expand to hundreds or even thousands of repeats in certain human disorders (e.g., fragile X syndrome, myotonic dystrophy, Huntington disease) by un-clarified mechanisms [4,5]. In contrast, minisatellite repeats (also known as variable number of tandem repeats; VNTR) are composed of longer repetitive units of 5 or 6–100 nucleotides and are found in arrays expanded up to 10–20 kbp. As opposed to microsatellites, only a few 1000 such loci are present in the genome [6]. Although the biological significance of minisatellite repeats remains largely unknown, some repeat regions are known to be hotspots for meiotic recombination [7,8]. They are also occasionally found in fragile chromosomal sites and could serve as targets for genomic recombination or chromosomal breakage [9,10].

Genetic alteration at a few specific microsatellite and minisatellite repeats results in several human disorders. Alterations in microsatellite repeats are frequently found in cancer cells, and are caused by both genetic and functional alterations of genes encoding mismatch repair proteins, including MLH1, MSH2, MSH6, PMS1, PMS2 and MLH3 [11,12]. Mutations in mismatch repair genes, as occurs in hereditary non-polyposis colorectal cancers, lead to microsatellite instabililty (MSI). MSI can, in turn, cause frame-shift mutations in long tracts of mononucleotides; examples of this have been observed in the transforming growth factor-β type II, BAX and insulin-like growth factor II receptor genes [11,12]. Under mismatch repair-deficient conditions, microsatellite repeats are altered by the insertion or deletion of small numbers of mono- or di-nucleotide repeat units [11,12].

Alteration at certain minisatelllite loci are also implicated in genetic predisposition to some human disorders, such as insulin-dependent diabetes mellitus type 2 (IDDM-2) [13]. A rare allele of the Ha-*ras*-VNTR, which is located in the 3′ region of the Ha-*ras* gene, appears to be associated with various cancers including breast, colon, urinary bladder and acute leukemia [14].

In this review article, characteristic structural features of G-rich short tandem repeats are summarized, and molecular mechanisms involved in maintaining genomic stability at these G-rich repetitive sequences are discussed.

## 2. Minisatellite Pc-1 and Pc-1-like repeats

The mouse Pc-1 minisatellite (also known as the expanded simple tandem repeat *Ms6-hm*) consists of a tandem array of G-rich repeats $d(GGCAG)_n$, flanked with locus specific sequence [15]. The length of the repeat arrays vary widely among mouse strains [16,17]. Pc-1 was observed to be a recombination hotspot at meiosis, with a germ-line mutation rate as high as 10% per gamete [18,19]. In normal somatic cells, however, the repeats are relatively stable and the mutation rate has been estimated to be around $2 \times 10^{-9}$ per cell division [17]. To date, many hypervariable minisatellites, consisting of G-rich repeat units similar to Pc-1, have been identified in the mouse and other mammalian genomes. Mouse loci Pc-2 and mo-1 are composed of d(GGCAGG) and d(CTGGGCAGGGAGGA) repeats [16], and human 33.6 and 33.15 minisatellites consist of d(AGGGCTGGAGG) and d(AGAGGTGGGCAGGTGG) repeats, respectively [20]. Since the core sequences of G-rich minisatellites share high similarity among the repeats, more than 20–30 bands are easily detected in mouse and human genomes by a low-stringent DNA fingerprint analysis using Pc-1 as a probe [21,22]. These Pc-1-like repeats are also stable in normal somatic cells, although there are several conditions that can induce mutational instability at this locus [19,21–23].

## 3. Size alteration of G-rich minisatellite repeats in the genome

Minisatellite repeats are generally stable in somatic cells compared to germ cells as described above [24–26], but alterations at minisatellite regions can be induced both in cell cultures and in vivo [21,27–31]. When culture cells are exposed to a variety of chemical carcinogens, ultraviolet irradiation or ionizing radiation, DNA fingerprint analysis reveals alterations of the banding pattern of genomic regions containing tandem repeats. We hereafter refer to these genetic alterations in minisatellite regions as 'minisatellite mutations', which may include, for example, changes in non-repeat flanking regions and restriction endonuclease sites, deletions/insertions within the repeat regions, and recombination events. Additionally, these types of mutations are frequently observed in sporadic human colorectal and gastric cancers, the incidence being 56% and 25%, respectively [22].

Minisatellite instability is an interesting phenomenon that is suggestive of a novel type of genomic instability [9,23]. As we reported previously, minisatellite instability was observed in a Scid fibroblast cell line transformed by simian virus 40 (SV40) large tumor antigen, SC3VA2, and an embryonal Scid fibroblast cell line, SC1K [23]. Both of these cell lines are deficient in DNA-dependent protein kinase (DNA-PK) activity. A considerable number of size alterations in minisatellites were observed in subclones of both the SC3VA2 and

SC1K cells (45 ± 6% and 37 ± 3%, respectively). These findings were corroborated in several sets of clones. In contrast, cells derived from the RD13B2 cell line, which was established from SC3VA2 by introducing a fragment of the human chromosome 8q, which includes the DNA-PK catalytic subunit, showed a very low frequency of minisatellite mutation (3 ± 3%). Although the underlying molecular mechanisms for this instability have not yet been fully elucidated, a lack of correlation between the presence of minisatellite mutations and microsatellite mutation/instability [21–23] suggests that the mismatch repair system is not involved. Therefore, it appears that another, as yet undiscovered mechanism specific to minisatellite instability, is at play in this case.

## 4. Formation of G4′ structure by d(GGCAG)$_n$ in vitro

We used structural analyses to investigate whether higher order structures occur in G-rich repetitive sequences to gain further insight into the molecular mechanisms operating in induced instability at these sites. The formation of secondary structures is likely due to the G-rich nature of the repeats. For example, a triple-stranded DNA between d(GGA:TCC) repeats and d(GGA) repeat oligonucleotides forms a D-loop-like higher order structure [32]. The d(CTG:CAG) repeats from the myotonic dystrophy locus form slipped-strand DNA (S-DNA) [33]. A polymorphic G-rich minisatellite, d(ACAGGGGTGTGGGG)$_n$, located in the promoter region of the human insulin gene [34], the G-rich immunoglobulin switch repeats [35], and the G-rich sequence d(G$_{16}$CG(GGT)$_2$GG) in the promoter region of the chicken β-globin gene are also known to form unusual DNA structures, such as four-stranded quadruplexes [36].

Based on these observations, genomic instability at G-rich minisatellites may also be caused as a consequence of the formation of higher order structures of DNA. To investigate this hypothesis, we used the Pc-1 minisatellite repeat, d(GGCAG), as a representative G-rich minisatellite locus of the mouse genome. As shown in Fig. 1A, NMR analysis demonstrated that a synthetic oligonucleotide containing eight repeats of d(GGCAG) (d(GGCAG)$_8$) has the characteristic features necessary to form a quadruplex structure [37]. Circular dichroism (CD) spectrum analysis showed a specific positive CD band at 290–295 nm (Fig. 1B). A positive peak at 260 nm, which is characteristic of the quadruplex with a four-stranded (inter-strand) parallel quadruplex as detailed previously [34,38], was not observed. Furthermore, melting temperatures of oligonucleotides d(GGCAG)$_8$ and d(GGCAG)$_5$ determined by thermal CD melting curves under physiological conditions (100 mM KCl) were independent of DNA concentration. As a reference, a 10-fold dilution of DNA concentration for duplexed DNA causes a 6–8 °C drop of the melting temperature [39]. Taken together, the data indicate that d(GGCAG)$_8$ and d(GGCAG)$_5$ form intra-strand folded-back quadruplex structures (also called a G4′-structure) as opposed to four-stranded quadruplex structures (G4-structure) [37]. The CD spectrum of d(GGCAG)$_3$ was different from those of d(GGCAG)$_5$ and d(GGCAG)$_8$, and gave a broad positive band around 280 nm, but not the typical positive band of the G4′ structure at 290–295 nm (data not shown). Taken together, intra-molecular G:G:G:G tetrad formation with four d(GGG) sequences in d(GGCAG)$_n$

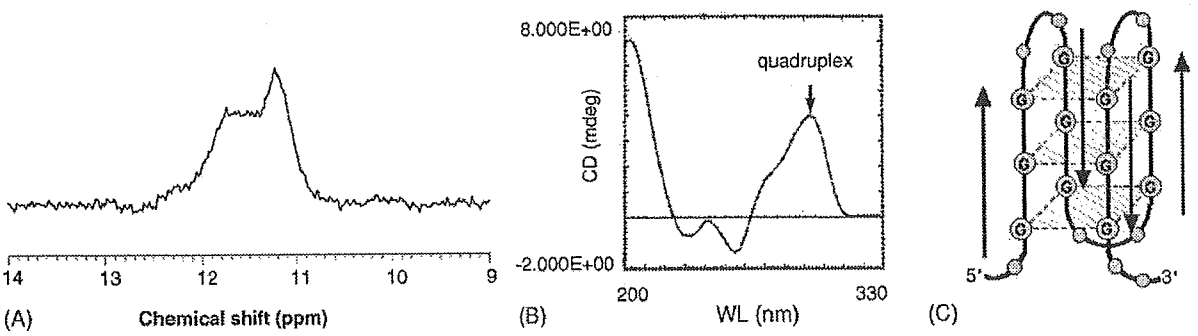


Fig. 1. NMR and CD analysis of d(GGCAG)$_8$ (A) Imino proton NMR spectrum of d(GGCAG)$_8$ in $^2H_2O$. The appearance of signals around at 11–12 ppm even in $^2H_2O$ at 5 °C is characteristic of a quadruplex structure with guanine-quartets. (B) A positive CD band at 290–295 nm is characteristic of a quadruplex with a folded-back structure. A four-stranded parallel quadruplex does not give a positive band at 290–295 nm, but gives one at 260 nm. (C) Schematic diagram illustrating how the d(GGCAG) repeats form a G4′ structure is depicted, although we have not determined yet whether the G4′ structure formed by d(GGCAG) repeats is a "chair" or "basket" type quadruplex [79]. In this model, one G-stretch is aligned anti-parallel to the next G-stretch, and four guanine residues are arranged in a square-planar array.

repeats ($n \geq 5$) could explain the formation of a G4′ structure, as depicted in Fig. 1C.

Although we have not yet proven the in vivo formation of a G4′ structure by d(GGCAG) minisatellite repeats, recent studies by Duquette et al. [40] and Paeschke et al. [41] strongly support the formation of G-quadruplex DNA structures in vivo. Duquette et al. demonstrated the formation of G-loops, novel structures containing G4 DNA, in *E. coli* using plasmid harboring the G-rich regions derived from the mammalian immunoglobulin S (switch) regions and GQN1, an endonuclease that specifically cleaves G4 DNA in the single stranded region 5′ of the stacked G-quartet [42]. Paeschke et al. beautifully demonstrated G-quadruplex formation in the macronucleus of the ciliated *Stylonychia* [41] by the use of antibodies specific for telomeric G-quadruplex DNA [43]. These findings support the hypothesis that G4′ DNA structures form at Pc-1 and Pc-1-like minisatellite repeats in vivo as well. However, further studies are required to validate this hypothesis.

## 5. DNA synthesis arrest at the d(GGG) sites in vitro

Several studies have demonstrated the ability of d(GGG) sites to cause DNA synthesis arrest. For example, the in vitro DNA synthesis assay showed DNA synthesis arrest at the first d(GGG) site of a single-stranded phagemid (pYA-3) carrying 12 repeats of d(GGCAG), with additional weaker stops at the second, third and fourth d(GGG) sequences (Fig. 2). A primer extension reaction using a synthetic oligonucleotide containing $d(GGCAG)_{15}$ gave similar results (further described below). Inhibition of in vitro DNA synthesis by the formation of G-quadruplex was also reported previously using synthetic oligonucleotides of human $[d(TTAGGG)_n]$, the *Tetrahymena* $[d(TTGGGG)_n]$ telomeric sequences [44] and the G-rich sequence found in the promoter region of the chicken β-globin gene, $d(G_{16}CG(GGT)_2GG)$ [36,45].

Recently, we found that DNA replication of plasmid in *E. coli* was substantially affected by the presence of d(GGCAG) repeats (unpublished observations). Considering both in vitro and in vivo data, it appears that the presence of the G-rich repeats in the template may inhibit normal DNA synthesis (replication). If this is the case, the lack of repeat instability in somatic cells suggests that somatic cells may have specialized machinery to compensate for the formation of unfavourable higher order DNA structures and to allow replication through the G-rich repeats.

