo-dG, and aminofluorene-dG (32). In the case of dG-C8-PhIP, however, PCNA was unable to promote the bypass synthesis of pol δ beyond the lesion (Fig. 3, *lane 12*). Extension reaction from the longer 22-mer primer, p22, also paused completely just before the PhIP adduct in the presence or absence of PCNA (Fig. 3, *lanes 5* and *6*). These results strongly suggest that the dG-C8-PhIP adduct on genome DNA in the living cells induces the complete block of replication forks including pol δ, PCNA, and pol α.

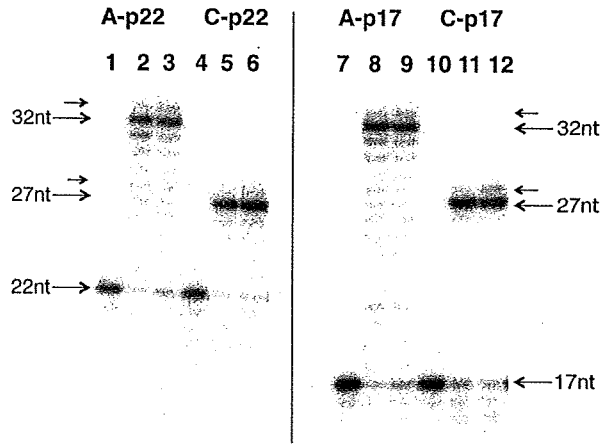


FIGURE 3. ***In vitro* DNA synthesis using pol δ in the presence or absence of PCNA.** Gel electrophoresis indicating the primer extensions obtained using the 32-mer oligonucleotide templates, p32A (*A*), and p32C (*C*), which have no PhIP adduct, and a PhIP adduct on the second G within the triple G sequence, respectively. The 3′ complementary 22- and 17-mer sequences, p22 and p17, were used as the extension primer. The final concentration of each template-primer complex was 12.5 nM. Concentrations of pol δ were 0 (*lanes 1, 4, 7,* and *10*) and 16 nM (*lanes 2, 3, 5, 6, 8, 9, 11,* and *12*). Concentrations of PCNA as a trimer were 0 (*lanes 1, 2, 4, 5, 7, 8, 10,* and *11*) and 20 nM (*lanes 3, 6, 9,* and *12*). *Large arrows* indicate the positions of primers (17- or 22-mer), full-length products (32-mer), and the products pausing just before the PhIP adduct (27-mer). *Small arrows* indicate the minor products that incorporated an additional 1 nucleotide (*nt*) to a full-length product or the pausing product.

To examine the translesional DNA synthesis by Y-family DNA polymerases—Translesional DNA synthesis at the dG-C8-PhIP adduct by the Y-family DNA polymerases, pol η, pol κ, pol ι, and REV1 was next examined. Two substrates, C-p27 and C-p28, and their counterparts without a PhIP adduct, A-p27 and A-p28, were used in these experiments (Fig. 4). Substrate C-p27 was prepared by annealing the p32C template (see Table 1) to its 3′-complimentary 27-mer sequence, p27, and was used to identify the nucleotides that are inserted opposite the dG-C8-PhIP adduct (Fig. 4). Similarly, substrate C-p28 was used to analyze the extension reaction from the 3′-end of the dC bases opposite the dG-C8-PhIP adduct (Fig. 4). We found that recombinant human DNA polymerase η (pol η) could insert a dC opposite the dG-C8-PhIP adduct, although at low efficiency compared with control experiments without the PhIP adduct (Fig. 5, *A* and *B*). Extension reactions catalyzed by pol η from the 3′-end of dC opposite the adduct were barely detectable (Fig. 5*D*), although an excessive amount of pol η produced byproducts that incorporated a mismatch nucleotide, dG, dA, or dT (supplemental Fig. S4). In the case of dG, incorporation of one to three dG nucleotides was observed (supplemental Fig. S4). In control experiments without the PhIP adduct, minor products were produced that incorporated mismatch nucleotides, in addition to a major product that incorporated a dC (Fig. 5*C*).

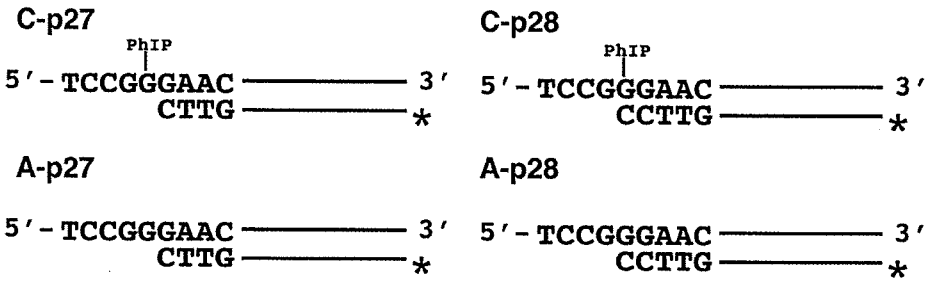


FIGURE 4. **Template-primer complexes.** Substrates C-p27 and C-p28 (series-C) have a PhIP adduct on the second dG within a GGG sequence. Substrates A-p27 and A-p28 (series-A) are control substrates without a PhIP adduct. The corresponding 3′ complimentary 27- and 28-mer sequences, p27 and p28, were used as extension primers. The template-primer complexes, C-p27 and C-p28, were used to monitor the nucleotide insertions into the site opposite dG-C8-PhIP and the extension reactions from the 3′-dC opposite dG-C8-PhIP, respectively.

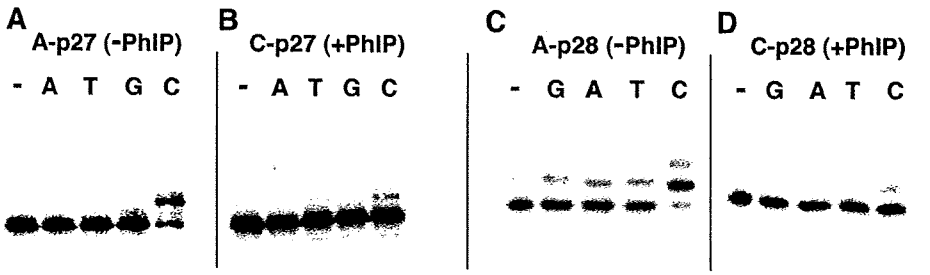


FIGURE 5. **Translesional DNA synthesis by pol η using substrates C-p27 and C-p28.** Control reactions were performed using substrates without the PhIP adduct, A-p27 (*A*) and A-p28 (*C*). An insertion reaction was performed with substrate C-p27 (*B*) and an extension reaction with substrate C-p28 (*D*). A single dNTP (G, A, T, C) was added into the reaction mixture as indicated by *G, A, T,* and *C above* each lane. The lanes indicated by − are controls without any nucleotides. Concentrations of pol η and each dNTP were 1.9 nM and 100 μM, respectively.

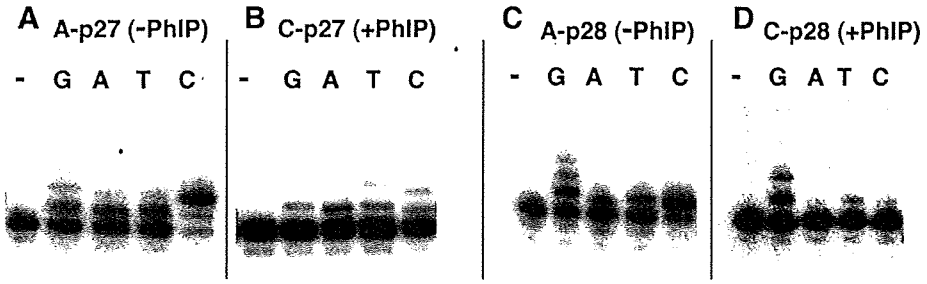


FIGURE 6. **Translesional DNA synthesis by pol κ using substrates C-p27 and C-p28.** Control reactions were performed using substrates without the PhIP adduct, A-p27 (*A*) and A-p28 (*C*). An insertion reaction was performed with substrate C-p27 (*B*) and an extension reaction with substrate C-p28 (*D*). A single dNTP (G, A, T, C) was added into the reaction mixture as indicated by *G, A, T,* and *C above* each lane. The lanes indicated by − are controls without any nucleotides. The concentrations of pol κ were 250 (*A* and *C*), 500 (*B*), and 1000 nM (*D*), respectively. The concentration of each dNTP was 100 μM.

