PKD1 Unusual DNA Conformations Are Recognized by Nucleotide Excision Repair*

Received for publication, January 30, 2001, and in revised form, February 23, 2001
Published, JBC Papers in Press, February 27, 2001, DOI 10.1074/jbc.M100845200

The 2.5-kilobase pair poly(purine-pyrimidine) (poly(R·Y)) tract present in intron 21 of the polycystic kidney disease 1 (PKD1) gene has been proposed to contribute to the high mutation frequency of the gene. To evaluate this hypothesis, we investigated the growth rates of 11 Escherichia coli strains, with mutations in the nucleotide excision repair, SOS, and topoisomerase I and/or gyrase genes, harboring plasmids containing the full-length tract, six 5′-truncations of the tract, and a control plasmid (pSPL3). The full-length poly(R·Y) tract induced dramatic losses of cell viability during the first few hours of growth and lengthened the doubling times of the populations in strains with an inducible SOS response. The extent of cell loss was correlated with the length of the poly(R·Y) tract and the levels of negative supercoiling as modulated by the genotype of the strains or drugs that specifically inhibited DNA gyrase or bound to DNA directly, thereby affecting conformations at specific loci. We conclude that the unusual DNA conformations formed by the PKD1 poly(R·Y) tract under the influence of negative supercoiling induced the SOS response pathway, and they were recognized as lesions by the nucleotide excision repair system and were cleaved, causing delays in cell division and loss of the plasmid. These data support a role for this sequence in the mutation of the PKD1 gene by stimulating repair and/or recombination functions.

Polycystic kidney disease (PKD) encompasses a family of closely related syndromes characterized by intraparenchymal renal cysts that are lined by a single layer of epithelial cells. The several forms of PKD include autosomal dominant polycystic kidney disease (ADPKD), which is one of the most common inherited human disorders (~1 per 500 worldwide). Affected individuals typically develop large cystic kidneys, but hepatic cysts, intracranial aneurysms, and cardiac valvular abnormalities are extra-renal manifestations often associated with this disorder. Approximately half of ADPKD patients develop end stage renal disease requiring renal replacement therapy and compose ~5% of the chronic dialysis population in the United States (reviewed in Refs. 1–5).

The genetic defect in >85% of ADPKD cases is mutations in PKD1, a gene that encodes a transcript of 14 kilobases from 46 exons spanning 50 kilobases on chromosome 16p13.3 (6). The 5′ portion of the gene (exons 1–34) is duplicated with more than 95% homology in at least three other copies on chromosome 16p13.1, from which transcripts of 20, 17, and 8.5 kilobases are released (7, 8). However, it is unclear whether these PKD1-like mRNAs are translated into proteins. Polycystin-1, the product of PKD1, is thought to be involved in cell-cell and cell-matrix interactions (9–11) and in calcium-permeable non-selective cation currents (12).

Genotypic analyses of PKD1 microsatellite markers revealed that cysts originate from a single cell that, in a number of cases, underwent loss of heterozygosity. These results led to the proposal that cysts form by a “two-hit” process including a germ line mutation and a subsequent somatic mutation on the functional allele leading to loss of function (13–15). The rate of somatic mutations must be high given the frequent occurrence of ADPKD and the very large number of cysts observed, suggesting the existence of a local hot spot for mutagenesis. The identification of mutations in PKD1 has been hampered by the presence of the homologous genes; several of the 92 mutations identified so far are confined to the 3′-unduplicated region. Therefore, it is unclear whether the frequency of mutations may vary throughout the 50 kilobases of the gene (6–8, 13, 17, 89–106).

The PKD1 gene is intersected in intron 21 by an extraordinary 2.5-kbp poly(purine-pyrimidine) (poly(R·Y) tract) (16), 1 of the 10 longest sequences of this kind. This poly(R·Y) tract is 66% G-C-rich with 95% C + T in the sense strand and is partly repeated in introns 1 and 22. These poly(R·Y) sequences, which are also present in the PKD1-like homologues, have been proposed to contribute to the high mutation rate of PKD1 (17, 18).

A computer search of the 2.5-kbp R·Y sequence revealed 23 mirror repeat sequences, which would be expected to adopt three-stranded DNA structures (intramolecular triplexes) (16) with stems of at least 10 bp. Also, 163 direct repeat sequences were identified, which may adopt slipped, mispaired conformations (16). These structures can block transcription and/or replication and thus induce repair functions (19–23).

Herein, the growth rates of various repair and DNA topological mutant strains of Escherichia coli that harbored plasmids containing the 2.5-kbp poly(R·Y) tract from intron 21 were analyzed. Unusual intramolecular DNA conformations, formed under the influence of negative supercoiling, induced the SOS system and were recognized as lesions by the NER pathway. Delays in cell division were observed as well as loss of the
plasmid, suggesting that such DNA conformations were substrates for strand breaks and endonuclease activities.