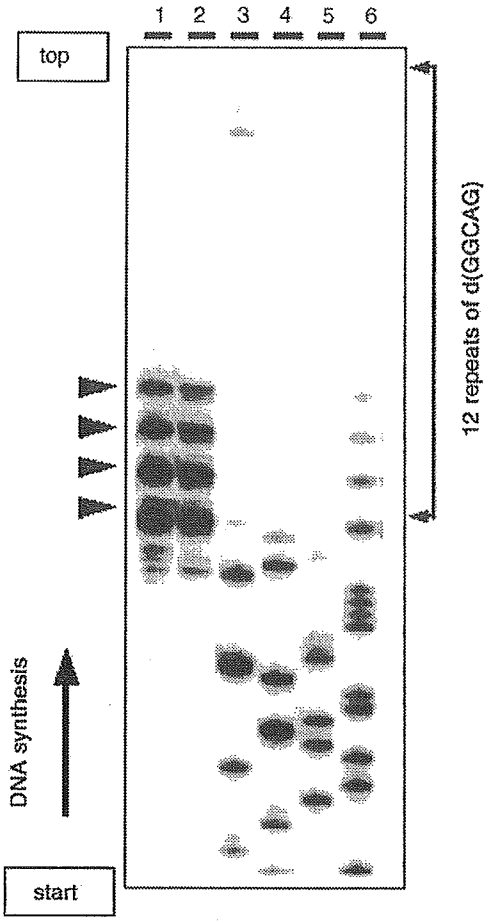


Fig. 2. In vitro DNA synthesis assay using a single-stranded phagemid carrying a $d(GGCAG)_{12}$ repeat. Primer extension reaction was performed as follows using the single-stranded phagemid pYA-3 carrying $d(GGCAG)_{12}$ [37]. The pUC/M13-M4 forward primer (Takara) labeled with $^{32}P$ at the 5′-end (600 fmol) was annealed with single-stranded pYA-3 (250 fmol) in 5 μl of buffer (40 mM Tris–HCl pH 7.5, 20 mM $MgCl_2$, 50 mM KCl, 5 μM dNTPs) at 72 °C for 5 min, followed by slow cooling to 37 °C. The reaction was started by the addition of 0.5 μl of Sequenase (0.8 U) or the Klenow Fragment (1.0 U), and the mixture was further incubated at 37 °C for 10 min. The reaction was stopped by adding 3.5 μl of the stop solution (10 mM NaOH, 95% formamide, 0.05% bromophenol blue, 0.05% xylene cyanole) and after incubation at 95 °C for 2 min, 3 μl aliquots of the sample DNA were electrophoresed in an 8% polyacrylamide gel containing 7 M urea. Polyacrylamide gel electrophoresis (PAGE) analysis of primer extension reaction with Sequenase (lane 1) or Klenow (lane 2) clearly shows the arrest of in vitro DNA synthesis at the first to fourth d(GGG) sites (arrow heads). Lanes 3–6 show the sequencing reactions with Sequenase in the presence of ddATP, ddTTP, ddCTP and ddGTP, respectively, using 200 fmol of single-stranded pYA-3 plasmids for each reaction. An arrow at the left indicates the direction of DNA synthesis.

## 6. Isolation of G-rich minisatellite binding proteins

To elucidate the underlying molecular mechanisms for maintenance of genomic stability at genomic G-

Table 1
Isolation of minisatellite binding proteins from NIH3T3 cell extract [48]

| ID | MW (kDa) | Common name |
|---|---|---|
| MNBP-A | 30 | hnRNP A3 |
| MNBP-B | 24 | hnRNP A1/UP1 |
| MNBP-D | 29 | ND[a] |
| MNBP-E | 98 | LRP130[b] |
| MNBP-F | 130 | LRP130 |
| MNBP-G | 110 | Tudor-SN/SND1 |

[a] ND: not determined.
[b] Probably a proteolytic product of MNBP-F, LRP130.

rich minisatellite sequences, we carried out an electrophoretic mobility shift assay (EMSA) to identify minisatellite binding proteins (MNBPs) using cell-free extracts from NIH3T3 cells treated with okadaic acid (Table 1). Okadaic acid (OA) is a known tumor promoter [46] and a specific inhibitor of the mammalian serine/threonine protein phosphatases [47]. OA is capable of inducing minisatellite mutations in NIH3T3 cells [21]. Two proteins, hnRNP A3 (MNBP-A) and hnRNP A1 (MNBP-B), were identified from OA-treated NIH3T3 cells as MNBPs for the G-rich strand of the Pc-1 minisatellite [48]. In addition, three proteins, MNBP-D, LRP130 (MNBP-E and -F) and Tudor-SN/SND1 (MNBP-G), were identified as C-rich strand binding proteins [48], whose functions and biological consequences will be briefly discussed later.

Binding of hnRNP A1 in cell-free extracts to the G-rich repeat was enhanced from cultures treated with 5 nM of OA. In contrast, hnRNP A3 did not show any such enhancement after OA treatment, suggesting some mechanistic differences in DNA recognition and binding between the two hnRNPs. Sequence requirements for the high-binding affinity of hnRNP A1 and hnRNP A3 were evaluated by an oligonucleotide competition assay using EMSA, and are summarized in Table 2. hnRNP A1 binds more widely to d(GGCAG)$_5$ and d(GGCAG)-like repeats, including G5$_{TEL}$, G5$_{+T}$ and poly(dG), with a $K_d$ value of nM magnitudes. Interestingly, the affinity of hnRNA3 with d(GGCAG)$_5$ was about 50-fold weaker than that of hnRNP A1, with IC$_{50}$ values for the formation of d(GGCAG)$_8$-protein complexes being 300 and 6 nM, respectively (see footnote for Table 2). This suggests that tandem arrays ($n > 5$ units) are required for sufficient binding of hnRNP A3 with Pc-1 d(GGCAG)$_n$ repeats. In contrast, competition with telomere repeats (G5$_{TEL}$) and telomere-like repeats (G5$_{+T}$) showed much stronger inhibitory activity for hnRNP A3 binding to d(GGCAG)$_8$, the value of which was almost comparable to that observed for hnRNP A1. G5$_{Pc-2}$, which contains a

Table 2
IC$_{50}$ values for formation of d(GGCAG)$_8$–protein complexes with hnRNP A1, hnRNP A3 and UP1 [48,52]

| Competitor | hnRNP A1 | hnRNP A3 | UP1 |
|---|---|---|---|
| G5$_{TEL}$ ((GTTAGG)$_5$) | 3 nM | 5 nM | 3 nM |
| G5$_{+T}$ ((GTCAGG)$_5$) | 1 nM | 7 nM | 4 nM |
| G5$_{Pc-2}$ ((GGCAGG)$_5$) | 9 nM | 20 nM | 40 nM |
| G5 ((GGCAG)$_5$) | 6 nM | 300 nM | 20 nM |
| Poly(dG) ((dG)$_{25}$) | 9 nM | 1 μM | 100 nM |
| G5$_{+C}$ ((GCCAGG)$_5$) | >1 μM | >1 μM | >1 μM |

IC$_{50}$ values were estimated from the inhibition curves generated by plotting the relative amount of complex against the concentration of a competitor, and expressed as the concentrations of individual unlabeled competitor oligonucleotides at which the labeled d(GGCAG)$_8$–protein complex yield was halved. The concentration of labeled d(GGCAG)$_8$ used in the competition reaction is 2 nM in all experiments.

G-insertion in the d(GGCAG) repeat of Pc-1, and is similar to minisatellite Pc-2, showed a slightly lower binding activity than the d(GGCAG) repeats. G5$_{+C}$ did not bind to either hnRNP A1 or hnRNP A3. A large difference was also observed between hnRNP A1 and hnRNP A3 with poly (dG) probes. Binding of hnRNP A3 to poly(dG) was 100-fold weaker than hnRNP A1.

Based on our data, minisatellite Pc-1 binding proteins are able to bind not only to Pc-1 but also to other minisatellites having Pc-1-like sequences. Sequence requirements for DNA binding of hnRNP A3 seem to be stricter than those for hnRNP A1, and this suggests that hnRNP A3 may serve a different function from hnRNP A1 by binding to separate regions of the genome. Both hnRNP A1 and hnRNP A3 may be telomere binding proteins, and elucidation of the biological functions of hnRNP A3 is currently ongoing in our laboratory.

## 7. hnRNP A1 and UP1 bind to G-rich repetitive sequences and unfold G4′ structures

To clarify the consequence of hnRNP A1 binding to G-rich repetitive sequences, we investigated its effect on the stability of the G4′ structure. Recombinant hnRNP A1 and unwinding protein 1 (UP1), a proteolytic product of hnRNP A1 lacking the C-terminal portion of hnRNP A1 [49–51], were expressed in *E. coli* as GST fusion proteins and purified, by releasing a GST tag, for use in G4′ binding assays. DNA binding affinity and sequence specificity of UP1 is almost equivalent to those of GST-UP1 and GST-hnRNP A1, but is slightly lower than that of hnRNP A1 (Table 2).

By EMSA, the intramolecular quadruplex (G4′ form) of d(GGCAG)$_8$ shows faster mobility than the single-stranded form on polyacrylamide gel electrophoresis (PAGE). Both UP1 and GST-hnRNP A1 bind to both

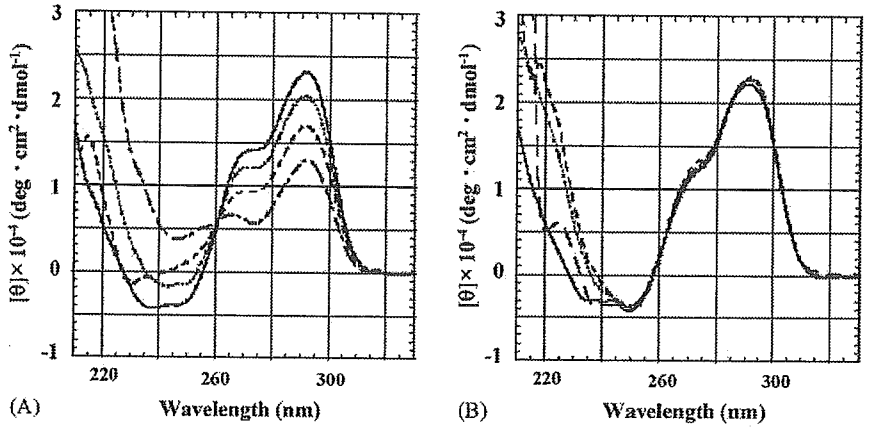


Fig. 3. CD spectra indicate UP1 unfolds the G4′ structure of d(GGCAG) repeats. CD spectra of $d(GGCAG)_5$ with GST-UP1 (A) and GST (B). For each CD spectrum, the spectrum of the corresponding protein was subtracted. Solid, the DNA:protein molar ratio of 1:0; dotted, 1:0.5; dashed, 1:1; centered, 1:2. The CD band at 290–295 nm was decreased by addition of GST-UP1 in a dose-dependent manner. GST alone did not induce any change in the levels of the peak.

G4′ and single-stranded forms of $d(GGCAG)_8$ [52], and decrease, in a dose dependent manner, the amount of the G4′ form (unpublished data). The characteristic CD band at 290–295 nm of the G4′ structure was decreased by addition of GST-UP1 in a concentration-dependent manner (Fig. 3A), while the addition of GST did not induce any change (Fig. 3B). These data clearly indicate that hnRNP A1 and UP1 unfold the G4′ structure at the d(GGCAG) repeats.