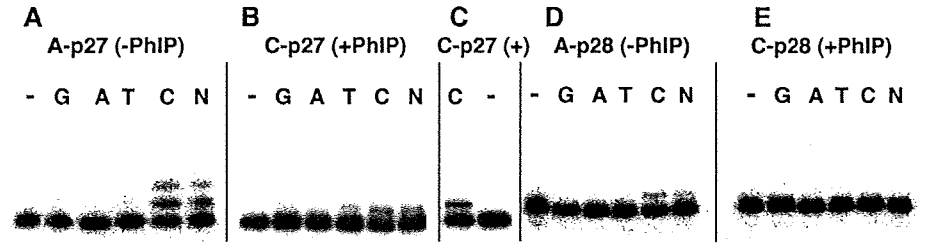


FIGURE 7. **Translesional DNA synthesis by REV1 using substrates C-p27 and C-p28.** Control reactions were performed using substrates without the PhIP adduct, A-p27 (*A*) and A-p28 (*D*). Insertion reactions were performed with substrate C-p27 (*B* and *C*) and an extension reaction with substrate C-p28 (*E*). A single dNTP (G, A, T, and C) or a mixture of each was added into the reaction mixture as indicated by *G, A, T, C,* and *N above* each lane. The lanes indicated by − are controls without any nucleotides. The concentrations of REV1 were 5.2 (*A* and *D*) and 26 nM (*B, C,* and *E*), respectively. The concentrations of each dNTP were 100 μM (*A, B, D,* and *E*) and 320 μM (*C*), respectively. The *N* mixture contained each dNTP at a concentration of 25 μM.

We next examined translesional DNA synthesis beyond the PhIP adduct using a truncated form of human DNA polymerase κ containing the N-terminal 559 amino acids. One or two dCs were inserted opposite the dG-C8-PhIP adduct by this polymerase, and misinsertions of three other nucleotides were also observed to a certain extent (Fig. 6*B*). pol κ incorporated two dCs and misincorporated dG, dA, and dT into the A-p27 substrate without the PhIP adduct at a low efficiency (Fig. 6*A*). Misincorporations of dG, dA, and dT into the A-p28 substrate without the adduct were also observed (Fig. 6*C*). In the case of the extension reaction from 3′-dC opposite the dG-PhIP adduct, pol κ also incorporated dC and misincorporated dT into the C-p28 substrate at low efficiency (Fig. 6*D*). Interestingly, one- and two-base incorporations of dG into the substrate C-p28 by pol κ dominated the incorporation of a dC (Fig. 6*D*). In the extension reaction with pol κ in the presence of all four dNTPs, fragments of 29 and 30 nucleotides were observed as major products, and a small amount of the 31-nucleotide fragment was observed (see supplemental Fig. S5, *lane 6*). Full-length products of 32 nucleotides were observed only when an excess amount of pol κ was present (data not shown). This poor extension activity of pol κ after adding two nucleotides was probably caused by the shortness (~4 nucleotides) of the 5′ region to the lesion in the template oligonucleotide. Extension with pol κ, pol η, and pol δ from the mismatched primers, where the 3′-terminal nucleotide of the p28 primer, dC, was substituted with another nucleotide, could not be observed (data not shown). REV1 inserted a dC opposite the PhIP adduct at a higher efficiency compared with pol κ and pol η (Fig. 7, *B* and *C*). REV1 was, however, unable to catalyze the extension reaction from the dC opposite the PhIP adduct in C-p28 (Fig. 7*E* and supplemental Fig. S5, *lane 5*). REV1 incorporated only dC nucleotides into A-p27 and A-p28 substrates without the adduct (Fig. 7, *A* and *D*). Neither nucleotide insertion nor extension reactions for the templates containing the PhIP adduct were detected using human pol ι (data not shown).

Kinetic Analyses of Translesional DNA Synthesis by pol κ and REV1—To evaluate translesional DNA synthesis beyond the dG-C8-PhIP adduct in further detail, additional quantitative analyses for pol κ and REV1 were performed. Insertion reactions catalyzed by pol κ for dC (Fig. 8, *B, lanes 2–5,* and *C, closed diamonds*) and dG (Fig. 8, *B, lanes 6–9,* and *C, closed triangles*) into substrate C-p28 were analyzed in the same way. Kinetic parameters for pol κ were determined using steady-state kinetic assays (Table 2). The catalytic efficiency (k_{cat}/K_m) of dC insertion into C-p28 (0.039 min^{-1} mM^{-1}) was found to be 4-fold greater than that into C-p27 (0.011 min^{-1} mM^{-1}). These results indicate that pol κ catalyzes the extension reaction from the 3′-terminal of dC opposite the dG-C8-PhIP with a higher efficiency than the insertion reaction opposite the adduct. The k_{cat}/K_m values of the dC insertion opposite the adduct were roughly 4 orders of magnitude less than those into counterparts without the adduct (see Table 2). The k_{cat}/K_m value of the dG incorporation into C-p28 was slightly higher than that of dC, and more than 8-fold higher than that of dG into C-p27 (see Table 2). This result indicates that pol κ skipped over the dG site just 5′ of dG-C8-PhIP on the template and incorporated dG opposite dC on the template strand of substrate C-p28 with a high efficiency. The k_{cat}/K_m values of the dC incorporation into D-p27 (0.19 min^{-1} mM^{-1}) were over 4-fold greater than into C-p28 (0.039 min^{-1} mM^{-1}) and over 8-fold higher than that of dG into B-p29 (0.023) (see supplemental Table S1). These data indicate that the efficiencies of the extension reaction by pol κ are the highest for template p32D containing the PhIP adduct in the third G of the triple G run, next for template p32C containing the PhIP/adduct in the second G, and lowest for template p32B containing the PhIP adduct in the first G.

Even at higher concentrations of dNTPs, extension reactions catalyzed by REV1 for substrate C-p28 could not be monitored (Table 3, Fig. 7*E*). The k_{cat}/K_m value of the dC incorporation by REV1 into substrate C-p27 was more than 2,000 times greater than that by pol k, and 1/44 of the values for counterparts with-

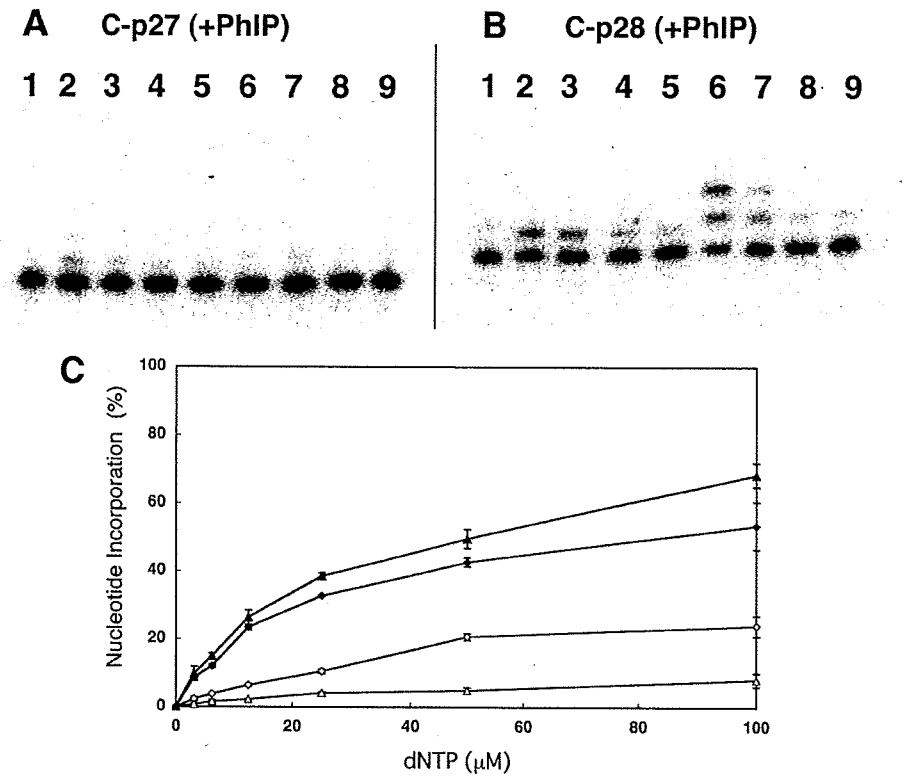


FIGURE 8. **Translesional DNA synthesis by pol κ.** Nucleotide incorporation by pol κ for substrates C-p27 (*A*) and C-p28 (*B*). Either dCTP (*lanes 2-5*) or dGTP (*lanes 6-9*) was added into the reaction mixture. *Lane 1* indicates a control without any nucleotides. The concentration of pol κ was 910 nM. The concentrations of dCTP or dGTP, respectively, were 25 (*lanes 2* and *6*), 12.5 (*lanes 3* and *7*), 6.25 (*lanes 4* and *8*), and 3.13 μM (*lanes 5* and *9*). *C*, incorporation efficiencies of dCTP and dGTP into substrate C-p27 and C-p28. Incorporations of dCTP into C-p27, dGTP into C-p27, dCTP into C-p28, and dGTP into C-p28 are indicated by *open diamonds*, *open triangles*, *closed diamonds*, and *closed triangles*, respectively. Each data point represents the mean of two separate experiments. The *error bars* represent residuals.