**Experimental Procedures**

E. coli Strains

JTT1 (gal-25, lac-, pyrF287, fur-1, rpsL195 (strR), iclR7 (const), trpR72 (Am)), RS2, JTT1 (topA10), and S72, JTT1 (topA10, gyrB226) were obtained from the E. coli Genetic Stock Center at Yale University (New Haven, CT). Strains KMBL1001 (no known mutations), CS5428: KMBL1001 (ΔsurA::cam), CS5429:KMBL1001 (ΔsurB::cam), CS5430: KMBL1001 (ΔsurC::cam), and CS5431:KMBL1001 (ΔsurD::tet) (24) were obtained from Dr. Nora Goosen (Leiden Institute of Chemistry, The Netherlands). Strains JCC123 (GY6781; λ [sfiA::lacZ], Δpro-lac, gallac, rpsL, mal::Tn9, LexAind1), JCC510 (GY7476; λ [p(sfiA::lacZ) CI-]), ΔlacI, Δpro-lac, rpsL), and JCC523 (GY5452; recA441, sulA II, ΔlacI169, thi, leuB6, hie4, argE3, ilvTS, gaiK2, rpsL37, lexA71; Tn5) were kind gifts of Dr. Benedicte Michel (Institut National de la Recherche Agronomique, France).

**Plasmids**

The cloning of pBS4.0 was reported (16). Plasmids containing truncated 5’-end of the 2.5-kbp poly(R-Y) tract were constructed and sequenced by treating pBS4.0 with exonuclease III (Erase-A-base kit, Promega, Madison, WI). The cleavage products were subcloned in the pBluescript KS- (Stratagene, La Jolla, CA). Six clones were selected for further study, pBS1.8, pBS1.5, pBS1.4, pBS1.3, pBS1.0, and pBS0.8, where the numerals indicate the size of the kanm-1::SacI insert (Fig. 2A). The identity of these clones was verified by DNA sequencing (see legend to Fig. 2A). Like pBS4.0, pBSPL3 (GenBank™ accession number U19867) is derived from pBluescript and contains, in addition to the vector backbone, coding sequences for the human immunodeficiency virus envelope protein gp160 used for exon trapping (Life Technologies, Inc.).

**Chemicals**

Novobiocin (Sigma), netropsin (Roche Molecular Biochemicals), and actinomycin D (Sigma) were dissolved in buffer A (0.1 M NaCl, 1 mM sodium phosphate, pH 6.4) at ~10 mg/ml, and their concentrations were determined spectrophotometrically from their extinction coefficients (ε260 = 6 × 10^4 (25), ε293 = 2.92 × 10^4 (26), and ε344 = 2.81 × 10^4 (25)), respectively.

**Results**

**Growth Curves**

The doubling time (t_d) during the exponential phase of growth was obtained from the relation t_d = ln 2/k, where k was the slope as shown in Equation 1.

\[
\ln(CFU/ml) = \ln(CFU/ml) + k t
\]

The standard error of t_d (E_d) was derived from the standard error of k (E_k) according to E_d = t_d/E_d.

The Maximum Fold Decrease in CFU/ml (MDC)—During the first few hours of growth, the number of CFU/ml either decreased or increased non-exponentially. These data were fit by high degree polynomials. MDC was obtained by dividing the CFU/ml at time 0 ((CFU/ml)_{0j}) by the CFU/ml at the curve minimum ((CFU/ml)_{min}). When there was no decrease in CFU/ml, (CFU/ml)_{min} = (CFU/ml)_{0j} and MDC = 1.

**Duration of the Non-exponential Phase of Growth (t_A) —** We define t_A as the time it would have taken for the bacterial population to reach (CFU/ml)_{0j} if it grew exponentially from the theoretical (CFU/ml)_{0j} given by Equation 1. Accordingly, t_A = ln((CFU/ml)_{0j}/(CFU/ml)_{0j})/k.

**Topology of Closed Circular DNA**

The following relations were used to analyze the superhelical density of plasmids. For relaxed, circular DNA L_K = θ (t + W_{0j}), where W_{0j} = n/h^2, n is the total number of bp and h^2 the helical pitch.

Addition of a ligand that binds DNA and alters the superhelical density increases the number of strand breaks, suggesting that such DNA conformations are substrates for strand breaks and endonuclease activities.
PKD1 Unusual DNA Conformations Are Genotoxic in E. coli

The results (Table II) reveal the following. First, the extent of DNA damage-induced SOS response (Table I). The parameters (MDC, \( t_d \), and \( t_L \)) obtained in E. coli strains harboring pBS4.0 were compared with those obtained for pSPL3, a control chosen because of its size (6.0 kb). Longer direct repeats are common as shown in row \( a \). The repeat also contains close runs of guanines throughout its length that may form four-stranded structures (G-guadruplets) (33).

