Original Paper

Pathobiology

Pathobiology 2000;68:283–290

Received: September 19, 2000
Accepted: March 19, 2001

Investigations into the Molecular Effects of Single Nucleotide Polymorphism

Horst D. Lohrer[a] Uwe Tangen[b]

[a]Radiobiological Institute, Ludwig Maximilians University, Munich, and [b]German National Research Centre for Information Technology, Sankt Augustin, Germany

Key Words
Ataxia telangiectasia · Genome instability · Mutation · Pattern recognition · Proliferating cell nuclear antigen · Single nucleotide polymorphism · Somatic mutation disease · *Thy-1*

Abstract
Objectives: DNA sequences are very rich in short repeats and their pattern can be altered by point mutations. We wanted to investigate the effect of single nucleotide polymorphism (SNP) on the pattern of short DNA repeats and its biological consequences. ***Methods:*** Analysis of the pattern of short DNA repeats of the *Thy-1* sequence with and without SNP. Searching for DNA-binding factors in any region of significance. ***Results:*** Comparing the pattern of short repeats in the *Thy-1* gene sequences of Turkish patients with ataxia telangiectasia (AT) with the 'wild type' sequence from the DNA database, we identified a missing 8-bp repeat element due to an SNP in position 1271 (intron II) in AT-DNA sequences. Only the mutated sequence had the potential for the formation of a stem loop in DNA or pre-mRNA. In super-shift experiments we found that DNA oligomers covering the area of this SNP formed a complex with proteins amongst which we identified the proliferating cell nuclear antigen (PCNA) protein. ***Conclusion:*** SNPs have the potential to alter DNA or pre-mRNA conformation. Although no SNP-depeding formation of the DNA-protein complex was evident, future investigations could reveal differential molecular mechanisms of cellular regulation.

Copyright © 2001 S. Karger AG, Basel

Introduction

The search for mutations in genomic DNA is mostly restricted to the identification of alterations in the coding sequence or the regulatory region of genes, which influence the expression of a phenotype. As a consequence, mutations in introns and intergenic sequences are rarely investigated and are mostly viewed as part of the spectrum of alterations in the DNA as proposed by the quasispecies model [1]. About 90% of the human genetic variation has been ascribed to single nucleotide polymorphisms (SNPs), which occur at frequencies of more than 1% [2]. The most important SNPs seem to be the usually rare non-synonymous mutations, which affect the amino acid sequence of a gene product and, because of linkage disequilibrium, can be genetic markers for mutations

KARGER

Fax + 41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

© 2001 S. Karger AG, Basel
1015–2008/00/0686–0283$17.50/0

Accessible online at:
www.karger.com/journals/pat

Dr. Horst D. Lohrer
GS – Forschungszentrum für Umwelt und Gesundheit
PO Box 1129
D–85764 Neuherberg (Germany)
Tel. +49 89 31872621, Fax +49 89 31873433, E-Mail lohrer@gsf.de

responsible for disease. However, SNPs in non-coding sequences could interfere with nucleotide correlation [3], the complexity and secondary struture of DNA and the mediation of the dynamics of the genome [4].

It had been shown earlier that DNA sequences contained a dense pattern of short repeat elements made of 3–25 bp [5–7], which were shown to stabilize the loop structures in non-coding sequences of pre-mRNA and to determine the half-life and expresssion of mRNA [8]. Also, the analysis of radiation-induced HGPRT deletion mutants revealed that the deleted sequences were flanked by short direct repeats [9]. Increased dynamics of the genome have been found in cells from patients suffering from somatic mutation disease (SMD) [10]. These cells express increased rates of mutations, sister chromatid exchanges, translocations and chromosome instability. Ataxia telangiectasia (AT) is one of the best described SMDs and is characterized by an increased cancer risk and, on the cellular level, by hypersensitivity to ionizing radiation, chromosome instability, hypermutability and a high rate of intra-chromosomal recombination [for reviews, see ref. 11, 12].

We hypothesized that these repeat patterns might possess a function in the stability of the genome, and hence could provide a signal to link the DNA sequence with the metabolism of the DNA. Thus, a factor capable of first binding to DNA and second recruiting DNA-metabolizing enzymes must be associated with these repeat sequences. Amongst the plethora of DNA-binding proteins the proliferating cell nuclear antigen (PCNA) seemed to be the one with the greatest potential to fulfil that role. The ring-shaped homo-trimer structure is capable of encircling double-stranded DNA, forming a sliding clamp [for a review, see ref. 13]. PCNA is an auxiliary protein of DNA polymerase δ [14, 15] and ε [16]. It plays an essential role in nucleotide excision repair [17] as well as in DNA replication and DNA synthesis [18]. Protein-protein interactions have been described with Gadd45 [19], maturation factor 1 (Fen1) [20], the cyclin-dependent kinase inhibitor $p21^{WAF1/Cip1}$ [21], DNA-(cytosin-5) methyltransferase (MCMT) [22] and DNA repair endonuclease XPG [23]. Interestingly, homologous regions of Fen1, $p21^{WAF1/Cip1}$, MCMT and XPG compete for binding to the same site on PCNA [24], which could have immense implications for a regulatory network, linking these diverse protein functions.