TABLE 2

k_{cat}/K_m values for pol κ

Substrate	K_m	k_{cat}	k_{cat}/K_m
	μM	$\times 10^{-3}$ min^{-1}	min^{-1} mM^{-1}
C-p27			
dCTP	70	0.76	0.011
dGTP	47	0.24	0.0050
C-p28			
dCTP	8.0	0.32	0.039
dGTP	11	0.48	0.042
A-p27			
dCTP	0.035	4.4	130
dGTP	0.26	1.3	5.0
A-p28			
dCTP	0.027	3.7	140
dGTP	2.1	8.8	4.1

TABLE 3

k_{cat}/K_m values for dCTP-insertion by REV1

Substrate	K_m	k_{cat}	k_{cat}/K_m
	μM	$\times 10^{-3}$ min^{-1}	min^{-1} mM^{-1}
C-p27	12	320	27
C-p28	ND[a]	ND	ND
A-p27	0.36	390	1100

[a] ND, not detectable.

out the adduct (Table 3). The k_{cat}/K_m values of the dC insertion by REV1 into three substrates, B-p28, C-p27 and D-p26, were 39, 27, and 73 min^{-1} mM^{-1}, respectively. Thus, the insertion reaction catalyzed by REV1 among the three templates was the most efficient for template p32D containing the PhIP adduct at the third G, similar to the extension reaction by pol κ.

DISCUSSION

In Vitro TLS Analysis Reconstituting PhIP-induced Mutations—HCAs are food-borne carcinogens produced when cooking meat (1, 9, 33). The most significant aspect of these molecules is that they exist normally in cooked food and are thus ubiquitous carcinogens (32). The mutagenicity and carcinogenicity of HCAs are mainly attributed to C8- and N2-dG adducts (9). Both excision repair and translesional DNA synthesis play critical roles in the mutagenesis steps induced by HCAs. However, despite the importance of HCAs as common environmental mutagens, there have been very few previous reports regarding the stalling of DNA polymerases and TLS caused by the DNA adducts they form. This is mainly because of the difficulty in preparing template DNA with introduced HCA adducts at specific sites. Choi *et al.* (34) have recently undertaken a biochemical study of TLS at adducts of the HCA 2-amino-3-methylimidazo[4,5-*f*]quinoline (IQ) using purified human polymerases. In our current study of TLS, we describe our findings for adducts of PhIP, the most abundant HCA in cooked foods (4).

A rat colon cancer model induced by PhIP shows profiles of cancer development similar to the multistep model of colon carcinogenesis in humans (35). In this rat model, *p53* and K-*ras* mutations are rarely observed, whereas mutations in *Apc* and its downstream gene *β-catenin* have been frequently observed (21, 36–38). Hence, mutations in *Apc* or *β-catenin* have been speculated to play a critical role in PhIP-induced colon carcinogenesis. Five mutations in the *Apc* gene were previously detected in four of eight PhIP-induced rat colon tumors, and all of these mutations involved a single guanine deletion in the 5′-GGGA-3′ sequence (21). This characteristic mutation induced by PhIP, 5′-GGGA-3′ to 5′-GGA-3′, was also observed in other *in vivo* mutation analyses using transgenic animals harboring introduced reporter genes, such as *lacI* (18–20). Hence, the 5′-TCCGGGAAC-3′ sequence corresponding to a mutation hot spot within the rat *Apc* gene, which we utilized to introduce the PhIP adduct and employed as the template for *in vitro* DNA synthesis analyses, could be a suitable model for revealing the molecular mechanisms associated with PhIP-induced mutations.

As discussed later, our results indicate a possible molecular mechanism for the 5′-GGGA-3′ to 5′-GGA-3′ mutation induced by PhIP.

DNA Polymerases Involved in TLS through the dG-PhIP Adduct—TLS through many DNA lesions requires the action of two different polymerases, an "inserter" and an "extender," the former to perform nucleotide insertions opposite the lesion site and the latter for subsequent extensions (39). The catalytic efficiency of the dCTP-insertion reaction opposite the dG-PhIP adduct by REV1 was found to be more than 2,000-fold greater than that by pol κ (see Tables 2 and 3). This result strongly suggests that REV1 functions *in vivo* as an inserter polymerase for TLS through the dG-PhIP adduct. This insertion step by REV1 is also error free. REV1 has been reported previously to insert dCTP opposite abasic sites and various N2-dG adducts (26, 39–41). However, our current study is the first to show that REV1 inserts dCTP opposite a large size C8-dG adduct. We used a shorter (C-terminal deleted) form of pol κ in our current experiments and an intact pol κ may be more effective for this insertion reaction. As for pol η, a detailed kinetic analysis was not performed. Hence, the possibility cannot be excluded that pol κ and pol η also function as inserter polymerases.

In addition to the Y-family DNA polymerases, DNA polymerase ζ (pol ζ), belonging to the B-family DNA polymerases, is considered to be involved in TLS through various lesions as an extender DNA polymerase (39, 42, 43). We have not carried out a primer extension assay with pol ζ and thus the possibility cannot be completely excluded by our current data that pol ζ functions *in vivo* as an extender polymerase for TLS through the dG-PhIP adduct. In our present study, we provide evidence that pol κ can extend from dC opposite the dG-C8-PhIP adduct *in vitro*. It is, therefore, possible that pol κ, at least partially, functions as an extender polymerase *in vivo* for TLS through the dG-PhIP adduct. Further study about cooperation between two or more DNA polymerases, including pol ζ, is necessary to verify which DNA polymerases are involved in the bypass synthesis through the PhIP lesion.

The catalytic efficiency of pol κ for a dGTP insertion into substrate C-p28 was a little higher than that for dCTP insertions (see Table 2 and Fig. 6*D*). The former generates a single guanine deletion, and the latter is an error-free extension. Consequently, our data suggest that the extension reaction with pol κ from the nucleotide opposite the dG-C8-PhIP adduct causes frequent single-guanine deletions from the GGG stretch. It has been reported that one characteristic feature of pol κ homologs, from bacteria to humans, is their propensity to generate single-base deletions (44–47). The crystal structure of Dpo4, a thermophilic archaea homolog of pol κ, in ternary complexes with DNA and an incoming nucleotide supports the model that a single base deletion by pol κ is generated through a misaligned intermediate complex where the template dG forms an extrahelical looped out structure and the incoming dGTP skips this extrahelical base and pairs with the next template base dC (48) (see supplemental Fig. S6). It is reasonable to speculate therefore that, in the case of TLS through dG-C8-PhIP, mammalian pol κ generates the single guanine deletion via a similar intermediate where the PhIP-adducted dG is looped out and template-primer slippage occurs. However, further analyses for determining whether the one-base skipping of pol κ beyond the lesion observed by us is dependent on the nucleotide placed 5′ to the lesion or not, are necessary to clarify the detailed molecular mechanism underlying one base skipping of pol κ.

Molecular Mechanisms Underlying Mutation Induction by PhIP—We have demonstrated herein by *in vitro* DNA synthesis analyses using oligonucleotide templates containing dG-PhIP that: 1) replicative DNA polymerases stall at the PhIP adduct and cannot perform translesional DNA synthesis beyond this point; 2) REV1 inserts a dC opposite the dG-PhIP with a much higher efficiency than other TLS polymerases, including pol κ and pol η; and 3) pol κ has a potential ability to catalyze an extension reaction from the 5′-dC opposite the adduct and often skips over one dG in the template during this extension step. A working model for the induction of mutations at the PhIP adducts based on the results shown in the present study is illustrated in supplemental Fig. S6. This model could be adopted for other sequences containing a G repeat stretch longer than GGG.

REFERENCES

1. Nagao, M. (2000) in *Food Borne Carcinogens: Heterocyclic Amines* (Nagao, M., and Sugimura, T., eds) pp. 163–196, John Wiley & Sons Ltd., Chichester, UK
2. Schut, H. A., and Snyderwine, E. G. (1999) *Carcinogenesis* **20**, 353–368
3. Felton, J. S., Knize, M. G., Shen, N. H., Lewis, P. R., Andresen, B. D., Happe, J., and Hatch, F. T. (1986) *Carcinogenesis* **7**, 1081–1086
4. Felton, J. S., Jagerstad, M., Knize, M. G., Skog, K., and Wakabayashi, K. (2000) in *Food Borne Carcinogens: Heterocyclic Amines* (Nagao, M., and Sugimura, T., eds) pp. 31–71, John Wiley & Sons Ltd., Chichester, UK
5. Holme, J. A., Wallin, H., Brunborg, G., Søderlund, E. J., Hongslo, J. K., and Alexander, J. (1989) *Carcinogenesis* **10**, 1389–1396
6. Felton, J. S., and Knize, M. G. (1991) *Mutat. Res.* **259**, 205–217
7. Ohgaki, H., Takayama, S., and Sugimura, T. (1991) *Mutat. Res.* **259**, 399–410
8. Ito, N., Hasegawa, R., Sano, M., Tamano, S., Esumi, H., Takayama, S., and Sugimura, T. (1991) *Carcinogenesis* **12**, 1503–1506
9. Sugimura, T., Wakabayashi, K., Nakagama, H., and Nagao, M. (2004) *Cancer Sci.* **95**, 290–299
10. Imaida, K., Hagiwara, A., Yada, H., Masui, T., Hasegawa, R., Hirose, M., Sugimura, T., Ito, N., and Shirai, T. (1996) *Jpn. J. Cancer Res.* **87**, 1116–1120
11. Frandsen, H., Grivas, S., Andersson, R., Dragsted, L., and Larsen, J. C. (1992) *Carcinogenesis* **13**, 629–635
12. Lin, D., Kaderlik, K. R., Turesky, R. J., Miller, D. W., Lay, J. O., Jr., and Kadlubar, F. F. (1992) *Chem. Res. Toxicol.* **5**, 691–697
13. Snyderwine, E. G., Davis, C. D., Nouso, K., Roller, P. P., and Schut, H. A. (1993) *Carcinogenesis* **14**, 1389–1395
14. Schut, H. A., and Herzog, C. R. (1992) *Cancer Lett.* **67**, 117–124
15. Endo, H., Schut, H. A., and Snyderwine, E. G. (1994) *Cancer Res.* **54**, 3745–3751
16. Morgenthaler, P. M., and Holzhäuser, D. (1995) *Carcinogenesis* **16**, 713–718
17. Yadollahi-Farsani, M., Gooderham, N. J., Davies, D. S., and Boobis, A. R. (1996) *Carcinogenesis* **17**, 617–624
18. Okonogi, H., Stuart, G. R., Okochi, E., Ushijima, T., Sugimura, T., Glickman, B. W., and Nagao, M. (1997) *Mutat. Res.* **395**, 93–99
19. Lynch, A. M., Gooderham, N. J., Davies, D. S., and Boobis, A. R. (1998) *Mutagenesis* **13**, 601–605
20. Okochi, E., Watanabe, N., Shimada, Y., Takahashi, S., Wakazono, K., Shirai, T., Sugimura, T., Nagao, M., and Ushijima, T. (1999) *Carcinogenesis* **20**, 1933–1988
21. Kakiuchi, H., Watanabe, M., Ushijima, T., Toyota, M., Imai, K., Weisburger, J. H., Sugimura, T., and Nagao, M. (1995) *Proc. Natl. Acad. Sci.*