In summary, although many mirror and direct repeats are present, their locations are clustered, suggesting highly specific evolutionary duplication events. To the best of our knowledge, this tract contains the highest density of unorthodox simple sequence repeat features (mirror, direct repeats, and RY strand bias) of any known sequence of this length.

Previous analyses revealed that under superhelical stress the 2.5-kbp poly(RY) tract was cleaved by the single-stranded specific nuclease P1 at four locations, suggesting the formation of unusual structures (19). The location of the four P1 nuclease-sensitive sites is not coincident with the position of the direct repeats with ≤5 bp (Fig. 1); however, sites 1 and 2 do superimpose with clusters of overlapping mirror repeats, as well as with two of the longest direct repeats. This suggests that the unusual DNA structures formed by the poly(RY) sequence in vitro include intramolecular triplexes and mismatched loops.

Length-dependent Delay in Cell Growth—To determine whether such structures form and influence cell physiology, we investigated the effect of the full-length tract and its truncations (Fig. 2A) on the growth rate of wild-type E. coli KMBL 1001. Fig. 2B shows the normalized growth curves obtained with each of the plasmids. The doubling time was lengthened in proportion with the length of the inserts from 21.9 min for pBS0.8 to 24.4 min for pBS4.0. Four plasmids (pBS4.0, pBS1.8, pBS1.5, and pBS0.8) caused a decrease in CFU/ml during the first few hours of growth (Fig. 2B, inset). This loss of viable cells was largest for pBS4.0 that had the full-length poly(RY) tract, and it decreased in proportion to the length of the inserts. Finally, the duration of the non-exponential phase of growth (lag period) also lengthened in accordance with the increasing insert lengths, from 0.8 h for pBS0.8 to 1.8 h for pBS4.0.

Thus, we conclude that the poly(RY) sequence compromised the viability of a number of cells in a length-dependent manner and contributed to the lengthening of the population doubling time as well as the lag period, suggesting that the magnitude of these effects correlated with the number of unusual DNA structures that formed.

DNA Topology, Repair Functions, and SOS Response—To identify genetic factors associated with cell death, growth curves were determined for E. coli mutant strains that affected DNA supercoiling density, nucleotide excision repair (NER), and DNA damage-induced SOS response (Table I). The parameters (MDC, \( t_d \), and \( t_L \)) obtained in E. coli strains harboring pBS4.0 were compared with those obtained for pSPL3, a control chosen because of its size (6.0 versus 7.0 kbp for pBS4.0), the absence of long repetitive sequences, and because, like pBS4.0, it was a derivative of pBluescript.

The results (Table II) reveal the following. First, the extent of cell loss mediated by pBS4.0 correlated with the steady-state level of negative supercoiling (as determined in vivo for pUC19, not shown) because the magnitude of MDC followed the order topA10 > wt > topA10gyrB226. Since the number and stability repeats are mostly CCT and are localized within the 5′-half of the tract, with the exclusion of the very 5′-end (row b). Pentanucleotide repeats are of three types as follows: CTCCC, CTCTT, and CCCAT. The first type (row c) can be further subdivided into three clusters according to the following reading frames: 8 CTCCC tandem repeat units (TRU) at the 5′-end, 5 CCTCC TRU in the middle, and a cluster of 19 TCCCC TRU at the 3′-end. The CTCTT direct repeats are positioned exclusively in the middle of the tract, interspersed by a single CCCAT tandem repeat (row d). Longer direct repeats are common as shown in row \( c \). The tract also contains close runs of guanines throughout its length that may form four-stranded structures (G-guadruplets) (33).

In summary, although many mirror and direct repeats are present, their locations are clustered, suggesting highly specific evolutionary duplication events. To the best of our knowledge, this tract contains the highest density of unorthodox simple sequence repeat features (mirror, direct repeats, and RY strand bias) of any known sequence of this length.

Previous analyses revealed that under superhelical stress the 2.5-kbp poly(RY) tract was cleaved by the single-stranded specific nuclease P1 at four locations, suggesting the formation of unusual structures (19). The location of the four P1 nuclease-sensitive sites is not coincident with the position of the direct repeats with ≤5 bp (Fig. 1); however, sites 1 and 2 do superimpose with clusters of overlapping mirror repeats, as well as with two of the longest direct repeats. This suggests that the unusual DNA structures formed by the poly(RY) sequence in vitro include intramolecular triplexes and mismatched loops.