We found that an 8-bp repeat element in the *Thy-1* sequence of Turkish AT patients was missing, which was attributed to a point mutation. The repeat sequence lies in an area of similarities to repetitive elements and potentially alters the loop structures of single-stranded DNA and pre-mRNA. Unfortunately, no cells carrying this mutation were available. Hence we were limited to structural analysis in our search for the biological effects of this mutation. In super-shift experiments we identified PCNA complexed with other proteins with double-stranded DNA oligomers containing the sequence of the 8-bp repeat element. However, the binding of the protein complex was independent of the confirmation of the DNA.

Materials and Methods

Sequence Analysis

The sequence of *Thy-1* [25] was obtained from GENBANK (accession No. HUMTHY1A) and the mutations identified in AT patients were communicated by Teraoka [26]. Short sequence repeats were identified using an algorithm described ealier [27]. Repeats of less than 8 bp were not analysed. For the analysis of the hypothetical structure of the single-stranded DNA, the program 'RNA fold' [28, 29] was employed.

Cells and Tissue Culture

Mormal primary MRC5 cells, primary AT4BI and AT7, SV40 transformed MRC5BIVA, AT5BIVA (a generous gift from C. Arlett, Brighton, UK) and GM637, HT29 and atxbc were grown under normal culture conditions in EMEM medium with Earle's salts (Northumbria Biological Ltd, nbl) containing 10% fetal bovine serum, 0.2 U/ml penicillin and 0.2 µg/ml streptomycin.

Preparation of Protein Extracts

4×10^6 were grown in 75-cm^2 flasks, washed twice with ice-cold PBS, scraped off and transferred to Eppendorff tubes. The cells were pelleted for 2 min in the cold room and resuspended in 400 µl of buffer A (10 m*M* HEPES-KOH pH 7.8, 1.5 m*M* $MgCl_2$, 10 m*M* KCl, 0.5 m*M* DTT, 0.2 m*M* PMSF) and incubated on ice for 10 min. After the swelling period the cells were spun down for 1 min in a cold room and resuspended in 50 µl of buffer C (20 m*M* HEPES-KOH pH 7.8, 25% glycerol, 420 m*M* NaCl, 1.5 m*M* $MgCl_2$, 0.2 m*M* EDTA, 0.5 m*M* DTT, 0.2 m*M* PMSF). High salt extraction was continued on ice for 20 min before the nuclei were spun down for 2 min in a cold room, the supernatants transferred into fresh Eppendorff tubes and stored at –70 °C [30].

5′-End-Labelling of DNA Oligomers with T4 Polynucleotide Kinase

20 pmol of each oligomer were directly phosphorylated using [γ^{32}P]dATP (Amersham) and a commercial end-labelling kit (Boehringer). The end-labelled oligomers were ethanol-precipitated, washed twice with ice-cold 80% ethanol and redissolved in 40 µl double-distilled water. The efficiency of the labelling process was determined by counting the cpm of 1 µl solution in a United Technologies/Packard 2000 Series scintillation counter.

For the preparation of double-stranded oligomers, equal quantities (pmol) of labelled complementary oligomers were mixed, incubated for 10 min at 65 °C in a waterbath and allowed to slowly cool down to room temperature. For storage the double-stranded oligomers were kept at –20 °C.

Sequences of the DNA oligomers: thy001 (position 1265–1289): GAG ACC AGG CTG TTC CTG GTC CCA G; thy002 (position 1289–1265): CTG GGA CCA GGA ACA GCC TGG TCT C; thy003 (position 1253–1276): GGC TGG TTC TGT GAG ACC AGG CTG; thy004 (position 1276–1253): CAG CCT GGT CTC ACA GAA CCA GCC; thy005 (position 1265–1289): GAG ACC GGG CTG TTC CTG GTC CCA G; thy006 (position 1289–1265): CTG GGA CCA GGA ACA GCC CGG TCT C; thy007 (position 1253–1276): GGC TGG TTC TGT GAG ACC GGG CTG; thy008 (position 1276–1253): CAG CCC GGT CTC ACA GAA CCA GCC; random oligomer (double-stranded DNA): TGT TCC GAA TGG CCA ACA TTC ACT CCT.

Gel Retardation Assay

If not indicated otherwise, 2 µl of the protein extract, 10 µl of 2 × buffer B (20 m*M* HEPES-KOH pH 7.8, 10 m*M* $MgCl_2$, 2 m*M* EDTA, 2 m*M* DTT, 10% glycerol), 2 µl dI-dC (1 µg/µl, Boehringer), 4 µl sterile double-distilled water and 2 µl of labelled double-stranded oligomer were mixed and incubated for 35 min at 22 °C. Following the incubation period the samples and 1 µl of 2 × loading buffer were applied to a 0.5 × TBE, 3.7% polyacrylamide gel (19:1 mono- to bis-acrylamide) and electrophoresed with 150 V until the dye front and free oligomers migrated out of the gel. The gels were vacuum-dried and autoradiographed [31].