U.S.A. **92,** 910–914
22. Takamura-Enya, T., Ishikawa, S., Mochizuki, M., and Wakabayashi, K. (2006) *Chem. Res. Toxicol.* **19,** 770–778
23. Masuda, Y., Suzuki, M., Piao J., Gu, Y., Tsurimoto, T., and Kamiya, K. (2007) *Nucleic Acids Res.* **35,** 6904–6916
24. Masutani, C., Kusumoto, R., Iwai, S., and Hanaoka, F. (2000) *EMBO J.* **19,** 3100–3109
25. Niimi, N., Sassa, A., Katafuchi, A., Grúz, P., Fujimoto, H., Bonala, R. R., Johnson, F., Ohta, T., and Nohmi, T. (2009) *Biochemistry* **48,** 4239–10234246
26. Masuda, Y., and Kamiya, K. (2002) *FEBS Lett.* **520,** 88–92
27. Masuda, Y., Ohmae, M., Masuda, K., and Kamiya, K. (2003) *J. Biol. Chem.* **278,** 12356–12360
28. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual,* 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
29. Fukuda, H., and Ohtsubo, E. (1997) *Genes Cells* **2,** 735–751
30. Fukuda, H., Katahira, M., Tsuchiya, N., Enokizono, Y., Sugimura, T., Nagao, M., and Nakagama, H. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99,** 12685–12690
31. Sugiyama, H., and Saito, I. (1996) *J. Am. Chem. Soc.* **118,** 7063–7068
32. Mozzherin, D. J., Shibutani, S., Tan, C. K., Downey, K. M., and Fisher, P. A. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94,** 6126–6231
33. Sugimura, T., and Adamson, R. H. (2000) in *Food Borne Carcinogens: Heterocyclic Amines* (Nagao, M., and Sugimura, T., eds) pp. 1–4, John Wiley & Sons Ltd., Chichester, UK
34. Choi, J. Y., Stover, J. S., Angel, K. C., Chowdhury, G., Rizzo, C. J., and Guengerich, F. P. (2006) *J. Biol. Chem.* **281,** 25297–25306
35. Nakagama, H., Ochiai, M., Ubagai, T., Tajima, R., Fujiwara, K., Sugimura, T., and Nagao, M. (2002) *Mutat. Res.* **506–507,** 137–144
36. Nagao, M. (1999) *Mutat. Res.* **431,** 3–12
37. Nagao, M., Ushijima, T., Toyota, M., Inoue, R., and Sugimura, T. (1997) *Mutat. Res.* **376,** 161–167
38. Dashwood, R. H., Suzui, M., Nakagama, H., Sugimura, T., and Nagao, M. (1998) *Cancer Res.* **58,** 1127–1129
39. Prakash, S., Johnson, R. E., and Prakash, L. (2005) *Annu. Rev. Biochem.* **74,** 317–353
40. Nelson, J. R., Lawrence, C. W., and Hinkle, D. C. (1996) *Nature* **382,** 729–731
41. Haracska, L., Prakash, S., and Prakash, L. (2002) *J. Biol. Chem.* **277,** 15546–15551
42. Johnson, R. E., Washington, M. T., Haracska, L., Prakash, S., and Prakash, L. (2000) *Nature* **406,** 1015–1019
43. Haracska, L., Unk, I., Johnson, R. E., Johansson, E., Burgers, P. M., Prakash, S., and Prakash, L. (2001) *Genes Dev.* **15,** 945–954
44. Kim, S. R., Maenhaut-Michel, G., Yamada, M., Yamamoto, Y., Matsui, K., Sofuni, T., Nohmi, T., and Ohmori, H. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94,** 13792–13797
45. Kobayashi, S., Valentine, M. R., Pham, P., O'Donnell, M., and Goodman, M. F. (2002) *J. Biol. Chem.* **277,** 34198–34207
46. Ogi, T., Kato, T., Jr., Kato, T., and Ohmori, H. (1999) *Genes Cells* **4,** 607–618
47. Ohashi, E., Bebenek, K., Matsuda, T., Feaver, W. J., Gerlach, V. L., Friedberg, E. C., Ohmori, H., and Kunkel, T. A. (2000) *J. Biol. Chem.* **275,** 39678–39684
48. Ling, H., Boudsocq, F., Woodgate, R., and Yang, W. (2001) *Cell* **107,** 91–102

Published online 12 March 2009

Nucleic Acids Research, 2009, Vol. 37, No. 9 2841–2853
doi:10.1093/nar/gkp138

Protein hnRNP A1 and its derivative Up1 unfold quadruplex DNA in the human *KRAS* promoter: implications for transcription

Manikandan Paramasivam[1], Alexandro Membrino[1], Susanna Cogoi[1], Hirokazu Fukuda[2], Hitoshi Nakagama[2,] and Luigi E. Xodo[1,*]

[1]Department of Biomedical Science and Technology, School of Medicine, P.le Kolbe 4, 33100 Udine, Italy and [2]Biochemistry Division, National Cancer Center Research Institute, 1-1, Tsukiji 5, Chuo-ku, Tokyo 104-0045, Japan

Received December 9, 2008; Revised and Accepted February 18, 2009

ABSTRACT

The promoter of the human *KRAS* proto-oncogene contains a structurally polymorphic nuclease hypersensitive element (NHE) whose purine strand forms a parallel G-quadruplex structure (called 32R). In a previous work we reported that quadruplex 32R is recognized by three nuclear proteins: PARP-1, Ku70 and hnRNP A1. In this study we describe the interaction of recombinant hnRNP A1 (A1) and its derivative Up1 with the *KRAS* G-quadruplex. Mobility-shift experiments show that A1/Up1 binds specifically, and also with a high affinity, to quadruplex 32R, while CD demonstrates that the proteins strongly reduce the intensity of the 260 nm-ellipticity—the hallmark for parallel G4-DNA—and unfold the G-quadruplex. Fluorescence resonance energy transfer melting experiments reveal that A1/Up1 completely abrogates the cooperative quadruplex-to-ssDNA transition that characterizes the *KRAS* quadruplex and facilitates the association between quadruplex 32R and its complementary polypyrimidine strand. When quadruplex 32R is stabilized by TMPyP4, A1/Up1 brings about only a partial destabilization of the G4-DNA structure. The possible role played by hnRNP A1 in the mechanism of *KRAS* transcription is discussed.