Length-dependent Delay in Cell Growth—To determine whether such structures form and influence cell physiology, we investigated the effect of the full-length tract and its truncations (Fig. 2A) on the growth rate of wild-type E. coli KMBL1001. Fig. 2B shows the normalized growth curves obtained with each of the plasmids. The doubling time was lengthened in proportion with the length of the inserts from 21.9 min for pBS0.8 to 24.4 min for pBS4.0. Four plasmids (pBS4.0, pBS1.8, pBS1.5, and pBS0.8) caused a decrease in CFU/ml during the first few hours of growth (Fig. 2B, inset). This loss of viable cells was largest for pBS4.0 that had the full-length poly(RY) tract, and it decreased in proportion to the length of the inserts. Finally, the duration of the non-exponential phase of growth (lag period) also lengthened in accordance with the increasing insert lengths, from 0.8 h for pBS0.8 to 1.8 h for pBS4.0.

Thus, we conclude that the poly(RY) sequence compromised the viability of a number of cells in a length-dependent manner and contributed to the lengthening of the population doubling time as well as the lag period, suggesting that the magnitude of these effects correlated with the number of unusual DNA structures that formed.

DNA Topology, Repair Functions, and SOS Response—To identify genetic factors associated with cell death, growth curves were determined for E. coli mutant strains that affected DNA supercoiling density, nucleotide excision repair (NER), and DNA damage-induced SOS response (Table I). The parameters (MDC, \( t_d \), and \( t_L \)) obtained in E. coli strains harboring pBS4.0 were compared with those obtained for pSPL3, a control chosen because of its size (6.0 versus 7.0 kbp for pBS4.0), the absence of long repetitive sequences, and because, like pBS4.0, it was a derivative of pBluescript.

The results (Table II) reveal the following. First, the extent of cell loss mediated by pBS4.0 correlated with the steady-state level of negative supercoiling (as determined in vivo for pUC19, not shown) because the magnitude of MDC followed the order topA10 > wt > topA10gyrB226. Since the number and stability
PKD1 Unusual DNA Conformations Are Genotoxic in E. coli

Table I

| Strain   | Genotype                                      | Relevant features                      |
|----------|-----------------------------------------------|----------------------------------------|
| JTT1     | gal-35, λ-, pgyR287, far-1, rpsL195(strR), nicR7(Const), trpR72(Am) | Wild type                              |
| RS2      | JTT1 except topA10                           | ~15% more negatively supercoiled than JTT1 |
| SD7      | JTT1 except topA10, gvrB226                  | ~10% less negatively supercoiled than JTT1 |
| KMBL1001 |                                  | Wild type                              |
| CS5428   | KMBL1001 except ωvrA::cam                    | Deficient in helicase activities in NER |
| CS5429   | KMBL1001 except ωvrB::cam                    | Deficient in endonuclease activities in NER |
| CS5430   | KMBL1001 except ωvrC::cam                    | Deficient in endonuclease activity in NER |
| CS5431   | KMBL1001 except ωvrD::tet                    | Deficient in helicase activity; constitutive induction of SOS response |
| JJC510   | Δ[p(sfiA::lacZC(ind2)], Δlac-pro, rpsL       | Wild type                              |
| JJC123   | JJC510 except mal::Tn9, gal-1, lexAind1      | Non-cleavable LexA repressor; non-inducible SOS response |
| JJC523   | recA441, lexA71::Tn5                         | Constitutive induction of SOS response  |

Table II

Parameters for the growth curves

| E. coli strain | pBS4.0 | pSPL3 |
|----------------|--------|-------|
|                | MDC    | t2    | tA    | MDC   | t2    | tA    |
| JTT1 (wt)      | 2778   | 31.8  | 8.9   | 1.2   | 22.5  | 1.2   |
| RS2 (topA10)   | 5004   | 42.3  | 12.4  | 3.1   | 27.0  | 2.4   |
| SD7 (topA10, gvrB226) | 238   | 32.8  | 8.2   | 1.1   | 27.4  | 0.9   |
| KMBL1001 (wt)  | 477    | 24.1  | 7.2   | 1     | 21.1  | 1.6   |
| CS5428 (ΔωvrA) | 101    | 21.8  | 3.7   | 15.8  | 22.4  | 2.4   |
| CS5429 (ΔωvrB) | 1      | 20.8  | 1.1   | 1.1   | 22.3  | 1.4   |
| CS5430 (ΔωvrC) | 3.4    | 23.7  | 3.1   | 1     | 20.3  | 1.5   |
| CS5431 (ΔωvrD)| 626    | 26.3  | 6.2   | 171   | 20.8  | 4.8   |
| JJC510 (wt)    | 10.0   | 21.6  | 3.1   | 3.3   | 21.7  | 1.5   |
| JJC123 (lexAind1)| 26.8  | 38.9  | 5.3   | 1     | 29.1  | 1.1   |
| JJC523 (recA441, lexA71)| 3.6   | 30.7  | 4.1   | 2.1   | 25.6  | 1.9   |

* Significant cell lysis was observed.

of underwound unusual DNA structures increase with the extent of negative supercoiling (20–23, 28, 34, 35), we conclude that supercoiling stabilized their formation and that this was critical for the loss of cell viability. Furthermore, the differences in doubling time between pSPL3 and pBS4.0 in the three topA gyr strains (15.3, 9.3, and 5.4 min, respectively) also depended upon negative supercoiling, suggesting that the formation of unusual DNA structures interfered with cell division.