For super-shift experiments, 2 µl of the protein extract, 10 µl of 2 × buffer B, 2 µl of dI-dC and 2 µl of double-distilled water were incubated on ice with 2 µl of antibody solution (PCNA antibody PC10 from DAKO A/S, Denmark, or control antibodies) for 45 min. Thereafter the labelled oligomers were added and the mix kept on ice for another 10 min. Only then the samples were incubated at 22 °C for 35 min and processed as described for band shift experiments.

DNA Preparation and SSCP Analysis

Genomic DNA was extracted using standard protocols. Amplification of the *Thy-1* sequences flanking the polymorphism at position 1271 was achieved in polymerase chain reactions (PCR) employing primers thy1b (position 1181–1201, TCA TGA TGT TTA GTA AGG TT) and thy1c (reverse complementary, position 1370–1351, AAG ACG GAT CAG GAC TCA GG). DNA amplification was performed in a total volume of 10 µl containing 30–50 ng DNA template in 10 m*M* Tris-HCl (pH 8), 50 m*M* KCl, 1.5 m*M* $MgCl_2$, 100 µ*M* of dCTP, dTTP, dATP, dGTP, 0.5 U Taq-DNA polymerase (Life Technologies, Karlsruhe, Germany) and 0.3 nmol of each primer. PCR was based on 36 cycles of 94 °C for 45 s, 55 °C for 1 min and 71 °C for 45 s with an initial denaturation step of 94 °C for 4 min and a final extension step of 71 °C for 7 min. For radioactive labelling 0.3 µl alpha-^{32}P-dCTP (3,000 Ci/m*M*, from Amersham Pharmacia Biotech) was added. 10 µl of each PCR amplification was transferred into 90 µl of dilution buffer (0.1% SDS, 10 m*M* EDTA), mixed and 6 µl thereof added to 6 µl of loading buffer (formamide, 0.1% bromophenol blue, 0.1% xylene cyanol, 10 m*M* EDTA). Finally the samples were heated to 95 °C for 3 min, chilled on ice and loaded onto a native polyacrylamide gel (5% acrylamide containing 5% glycerol, 0.5 × TBE). The fragments were separated by electrophoresis in a standard sequencing apparatus at 12 W for 4 h. Gels were fixed (10% acetic acid in methanol) for 15 min, dried at 90 °C under vacuum and exposed to X-ray film for up to 3 days.

Results

Analysis of the Pattern of Short DNA Repeat Sequences

Based on the analysis of 62 Turkish AT individuals and 35 unaffected individuals 4 point mutations were previously described in the *Thy-1* gene of AT patients: positions 162 (C to T), 197 (G to A), 240 (G to A) in intron I and the position 1271 (A to G) in intron II [26, 32]. The 'wild type' sequence of the *Thy-1* gene was retrieved from the database (HUMTHY1A) and a copy of that sequence was amended with the 4 point mutations identified in AT patients to yield the 'AT' sequence of *Thy-1.*

Prior to the analysis of sequence patterns, the *Thy-1* sequence of the AT patients in 5′–3′ orientation was catenated to its reverse complementary strand, as was the 'wild type'. Disregarding all repeats of less than 8 bp, a total of 52 repeats were found in the AT transgene compared to a total of 53 repeats in the control (data not shown). The excess repeat sequence GACCAGGC (transgene positions 1267–1274 and 4601–4608) from the control transgene was lost in the AT transgene due to the transition (A to G) in position 1271. Hence, despite our restricted pattern analysis (sequence of the repeat had to be 8 bp or longer), 1 out of 4 SNPs in the *Thy-1* gene of Turkish AT patients had abrogated an inverted repeat.

Figure 1a shows the *Thy-1* sequence from the Turkish AT families from position 1161–1400, which became the focus of our investigation. When the 'wild type' *Thy-1* sequences from position 1251 to 1290 were screened against the database of repetitive sequences (BLASTA and FASTA programs, HGMP, Cambridge [33]), similarities to sequences of the rodent b1 repetitive sequence (18 out of 24 bp, boxed with dark-grey shading) and to human *Alu* repetitive sequences (14 out of 15 bp, boxed with light-grey shading) were identified. The same analysis performed on the mutant 'AT' *Thy-1* segment identified no similarities as the mutation abolished the 8-bp minimal identity demanded by the algorithm (data not shown).