INTRODUCTION

The mammalian *KRAS* gene encodes for a guanine nucleotide-binding protein of 21 kDa that activates several cellular pathways controlling important events such as proliferation, differentiation and signalling (1). The Ras proteins behave as a molecular switch cycling between inactive GDP-bound and active GTP-bound states. The state of nucleotide occupancy is regulated by specific proteins named guanine nucleotides exchange factors (GEFs) and GTPases activating proteins (GAPs) (1,2). The *RAS* genes are frequently mutated in solid and haematological neoplasias with single point mutations at exons 12, 13 and 61 (3,4). The most common mutated *RAS* gene in solid tumours is *KRAS*, with a 90% incidence in pancreatic adenocarcinomas (5,6). As the mutated Kras protein has a defective GTPase activity, it is not inactivated by GAPs (7). It remains locked into the GTP-bound active state which continuously transmits to the nucleus mitotic signals that contribute to the neoplastic phenotypes in cancer cells (8–10). As pancreatic adenocarcinomas are refractory to conventional treatments, the discovery of new drugs capable to sensitize tumour cells to chemotherapy is being pursued in many laboratories. In our laboratory, we focused on *KRAS* and in order to design anti-*KRAS* drugs we investigated how the transcription of this proto-oncogene is controlled. Previous studies have shown that a nuclease hypersensitive element (NHE), located in the *KRAS* promoter upstream of the transcription start between -327 and -296, is responsible for most of the transcription activity (11). Earlier we reported that the purine strand of NHE is structurally polymorphic, as its tract of sequence recognized by nuclear proteins is able to fold into stable G-quadruplex structures (12,13). Using the purine strand of NHE (called 32R) in quadruplex conformation as a bait, we pulled down from a pancreatic nuclear extract three proteins with affinity for the *KRAS* quadruplex. By SDS–PAGE and mass spectrometry, we identified these proteins as poly[ADP-ribose] polymerase 1 (PARP-1), ATP-dependent DNA helicase 2, subunit 1 (Ku70) and heterogeneous ribonucleoprotein A1 (hnRNP A1) (13). Protein hnRNP A1 (from now on A1) is a member of the heterogeneous ribonucleoprotein family, which is highly abundant in the nucleus of actively

*To whom correspondence should be addressed. Tel: +39 432 494395; Fax: +39 432 494301; Email: luigi.xodo@uniud.it

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

growing mammalian cells (14,15). All members of the hnRNP family are characterized by two highly conserved RNA recognition motifs (RRMs) at the *N*-terminus and by a glycine-rich domain at the *C*-terminus (16,17). Although a recent structure of a co-crystal of Up1 (a proteolitic portion of A1 retaining binding activity) bound to the telomeric repeat $(TTAGGG)_2$ suggests that both RRMs interact with DNA (18), it has been reported that only one motif (RRM1) is sufficient for strong and specific binding to single-stranded telomeric DNA (19) and that its sub-element RNP11 mediates destabilization of quadruplex $(CGG)_n$ (20). Proteins hnRNP play various roles in mRNA metabolism (14,15) and in the biogenesis of telomeres (21). As protein A1 (and its derivative Up1) was reported to have a telomere-lengthening effect in erytroleukemia cells (21,22), it is suspected to function as an auxiliary factor of the telomerase holoenzyme (23). Considering that the 3′ G-rich repeats of the telomeres are folded in stable G-quadruplex structures, it has been hypothesized that A1 stimulates telomere elongation by disrupting high-order structures formed by the telomere repeats. Indeed, Up1 was reported to destabilize the bimolecular quadruplex formed by human telomere repeats d(TTAGGGTTAGGG), $d(TTAGGG)_4$ and the intramolecular quadruplex of $d(GGCAG)_5$ (23–25).

Since we discovered that A1 is associated to the *KRAS* promoter, in this study we have investigated the interaction between recombinant A1/Up1 and the *KRAS* G-quadruplex. Electrophoretic mobility shift assay (EMSA) showed that A1/Up1 binds to the *KRAS* quadruplex with high affinity and specificity, while CD and fluorescence resonance energy transfer (FRET) experiments revealed that A1/Up1 destabilizes this non B-DNA structure of the *KRAS* promoter. The results of our study support a transcription mechanism in which A1 should function as a G-quadruplex destabilizing protein, as it seems to occur in the G-rich 3′ overhang strand of the telomeres (23). In conclusion, this study sheds some light on the mechanism of *KRAS* transcription regulation and may be useful for the rationale design of anticancer drugs specific for oncogenic *KRAS*.

MATERIALS AND METHODS

DNA and proteins hnRNP A1/Up1

The oligonucleotides used in this study (Table 1) were obtained from MWG (Germany) and Microsynth (Switzerland). They have been purified by 20% PAGE (acrylamide: bisacrylamide, 19:1) in TBE, under denaturing conditions (7 M urea, 55°C). The bands were excised from the gel and eluted in water. The DNA solutions were filtered (Ultrafree-DA, Millipore) and precipitated. DNA concentration was determined from the absorbance at 260 nm of the oligonucleotides diluted in milli Q water, using as extinction coefficients 7500, 8500, 15 000 and 12 500 $M^{-1} cm^{-1}$ for C, T, A and G, respectively. Dual-labelled F-32R-T (5′ end with FAM, 3′ end with TAMRA) were HPLC-purified.

Recombinant proteins Up1 and A1 tagged to GST were expressed in *Escherichia coli* BL21 using plasmids pGEX-Up1 and pGEX-hnRNP A1. After transformation, the bacteria were grown for 2 h at 37°C with 50 μg/ml ampicillin to an A_{600} of 0.5–2.0 prior to induction with IPTG (100 μM final concentration). Cells were allowed to grow for 7 h before harvesting. The cells were centrifuged at 5000 r.p.m., 4°C. After centrifugation the supernatant was removed carefully and the cells washed twice with PBS. The pellet was re-suspended in a solution of PBS with PMSF 100 mM and DTT 1 M. The bacteria were lysed by sonication, added with Triton X-100 (1% final concentration) and incubated for 30 min on a shaker at room temperature. The lysate was then centrifuged for 10 min at 4°C at 10 000 r.p.m. Glutathione Sepharose 4B (GE Healthcare) (50% slurry in PBS) was added to the supernatant and incubated for 30 min at 4°C on a shaker. The mix was centrifuged for 5 min at 500 g and the pellet was washed 5 times in PBS and eluted with elution buffer containing 20 mM NaCl, 20 mM reduced glutathione, 200 mM Tris–HCl, pH 9.5 for A1 elution and pH 7.5 for Up1 elution. Alternatively, to remove the GST tag, the mix was centrifuged for 5 min at 500*g*, washed with PreScission Cleavage buffer (GE Healthcare) and centrifuged 5 min at 500*g*. The pellet was incubated for 4 h at 4°C with PreScission protease to cleave the GST tag from the purified proteins. After PreScission cleavage, the A1 or Up1 moieties were detached from GST which remained bound to the Gluthatione Sephadex beads. The reaction mixtures were centrifuged for 5 min at 500*g*, 4°C, and the untagged proteins collected from the supernatant. Finally, the purification of tagged and untagged Up1 and A1 proteins were checked by SDS–PAGE.

CD and fluorescence experiments

CD spectra have been obtained with a JASCO J-600 spectropolarimeter equipped with a thermostatted cell holder. CD experiments were carried out with oligonucleotides (3 μM) in 50 mM Tris–HCl, pH 7.4, 100 mM KCl. Spectra were recorded in 0.5 cm quartz cuvette. A thermometer inserted in the cuvette holder allowed a precise measurement of the sample temperature. The spectra were calculated with J-700 Standard Analysis software (Japan Spectroscopic Co., Ltd) and are reported as ellipticity (mdeg) versus wavelength (nm). Each spectrum was recorded three times, smoothed and subtracted to the baseline.

Fluorescence measurements were carried out with a Microplate Spectrofluorometer System (Molecular Devices) using a 96-well black plate, in which each well contained 50 μl of 200 nM dual-labelled F-32R-T in 50 mM Tris–HCl, pH 7.4 and KCl as specified in the figure captions. Before adding the protein, the samples were incubated for 24 h at room temperature in the specified buffer. The protein (Up1, A1 or BSA) was added 30 min before fluorescence analysis. The emission spectra were obtained by setting the excitation wavelength at 475 nm, the cut-off at 515 nm and recording the emission from 500 to 650 nm. Upon addition of KCl, F-32R-T assumes a folded quadruplex conformation and FRET is expected between the 5′ and 3′ fluorophores. The emission

intensity of the donor (FAM) decreases while the intensity of the acceptor increases, correspondingly, as K^+ is added to the sample solution. The energy transfer from the donor to the acceptor and vice versa can be empirically represented by the parameter P:

$$P = \frac{I_D}{(I_D + I_A)}$$

where I_D and I_A are the intensities of the donor and acceptor (26,27). Fluorescence melting experiments were performed on a real-time PCR machine (iQ5, BioRad), using a 96-well plate filled with 50 µl solutions of dual-labelled F-32R-T. The protocol used for the melting experiments is the following: (i) equilibration step of 5 min at low temperature (15°C); (ii) stepwise increase of the temperature of 1°C per min for 76 cycles to reach 95°C. All samples in the wells were melted in 76 min.

Kinetic experiments were carried out using the iQ5 real-time machine. Oligonucleotide F-32R-T (200 nM) in 100 mM KCl, i.e. in the quadruplex conformation, was mixed with the complementary 32Y strand and the increase at 525 nm of the fluorescence was measured as a function of time. The experiment was also performed adding to F-32R-T a mixture containing 32Y (8-fold) and Up1 (400 nM). The increase of fluorescence $\Delta F = F - F_0$, where F_0 and F is the fluorescence at 525 nm (FAM) at $t = 0$ and at any time t, was best-fitted to a single or double-exponential curve. The half-life of the reaction is given by $t_{1/2} = 0.693/k$.