Second, E. coli NER mutants revealed that UvrB and UvrC were necessary for the loss of cells. In fact, the MDC was negligible for both ΔωvrB and Δωvr strains compared with their isogenic wild-type KMBL1001. In addition, for ΔωvrB there was no lengthening of the doubling time, whereas ΔωvrA was shortened to ~15 min from ≥13 h observed for all the other strains. The second significant result from this analysis was the considerable loss of cells in the ΔωvrA and ΔωvrD strains harboring only the control plasmid pSPL3. Because deficiency in UvrD is known to lead to constitutive expression of the SOS response (36–38), a likely explanation is that the SOS system may induce apoptosis during the first period of cell population growth.

Third, we investigated the E. coli SOS strains. During the preparation of JJC523 (ΔlexA71) “founder cells” harboring pBS4.0, we observed significant cell lysis that yielded ~1 × 10⁸ CFU/ml from overnight cultures compared with ≥1 × 10⁹ obtained normally with the same strain harboring pSPL3 or all the other strains with either pBS4.0 or pSPL3. This result clearly indicates that pBS4.0 caused extensive cell death when associated with a fully active SOS response. The growth curve started from the surviving cells did not show a large decrease in cell count. The MDC value with pBS4.0 for the wild-type JJC510 was much lower than for the other two wild-type strains, JTT1 and KMBL1001. We observed that the values of MDC (and therefore tA) for the three wild-type strains, for topA10, and for topA10 gyrB226, varied considerably between experiments. Further analyses indicated that most of this variability arose during the preparation of founder cells used to start the growth curves. At that stage a variable number of cells lost their plasmid; these cells then succumbed at the beginning of the growth culture, when MDC was measured. In the lexAind1 strain JJC123 harboring pBS4.0, where the SOS system is not inducible, the MDC value was comparable to that of its isogenic wild-type strain, JJC510.

In summary, the data indicate that the poly(R·Y) sequence from intron 21 of PKD1 formed unusual DNA structures as a consequence of negative superhelical density and that the detrimental effects on growth depended upon activation of the SOS system. Furthermore, both UvrB and UvrC were necessary to elicit the pBS4.0-dependent lethality.

Integrity of the poly(R·Y) Sequence in Surviving Cells—Previous work showed that long repetitive sequences such as (CTG·CGG)₃ and (CGG·CCG)₃ are unstable and that the average length of the repeats shortens as the cell population ages (39–42). Therefore, we investigated whether the surviving cells represented a subpopulation that had deleted part, or all, of the poly(R·Y) tract. pBS4.0 was isolated from all of the strains and analyzed in two ways. First, the 4.0-kbp PKD insert was excised from the vector, and the insert and vector were separated by agarose gel electrophoresis. Most lanes showed considerable smearing throughout their length. How-
The circles shown. Note that the and 30 selected on Ap plates. A few colonies were suspended in 1 ml of LB agarose gel electrophoresis (1% w/v) in 90 mM Tris borate, 2 mM EDTA, in the absence or in the presence of 5 μM novobiocin and analyzed by agarose gel electrophoresis (1% w/v) in 90 mM Tris borate, 2 mM EDTA, pH 8.0, in the presence of chloroquine. The gel negatives were scanned with a PhosphorImager and the data were smoothed by a low pass filter that used a fast Fourier transform to remove abnormal high frequencies. The y axis is the area of each peak in pixels as given by the PhosphorImager. The negative superhelical density (x axis) was obtained as described (35). Data were fit to a three-parameter Gaussian curve. Filled circles, no drug; open circles, with drug.

FIG. 2. Negative superhelical density in vivo. Topoisomer distributions of pUC19 were isolated from E. coli strain KMB1001 grown to an A$_{600}$ of 0.5 in the absence or in the presence of 5 μM novobiocin and analyzed by agarose gel electrophoresis (1% w/v) in 90 mM Tris borate, 2 mM EDTA, pH 8.0, in the presence of chloroquine. The gel negatives were scanned with a PhosphorImager and the data were smoothed by a low pass filter that used a fast Fourier transform to remove abnormal high frequencies. The y axis is the area of each peak in pixels as given by the PhosphorImager. The negative superhelical density (x axis) was obtained as described (35). Data were fit to a three-parameter Gaussian curve. Filled circles, no drug; open circles, with drug.