Analysis of the Potential Secondary Structure of Thy-1 *Sequences*

The *Thy-1* 'wild type' sequences and the mutated sequence were analysed for their potential to form secondary structures using the 'RNA fold' program [29]. As depicted in figure 1b, a triple-loop structure could be formed starting with position 1244 and extending to position 1308 only with the 'AT' *Thy-1* sequence due to stabilisation by 8 G/C base pairs, including G of position 1271

	1161	TCCCTCAGCC	AGTTTTCTTG	TCATGATGTT	TAGTAAGGTT
	1201	TTCATAAGAT	GATATGTGTG	CAAGAGATCA	GTAATCTGCA
	1241	AATgggAAAG	ATGgcTGGTT	CTGTGAGACC	gggCTGTTCC
	1281	TGGTcccAGC	TAAGACATTg	cAGTAcccAC	CTCCCAAAGG
	1321	GAGTACACCC	TTGCTTTGGG	CCTGTGCCTG	CCTGAGTCCT
a	1361	GATCCGTCTT	CCTTCCTACC	CTGCCCCCGG	CCCCCTTCTC

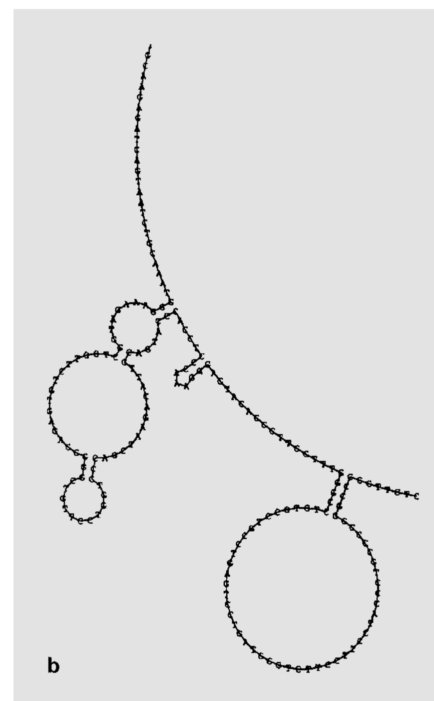

Fig. 1. *Thy-1* sequence from Turkish AT patients and the potential formation of a stem loop. **a** Sequence of *Thy-1* from AT patients from position 1161 to 1400. Lower-case letters indicate the double-stranded areas of the stem loop. The single nucleotide polymorphism in position 1271 found in AT patients (A to G) is indicated in bold. Light shading shows areas of *Alu* repeat homology, darker shading homologies to b1 repeats. **b** Stem loop structures of the *Thy-1* gene from AT patients spanning the sequence from position 910 to 1400 (5′–3′ anticlockwise). The formation of the most distal loop of the triple loop structure is stabilized by the A to G mutation in position 1271.

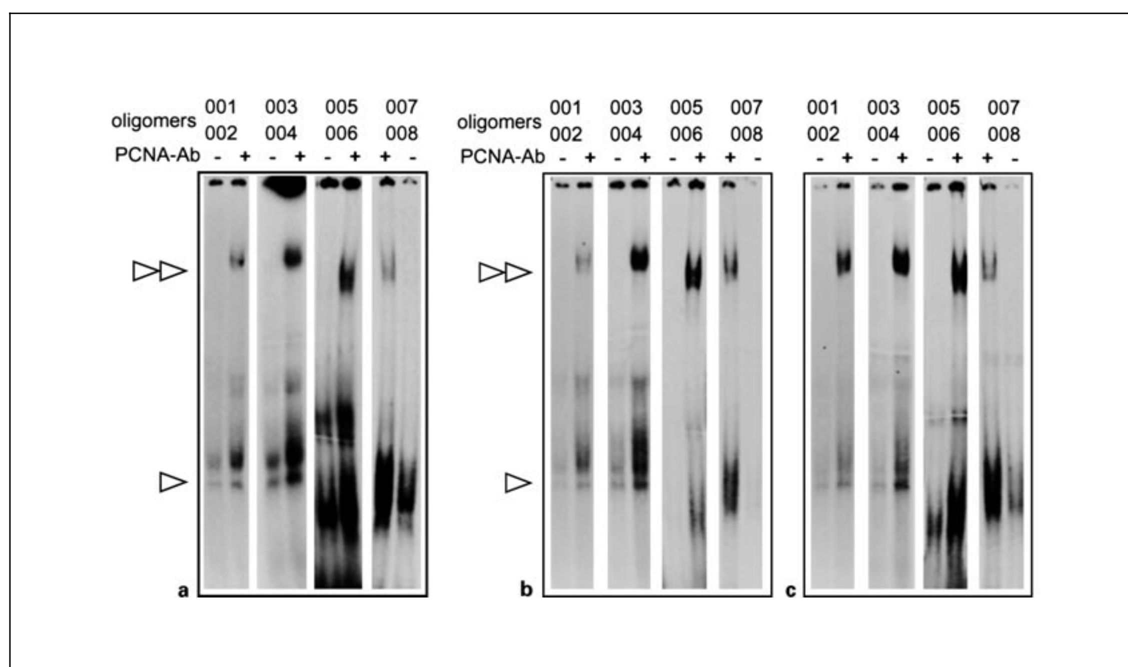

Fig. 2. Gel retardation assay. Nuclear extract from the cell line AT5BIVA (**a**), and primary cells AT4BI (**b**) and MRC5 (**c**) were incubated with radioactively labelled double-stranded oligomers of the 'wild-type' sequence thy001/thy002 or thy003/thy004 and with the 'AT' sequence thy005/thy006 or thy007/thy008. Samples pre-incubated with the monoclonal antibody PC10 against PCNA are marked (+). The single arrow indicates the position of the band shifts, the double arrow the position of the super-shifts. Free DNA oligomer was allowed to run off the gel.