Another member of the hnRNP family proteins, hnRNP D, also binds and destabilizes the quadruplex structure formed by telomeric $d(TTAGGG)_n$ repeats, and is suggested to be involved in maintenance of the telomere 3′-overhang [53]. LR1, a heterodimer of nucleolin and hnRNP D, binds to duplex DNA sites in the immunoglobulin heavy chain switch (S) regions, which are G-rich DNA regions conforming to the consensus sequence, dGGNCNAG(G/C)CTG(G/A) [54]. Dempsey et al. demonstrated LR1 binding to G4 DNA formed by G-G pairing in S region sequences, and suggested that LR1 may juxtapose donor and acceptor switch regions for recombination during S region transcription [54]. Two hnRNP-related telomeric DNA-binding proteins, uqTBP25 and qTBP42, which show close sequence similarity to hnRNP A1, hnRNP A2/B1 and/or hnRNP C, also bind to quadruplex telomeric DNA and increase the heat stability and resistance of the telomeric repeats to micrococcal nuclease digestion [55,56]. Weisman-Shomer et al. recently demonstrated that both uqTBP25 and qTBP42 destabilize the quadruplex (tetraplex) structures of $d(CGG)_n$, but quadruplex structures of telomeric and IgG sequences are resistant to destabilization by these two proteins, suggesting the presence of some structural differences among the quadruplexes [57].

## 8. Abrogation of DNA synthesis arrest at the G4′ structure by UP1

When template carrying $d(GGCAG)_{12}$ was used for an in vitro DNA synthesis assay, a primer extension reaction with *BcaBEST* DNA polymerase was obstructed mainly at the first d(GGG) site (Fig. 4A). A primer extension reaction with a synthetic oligonucleotide containing a $d(CAGGG)_{15}$ repeat also demonstrated that progression of DNA polymerase was obstructed mainly at the first d(GGG) site followed by additional weaker stops at the second, third, fourth, fifth and sixth d(GGG) sites (Fig. 4B). When UP1 was added to the reaction with an excess molar amount over the template, DNA synthesis arrest at the d(GGG) sites was reduced in a dose-dependent manner and the synthesis of the complete length of template DNA was considerably enhanced (Fig. 4A and B). Other DNA polymerases, such as Taq, Klenow fragment and human DNA polymerase α (pol α) also gave similar results [52]. The data suggest that UP1/hnRNP A1 may abrogate DNA synthesis arrests at d(GGG) sites by destabilizing the G4′ structure. UP1 does not require NTP hydrolysis energy to unfold the quadruplex as, for example, BLM and WRN helicases do to unwind tetra- or bi-molecular quadruplex DNA [58,59].

## 9. Destruction of the telomere G4′ structure by binding of UP1

UP1 also binds to $G5_{+TEL}$ and TRM4 [$d(TTAGGG)_4$], both of which contain four telomeric repeats [52,60]. Under physiological-like conditions, $d(TTAGGG)_4$ also gave a positive CD band at

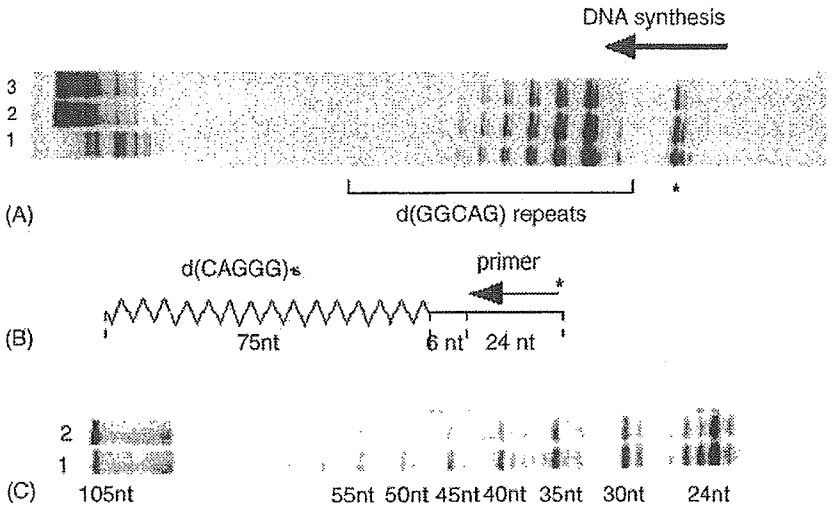


Fig. 4. Effect of UP1 on the arrest of DNA synthesis at d(GGCAG) repeats. (A) Denaturing PAGE analysis of primer extension reactions was carried out with or without UP1 using the single-stranded phagemid pYA-3 carrying $d(GGCAG)_{12}$ as described above (see legend of Fig. 2.) with some modifications [52]. *BcaBEST* DNA polymerase was used for the reaction. Single-stranded pYA-3 with $d(GGCAG)_{12}$ repeat (final 100 nM) in TE buffer containing 100 mM KCl was stored at 7 °C overnight to allow sufficient formation of the quadruplex structure, after heat denaturation at 95 °C for 3 min. Five μl of DNA samples was mixed with 7.5 μl of Milli-Q water and 5 μl of the pUC/M13 (−40) forward primer (Promega) labeled with $^{32}P$ at the 5′-end (0.5 μM) in TE buffer. The mixture was heated at 60 °C for 7 min and incubated at 37 °C for more than 30 min to allow sufficient quadruplex formation. An aliquot of 1.75 μl of this primer-annealed template (final 23 nM) was mixed with 0.75 μl of 10× *BcaBEST* buffer (200 mM Tris–HCl, pH 8.5/100 mM $MgCl_2$) and 0.5 μl of dNTPs mixture (50 μM each). After addition of 4 μl GST-UP1 suspended in a reaction buffer (20 mM sodium phophate, pH 7.0/0.5 mM DTT), the mixture was incubated at 37 °C for 5 min, and then the primer extension reaction was carried out at 37 °C for 8 min in the presence of *BcaBEST* DNA polymerase (final 66 units/ml). Final concentrations of UP1 added in the reaction are 0 M (lane 1), 6.0 μM (lane 2) and 48 μM (lane 3), respectively. Horizontal arrows indicate the direction of the DNA synthesis, and an asterisk (*) indicates the pausing of the primer extension at d(GGG) present in the multiple cloning sites of the vector plasmid. (B) and (C) A primer extension reaction using a synthetic oligonucleotide pSub15 as a template was performed in the presence and absence of GST-UP1 as described above with some modification. A schematic representation of a 105-mer template pSub15 with $d(CAGGG)_{15}$ repeats used for the primer extension reaction is depicted (B). A mixture of the pSub15 and $^{32}P$-labeled M13–20 (Takara) primer (final 100 nM each) was heated in TE buffer containing 150 mM KCl at 95 °C for 5 min and then at 72 °C for 5 min, followed by gradual cooling to room temperature, and stored at 7 °C overnight. The concentrations of *BcaBEST* DNA polymerase, the primer-annealed template, and dNTPs were 18 units/ml, 10 nM, and 1.7 μM, respectively. The concentrations of GST-UP1 were 0 M (lane 1 in (C)) and 750 nM (lane 2 in (C)). Fragment sizes are indicated in nucleotides (nt).

290–295 nm, similar to d(GGCAG) repeats, indicating the formation of a G4′ structure [52]. In contrast, four-stranded parallel quadruplex DNA was not observed except at a much higher DNA concentration (data not shown). In vitro DNA synthesis using synthetic oligonucleotides and several DNA polymerases, including human pol α, was also strongly inhibited at the d(GGG) sites of the telomeric repeats (our unpublished observation). The CD band at 290–295 nm, which is specific for the G4′ structure, was also decreased by addition of GST-UP1 in a dose-dependent manner as in the case of the d(GGCAG) repeats. GST alone did not induce any change in the levels of the band peak [52].

In hnRNP A1-deficient cells, telomere repeats become significantly shorter than in cells with normal levels of hnRNP A1 expression [61]. Restoration of hnRNP A1 expression dramatically increased the average size of the terminal repeat fragment length of the telomere. Interaction of telomerase with hnRNP A1 was also demonstrated by LaBranche et al. [61]. Physical interaction of hnRNP A1, telomeric DNA and human telomerase RNA was also demonstrated by Fiset and Chabot in vitro [62]. These results suggest that hnRNP A1 and its shortened derivative UP1 possibly play a key role in telomere maintenance by destroying the quadruplex structure of the telomere end in vivo. After destruction of the G4′ structure of telomere repeats, telomeric regions may become accessible for telomerase binding, enabling the proper maintenance of telomeres in cells [63].

## 10. UP1 unfolds the higher order DNA structures of d(CGG) triplet repeats

Another important functional aspect of UP1 is its role in the unfolding the higher order DNA structures of d(CGG) triplet repeats [64]. The $d(CGG)_n$ tract

forms hairpin, quadruplex, and homoduplex structures in vitro under physiological-like conditions [64–66]. In our study, the CD band analysis of $d(CGG)_{16}$ in either the presence or absence of 150 mM KCl showed a large negative peak at 255 nm and a weak positive peak at 280 nm, suggesting the formation of a non-B-type higher order DNA structure (Fig. 5A). The CD spectrum of oligonucleotide CGG16mut, having C to A substitutions at two sites in CGG16, represented a typical pattern of the B structure with comparable negative and positive peaks at 255 and 280 nm, respectively (Fig. 5B). Fragile X syndrome, the most common form of inherited mental retardation in humans, is caused by expansion of a d(CGG) triplet repeat in the 5′-untranslated region of the first exon of the FMR1 gene [67–69]. Although the molecular mechanisms underlying the unstable expansion of the repeats are yet to be elucidated, formation of higher order DNA structures by d(CGG) repeats as described above could be responsible for the induction of genomic instability and the expansion mutation at these repeats.

Several proteins have been identified and reported to unfold or destabilize the higher order structure of $d(CGG)_n$, including WRN helicase [59,70] and two telomeric DNA binding proteins, qTBP42 and uqTBP25 [57]. Here, we demonstrate that UP1 also

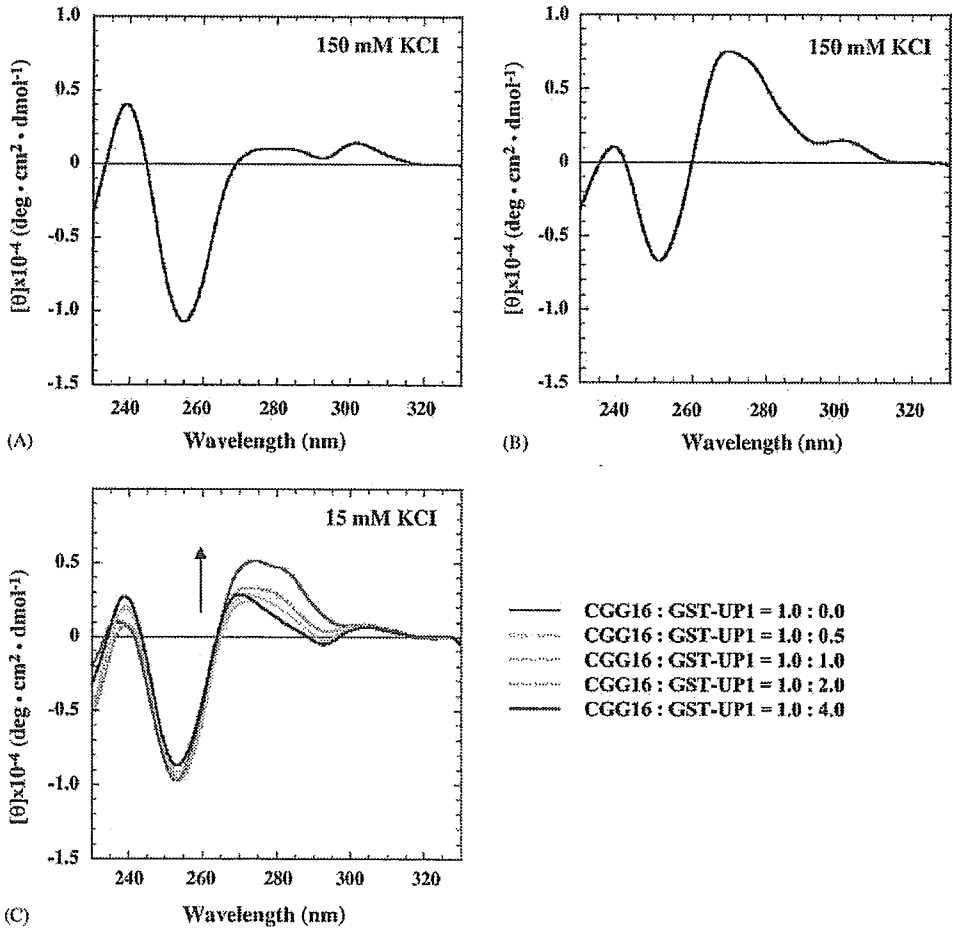


Fig. 5. CD spectra showing the effect of UP1 on higher order DNA structures of the d(CGG) repeat. (A) CD spectrum of $d(CGG)_{16}$ [CGG16] in the presence of 150 mM KCl. A large negative peak at 255 nm and a weak positive peak at 280 nm, which is suggestive of the formation of a non-B-type DNA structure, was observed. CGG16 also gave a similar CD pattern in the absence of KCl (data not shown). (B) CD spectrum of oligonucleotide CGG16mut in the presence of 150 mM KCl. CGG16mut, the sequence being $d(CGG)_5(AGG)(CGG)_4(AGG)(CGG)_5$, has C to A substitutions at two sites in CGG16. The spectrum represents a typical pattern of the B-structure with comparable negative and positive peaks at 255 and 280 nm, respectively. CGG16mut also showed the B-structure pattern in the absence of KCl. (C) CD spectra of CGG16 titrated with GST-UP1 in the presence of 15 mM KCl, the concentration at which the primer extension reaction was carried out with a synthetic 92-mer oligonucleotide pSubCGG16, containing a d(CGG) repeat, as a template. From each spectrum, the spectrum of the corresponding protein is subtracted, and CGG16:GST-UP1 molar ratios are indicated at the right bottom. Arrows indicate the changes of positive peaks of CD spectra in accordance with increased amounts of added GST-UP1 protein.

binds and unfolds the non-B structure of d(CGG) repeats (Fig. 5C). DNA synthesis is severely obstructed in vitro within the repetitive sequence, similar to the $d(GGCAG)_n$ repeats (data not shown). Addition of molar excess (compared to the template) of GST-UP1 reduces the arrest of DNA synthesis for several DNA polymerases in vitro, including human DNA polymerase α [64].