FIG. 3. Growth curves for E. coli strains transformed with pBS4.0 prepared in the presence of novobiocin, actinomycin D, or netropsin. E. coli cells were transformed with pBS4.0 and selected on Ap plates. A few colonies were suspended in 1 ml of LB medium, and 100–200 μl was transferred to several tubes containing 10 ml of LB medium with Ap and various concentrations of novobiocin or actinomycin D or netropsin. The cultures were grown overnight, and the number of CFU/ml was determined by plating aliquots on agar plates without Ap. A known number of CFU/ml was then transferred the next day to 1-liter cultures with Ap but without novobiocin or actinomycin D or netropsin, and the growth curves were determined. A. E. coli topA10 and 5 μM novobiocin; B. E. coli KMB1001 and 30 μM actinomycin D; C. E. coli KMB1001 and 4 μM netropsin; D. E. coli ΔuvrB and 4 μM netropsin. Filled circles, no drug; open circles, with drug. Only the data points from the early times are shown. Note that the x and y axes are not identical in the four panels. The Δt values (t$_{drug}$ − t$_{control}$) for A–D were 3.2, −0.9, 0.5, and −0.2 min, whereas the ratios MDC$_{drug}$/MDC$_{control}$ were 0.05, 0.05, 16.7, and 0.8.

However, no discrete products were visible besides the two expected bands, except 1 sample out of 33 where a recombination event took place. Second, these inserts and vectors were radioactively labeled, and their molar ratio was calculated. The ratios ranged from 0.47 ± 0.07 to 0.64 ± 0.06; however, no statistical significance was found among the mean values, and no correlation was observed between these ratios and the MDC values obtained for the same E. coli strains.

Thus, we conclude the following. First, loss of the full-length poly(R-Y) tract in favor of more stably transmitted deletion products was not observed. Second, both the smearing and the lower molar ratio of the PKD insert indicate that a proportion of pBS4.0 was in the process of being degraded at the time of plasmid isolation and that the starting point for such degradation was within the PKD insert. We suggest that pBS4.0 was either replicated and transmitted intact during cell division or cleaved at the poly(R-Y) tract and then rapidly degraded.

**Influence of Negative Supercoiling**—To verify further the differences observed in the topoisomerase I and gyrase E. coli mutants, as well as the ΔuvrB strain, MDC was evaluated in identical cells following alterations in their levels of negative supercoiling. We used novobiocin to inhibit the assembly of active gyrase (43, 44) and thus achieve a relaxation of the DNA in vivo. This shows the formation of unusual DNA structures and relieve the extent of founder cell loss. Fig. 3 shows that the population of more highly negative supercoiled topoisomers of pUC19 was progressively reduced in the presence of novobiocin, as expected.

Fig. 4A shows the growth curves for the topA10 strain carrying pBS4.0 in the absence or presence of novobiocin. Addition of novobiocin during the preparation of cells maintained their complete viability in the subsequent growth, whereas >90% was lost in its absence. Superimposable growth curves were found for 2–10 μM novobiocin. Cell loss was not prevented when novobiocin was added at the beginning of the growth curves rather than during the preparation of founder cells, indicating that the loss required an activity to take place at an early step.

An alternate strategy was also implemented to influence supercoiling, and hence DNA structure, by preparing cells in the presence of drugs (actinomycin D or netropsin) that bind directly to DNA and thereby influence its global topology as well as its conformations at specific sequences. Analyses of the topoisomer distributions of pUC19 in the presence of either drug confirmed their binding to the DNA (not shown).

Actinomycin D intercalates at G-C pairs, and it reduces the number of negative superhelical turns and stabilizes DNA in the right-handed duplex B-form (45–50). Due to its G-C richness, the poly(R-Y) tract offers numerous binding sites for the drug, whose activity is expected to decrease the number of invivable cells.

Fig. 4B shows the growth curves for the wild-type E. coli strain KMB1001 harboring pBS4.0 in the absence or presence of actinomycin D. As described for novobiocin, actinomycin D was only added during the preparation of cells. The extent of cell loss within the first 2 h decreased progressively in the presence of 5–30 μM actinomycin D, supporting the hypothesis that the formation of underwound non-B DNA structures was responsible for cell loss.

Netropsin binds to the minor groove of A-T pairs, requiring four or more such bp for optimal contacts, introduces additional negative supercoils, destabilizes triplex DNA, and increases the stiffness of the double helix (26, 51–58). We reasoned that netropsin would accentuate cell loss in two ways. First, it would bind to duplex B-DNA and induce additional negative supercoiling. Second, it would bind selectively to the vector sequences (of the 131 consecutive four A-T pairs only 7% are in the poly(R-Y) tract) and thus increase the stiffness of the vector. Thus, the superhelical density would be preferentially partitioned into the poly(R-Y) tract, which is the most flexible and writhed region of the plasmid (59).