(pairing bases were written in lower case letters in figure 1a).

Investigations into Factors Binding to DNA Regions with SNPs

Double-stranded oligomers of the sequence thy001/thy002, thy003/thy004, thy005/thy006 and thy007/thy008 were incubated with protein extracts from AT5BIVA cells (fig. 2a), AT4BI (fig. 2b) and MRC5 (fig. 2c), which resulted in a retardation of the migration of the oligos, as shown in figure 2, indicated by a single arrow-head (all lanes –PCNA-Ab). This complex formation was resistant to the addition of double-stranded oligomers of random sequence and the migration of the oligomers with random sequence showed no indication of retardation following pre-incubation with protein extract or PCNA antibody. However, addition of a 20-fold excess of unlabelled thy001/thy002 oligomers to the samples abolished the signal in the gel retardation assay (fig. 3).

The autoradiographs of band shifts shown in figure 2a (single arrow; lanes –PCNA-Ab) indicate that the DNA-binding activity was independent of the 'wild type' (oligos 001/002 and 003/004) or 'AT' sequence (oligos 005/006 and 007/008). The additional signals found using oligomers 005/006 were most likely due to overexpression of the gel. However, preliminary experiments revealed that 4 proteins can be found associated with the DNA oligomer (data not shown). Hence, a second explanation for the additional signals could be that these represent intermediate complexes of proteins and DNA oligomer. In order to exclude the possibility of proteins binding to DNA ends as cause for the band shift, we used protein extracts from cell lines defective in Ku protein (xrs-2 and xrs-4), which resulted in the same signals as observed before (data not shown).

PCNA Participates in the DNA Binding Complex

Pre-incubation of nuclear extracts with antibodies against p53, SV 40-T-antigen, ribonuclease reductase, DNA polymerase α and cyclin D1 had no effect on the position of the shifted signal (data not shown). Only pre-incubation with monoclonal antibody PC10 against PCNA consistently caused a significant super-shift of the signal. Figure 2a shows the autoradiographs of band shift experiments using all 4 double-stranded oligomers and protein extracts from AT5BIVA cells. Incubation of extract and oligomer alone (lanes: –PCNA-Ab) showed a band shift indicated by a single arrow. Pre-incubating with antibody PC10 against PCNA (lanes: +PCNA-Ab) produced a high-molecular weight super-shifted signal,

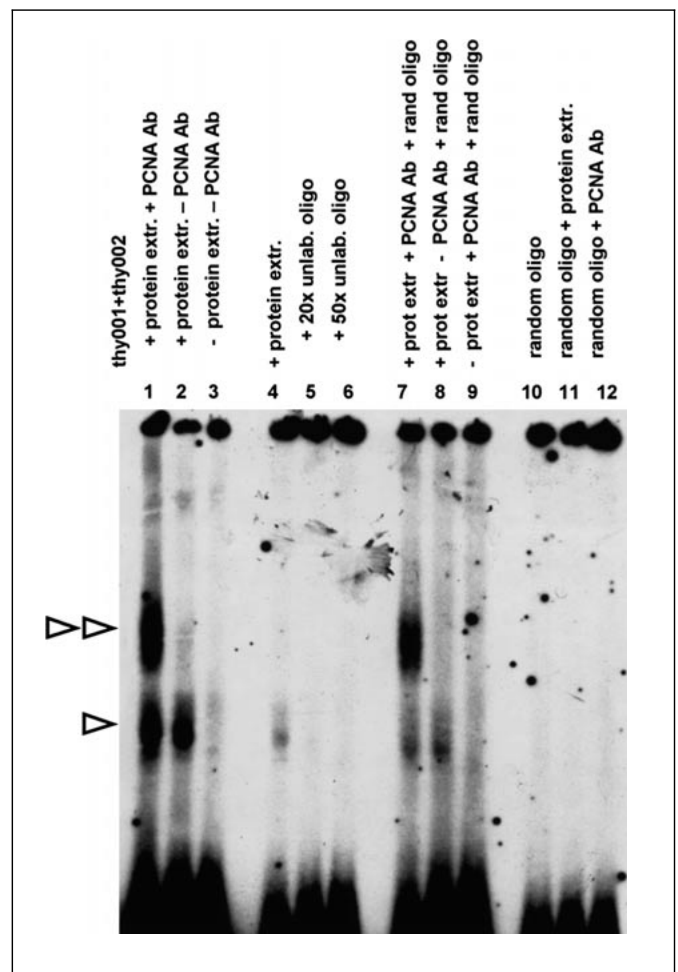

Fig. 3. Gel retardation assay and controls. Protein extract of the cell line AT5BIVA was incubated with radioactively labelled thy001/thy002 double-stranded oligo. Retardiation of the migration of the free oligomer (lane 3) was observed after incubation with protein extract (lane 2, single arrow) and further retardation was achieved following the addition of PCNA antibody to the incubation mix (lane 1, super-shift, double arrow). Addition of 20-fold or 50-fold excess of unlabelled oligomer (lanes 5 and 6, respectively) to the incubation with protein extract obliterates the band shift signal (lane 4). Addition of 50-fold excess of a random DNA-oligomer had no influence on the band shift or the super-shift (lanes 7–9) and the random DNA-oligomer had no protein-binding capacity (lanes 10–12).