## 11. Expression of hnRNP A1 and hnRNP A3 in sporadic human colorectal cancers

Both hnRNP A1 and A3 are over-expressed in human colorectal cancers [71]. In the case of hnRNP A1, quantitative gene expression analysis revealed that 60% (18/30) of sporadic human colorectal cancers showed over-expression of hnRNP A1 in cancer tissues by at least two-fold compared to their normal counterparts. Interestingly, 78% of cases at clinicopathological stage II showed increased expression of two-fold or greater; this is two-fold higher than that seen in the more advanced stage IV [71]. This may imply that the biological impact of hnRNP A1 over-expression is in the relatively early stages of colon carcinogenesis. The over-expression of hnRNP A1 could contribute to the maintenance of telomere repeats in cancer cells and allow enhanced cell proliferation.

## 12. C-rich strands of minisatellite binding proteins

C-rich binding proteins, LRP130 and Tudor-SN/SND1, were also isolated from OA-treated NIH3T3 cells using oligonucleotide $d(CTGCC)_8$ as a probe, as described earlier [48]. It has been demonstrated that LRP130 and Tudor-SN/SND1 have some sequence specificity for DNA binding [72,73]. Interestingly, both LRP130 and Tudor-SN/SND1 bind to C-rich RNA sequences in a sequence specific manner, and are mainly localized at perinuclear regions and in the cytoplasm [74]. This may suggest the involvement of LRP130 and Tudor-SN/SND1 in RNA metabolism, mRNA transportation and/or specific mRNA transcription of sequences harboring a portion of C-rich residues. At present, we do not have a clear idea of the biological functions of these C-rich DNA/RNA binding proteins. It is possible that these proteins cooperate with hnRNP A1 and A3 (that bind to the G-rich DNA sequences) and exert some novel control of, for example, transcriptional regulation, mRNA metabolism or some other biological pathways.

## 13. Implication of G4 DNA structures in promoter regions in gene transcription

Although the biological roles of G-rich repeat sequences capable of adopting G4 DNA structures are still largely elusive, one of the intriguing fields of G4 DNA structure research is in its possible involvement in regulation of gene transcription. The G4′-quadruplex structure of the promoter region of the *c-myc* oncogene has been found to function as a transcriptional repressor [75] and the expression of *c-myc* can be inhibited by ligand-mediated G4-quadruplex stabilization [76]. In the case of the human insulin gene, unusual DNA structures of 14 bp G-rich tandem repeats are also thought to affect transcriptional activity, possibly through formation of G-quartets [34,77]. The presence of a guanine-quadruplex forming region within a polypurine tract of the hypoxia inducible factor 1α (HIF1α) promoter was also recently demonstrated; this unusual DNA structure may be involved in regulation of basal HIF-1α expression [78].

## 14. Future perspectives

G-rich repetitive sequences are frequently observed in the genome and are a subset of triplet repeats and of minisatellite DNA. Because of the peculiar and specific DNA structures adopted by G-rich repetitive sequences, these genomic regions might induce DNA replication fork arrest, leading to size alterations of the repeats in vivo. Various cellular components, such as WRN, TBPs, UP1 and hnRNP A1, may work as guardians against G-rich repeat length instability. In addition, C-rich sequence binding proteins, LRP130 and Tudor-SN/SND1, may cooperate with G-rich binding proteins, and facilitate the resolution of G4′ structures. A fascinating and intriguing scenario for the biological function of these MNBPs is their involvement in transcription of either the G- or C-rich repeat strands. Indeed, most MNBPs isolated from NIH3T3 cells in our own studies as well as by others, bind to mRNA transcripts (data not shown). Binding of hnRNP A1 and hnRNP A3 to G- or C-rich transcripts may affect the transcriptional efficiency and/or the fate of the transcripts of the G-rich region in the genome [40]. Further studies should be conducted to gain more insight into the biological impact of G-rich repetitive sequences, which are widely spread throughout the genome.

## Acknowledgements

The authors thank Drs. Minako Nagao and Takashi Sugimura for their helpful comments and continuing

support for the research. This work was supported by a Grant-in-Aid for Cancer Research from the Ministry of Health, Labour and Welfare of Japan, and Grants-in-Aid for Scientific Research (for M.K. and for H.F.) and the Protein 3000 Project (for M.K.) from the Ministry of Education, Culture, Sports, Science and Technology of Japan. K.H., E.T. and K.N. are the recipients of a Research Resident Fellowship from the Foundation for Promotion of Cancer Research in Japan.

## References

[1] J.C. Venter, M.D. Adams, E.W. Myers, P.W. Li, R.J. Mural, G.G. Sutton, H.O. Smith, M. Yandell, C.A. Evans, R.A. Holt, The sequence of the human genome, Science 291 (2001) 1304–1351.

[2] M. Dewannieux, T. Heidmann, LINEs, SINEs and processed pseudogenes: parasitic strategies for genome modeling, Cytogenet. Genome Res. 110 (2005) 35–48.

[3] H. te Riele, N. Claij, Microsatellite instability in human cancer: a prognostic marker for chemotherapy, Exp. Cell Res. 246 (1999) 1–10.

[4] C.J. Cummings, H.Y. Zoghbi, Fourteen and counting: unraveling trinucleotide repeat diseases, Hum. Mol. Genet. 9 (2000) 909–916.

[5] N.A. Di Prospero, K.H. Fishbeck, Therapeutics development for triplet repeat expansion diseases, Nat. Rev. Genet. 6 (2005) 756–765.

[6] A.J. Jeffreys, M.J. Allen, J.A.L. Armour, A. Collick, Y. Dubrova, N. Fretwell, T. Guram, M. Jobling, C.A. May, D.L. Neil, R. Neumann, Mutation processes at human minisatellites, Electrophoresis 16 (1995) 1577–1585.

[7] W.P. Wahls, L.J. Wallace, P.D. Moore, Hypervariable minisatellite DNA is a hotspot for homologous recombination in human cells, Cell 60 (1990) 95–103.

[8] A.J. Jeffreys, J.K. Holloway, L. Kauppi, C.A. May, R. Neumann, M.T. Slingsby, A.J. Webb, Meiotic recombination hot spots and human DNA diversity, Philos. Trans. R. Soc. London B: Biol. Sci. 359 (2004) 141–152.

[9] O. Handt, G.R. Sutherland, R.I. Richards, Fragile sites and minisatellite repeat instability, Mol. Genet. Metab. 70 (2000) 99–105.

[10] S. Yu, M. Mangelsdorf, D. Hewett, L. Hobson, E. Baker, H.J. Eyre, N. Lapsys, D. Le Paslier, N.A. Doggett, G.R. Sutherland, R.I. Richards, Human chromosomal fragile site FRA16B is an amplified AT-rich minisatellite repeat, Cell 88 (1997) 367–374.

[11] A. de la Chapelle, Genetic predisposition to colorectal cancer, Nat. Rev. Cancer 4 (2004) 769–780.

[12] P. Peltomaki, Deficient DNA mismatch repair: a common etiologic factor for colon cancer, Hum. Mol. Genet. 10 (2001) 735–740.

[13] T.G. Krontiris, Minisatellites and human disease, Science 269 (1995) 1682–1683.

[14] T.G. Krontiris, B. Devlin, D.D. Karp, N.J. Robert, N. Risch, An association between the risk of cancer and mutations in the HRAS1 minisatellite locus, N. Engl. J. Med. 329 (1993) 517–523.

[15] R. Kelly, G. Bulfield, A. Collick, M. Gibbs, A.J. Jeffreys, Characterization of a highly unstable mouse minisatellite locus: evidence for somatic mutation during early development, Genomics 5 (1989) 844–856.

[16] K. Mitani, Y. Takahashi, R. Kominami, A GGCAGG motif in minisatellites affecting their germline instability, J. Biol. Chem. 265 (1990) 15203–15210.

[17] S. Suzuki, K. Mitani, K. Kuwabara, Y. Takahashi, O. Niwa, R. Kominami, Two mouse hypervariable minisatellites: chromosomal location and simultaneous mutation, J. Biochem. (Tokyo) 114 (1993) 292–296.

[18] M.N. Weitzmann, K.J. Woodford, K. Usdin, The mouse Ms6-hm hypervariable microsatellite forms a hairpin and two unusual tetraplexes, J. Biol. Chem. 273 (1998) 30742–30749.

[19] M. Yamauchi, M. Nishimura, S. Tsuji, M. Terada, M. Sasanuma, Y. Shimada, Effect of SCID mutation on the occurrence of mouse Pc-1 (Ms6-hm) germline mutations, Mutat. Res. 503 (2002) 43–49.

[20] A.J. Jeffreys, V. Wilson, S.L. Thein, Hypervariable 'minisatellite' regions in human DNA, Nature 314 (1985) 67–73.

[21] H. Nakagama, S. Kaneko, H. Shima, H. Inamori, H. Fukuda, R. Kominami, T. Sugimura, M. Nagao, Induction of minisatellite mutation in NIH 3T3 cells by treatment with the tumor promoter okadaic acid, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 10813–10816.

[22] H. Inamori, S. Takagi, R. Tajima, M. Ochiai, T. Ubagai, T. Sugimura, M. Nagao, H. Nakagama, Frequent and multiple mutations at minisatellite loci in sporadic human colorectal and gastric cancers—possible mechanistic differences from microsatellite instability in cancer cells, Jpn. J. Cancer Res. 93 (2002) 382–388.

[23] H. Imai, H. Nakagama, K. Komatsu, T. Shiraishi, H. Fukuda, T. Sugimura, M. Nagao, Minisatellite instability in severe combined immunodeficiency mouse cells, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 10817–10820.

[24] Y.E. Dubrova, Long-term genetic effects of radiation exposure, Mutat. Res. 544 (2003) 433–439.

[25] A.J. Jeffreys, K. Tamaki, A. MacLeod, D.G. Monckton, D.L. Neil, J.A. Armour, Complex gene conversion events in germline mutation at human minisatellites, Nat. Genet. 6 (1994) 136–145.

[26] C.L. Yauk, Advances in the application of germline tandem repeat instability for in situ monitoring, Mutat. Res. 566 (2004) 169–182.

[27] T. Kitazawa, R. Kominami, R. Tanaka, K. Wakabayashi, M. Nagao, 2-Hydroxyamino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine induction of recombinational mutations in mammalian cell lines as detected by DNA fingerprinting, Mol. Carcinogr. 9 (1994) 67–70.