Fig. 4C shows the growth curves for wild-type E. coli KMB1001 carrying pBS4.0 in the absence and presence of
netropsin. As described above, the drug was only added during the preparation of cells. The data show that netropsin strongly accentuated cell selection during the first 5 h. The growth curves were not affected by the addition of netropsin to *E. coli* KMBL1001 cells harboring pSPL3 or when the compound was added to *E. coli* KMBL1001 containing pBS4.0 only at the beginning of the growth curve culture.

To verify further whether cell loss was caused by the formation of non-B DNA structures alone or whether it required their recognition by the NER system, the experiment with netropsin was conducted in *E. coli* Δ*uvrB* carrying pBS4.0. Contrary to the wild-type strain, Δ*uvrB* cells were unaffected by the drug (Fig. 4D), proving that the recognition of unusual DNA structures by NER was indispensable for eliciting cell selection.

In summary, we conclude that optimal viability required the DNA to be maintained in an orthodox right-handed B-form. We also conclude that the reactions of the NER proteins on the undamaged non-B DNA structures, but not their formation alone, was indispensable for eliciting loss of cell viability.

**Loss of Plasmid**—The relationship between the SOS and NER pathways and cell viability became apparent when colonies derived from the exponential phase of the growth curves were analyzed. Samples of *E. coli* transformed with pSPL3 taken throughout the period of growth yielded colonies of essentially homogeneous size when plated without Ap. Alternatively, colonies derived from several *E. coli* strains transformed with pBS4.0 plated after the cultures grew for 3–5 h appeared heterogeneous in size. DNA isolated from small and large colonies revealed that pBS4.0 was only present in the small but not in the large colonies.

To determine the extent of plasmid loss for pSPL3 and pBS4.0, the number of Ap-resistant (ApR) and Ap-sensitive colonies for various *E. coli* strains was measured. Fig. 5A shows the logarithm of the ratio between the number of ApR colonies and the total number of colonies for *E. coli* strains harboring either of the two plasmids. For pSPL3 (Fig. 5A, hatched bars), the greatest differences were observed among the wild-type strains, whereas only small differences were seen between the wild-types and their respective mutants. Ten and 90% of the JTT1 and JJC510 cells, respectively, contained the plasmid, whereas less than 0.1% of cells retained pSPL3 by 10-, and 1000-fold in KMBL1001, JTT1, and JJC510, respectively.

As for pSPL3, pBS4.0 was least stable in wild-type KMBL1001 strain. Retention of pSPL3 was slightly greater (2–5-fold) in the Δ*uvrB* and Δ*uvrC* mutants (the Δ*uvrA* and Δ*uvrD* mutants were not tested). Addition of netropsin to KMBL1001 cells caused the dramatic reduction by 4 orders of magnitude in the number of plasmid-containing cells, whereas addition of actinomycin D or novobiocin slightly increased the retention of pSPL3 5- and 10-fold, respectively.

These results indicate that plasmid stability was improved when negative supercoiling was lessened, as expected. The *topA10* mutation increased pSPL3 instability by about 5-fold, whereas the stability increased 5-fold in the *topA10 gyrB226* strain, confirming that supercoiling was detrimental to plasmid stability. Inactivation or constitutive expression of the SOS system in the Δ*lexA71* and *lexAind1* strains, respectively, did not affect the retention of pSPL3.

In summary, these results with pSPL3 show that plasmid loss commonly took place during cell growth, that the extent of loss depended on the host genetic background, and that it correlated with negative supercoil density.
PKD1 Unusual DNA Conformations Are Genotoxic in E. coli

18603

The role of DNA topology (supercoiling) was evaluated with mutant strains and with a gyrase inhibitor (novobiocin) that affected supercoil density as well as with ligands (actinomycin D and netropsin) that bind directly to DNA and thereby influence its global topology and hence its conformations at specific sequences. The composite data show that unusual DNA structures (triplexes and slipped structures) in the 2.5-kbp poly(R-Y) tract of the PKD1 gene exist in E. coli and are responsible for the observed genotoxicity. The PKD1 gene is a human sequence on chromosome 16p13.3, and hence the types of structural transitions formed in this plasmid system may also occur in humans.

Fig. 6 outlines the main steps that may be involved in the process. The in vivo homeostatic control of negative supercoiling (71–73) (Fig. 6, A and B) generates plasmid topoisomeres in which pleonemotic conformations maintain the DNA under highly negative torsional stress (74, 75). The poly(R-Y) tract undergoes structural transitions (Fig. 6C) to various non-B structures (19–23) under the influence of both steady-state levels of negative supercoiling and waves of hyper-negative supercoiling generated by the passage of DNA helix-tracking enzymes such as during transcription, replication, or repair. Underwound unusual DNA structures should be dissipated as they are approached by an incoming polymerase due to positive supercoiling formed ahead of the complex (76–79) and thus are predicted to not interfere with replication. E. coli strains defective in the SOS-induced genes, specifically $\Delta$uvrA and $\Delta$uvrB (80), failed to show a lengthening of the population doubling time, in contrast with the strains with a functional SOS system. Thus, unusual DNA structures may have interfered with cell division only after DNA damage (i.e. strand break(s)) inflicted on them by NER (81) caused replication forks to collapse (82–85) (Fig. 6, D and E).