EMSA

Oligonucleotides 32R, *HRAS-1*, *HRAS-2*, *CMYC*, *CKIT*, *VEGF*, 32Y, Gmut1 and Gmut2 were end-labelled with [γ-^{33}P]ATP and T4 polynucleotide kinase. Duplex dsNHE was prepared annealing (10 min at 95°C, overnight at room temperature) a mixture containing equimolar amounts of radiolabelled 32R and complementary 32Y in 50 mM Tris–HCl, pH 7.4, 100 mM NaCl. Before EMSA, the quadruplex-forming oligonucleotides were allowed to form their structure in 50 mM Tris–HCl, pH 7.4, 100 mM KCl, 37°C (overnight incubation). Radiolabelled oligonucleotides (35 nM) were treated for 30 min at room temperature with different amounts of A1/Up1, (r ([protein]/[oligonucleotide]) ratios are specified in Figure 3) in 20 mM Tris–HCl, pH 8, 30 mM KCl, 1.5 mM $MgCl_2$, 1 mM DTT, 8% glycerol, 1% Phosphatase Inhibitor Coktail I (Sigma, Milan, Italy), 5 mM NaF, 1 mM Na_3VO_4, 2.5 ng/ml poly [dI-dC]. After incubation, the reaction mixtures were loaded in 8% TBE (1×) polyacrylamide gel, thermostatted at 16°C. After running the gel was dried and exposed to autoradiography (G E Healthcare, Milan) for 24–36 h at −80°C.

Polymerase-stop assay

A linear DNA fragment of 87 nt, containing the G-rich element of NHE, was used as a template for Taq polymerase primer-extension reactions. This DNA sequence was purified by PAGE. The template (100 nM) was mixed with the labelled primer (50 nM) in 100 mM KCl, Taq buffer 1× and overnight incubated at 50°C. The primer extension reactions were carried out for 1 h, by adding 10 mM DTT, 100 µM dATP, dGTP, dTTP, dCTP and 3.75 U of Taq polymerase (Euro Taq, Euroclone, Milan). The reactions were stopped by adding an equal volume of stop buffer (95% formamide, 10 mM EDTA, 10 mM NaOH, 0.1% xylene cyanol, 0.1% bromophenol blue). The products were separated on a 15% polyacrylamide sequencing gel prepared in TBE, 8 M urea. The gel was dried and exposed to autoradiography. Standard dideoxy sequencing reactions were performed to detect the points in which DNA polymerase I was arrested.

RESULTS

We previously demonstrated that the G-rich strand of NHE can form G-quadruplex structures (13,28). By means of CD and DMS-footprinting experiments we found that the G-tract called 32R forms a parallel G-quadruplex characterized by three G-tetrads (T_m of 70°C in 100 mM KCl) (Figure 1). Pull-down assays with a pancreatic nuclear extract combined to mass spectrometry showed that quadruplex 32R binds to three proteins: PARP-1 (116 kDa), Ku70 (72 kDa) and A1 (34 kDa) (13). Since A1 is involved in the biogenesis of the telomeres as a G4-DNA destabilizing protein (23) and is able to disrupt the secondary structures of the hypervariable minisatellite sequence $d(GGCAG)_5$ (24), we asked whether A1/Up1 can have a similar functional role in the human *KRAS* promoter. To address this question, recombinant A1 and its derivative Up1 were expressed in *Escherichia coli* as proteins fused to GST and purified by affinity chromatography with glutathione sepharose 4B. The GST moiety was removed with a pre-scission protease and recombinant tagged and untagged proteins were obtained with a high purity level (Figure 2). Up1 is a proteolitic fragment (195 aa) of A1 (319 aa) that retains the two RNA-recognition motifs (RRMs) responsible for binding to nucleic acids (18,22).

The interaction between A1/Up1 and a variety of DNA substrates, some of which were structured in G4-DNA and some not, was analysed by EMSA. ^{33}P-labelled 32R (35 nM) was first incubated for 24 h in 100 mM KCl to allow quadruplex formation, then incubated for 30 min with increasing amounts of Up1 or A1: r ([protein]/[32R]) = 0, 0.5, 1, 2, 5, 10, 20, 50, 100. As preliminary experiments showed that GST-tagged and untagged proteins behave in the same way, we performed EMSA with the tagged proteins. Figure 3a and b shows that quadruplex 32R forms with A1/Up1 a DNA–protein complex that, being detected even at $r = 0.5$, should have a 1 : 1 stoichiometry. In addition, for $r > 20$, another slow-migrating DNA–protein complex appears in the gel, most likely due to a 1:2 complex. When r was increased to 200 and the samples run in a longer gel, 32R migrated essentially as 1:2 complex (Figure 3c). The formation of two DNA–protein complexes by A1/Up1 is in keeping with the results of Zhang *et al.* (23) and the crystal structure of $d(TTAGGG)_2$ bound to Up1 (18). Since a tract of 12 nt functions as a minimum binding unit, 32R has

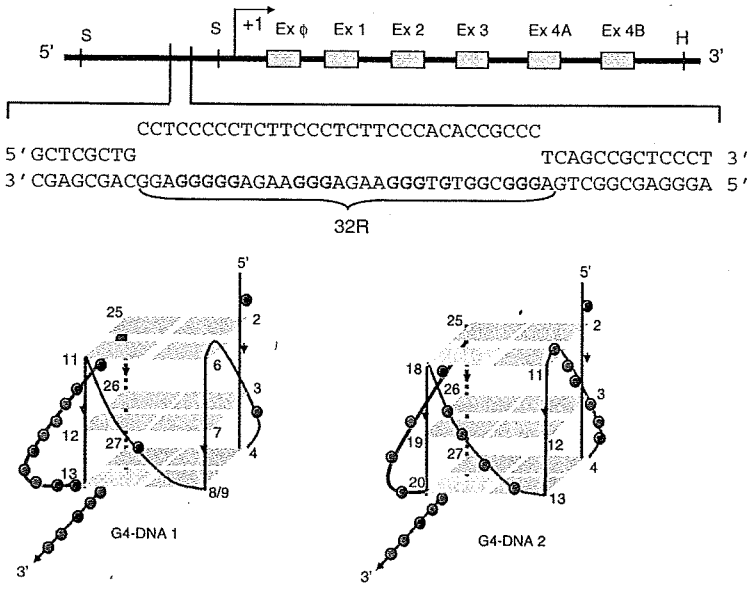


Figure 1. Sequence of the nuclease hypersensitive element (NHE) in the human *KRAS* promoter. The G-rich sequence 32R forms a G-quadruplex whose putative structure, consistent with CD and dimethyl sulfate footprinting, is G4-DNA1, which is characterized either by a flipped-out thymidine connecting G7 to G9 or a GGGT triad (13). The expected G4-DNA2 structure is not supported by dimethyl sulfate footprinting. The nucleotides of 32R (Table 1) are numbered from the 5′-end.

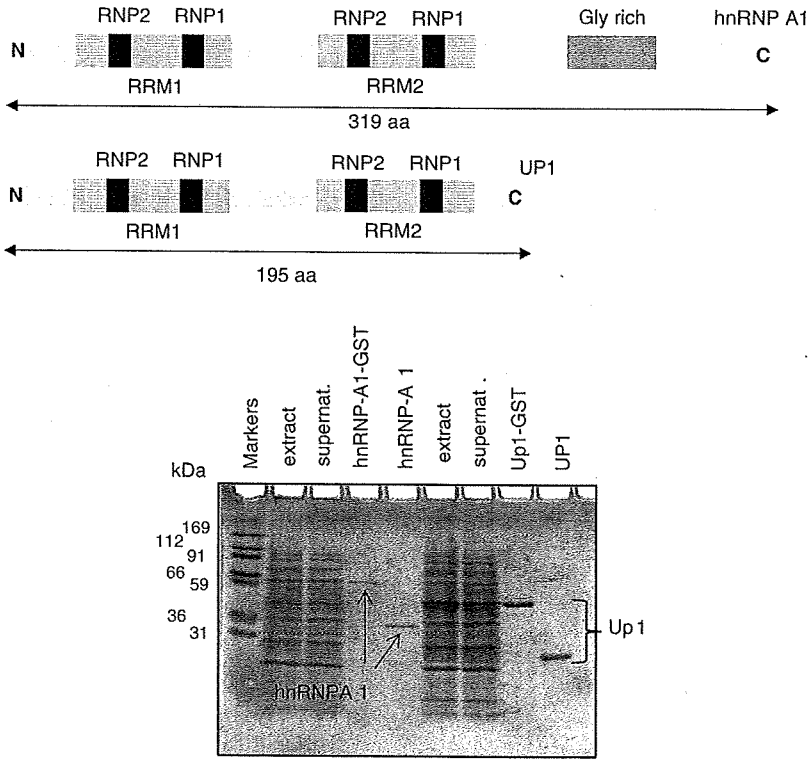


Figure 2. Schematic representations of proteins hnRNP A1 and Up1. The two RNA-recognition motifs (RRMs), that mediate ssDNA binding, contain each two conserved RNP2 and RNP1 submotifs. Up1 encompasses the amino-terminal two-third of the hnRNPA1 sequence. SDS PAGE of GST-tagged and untagged hnRNP A1 and Up1, after gluthatione sepharose 4B purification. Lane 1, protein markers; lane 2, total extract (hnRNP A1); lane 3, supernatant; lane 4, purified GST-tagged hnRNP A1; lane 5, purified untagged hnRNP A1; lane 6, total extract (Up1); lane 7, supernatant; lane 8, purified GST-tagged Up1; lane 9, purified untagged Up1.