[28] B.J. Ledwith, D.J. Joslyn, P. Troilo, K.R. Leander, J.H. Clair, K.A. Soper, S. Manam, S. Prahalada, M.J. van Zwieten, W.W. Nichols, Induction of minisatellite DNA rearrangements by genotoxic carcinogens in mouse liver tumors, Carcinogenesis 16 (1995) 1167–1172.

[29] B.J. Ledwith, R.D. Storer, S. Prahalada, S. Manam, K.R. Leander, M.J. van Zwieten, W.W. Nichols, M.O. Bradley, DNA fingerprinting of 7,12-dimethylbenz[*a*]anthracene-induced and spontaneous CD-1 mouse liver tumors, Cancer Res. 50 (1990) 5245–5249.

[30] Y. Matsumura, D. Tarin, DNA fingerprinting survey of various human tumors and their metastases, Cancer. Res. 52 (1992) 2174–2179.

[31] S.L. Thein, A.J. Jeffreys, H.C. Gooi, F. Cotter, J. Flint, N.T. O'Connor, D.J. Weatherall, J.S. Wainscoat, Detection of somatic changes in human cancer DNA by DNA fingerprint analysis, Br. J. Cancer 55 (1987) 353–356.

[32] Y. Mishima, T. Suda, R. Kominami, Formation of a triple-stranded DNA between d(GGA:TCC) repeats and d(GGA) repeat oligonucleotides, J. Biochem. (Tokyo) 119 (1996) 805–810.

[33] C.E. Pearson, Y.H. Wang, J.D. Griffith, R.R. Sinden, Structural analysis of slipped-strand DNA (S-DNA) formed in $(CTG)_n$. $(CAG)_n$ repeats from the myotonic dystrophy locus, Nucl. Acids Res. 26 (1998) 816–823.

[34] P. Catasti, X. Chen, R.K. Moyzis, E.M. Bradbury, G. Gupta, Structure–function correlations of the insulin-linked polymorphic region, J. Mol. Biol. 264 (1996) 534–545.

[35] D. Sen, W. Gilbert, Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis, Nature 334 (1988) 364–366.

[36] K.J. Woodford, R.M. Howell, K. Usdin, A novel K(+)-dependent DNA synthesis arrest site in a commonly occurring sequence motif in eukaryotes, J. Biol. Chem. 269 (1994) 27029–27035.

[37] M. Katahira, H. Fukuda, H. Kawasumi, T. Sugimura, H. Nakagama, M. Nagao, Intramolecular quadruplex formation of the G-rich strand of the mouse hypervariable minisatellite Pc-1, Biochem. Biophys. Res. Commun. 264 (1999) 327–333.

[38] P. Balagurumoorthy, S.K. Brahmachari, D. Mohanty, M. Bansal, V. Sasisekharan, Hairpin and parallel quartet structures for telomeric sequences, Nucl. Acids Res. 20 (1992) 4061–4067.

[39] M. Katahira, M. Kanagawa, H. Sato, S. Uesugi, S. Fujii, T. Kohno, T. Maeda, Formation of sheared G:A base pairs in an RNA duplex modelled after ribozymes, as revealed by NMR, Nucl. Acids Res. 22 (1994) 2752–2759.

[40] M.L. Duquette, P. Handa, J.A. Vincent, A.F. Taylor, N. Maizels, Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA, Genes Dev. 18 (2004) 1618–1629.

[41] K. Paeschke, T. Simonsson, J. Postberg, D. Rhodes, H.J. Lipps, Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo, Nat. Struct. Mol. Biol. 12 (2005) 847–854.

[42] H. Sun, A. Yabuki, N. Maizels, A human nuclease specific for G4 DNA, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 12444–12449.

[43] C. Schaffitzel, I. Berger, J. Postberg, J. Hanes, H.J. Lipps, A. Pluckthun, In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with Stylonychia lemnae macronuclei, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 8572–8577.

[44] H. Han, L.H. Hurley, M. Salazar, A DNA polymerase stop assay for G-quadruplex-interactive compounds, Nucl. Acids Res. 27 (1999) 537–542.

[45] R.M. Howell, K.J. Woodford, M.N. Weitzmann, K. Usdin, The chicken beta-globin gene promoter forms a novel "cinched" tetrahelical structure, J. Biol. Chem. 271 (1996) 5208–5214.

[46] M. Suganuma, H. Fujiki, H. Suguri, S. Yoshizawa, M. Hirota, M. Nakayasu, M. Ojika, K. Wakamatsu, K. Yamada, T. Sugimura, Okadaic acid: an additional non-phorbol-12-tetradecanoate-13-acetate-type tumor promoter, Proc. Natl. Acad. Sci. U.S.A. 85 (1988) 1768–1771.

[47] A. Takai, C. Bialojan, M. Troschka, J.C. Ruegg, Smooth muscle myosin phosphatase inhibition and force enhancement by black sponge toxin, FEBS Lett. 217 (1987) 81–84.

[48] H. Fukuda, T. Sugimura, M. Nagao, H. Nakagama, Detection and isolation of minisatellite Pc-1 binding proteins, Biochim. Biophys. Acta 1528 (2001) 152–158.

[49] G. Herrick, B. Alberts, Purification and physical characterization of nucleic acid helix-unwinding proteins from calf thymus, J. Biol. Chem. 251 (1976) 2124–2132.

[50] B.M. Merrill, M.B. LoPresti, K.L. Stone, K.R. Williams, High pressure liquid chromatography purification of UP1 and UP2, two related single-stranded nucleic acid-binding proteins from calf thymus, J. Biol. Chem. 261 (1986) 878–883.

[51] O. Valentini, G. Biamonti, M. Pandolfo, C. Morandi, S. Riva, Mammalian single-stranded DNA binding proteins and heterogeneous nuclear RNA proteins have common antigenic determinants, Nucl. Acids Res. 13 (1985) 337–346.

[52] H. Fukuda, M. Katahira, N. Tsuchiya, Y. Enokizono, T. Sugimura, M. Nagao, H. Nakagama, Unfolding of quadruplex structure in the G-rich strand of the minisatellite repeat by the binding protein UP1, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 12685–12690.

[53] Y. Enokizono, Y. Konishi, K. Nagata, K. Ouhashi, S. Uesugi, F. Ishikawa, M. Katahira, Structure of hnRNP D complexed with single-stranded telomere DNA and unfolding of the quadruplex by heterogeneous nuclear ribonucleoprotein D, J. Biol. Chem. 280 (2005) 18862–18870.

[54] L.A. Dempsey, H. Sun, L.A. Hanakahi, N. Maizels, G4 DNA binding by LR1 and its subunits, nucleolin and hnRNP D, A role for G–G pairing in immunoglobulin switch recombination, J. Biol. Chem. 274 (1999) 1066–1071.

[55] R. Erlitzki, M. Fry, Sequence-specific binding protein of single-stranded and unimolecular quadruplex telomeric DNA from rat hepatocytes, J. Biol. Chem. 272 (1997) 15881–15890.

[56] G. Sarig, P. Weisman-Shomer, R. Erlitzki, M. Fry, Purification and characterization of qTBP42, a new single-stranded and quadruplex telomeric DNA-binding protein from rat hepatocytes, J. Biol. Chem. 272 (1997) 4474–4482.

[57] P. Weisman-Shomer, Y. Naot, M. Fry, Tetrahelical forms of the fragile X syndrome expanded sequence d(CGG)(*n*) are destabilized by two heterogeneous nuclear ribonucleoprotein-related telomeric DNA-binding proteins, J. Biol. Chem. 275 (2000) 2231–2238.

[58] H. Sun, J.K. Karow, I.D. Hickson, N. Maizels, The Bloom's syndrome helicase unwinds G4 DNA, J. Biol. Chem. 273 (1998) 27587–27592.

[59] M. Fry, L.A. Loeb, Human werner syndrome DNA helicase unwinds tetrahelical structures of the fragile X syndrome repeat sequence $d(CGG)_n$, J. Biol. Chem. 274 (1999) 12797–12802.

[60] Y. Enokizono, A. Matsugami, S. Uesugi, H. Fukuda, N. Tsuchiya, T. Sugimura, M. Nagao, H. Nakagama, M. Katahira, Destruction of quadruplex by proteins, and its biological implications in replication and telomere maintenance, Nucl. Acids Res. Suppl. (2003) 231–232.

[61] H. LaBranche, S. Dupuis, Y. Ben-David, M.R. Bani, R.J. Wellinger, B. Chabot, Telomere elongation by hnRNP A1 and a derivative that interacts with telomeric repeats and telomerase, Nat. Genet. 19 (1998) 199–202.

[62] S. Fiset, B. Chabot, hnRNP A1 may interact simultaneously with telomeric DNA and the human telomerase RNA in vitro, Nucl. Acids Res. 29 (2001) 2268–2275.

[63] L.P. Ford, W.E. Wright, J.W. Shay, A model for heterogeneous nuclear ribonucleoproteins in telomere and telomerase regulation, Oncogene 21 (2002) 580–583.

[64] H. Fukuda, M. Katahira, E. Tanaka, Y. Enokizono, N. Tsuchiya, K. Higuchi, M. Nagao, H. Nakagama, Unfolding of higher DNA structures formed by the d(CGG) triplet repeat by UP1 protein, Genes Cells 10 (2005) 953–962.

[65] M. Fry, L.A. Loeb, The fragile X syndrome $d(CGG)_n$ nucleotide repeats form a stable tetrahelical structure, Proc. Natl. Acad. Sci. U.S.A. 91 (1994) 4950–4954.

[66] P. Fojtik, I. Kejnovska, M. Vorlickova, The guanine-rich fragile X chromosome repeats are reluctant to form tetraplexes, Nucl. Acids Res. 32 (2004) 298–306.

[67] D.C. Crawford, J.M. Acuna, S.L. Sherman, FMR1 and the fragile X syndrome: human genome epidemiology review, Genet. Med. 3 (2001) 359–371.

[68] J.L. Mandel, Questions of expansion, Nat. Genet. 4 (1993) 8–9.

[69] F. Rousseau, D. Heitz, J.L. Mandel, The unstable and methylatable mutations causing the fragile X syndrome, Hum. Mutat. 1 (1992) 91–96.

[70] A.S. Kamath-Loeb, L.A. Loeb, E. Johansson, P.M. Burgers, M. Fry, Interactions between the Werner syndrome helicase and DNA polymerase delta specifically facilitate copying of tetraplex and hairpin structures of the d(CGG)$_n$ trinucleotide repeat sequence, J. Biol. Chem. 276 (2001) 16439–16446.

[71] M. Ushigome, T. Ubagai, H. Fukuda, N. Tsuchiya, T. Sugimura, J. Takatsuka, H. Nakagama, Up-regulation of hnRNP A1 gene in sporadic human colorectal cancers, Int. J. Oncol. 26 (2005) 635–640.

[72] H. Fukuda, N. Tsuchiya, M. Sato, A. Yamaguchi, N. Tanaka, M. Nagao, H. Nakagama, DNA-binding activity of p100, a transcriptional coactivator, to single-stranded C-rich sequences, Proc. Jpn. Acad. 79 (Ser(B)) (2003) 120–123.

[73] N. Tsuchiya, H. Fukuda, T. Sugimura, M. Nagao, H. Nakagama, LRP130, a protein containing nine pentatricopeptide repeat motifs, interacts with a single-stranded cytosine-rich sequence of mouse hypervariable minisatellite Pc-1, Eur. J. Biochem. 269 (2002) 2927–2933.

[74] N. Tsuchiya, H. Fukuda, K. Nakashima, M. Nagao, T. Sugimura, H. Nakagama, LRP130, a single-stranded DNA/RNA-binding protein, localizes at the outer nuclear and endoplasmic reticulum membrane, and interacts with mRNA in vivo, Biochem. Biophys. Res. Commun. 317 (2004) 736–743.

[75] T. Simonsson, P. Pecinka, M. Kubista, DNA tetraplex formation in the control region of c-myc, Nucl. Acids Res. 26 (1998) 1167–1172.

[76] T. Lemarteleur, D. Gomez, R. Paterski, E. Mandine, P. Mailliet, J.F. Riou, Stabilization of the c-myc gene promoter quadruplex by specific ligands' inhibitors of telomerase, Biochem. Biophys. Res. Commun. 323 (2004) 802–808.