indicated by the double arrow. In addition to the SV40-transformed cell line AT5BIVA, we also employed the primary AT4BI cells from an AT patient and MRC5 as normal control cells. As depicted in figure 2b, c, the band shift signals using protein extracts of these primary cells were fainter than those of the transformed cell line AT5BIVA.

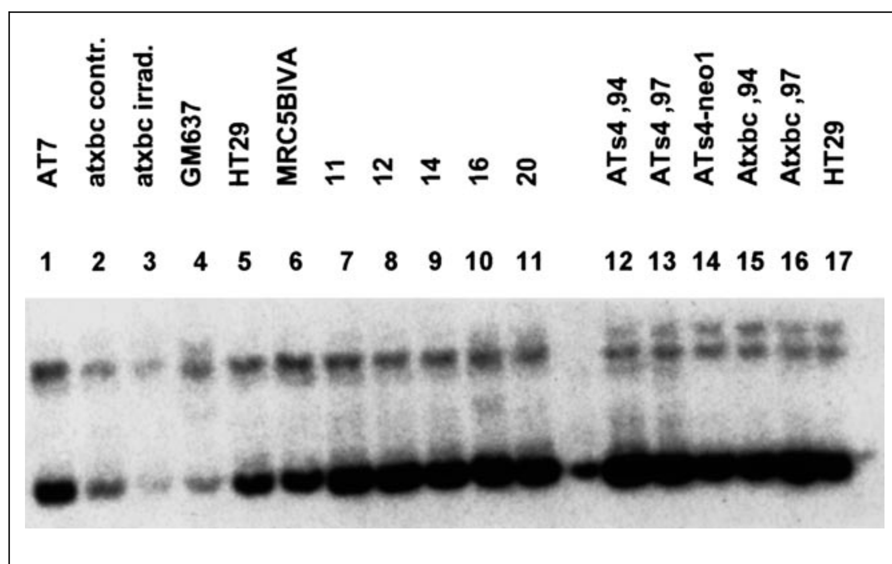

Fig. 4. SSCP analysis of *Thy-1* sequences harbouring the polymorphism at position 1271. Conformational analysis of PCR-amplified DNA: of the cell lines AT7, atxbc, GM637, HT29, MRC5BIVA (lanes 1–6); from blood samples from healthy donors; 11, 12, 14, 16, and 20 (lanes 7–11); from frozen cell samples representing at least 3 years of continuous culture ATs4,94, ATs4,97 (lanes 12 and 13, respectively), atxbc,94, atxbc,97 (lanes 15 and 16, respectively); from ATs4neo1, a subclone of ATs4 (lane 14) and from HT29 (lane 17).

Nevertheless, pre-incubation with PC10 antibody (+PCNA-Ab) produced a strong super-shifted signal.

Estimation of the Frequency of the Polymorphism at Position 1271 of Thy-1

A series of cell lines available in our laboratory and blood samples from 5 healthy volunteers were investigated for the polymorphism at position 1271 of the *Thy-1* gene. None of the cell lines AT7 and ATs4 (from patients with AT, X-ray-sensitive), ATs4-neo1 (G418-resistant pSV2neo transfectant of ATs4), atxbc [34] (X-ray-resistant derivative of ATs4), GM637 (normal), HT29 (colon adenoma, mismatch repair deficient), nor the DNA samples of the 5 volunteers (No. 11, 12, 14, 16 and 20) showed any indications of a polymorphism in SSCP experiments (fig. 4). DNA preparations from frozen ATs4 cells with at least 3 years of permanent culture between them did not result in the accumulation of the '1271' polymorphism in the cell population. The same was true for DNA preparations from two cultures of atxbc cells, one irradiated once per month with 2 Gy for 3 years and the other cultured in parallel under normal conditions. This small study suggested that the occurrence of the polymorphism in position 1271 of the *Thy-1* gene can be expected to be less than 10% of the population, that this locus is resistant to treatment of the cells with mutagens and is stable in cells with deficient DNA repair mechanisms.

Discussion

The comparison of the *Thy-1* DNA sequence from a family of Turkish AT patients with the normal control sequence revealed the abrogation of an 8-bp inverse repeat in the mutated 'AT' sequence, caused by the A to G transition at position 1271. Computer analysis of the 'wild type' *Thy-1* sequence from position 1251 to 1290 revealed similarities to *Alu* sequences (position 1253–1276) and to b1 repetitive elements (position 1265–1290).