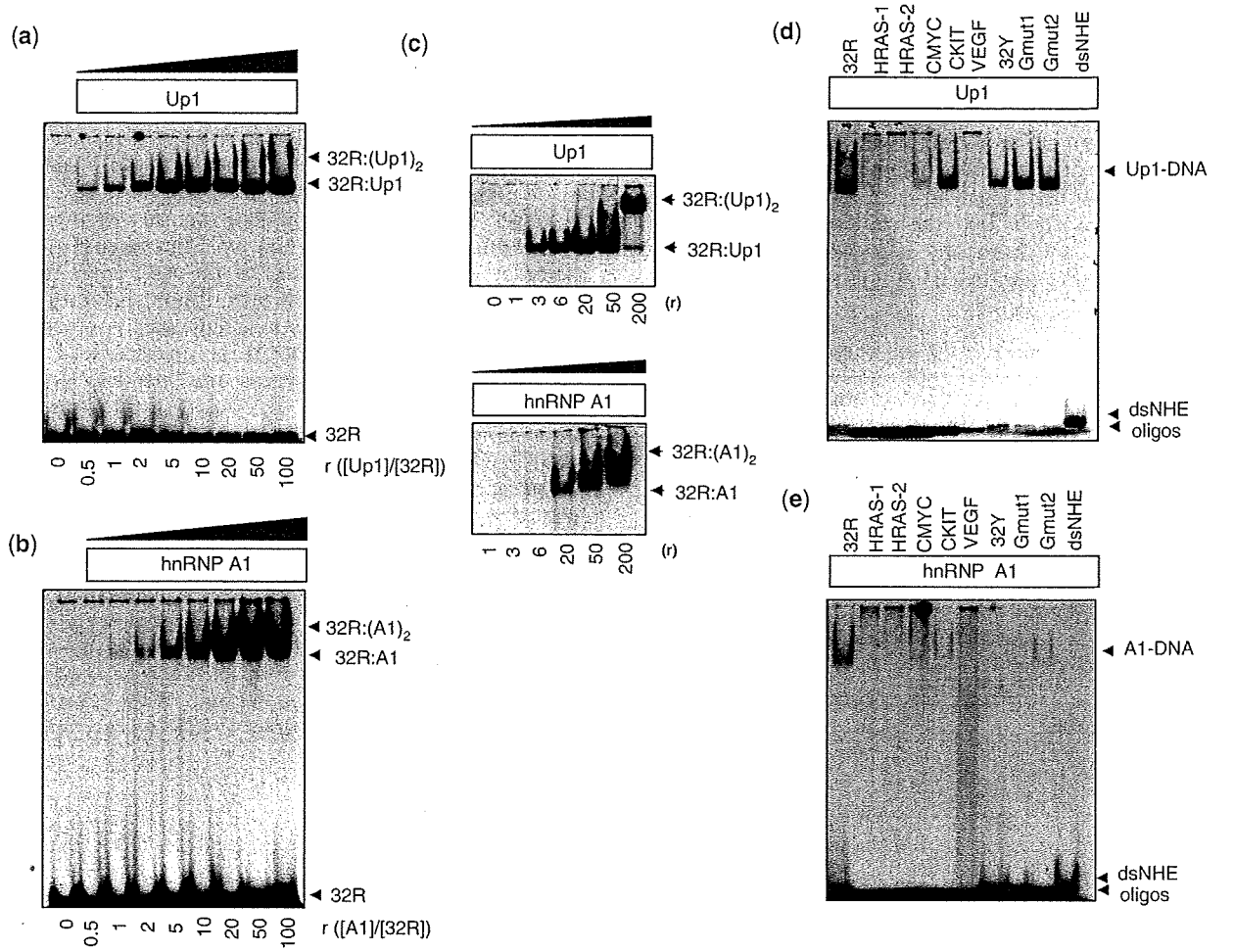


Figure 3. (**a**, **b**) EMSA of 35 nM ^{33}P-labelled quadruplex 32R after 30 min incubation with increasing amounts of Up1 or A1 at the specified r values, in 20 mM Tris–HCl, pH 8, 30 mM KCl, 1.5 mM $MgCl_2$, 1 mM DTT, 8% glycerol, 1% Phosphatase Inhibitor Coktail 1 (Sigma), 5 nM NaF, 1 mM Na_3VO_4, 2.5 ng/μl poly dI-dC, for 25°C. The analyses were carried out in 8% polyacrylamide gel (29:1) in TBE (1×) at 16°C. Before the EMSA, 32R was incubated overnight in 100 mM KCl to get it in the G-quadruplex conformation; (**c**) EMSA as in (a, b) but with r values up to 200; (**d**, **e**) EMSA of A1/Up1 mixed to various DNA substrates [G-quadruplexes 32R, *HRAS*-1, *HRAS*-2, *CMYC*, *CKIT*, *VEGF*, dsNHE (32R:32Y) and unstructured oligonucleotides Gmut1, Gmut2, 32Y]. PAGE carried out in 8% polyacrylamide gel (29:1) in TBE (1×) at 16°C.

potentially two binding sites, which can in principle form two DNA–protein complexes by binding one or two protein molecules. By quantifying the intensity of the electrophoretic bands, we roughly estimated that the dissociation constant K_d of the 1:1 complex is about 50 nM for Up1 and 200 nM for A1. We also tested the binding specificity of A1/Up1 for a variety of well known G-quadruplex structures obtained from *CMYC*, *CKIT*, *VEGF* and *HRAS* promoter sequences (29–32) (for *HRAS* quadruplexes, see Supplementary Data S1) (Figure 3d and e). The various DNA substrates have been ^{33}P-labelled and treated with an excess of protein ($r = 50$). It can be seen that A1 shows good specificity for the *KRAS* quadruplex, as it does not bind to the other quadruplex-forming sequences, unstructured oligonucleotides Gmut1, Gmut2, 32Y (the complementary polypyrimidinic NHE strand) and dsNHE (32R:32Y) (Table 1). Instead, protein Up1, besides quadruplex 32R, shows affinity also for the *CKIT* quadruplex and unstructured oligonucleotides.

To analyse the effect of A1/Up1 on the *KRAS* G-quadruplex, we could not employ electrophoresis because the mobility between an *intra*molecular quadruplex and its unfolded form is not very different. Therefore, we used spectroscopic techniques such as circular dichroism (CD) and FRET. Figure 4 shows that in 100 mM KCl, 32R is characterized by a CD signature typical of a parallel G-quadruplex: a strong and positive ellipticity at 260 nm and a weak and negative ellipticity at 240 nm (33). When quadruplex 32R is denatured by increasing the temperature, the positive 260 nm band is dramatically reduced and its spectrum becomes similar to that of unstructured oligonucleotides (data not shown). Thus, the structural transition from quadruplex-to-ssDNA is accompanied by a strong reduction of the 260 nm ellipticity. A similar transition was obtained by adding to

quadruplex 32R increasing amounts of A1/Up1 (r = 1, 2, 4, 6). It can be seen that the protein causes a progressive reduction of the 260 nm ellipticity, indicating that the G4-DNA structure is unfolded by the protein. As a control, we treated quadruplex 32R with an unrelated protein, the trypsinogen inhibitor, and found that the 260 nm ellipticity was not affected and remained constant at all protein concentrations used. The CD spectra of Up1 at increasing concentrations show that the protein between 240 and 320 nm does not have any negative band, but below 240 nm it shows a negative band typical of the polypeptide backbone. The CD data showing G-quadruplex unfolding are in keeping with those previously obtained with the telomeric TTAGGG repeat (24,25) and the hypervariable minisatellite sequence d(GGCAG)$_5$ treated with Up1 (24).

Table 1. Oligonucleotides (5′ > 3′) used in this study

AGGGCGGTGTGGGAAGAGGGAAGAGGGGGAGG	32R
F-AGGGCGGTGTGGGAAGAGGGAAGAGGGGGAGG-T	F-32R-T
AGGGAGGGCGCTGGGAGGAGGG	*CKIT*
GGGCGGGCCGGGGGCGGGTCCCGGCGGGG	*VEGF*
TGGGGAGGGTGGGGAGGGTGGGGAAGG	*CMYC*
TCGGGTTGCGGGCGCAGGGCACGGGCG	*HRAS*-1
CGGGGCGGGGCGGGGGCGGGGGCG	*HRAS*-2
GCGGTGTGTGAAGAGTGAAGAGTGGGATGCAG	Gmut1
GCATTCTGATTACACGTATTACCTTCACTCCA	Gmut2
CCTCCCCCTCTTCCCTCTTCCCACACCGCCCT	32Y
GTACTACACTTGATA	primer
ACCTTGATGAATCCAGGGCGGTGTGGGAAGAG	template
GGAAGAGGGGGAGGAATCGCTACCGTTAAGCA	
TCGATCATATCAAGTGATAGTAC	

F: FAM; T: TAMRA.