[77] A. Lew, W.J. Rutter, G.C. Kennedy, Unusual DNA structure of the diabetes susceptibility locus IDDM2 and its effect on transcription by the insulin promoter factor Pur-1/MAZ, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 12508–12512.

[78] R. De Armond, S. Wood, D. Sun, L.H. Hurley, S.W. Ebbinghaus, Evidence for the presence of a guanine quadruplex forming region within a polypurine tract of the hypoxia inducible factor 1alpha promoter, Biochemistry 44 (2005) 16341–16350.

[79] V.M. Marathias, P.H. Bolton, Determinants of DNA quadruplex structural type: sequence and potassium binding, Biochemistry 38 (1999) 4355–4364.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Mutation Research 603 (2006) 186–192
Genetic Toxicology and Environmental Mutagenesis
www.elsevier.com/locate/gentox
Community address: www.elsevier.com/locate/mutres

# Detection of 4-oxo-2-hexenal, a novel mutagenic product of lipid peroxidation, in human diet and cooking vapor

Kazuaki Kawai [a], Koji Matsuno [b], Hiroshi Kasai [a,*]

[a] *Department of Environmental Oncology, Institute of Industrial Ecological Sciences, University of Occupational and Environmental Health, 1-1 Iseigaoka, Yahatanishi-ku, Kitakyushu 807-8555, Japan*
[b] *Bio-information Research Center, University of Occupational and Environmental Health, 1-1 Iseigaoka, Yahatanishi-ku, Kitakyushu 807-8555, Japan*

Received 22 September 2005; received in revised form 28 November 2005; accepted 29 November 2005
Available online 19 January 2006

**Abstract**

Since the diet plays an important role in the development of human cancer, it is important to identify mutagens in foods. We have detected a novel mutagenic product, 4-oxo-2-hexenal (4-OHE), in a model lipid peroxidation reaction mixture [H. Kasai, M. Maekawa, K. Kawai, K. Hachisuka, Y. Takahashi, H. Nakamura, R. Sawa, S. Matsui, T. Matsuda, 4-Oxo-2-hexenal, a mutgen formed by ω-3 fat peroxidation, causes DNA adduct formation in mouse organs, Ind. Health 43 (2005) 699–701]. In the present study, the contents of 4-OHE in various food samples were determined by a GC/MS method. Commercial perilla oil (derived from the seed of *Perilla frutescens* var. *frutescens*), which is rich in linolenic acid triglyceride (TG), the edible part of broiled fish, and various fried foods contained 4-OHE in the range of 1–70 μg/g. Furthermore, from the ethyl acetate trap (extracts) of the smoke released during the broiling of fish, 4-OHE was also detected by GC/MS. These results provide a warning to humans, who may be exposed to this mutagen. The 4-OHE may be produced from ω-3 polyunsaturated fats, such as α-linolenic acid-, docosahexaenoic acid (DHA)- and eicosapentaenoic acid (EPA)-TG, which are more easily oxidized than ω-6 fats, such as linoleic acid-TG.


*Keywords:* 4-oxo-2-hexenal; Lipid peroxidation; DNA adduct; Human diet; Cooking vapor

## 1. Introduction

It is well known that the human diet is an important factor in the causes of cancers. Especially, epidemiological studies have suggested that a high-fat diet and the consumption of red meat are risk factors of various cancers, such as breast, colon and prostate cancer. Lipid peroxyl radicals from oxidized oils and heme-iron were implicated as a mechanism of high-fat diet-induced colon carcinogenesis [1]. Lipid peroxidation products, such as 4-hydroxy-2-nonenal, 4-hydroxy-2-hexenal and 4-oxo-2-nonenal, reportedly modify DNA bases covalently [2–4]. However, there have been few studies on their detection in foods [5]. We have recently detected 4-oxo-2-hexenal (4-OHE)-deoxyguanosine (dG)-adduct, a novel nucleoside derivative, when dG was reacted with products in model lipid peroxidation reactions [6,7]. Although several studies have been carried out with lipid peroxidation-derived aldehydes, none of the reports have

*Abbreviations:* 4-OHE, 4-oxo-2-hexenal; 4-ONE, 4-oxo-2-nonenal; SIM, selected ion monitoring; SIM-TIC, SIM mode total ion chromatogram

* Corresponding author. Tel.: +81 93 691 7469/7468; fax: +81 93 601 2199.
*E-mail address:* h-kasai@med.uoeh-u.ac.jp (H. Kasai).

1383-5718/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.mrgentox.2005.11.012

described the formation of 4-OHE. 4-OHE is probably generated by the oxidation of ω-3 fatty acids triglyceride (TG), which are commonly found in dietary fats, such as fish oil, perilla oil (derived from the seed of the plant *Perilla frutescens* var. *frutescens*, about 60% of perilla oil is α-linolenic acid TG), rapeseed oil, and soybean oil. Therefore, 4-OHE might be a major cytotoxic lipid peroxide in foods. We have also found that 4-OHE is mutagenic in the Ames test, and detected 4-OHE-DNA adduct formation in vivo [6,7]. Dietary exposure to 4-OHE may play an important role in human cancer development. Furthermore, epidemiological studies in China indicated that heated cooking oil vapors may be related to lung cancer risk [8,9]. Airborne 4-OHE produced by cooking may also be a risk factor for lung cancer in cooks [10,11]. In this study we determined the amounts of 4-OHE in food products and smoke condensates.

## 2. Materials and methods

### 2.1. Chemicals

4-Oxo-2-hexenal (4-OHE) diethylacetal was prepared by $MnO_2$ oxidation of 4-hydroxy-2-hexenal diethylacetal, which was synthesized according to the method of Esterbauer and Weger [12]. 4-OHE was obtained by acid treatment of 4-OHE diethylacetal. The structure of the synthetic 4-OHE thus obtained was confirmed by mass- and $^1H$ NMR-spectra. Detailed experimental conditions and spectral data will be published elsewhere [7]. Authentic 4-oxo-2-nonenal (4-ONE) was prepared by $MnO_2$ oxidation of 4-hydroxy-2-nonenal (a product of Calbiochem-Novabiochem Corp.; purchased from Wako Pure Chemical Ind. Ltd., Japan).

### 2.2. Ethyl acetate extracts of food

The food sample (1.0 g) was submerged in 2 ml of ethyl acetate in a screw-capped glass vial and extracted for 16 h at 4 °C. The resulting supernatant (40 μl) was diluted with 160 μl of ethyl acetate. The cooking oil was diluted 100-fold with ethyl acetate. These diluted solutions were used for the GC–MS analysis.

### 2.3. Smoke condensate

Fresh saury was obtained from a supermarket in Kitakyushu. One saury (about 180 g) was broiled until well-colored on a gas burner (about 8 min), using a grid. The smoke was collected by aspiration, using a funnel connected to a gas washing bottle filled with 100 ml of ethyl acetate. The smoke components were thus trapped by bubbling into ethyl acetate. The ethyl acetate extract was concentrated to a small volume under reduced pressure before the analysis by GC/MS.

### 2.4. Analytical methods

The standard 4-OHE solution in ethyl acetate (1 mg/ml) was diluted to various concentrations with ethyl acetate, and 2 μl of each was injected into the GC–MS apparatus. The calibration curve was obtained by the least square method. An authentic sample of 4-ONE was also analyzed by GC–MS, although the exact concentration was not determined due to the scarcity of the synthetic product. From each mass spectrum of 4-OHE and 4-ONE obtained by scanned GC–MS, the molecular ions ($M^+$, 112 and 154, respectively) and the major fragment ions ($M$–CHO, 83 and 139, respectively) were chosen as the identification ions for the GC–MS analysis in the selected ion monitoring (SIM) method. For SIM mode total ion chromatography (SIM-TIC), the identification ions at *m/z* 112, 154, 83 and 139 were used. 4-OHE and 4-ONE were identified by their retention times and by detecting the identification ions ($M^+$, $M$–CHO). The concentrations of 4-OHE were determined by use of a calibration curve, which was produced on the basis of the ion counts of the molecular ions ($M^+$). The relative amount of 4-ONE was estimated from the ratio of the molecular ion count of 4-ONE to that of 4-OHE obtained by the SIM–GC–MS analysis, based on the assumption that homologous compounds have similar dose-ion count responses in the calibration curve [13]. Knowing the 4-OHE concentration, we thus estimated the 4-ONE concentration.

The diluted ethyl acetate extracts of food samples were transferred into a 200 μl glass vial and sealed with gas-tight caps. For the analysis of 4-OHE and 4-ONE, 2 μl aliquots were injected into GC–MS.

### 2.5. GC–MS conditions

The GC–MS analysis was carried out on a Hewlett-Packard gas chromatograph (HP 6890) connected to a mass spectrometer (JEOL JMS-BU 20, Tokyo). A 30 m × 0.25 mm i.d. CP-CIL5CB capillary column with a 250 μm film thickness (Chromopack) was used for the sample separation. The injection temperature was 280 °C. The column oven temperature was programmed at an initial temperature of 60 °C for 1 min, then to rise from 60 to 150 °C at a rate of 10 °C/min., and finally to increase from 150 to 270 °C at a rate of 40 °C/min. The flow rate of the carrier gas (He) was 1.5 ml/min (splitless). The MS analysis was carried out using electron impact ionization at 70 eV. The temperatures of the interface and ion chamber were 280 and 270 °C, respectively.

## 3. Results

The mass spectrum of 4-OHE, measured by our GC–MS, is shown in Fig. 1. The molecular ion ($M^+$) at *m/z* 112 and the major fragment ion at *m/z* 83 ($M$–CHO) are visible. Consequently, 4-OHE was determined by a SIM method at *m/z* 112 (quantification ion) and confirmed by SIM at *m/z* 83 (data not shown). The detection

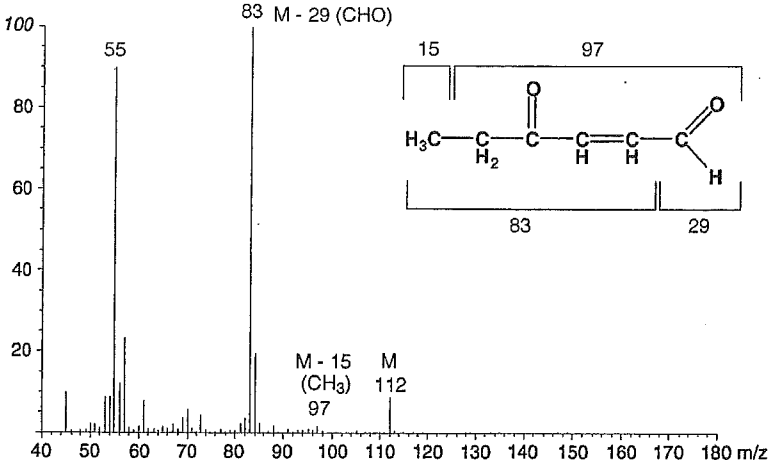


Fig. 1. Mass spectrum and specific mass fragments of 4-OHE.

limit was 0.01 μg/ml (S/N > 3), and the lower limit for the quantitative analysis was 0.05 μg/ml (S/N > 5).

The mass spectrum of 4-ONE (Fig. 2) shows the molecular ion at *m/z* 154 ($M^+$) and the major fragment ions at *m/z* 125 (*M*–CHO) and 139 (*M*–$CH_3$). For the semi-quantitative determination of 4-ONE, the SIM method at the molecular ion *m/z* 154 was used and confirmed by *m/z* 139, an associate ion for identification (data not shown).

Fig. 3A shows the results of the GC–MS analysis for the 4-OHE and 4-ONE present in broiled saury, as a typical example. It shows a SIM mode total ion chromatogram (SIM-TIC) and the mass chromatograph (SIM) at *m/z* values of 112 (the molecular ion) for 4-OHE and 154 (the molecular ion) for 4-ONE. Peaks 1 and 2 in Fig. 3 correspond to 4-OHE and 4-ONE, respectively. This was confirmed by the GC–MS chromatograms of authentic 4-OHE (Fig. 3B, B-1, B-2) and 4-ONE (Fig. 3B, B-3, B-4).