The analysis of the physico-chemical characteristics of the sequence (910–1,400 bp) revealed the potential for stem loop formation only in the mutated sequence (G in position 1271). Such stem loop formation could alter the processing and transcription rate of pre-mRNA [8] and hence influence cellular regulatory systems.

Since there were no cells available to investigate the biological implication of the 'AT' mutation, we were limited to collect evidence of molecular interactions of cellular factors with this sequence. Hence, two sets of double-stranded oligomers were designed to cover the region of interest: mutant 'AT' and 'wild-type' versions each of the *Alu* and b1 similarity regions. Band shift experiments proved the specific binding of proteins to all oligomers irrespective of the point mutation in position 1271. Since the sequence of the *Alu* set of oligomers overlapped with the b1 set, the sequence 5′-GAG ACC a/gGG CTG-3′ was common to both types of oligomers. We assumed this to the binding site for proteins, since controls in our experi-

ments made it unlikely that certain DNA conformations or proteins binding to DNA ends were responsible for binding to the oligomers.

Pre-incubation of protein extracts with antibody to PCNA caused further retardation of the migration during gel electrophoresis and hence a super-shift of the oligomers. In accordance with the band shift experiments, the migration of the DNA-protein complex was not influenced by the presence of A or G in position 1271 in the super-shift experiments.

The interaction of PCNA with other proteins in the regulation of cellular processes is well described. PCNA recruits DNA polymerases δ and ε in cycling cells and during DNA repair [35]. If p21$^{WAF1/Cip1}$ is found complexed with PCNA, it inhibits DNA synthesis and together with the inactivation of cyclins by p21 $^{WAF1/Cip1}$ it synchronizes the replication of the DNA with the cell cycle [36]. This regulation could also be assisted by the competitive binding of Fen1, MCMT, XPG, and p21$^{WAF1/Cip1}$ for PCNA [37]. A complex of PCNA and p21$^{WAF1/Cip1}$ prevents DNA methyltransferase from methylating GpC islands.

It is conceivable that the functional multiplicity of PCNA needs multiple determinants for the assembly of the appropriate protein complex. PCNA exists in the cell as a stable trimer encircling double-stranded DNA. Such a PCNA ring contains three potential binding sites for heterologous proteins like p21$^{WAF1/Cip1}$, Fen1, MCMT and XPG to bind simultaneously and regulate various processes of the DNA metabolism [37]. In addition, certain locations of the DNA in forming a complex with the protein-PCNA agglomeration could participate in the regulatory network. Then, depending on the association of PCNA with a particular DNA sequence through interaction with a protein complex, PCNA could recruit different proteins for differential activities like DNA repair, DNA replication or DNA methylation. We would like to suggest that the sequence-specific coexistence of PCNA together with other proteins might be a prerequisite for its activity in the replication process as opposed to DNA repair and methylation. However, it will be necessary to analyse the conjugating proteins in order to decide which function a given protein aggregation will assume.

Acknowledgements

The authors would like to thank C. Arlett, P. Cocannon, R. Gatti, HGMP Resource Centre, P.J. Lohrer, Y, Shiloh, S. Teraoka and J. Thacker for discussions and the provision of material. We are also grateful to the GMD (Sankt Augustin), the AT-Society and the Gray Laboratory (Northwood) for support of the project and the British Council for a travel grant for H.L.