The unfolding of the human *KRAS* quadruplex by A1/Up1 was also investigated by FRET, using the quadruplex-forming sequences tagged at the 5′ and 3′ ends with FAM (donor) and TAMRA (acceptor) (34). By exciting F-32R-T at 475 nm, the emission intensity of the donor at 525 nm decreases while the emission intensity of the acceptor at 580 nm increases, as the KCl concentration is increased from 0 to 140 mM (Supplementary Data S2). F-32R-T folded in the G-quadruplex conformation (T_m = 75°C in 140 mM KCl) is characterized by a P-value of 0.52 (see 'Materials and methods' section). This P-value is higher than that observed for the quadruplex formed by the human telomeric repeat d(GGGTTAG GGTTAGGGTTAGGG) (26), because F-32R-T forms a parallel quadruplex where the two fluorophores are at opposite ends of the structure (13). When the G-quadruplex is destabilized by scaling down the KCl concentration to zero or by adding the complementary 32Y strand, that transforms the G-quadruplex into a B-DNA duplex where the donor and acceptor are separated by about 115 Å, the donor fluorescence significantly increases (for instance, from spectrum 2 to spectrum 1, Figure 5a) and the P-value becomes 0.75. This means that the unfolding of quadruplex F-32R-T is accompanied

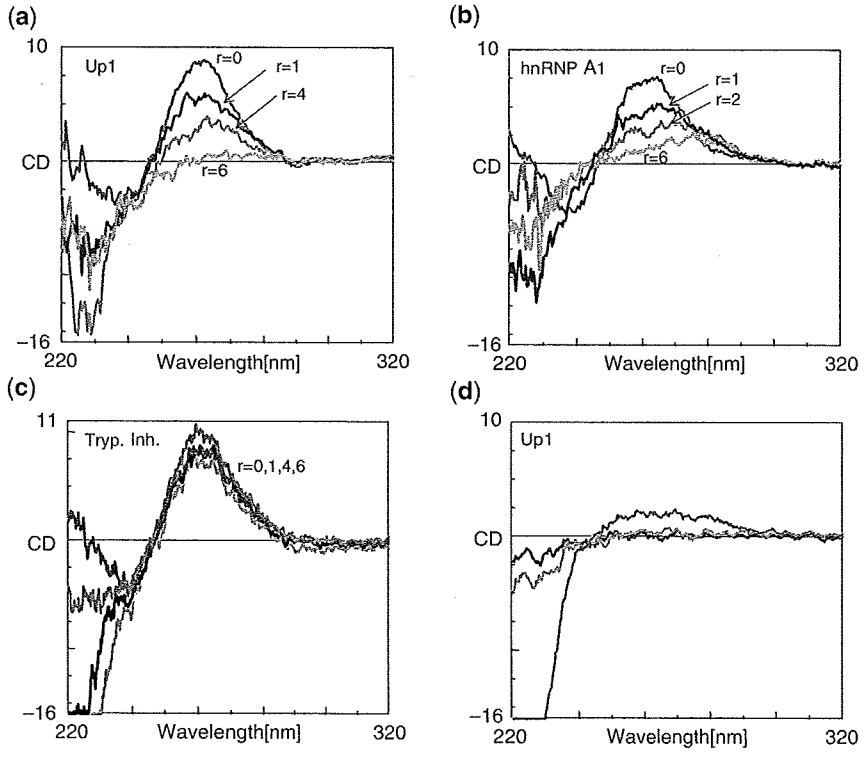


Figure 4. CD of 32R (2 μM) in 50 mM Tris pH 7.4, 100 mM KCl in the presence of increasing amounts of Up1 (r = 0, 1, 4, 6) (r = [protein]/[DNA]) (**a**); hnRNPA1 r = 0, 1, 4, 6, (**b**); trypsinogen inhibitor (TI) (r = 0, 1, 4, 6) (**c**). The CD of Up1 at three concentrations is reported (2, 4 and 8 μM) (**d**). Spectra have been recorded at room temperature with a path length cuvette of 0.5 cm. Ordinate reports ellipticity values in mdeg.

by a $\Delta P = 0.75 - 0.52 = 0.23$. We then asked if quadruplex F-32R-T is unfolded by A1/Up1. To choose at which ionic strength the FRET experiments in the presence of A1/Up1 should be performed, we measured the T_m of quadruplex F-32R-T in KCl and NaCl solutions (in 50, 100 and 140 mM KCl, T_m is 48, 70 and 75°C, respectively; in 100 mM NaCl, the T_m is 32°C). Figure 5a shows the effect on quadruplex F-32R-T in 50 mM KCl (T_m = 48°C,

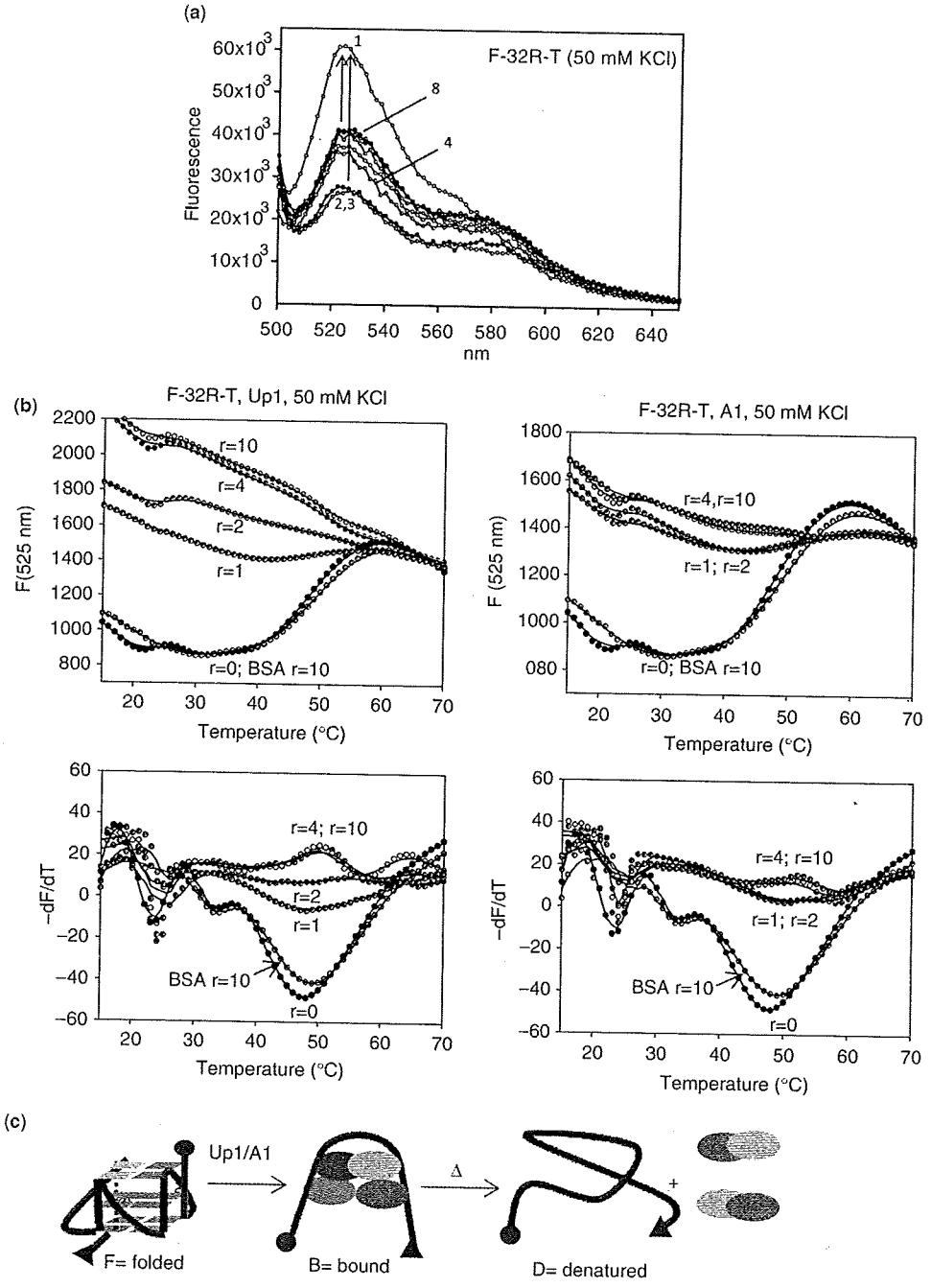


Figure 5. (a) Fluorescence spectra of 200 nM F-32R-T in water (spectrum 1) or 50 mM Tris-HCl, pH 7.4, 50 mM KCl in the absence (spectrum 2) or presence of BSA (r = 10, spectrum 3) or Up1 (r = 0.5, 1, 3, 6, 10, spectra 4–8); (b) row FRET-melting curves (F_{525} versus T) obtained with the iQ5 real-time PCR machine of quadruplex F-32R-T treated with A1/Up1 at various [protein]/[DNA] ratios, in 50 mM Tris pH 7.4, 50 mM KCl. As reference a melting curve of F-32R-T in the presence of BSA (r = 10) is reported. Bottom panels show the corresponding first derivative curves, dF_{525}/dT versus T. The G-quadruplex was incubated with the protein for 30 min prior to melting; (c) schematic representation of the U-shape structure of the DNA–protein complex between F-32R-T and Up1.