Table 1 summarizes the data for the 4-OHE contents in various foods. Higher levels of 4-OHE were found in broiled sardine and mackerel, than in the same fishes cooked by other methods. Furthermore, relatively high levels of 4-OHE were also observed in fried vegetables. On the other hand, most of the cooked meat samples did not contain detectable levels of 4-OHE, with the one exception of the beef fried in perilla oil. Concerning 4-ONE, since a highly pure synthetic standard is not available, we could not fully quantitate the amount of 4-ONE in food. However, by the ratio of the molecular ion counts of 4-OHE and 4-ONE, their approximate relative amounts were estimated. As a result, lower levels

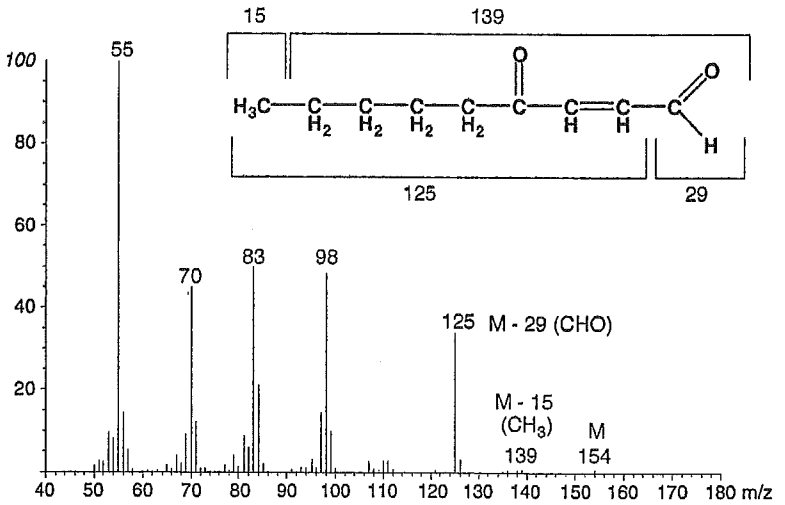


Fig. 2. Mass spectrum and specific mass fragments of 4-ONE.

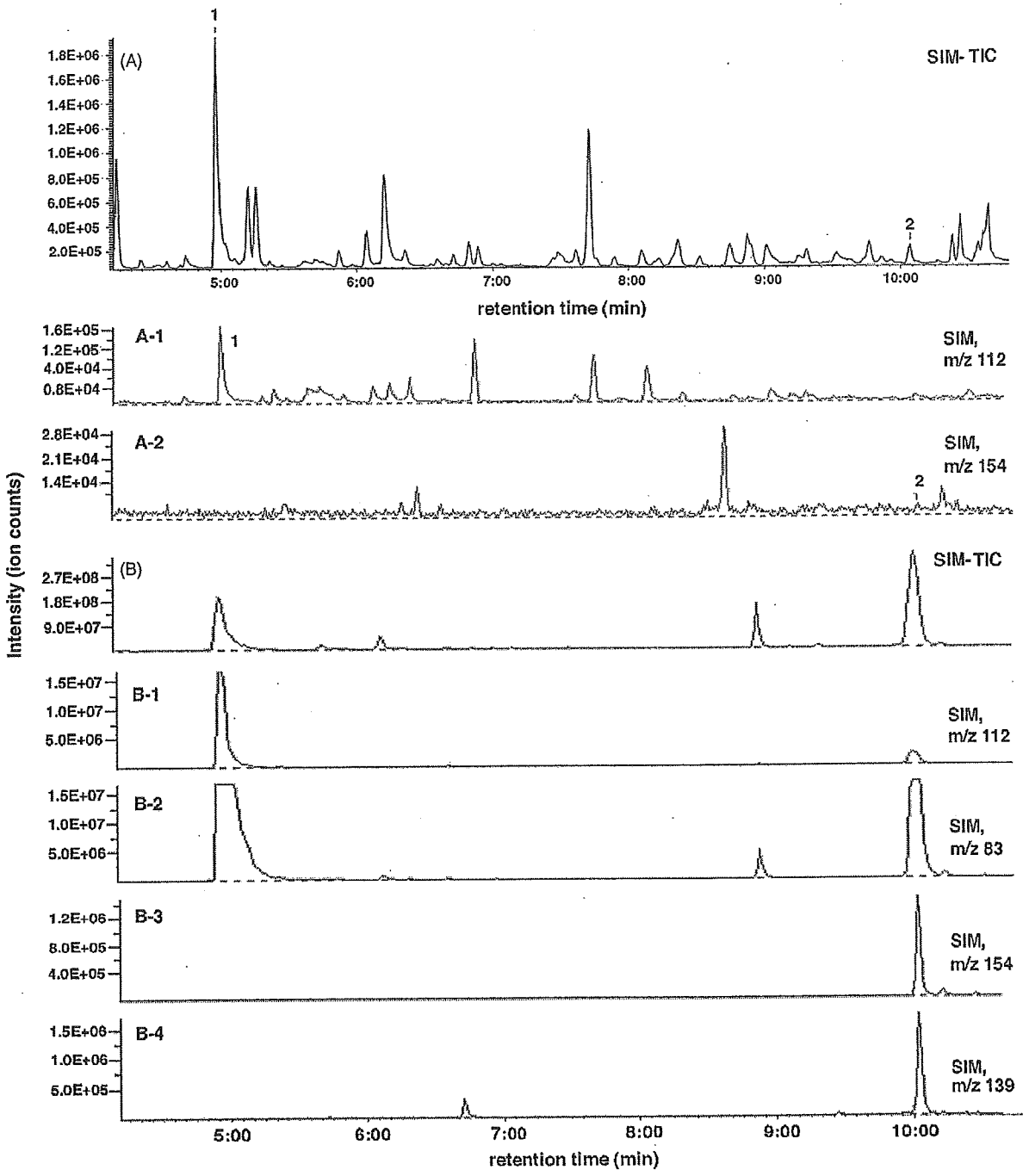


Fig. 3. (A) SIM-TIC and the mass chromatogram (SIM) of *m/z* 112 (A-1) and *m/z* 154 (A-2) of the ethyl acetate extract from broiled saury. Quantifications of peak 1 for 4-ONE ($t_R$ = 4:58, *m/z* 112), and peak 2 for 4-OHE ($t_R$ = 10:04, *m/z* 154). (B) SIM-TIC and mass chromatogram (SIM) of the mixture of authentic 4-OHE and 4-ONE. Identification of the peaks for 4-OHE ($t_R$ = 4:58, *m/z* 112: quantification ion and 83: identification ion) (B-1, B-2), and 4-ONE ($t_R$ = 10:04, *m/z* 154: quantification ion and 139: identification ion) (B-3, B-4).

of 4-ONE were detected, as compared to those of 4-OHE (Table 1).

The amount of 4-OHE in the smoke condensate obtained by broiling one saury was 31 μg. This value may be underestimated, because a portion of the 4-OHE is lost during the evaporation of ethyl acetate under reduced pressure. In contrast, 4-ONE was not detected in the smoke condensate.

Table 1
Amounts of 4-OHE and 4-ONE in cooked foods

| Food | Cooking method | 4-OHE (μg/g) | 4-ONE[a] |
|---|---|---|---|
| Cooking oil | | | |
| Perilla oil A | | 43.5 | N.D.[b] |
| Perilla oil B | | N.D. | N.D. |
| Fish | | | |
| Sardine (dried) | Broiled | 14.97 | N.D. |
| Sardine | Fritter (salad oil) | 7.09 | N.D. |
| | Boiled | 8.54 | N.D. |
| Salmon | Sauteed (salad oil, butter) | 3.67 | 1/5 |
| Saury | Broiled | 1.09–70.18 ($n$ = 8, mean 13.86) | N.D. |
| Mackerel | Broiled | 12.12 | 1/30 |
| | | 4.29 | N.D. |
| | Fried (salad oil) | 2.56 | N.D. |
| | Boiled | 5.49 | N.D. |
| Meat | | | |
| Chicken | Char-grilled | N.D. | N.D. |
| Pork | Fried (salad oil) | N.D. | N.D. |
| | Char-grilled | N.D. | N.D. |
| Beef | Fried (beef tallow oil) | N.D. | N.D. |
| | Fried (canola oil) | N.D. | N.D. |
| | Fried (salad oil) | N.D. | N.D. |
| | Fried (perilla oil A) | 45.88 | 1/60 |
| | Fried (perilla oil B) | N.D. | N.D. |
| | Fried (seasoned) | N.D. | N.D. |
| Vegetable | | | |
| Spinach | Fried (perilla oil A) | 5.28 | 1/5 |
| | Fried (perilla oil B) | 2.01 | 1/2 |
| | Fried (salad oil) | 4.40 | 1/2 |
| | Fried (canola oil) | N.D. | N.D. |
| Squash | Fried (salad oil) | 2.08 | N.D. |
| Other | | | |
| Snack (broad bean) | | N.D. | N.D. |
| Snack(corn) | | N.D. | N.D. |

[a] Approximate relative amount of 4-ONE to 4-OHE (molar ratio).
[b] N.D.: not detected.

## 4. Discussion

Our results represent the first description of the 4-OHE contents in food. Recently, 4-OHE was identified as a dG adduct in a dietary lipid peroxidation model (methyl linolenate + hemin) with dG [6,7]. We have also found that 4-OHE is mutagenic in the *Salmonella typhimurium* assay [6,7]. Furthermore, 4-OHE-DNA adducts were detected in the mouse gastrointestinal tract after oral administration of 4-OHE [6]. The 4-OHE may play an important role in human cancer development. From the aforementioned observations, it is important to determine the 4-OHE contents in food. In this study, we found that the amount of 4-OHE in most of the cooked fish was in the range of 2.6–70.2 μg/g food, though the amount was lower in cooked meat (barely detected). As an exception, we found an abundance of 4-OHE in fried beef, with frying performed in perilla oil (A). Actually, the perilla oil (A) itself contained a high amount of 4-OHE. The amounts of 4-OHE in food seem to correlate with the contents of ω-3 fats. Since 4-OHE is produced by the oxidation of α-linolenic acid TG, but not from linoleic acid TG, the 4-OHE may be formed by the oxidation of ω-3 fats. The presence of Fe or Cu in food ingredients may stimulate the formation of 4-OHE during cooking or preservation. It is well known that fish and perilla oil contain high levels of ω-3 fats. In addition, Sawa et al. [1] reported that lipid peroxyl radicals are formed from vegetable oils in the presence of heme-iron.

On the other hand, 4-hydroxy-2-nonenal (4-HNE) is the most widely studied aldehydic lipid peroxidation

product. 4-HNE is known to show cytotoxic effects. Recently, it was reported that 4-ONE is also a major product of lipid peroxidation [14]. In addition, 4-ONE is more reactive than 4-HNE toward DNA. Therefore, we also measured the amounts of 4-ONE in foods, and found that the levels of 4-ONE were lower than those of 4-OHE. This result is consistent with the fact that we did not detect the 4-ONE-dG adduct in a model reaction consisting of a mixture of methyl linoleate, hemin and dG (data not shown). As mentioned previously, 4-OHE is a lipid peroxidation product derived from ω-3 fats. In contrast, 4-ONE is produced from the ω-6 series of fats. The detection of a higher amount of 4-OHE than 4-ONE is also consistent with the fact that there are more ω-3 fats, such as α-linolenic acid-, docosahexaenoic acid (DHA)- and eicosapentaenoic acid (EPA)-TG, than ω-6 fats, such as linoleic acid TG, in most of the food samples we measured. Another explanation for the higher level of 4-OHE than 4-ONE is that the major ω-3 fats, α-linolenic acid-, EPA- and DHA-TG, in food are more rapidly oxidized than the major ω-6 fat, linoleic acid TG.

In the case of heterocyclic amines, which are well known carcinogens in the human diet, the most abundant, 2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine (PhIP), has been detected in cooked food in the range of 0.3–180 ng/g [15]. In comparison, the amounts of 4-OHE in the human diet are two to three orders of magnitude higher than those of heterocyclic amines.

Epidemiological [8,9] and experimental [10,11,16] studies have demonstrated that cooking oil vapors may be related to cancer risk. Some airborne samples taken in the cooking areas were mutagenic to the TA 98 and TA 100 Salmonella strains without metabolic activation [17]. Since 4-OHE was detected in the smoke condensate from saury and it is mutagenic in the Ames test without metabolic activation [6], it may play a role as a volatile mutagen produced during cooking, as well as a mutagen in food.