References

1 Eigen M, McCaskill J, Schuster P: Molecular quasi species. J Phys Chem 1988;92:6881–6891.
2 Sunyaev S, Ramenstky V, Bork P: Towards a structural basis of human non-synonymous single nucleotide polymorphisms. Trends Genet 2000;16:198–200.
3 Peng CK, Buldyrev SV, Goldberger AL, Havlin S, Sciortino F, Simons M, Stanley HE: Long-range correlation in nucleotide sequences. Nature 1992;536:168–170.
4 Yockey HP: Information theory and molecular biology. Cambridge, Cambridge University Press, 1992.
5 Nussinov R: Eukaryotic dinucleotide preference rules and their implications for degenerated codon usage. J Mol Biol 1981;149:125–131.
6 Tautz D, Trick M, Dover GA: Cryptic simplicity in DNA is a major source of genetic variation. Nature 1986;322:652–656.
7 Keese PK, Gibbs A: Origins of genes: 'Big bang' or continuous creation? Proc Natl Acad Sci USA 1992;89:9489–9493.
8 McCarthy JEG, Kollmus H: Cytoplasmic mRNA-protein interactions in eucaryotic gene expression. Trends Biochem Sci 1995;20:191–197.
9 Morris T, Thacker J: Formation of large deletions by illegitimate recombination in the *HPRT* gene of primary human fibroblasts. Proc Natl Acad Sci USA 1993;90:1392–1396.
10 Digweed M: Human genetic instability syndromes: Single gene defects with increased risk of cancer. Toxicol Lett 1993;67:259–281.
11 Rotman G, Shiloh Y: ATM: A mediator of multiple responses to genotoxic stress. Oncogene 1999;18:6135–6144.
12 Lohrer HD: Regulation of the cell cycle following DNA damage in normal and ataxia telangiectasia cells. Experientia 1996;52:316–328.
13 Tsurimoto T: PCNA, a multifunctional ring on DNA. Biochim Biophys Acta 1998;1443:23–39.
14 Bravo R, Frank R, Blundell P, MacDonald-Bravo H: Cyclin/PCNA is the auxiliary protein of DNA polymerase delta. Nature 1987;326:515–517.
15 Prelich G, Tan C-K, Kostura M, Mathews MB, So AG, Downey KM, Stillman B: Functional identity of proliferation cell nuclear antigen and a DNA polymerase delta auxiliary protein. Nature 1987;326:517–520.
16 Burgers PMJ: *Saccharomyces cerevisiae* replication factor C. II. Formation and activity of complexes with the proliferating cell nuclear antigen and with DNA polymerase delta and epsilon. J Biol Chem 1991;266:22698–22706.
17 Shivji MKK, Kenny MK, Wood RD: Proliferating cell nuclear antigen is required for DNA excision repair. Cell 1993;69:367–374.
18 Bravo R, MacDonald-Bravo H: Existence of two population of cyclin/proliferating cell nuclear antigen during the cell cycle. Association with DNA replication sites. J Cell Biol 1987; 105:1549–1554.
19 Smith MI, Chen I-T, Zhan Q, Bae I, Chen C-Y, Gilmer TM, Kastan MB, O'Connor PM, Fornace AJ Jr: Interaction of the p53-regulated protein Gadd45 with proliferating cell nuclear antigen. Science 1994;266:1376–1380.
20 Harrington JJ, Leiber MR: DNA structural elements required for FEN-1 binding. J Biol Chem 1995;270:4503–4508.

21 Li R, Waga S, Hannon GJ; Beach D, Stillman B: Differential effects by the p21 CDK inhibitor on PCNA-dependent DNA replication and repair. Nature 1994;371:534–537.

22 Chuang LSH, Ian H-I, Koh T-W, Ng H-H, Xu G, Li BFL: Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21^{WAF1}. Science 1997;277:1996–2000.

23 Gary R, Ludwig DL, Cornelius HL, MacInnes MA, Park MS: The DNA repair endonuclease XPG binds to proliferating cell nuclear antigen (PCNA) and shares sequence elements with the PCNA-binding regions of FEN-1 and cyclin-dependent kinase inhibitor p21. J Biol Chem 1997;272:24522–24529.

24 Warbrick E, Lane DP, Glover DM, Cox LS: Homologous regions of Fen1 and p21^{Cip1} compete for binding to the same site on PCNA: A potential mechanism to co-ordinate DNA replication and repair. Oncogene 1997;14:2313–2321.

25 Seki T, Spurr N, Obata F, Goyert S, Goodfellow P, Silver J: The human *Thy-1* gene: Structure and chromosomal location. Proc Natl Acad Sci USA 1985;82:6657–6661.

26 Teraoka S: Polymerase chain reaction applied to establishing linkage disequilibrium between ataxia-telangiectasia and Thy-1, and to the characterization of alpha tubulin isotypes in HL-60, a human leukemic cell line; PhD thesis, University of California, Los Angeles, 1992.

27 Tangen U: The extension of the quasi-species to functional evolution; PhD thesis, Institute for Molecular Biotechnology, Jena, 1994.

28 Zucker M, Stiegler P: Optimal computer folding of large RNA sequences using thermodynamic and auxiliary information. Nucleic Acids Res 1981;9:133–148.

29 McCaskill JS: The equilibrium function and base pair binding probabilities for RNA secondary structures. Biopolymeres 1990;29: 1105–1119.

30 Andrews NC, Faller DV: A rapid micropreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. Nucleic Acids Res 1991;19:2499.

31 Mathey-Prevot B, Andrews NC, Murphy HS, Kreissman SG, Nathan DG: Positive and negative elements regulate human interleukin 3 expression. Proc Natl Acad Sci USA 1990;87: 5046–5050.

32 Gatti RA, Shaked R, Wei S, Koyama M, Salser W, Silver J: DNA polymorphism in the human Thy-1 gene. Hum Immunol 1988;22:145–150.

33 Altschul SF, Gish W, Miller W, Myers EW, David J: Basic local alignment search tool. J Mol Biol 1990;215:403–410.

34 Lohrer HD, Tangen U, Anderson RF, Arrand JE: Characterization of a hybrid hamster-human cell line complemented for the ataxia telangiectasia DNA repair defects. Pathobiology 1994;62:140–148.

35 Prelich G, Stillman B: Coordinated leading and lagging strand synthesis during SV40 DNA replication in vitro. Cell 1988;53:117–126.

36 Waga S, Hannon GJ, Beach D, Stillman B: The p21 inhibitor of cyclin-dependent kinases controls DNA replication by interaction with PCNA. Nature 1994;369:574–578.

37 Warbrick E: PCNA binding through a conserved motif. Bioessays 1998;20:195–199.