Recently, it was suggested that the ingestion of ω-3 fats has desirable effects on human health [18], by preventing cardiovascular disease and cancer. In contrast, some studies [19] suggested that α-linolenic acid TG intake might be a risk factor. These discrepancies may be attributable to the 4-OHE content, depending upon the extent (degree) of lipid peroxidation.

## References

[1] T. Sawa, T. Akaike, K. Kida, Y. Fukushima, K. Takagi, H. Maeda, Lipid peroxyl radicals from oxidized oils and heme-iron: implication of a high-fat diet in colon carcinogenesis, Cancer Epidemiol. Biomarkers Prev. 7 (1998) 1007–1012.

[2] C.K. Winter, H.J. Segall, W.F. Haddon, Formation of cyclic adducts of deoxyguanosine with the aldehydes trans-4-hydroxy-2-hexenal and trans-4-hydroxy-2-nonenal in vitro, Cancer Res. 46 (1986) 5682–5686.

[3] H. Esterbauer, R.J. Schaur, H. Zollner, Chemistry and biochemistry of 4-hydroxynonenal, malonaldehyde and related aldehydes, Free Radic. Biol. Med. 11 (1991) 81–128.

[4] Y. Kawai, K. Uchida, T. Osawa, 2′-Deoxycytidine in free nucleosides and double-stranded DNA as the major target of lipid peroxidation products, Free Radic. Biol. Med. 36 (2004) 529–541.

[5] T. Sakai, S. Kuwazuru, K. Yamauchi, K. Uchida, A lipid peroxidation-derived aldehyde, 4-hydroxy-2-nonenal and omega 6 fatty acids contents in meats, Biosci. Biotechnol. Biochem. 59 (1995) 1379–1380.

[6] H. Kasai, M. Maekawa, K. Kawai, K. Hachisuka, Y. Takahashi, H. Nakamura, R. Sawa, S. Matsui, T. Matsuda, 4-Oxo-2-hexenal, a mutagen formed by ω-3 fat peroxidation, causes DNA adduct formation in mouse organs, Ind. Health 43 (2005) 699–701.

[7] M. Maekawa, K. Kawai, Y. Takahashi, H. Nakamura, T. Watanabe, R. Sawa, K. Hachisuka, H. Kasai. Identification of 4-oxo-2-hexenal and other direct mutagens formed in model lipid peroxidation reactions as dG-adducts, Chem. Res. Toxicol., in press.

[8] T.J. Wang, B.S. Zhou, J.P. Shi, Lung cancer in nonsmoking Chinese women: a case-control study, Lung Cancer 14 (1996) S93–S98.

[9] L. Zhong, M.S. Goldberg, Y.T. Gao, F. Jin, Lung cancer and indoor air pollution arising from Chinese-style cooking among nonsmoking women living in Shanghai, China, Epidemiology 10 (1999) 488–494.

[10] Y.H. Qu, G.X. Xu, J.Z. Zhou, T.D. Chen, L.F. Zhu, P.G. Shields, H.W. Wang, Y.T. Gao, Genotoxicity of heated cooking oil vapors, Mutat. Res. 298 (1992) 105–111.

[11] H. Chen, M. Yang, S. Ye, A study on genotoxicity of cooking fumes from rapeseed oil, Biomed. Environ. Sci. 5 (1992) 229–235.

[12] H. Esterbauer, W. Weger, Uber die wirkungen von aldehyden auf gesunde und maligne zellen 3. Mitt: Synthese von homologen 4-hydroxy-2-alkenalen, II, Monatsh. Chem. 98 (1967) 1994–2000.

[13] P. Manini, R. Andreoli, A. Mutti, E. Bergamaschi, W.M. Niessen, Determination of *n*-hexane metabolites by liquid chromatography/mass spectrometry. 2. Glucuronide-conjugated metabolites in untreated urine samples by electrospray ionization, Rapid Commun. Mass Spectrom. 12 (1998) 1615–1624.

[14] D. Rindgen, M. Nakajima, S. Wehrli, K. Xu, I.A. Blair, Covalent modifications to 2′-deoxyguanosine by 4-oxo-2-nonenal, a novel product of lipid peroxidation, Chem. Res. Toxicol. 12 (1999) 1195–1204.

[15] T. Sugimura, K. Wakabayashi, H. Nakagama, M. Nagao, Heterocyclic amines: mutagens/carcinogens produced during cooking of meat and fish, Cancer Sci. 95 (2004) 290–299.

[16] M. Rojas-Molina, J. Campos-Sanchez, M. Analla, A. Munoz-Serrano, A. Alonso-Moraga, Genotoxicity of vegetable cooking oils in the Drosophila wing spot test, Environ. Mol. Mutagen 45 (2005) 90–95.

[17] K. Teschke, C. Hertzman, C. Van Netten, E. Lee, B. Morrison, A. Cornista, G. Lau, A. Hundal, Potential exposure of cooks to airborne mutagens and carcinogens, Environ. Res. 50 (1989) 296–308.
[18] D. Kromhout, E.B. Bosschieter, C. de Lezenne Coulander, The inverse relation between fish consumption and 20-year mortality from coronary heart disease, N. Engl. J. Med. 312 (1985) 1205–1209.
[19] C. Stripp, K. Overvad, J. Christensen, B.L. Thomsen, A. Olsen, S. Moller, A. Tjonneland, Fish intake is positively associated with breast cancer incident rate, J. Nutr. 133 (2003) 3664–3669.

# Identification of 4-Oxo-2-hexenal and Other Direct Mutagens Formed in Model Lipid Peroxidation Reactions as dGuo Adducts

**Muneyuki Maekawa, Kazuaki Kawai, Yoshikazu Takahashi, Hikaru Nakamura, Takumi Watanabe, Ryuichi Sawa, Kenji Hachisuka, and Hiroshi Kasai**

Department of Environmental Oncology, Institute of Industrial Ecological Sciences, and Department of Rehabilitation Medicine, University of Occupational and Environmental Health (UOEH), 1-1, Iseigaoka, Yahatanishi-ku, Kitakyushu 807-8555, Japan, and Microbial Chemistry Research Center, 3-14-23, Kamiosaki, Shinagawa-ku, Tokyo 141-0021, Japan

Chemical Research in Toxicology®

Reprinted from
Volume 19, Number 1, Pages 130–138

# Identification of 4-Oxo-2-hexenal and Other Direct Mutagens Formed in Model Lipid Peroxidation Reactions as dGuo Adducts

Muneyuki Maekawa,[†,§] Kazuaki Kawai,[†] Yoshikazu Takahashi,[‡] Hikaru Nakamura,[‡] Takumi Watanabe,[‡] Ryuichi Sawa,[‡] Kenji Hachisuka,[§] and Hiroshi Kasai*[,†]

*Department of Environmental Oncology, Institute of Industrial Ecological Sciences, and Department of Rehabilitation Medicine, University of Occupational and Environmental Health (UOEH), 1-1, Iseigaoka, Yahatanishi-ku, Kitakyushu 807-8555, Japan, and Microbial Chemistry Research Center, 3-14-23, Kamiosaki, Shinagawa-ku, Tokyo 141-0021, Japan*

*Received August 31, 2005*

We searched for mutagens that react with 2′-deoxyguanosine (dGuo) in model systems of lipid peroxidation. To autoxidation systems of methyl linoleate (model of ω-6 fat), methyl α-linolenate (MLN) (model of ω-3 fat), and commercial salad oil, dGuo was added. The reaction mixtures were analyzed by HPLC. Six adducts were detected, and their structures were determined by $^{1}H$ and $^{13}C$ NMR, UV, and mass spectra and by comparison with synthetic authentic samples. The mutagens that reacted with dGuo to form these adducts were proposed as glyoxal, glyoxylic acid, ethylglyoxal, and 4-oxo-2-hexenal (4-OHE). The formation of 8-hydroxy-dGuo, an oxidized product of dGuo, was also detected in the model reaction mixtures. Among them, glyoxal and glyoxylic acid are known mutagens, while ethylglyoxal and 4-OHE, produced from MLN, have not been reported as mutagens thus far. We confirmed the mutagenic activity of 4-OHE with *Salmonella* strains, TA100 and TA104, without S9 mix. These compounds may be involved in lipid peroxide-related cancers.

## Introduction

As diet is one of the main causes of human cancer, it is important to identify mutagens in food to prevent cancer. Epidemiological studies suggest that a high-fat diet is a risk factor for various cancers, such as breast and prostate cancer (*1*, *2*). On the other hand, it is known that an elevated risk of colon cancer is associated with red meat intake (*3*), while the role of total fat and specific fatty acids in colon carcinogenesis is not clear (*4*). Brink et al. suggested that a high-intake diet of polyunsaturated fatty acids is associated with an increased risk of mutated K-ras colon tumors, based on a cohort study (*5*). We have been interested in the mutagens produced by lipid peroxidation, because foods contain various polyunsaturated fatty acids and heme iron (Fe), a catalyzer of lipid peroxidation (*6*). Lipid peroxidation may occur during the cooking or storage of foods. After red meat ingestion, heme (ferroprotoporphyrin) is released by hemoglobin digestion and may enter the colon (*7*), while some of the ingested fat may also pass through the colon as triglycerides or as free fatty acids (*8*). Therefore, autoxidation systems of various unsaturated fatty acids with hemin are interesting models of lipid peroxidation reactions, to clarify the combined effect of a high-fat and red meat diet. Although aldehyde type mutagens, such as malondialdehyde, acrolein, crotonaldehyde, and 4-hydroxynonenal, have been identified (*9*−*11*), there have been no detailed studies on other mutagens, particularly those formed from ω-3 polyunsaturated fatty acids. In this report, we describe our identification of four mutagens, including new mutagens, as 2′-deoxyguanosine (dGuo)[1] adducts in model lipid peroxidation reaction mixtures containing hemin and methyl linoleate (MLA), methyl α-linolenate (MLN), and salad oil.

## Experimental Procedures

**Chemicals.** dGuo and hemin were purchased from Sigma Chemical Co. (St. Louis, MO). MLA (purity > 99%) and MLN (purity > 99%) were obtained from Fluka (Buchs, Switzerland) and MP Biomedicals, LLC (Aurora, OH), respectively.

**Reaction of dGuo with Products Formed in Lipid Peroxidation Model Systems and Analysis of the Adducts.** dGuo (20 mg) and hemin (2.5 mg) were mixed in 50 mM phosphate buffer (pH 7.4) (10 mL) and were vigorously shaken with 1 mL of MLA, MLN, or commercial salad oil (rapeseed oil + soybean oil) in an open plastic tube (50 mL) to form a homogeneous emulsion. Occasionally, deionized water was added to maintain the volume of each reaction mixture. The reaction was continued for 3 days at room temperature. After centrifugation to separate the oil layer, 100 μL of the aqueous layer was injected into an HPLC column (CAPCELL PAK C18 MG, 5 μm, 4.6 mm × 250 mm, Shiseido Fine Chemicals, Japan) connected with a photodiode array UV detector (Hewlett-Packard 1100 HPLC Detection System). Used were the following linear gradients of acetonitrile concentrations in 0.1% acetic acid: 0−15 min, linear gradient of acetonitrile (0−5%); 15−28 min, linear gradient of acetonitrile (5−20%). Adduct samples for structure determination were isolated by repeated (about 50 times) rounds of HPLC, by injecting 1 mL of the reaction mixture into a wider column (10 mm × 250 mm).

**Spectra Measurements.** The mass spectrum (EI) of synthetic 4-OHE was measured on a JEOL JMS-BU20 spectrometer. The mass spectra (FAB) of the adducts were recorded with a JEOL

* To whom correspondence should be addressed. Tel: +81-93-691-7469. Fax: +81-93-601-2199. E-mail: h-kasai@med.uoeh-u.ac.jp.
† Department of Environmental Oncology, UOEH.
‡ Microbial Chemistry Research Center.
§ Department of Rehabilitation Medicine, UOEH.

[1] Abbreviations: dGuo, 2′-deoxyguanosine; DHA, docosahexaenoic acid; EPA, eicosapentaenoic acid; MLA, methyl linoleate; MLN, methyl α-linolenate; 4-OHE, 4-oxo-2-hexenal; 4-ONE, 4-oxo-2-nonenal.

10.1021/tx050236m CCC: $33.50 
Published on Web 12/21/2005