Experiments performed previously with plasmids containing (CTG-CAG) triplet repeats (39, 67, 86, 87) showed that the tract frequently deletes to discrete and stably transmitted new species in E. coli populations and that both the MMR and NER systems are involved (67, 86, 87). On the contrary, we show here that the poly(R-Y) tract from the PKD1 gene does not give rise to such discrete and stably transmitted deletions. A possible reason for the difference is that the presence of the triplet repetitive elements within the (CTG-CAG) tracts facilitates post-DNA damage repair and/or recombination and thus maintains plasmid viability, which is lost when similar damage is inflicted within the poly(R-Y) tract. We also show (Fig. 4) that the NER nuclease activity related to DNA safety is not directed to a particular DNA sequence but rather to DNA structures that may arise from several types of sequences. Therefore, we conclude that the NER system has the capacity of recognizing certain non-B DNA structures formed by undamaged bases, in addition to recognizing damaged bases (88). Future in vitro studies with the highly purified components of the NER system may provide further insights into the substrate requirements.

**Acknowledgments**—We thank Drs. Gregory M. Landes for advice; John Bouck (Baylor College of Medicine) for computer searches of the human genome; Nora Goosen (Leiden Institute of Chemistry, The Netherlands) and Benedicte Michel (Institut National de la Recherche Agronomique, France) for E. coli strains; and E. Lynn Zechiedrich (Baylor College of Medicine), Gregory G. Germino (The Johns Hopkins University), Ravi R. Iyer, and Richard R. Sinden for critically reading the manuscript, encouragement, and advice.

**REFERENCES**

1. Calvet, J. P. (1998) J. Nephrol. 11, 24–34
2. Mureia, N. S., Woychik, R. P., and Anver, E. D. (1998) Pediatr. Nephrol. 12, 721–726
3. Grunfeld, J.-P., Chauveau, D., Joly, D., Fonck, C., and Oualim, Z. (1999) J. Nephrol. 12, 842–846
4. Harris, P. C. (1999) Hum. Mol. Genet. 8, 1861–1866
5. Wu, G., and Song, S. (2000) Mol. Genet. Metabol. 69, 1–5
6. European Polycystic Kidney Disease Consortium (1994) Cell 77, 881–894

Discussion

To determine whether cell loss during the growth curves (Table II) could be explained by the behavior of cells that lack plasmid, untransformed wild-type E. coli JTT1, JJC510, KMB1001, and the $\Delta$uvrB mutant were prepared in the absence of Ap and then used to mock-start a growth curve in fresh medium with Ap. Fig. 5B shows that >98% of these cells lost their viability within the first 2 h of growth. This result is consistent with the hypothesis that cell death was mediated by the selection stress on cells that had lost their plasmid during the previous culture.

In summary, these data show that the poly(R-Y) sequence from intron 21 of the PKD1 gene contains unusual DNA structures that form under the influence of negative supercoiling, that these structures interact with components of the NER pathway, and that such interactions lead to plasmid loss.

**REFERENCES**

1. Calvet, J. P. (1998) J. Nephrol. 11, 24–34
2. Mureia, N. S., Woychik, R. P., and Anver, E. D. (1998) Pediatr. Nephrol. 12, 721–726
3. Grunfeld, J.-P., Chauveau, D., Joly, D., Fonck, C., and Oualim, Z. (1999) J. Nephrol. 12, 842–846
4. Harris, P. C. (1999) Hum. Mol. Genet. 8, 1861–1866
5. Wu, G., and Song, S. (2000) Mol. Genet. Metabol. 69, 1–5
6. European Polycystic Kidney Disease Consortium (1994) Cell 77, 881–894

**REFERENCES**

1. Calvet, J. P. (1998) J. Nephrol. 11, 24–34
2. Mureia, N. S., Woychik, R. P., and Anver, E. D. (1998) Pediatr. Nephrol. 12, 721–726
3. Grunfeld, J.-P., Chauveau, D., Joly, D., Fonck, C., and Oualim, Z. (1999) J. Nephrol. 12, 842–846
4. Harris, P. C. (1999) Hum. Mol. Genet. 8, 1861–1866
5. Wu, G., and Song, S. (2000) Mol. Genet. Metabol. 69, 1–5
6. European Polycystic Kidney Disease Consortium (1994) Cell 77, 881–894

**REFERENCES**

1. Calvet, J. P. (1998) J. Nephrol. 11, 24–34
2. Mureia, N. S., Woychik, R. P., and Anver, E. D. (1998) Pediatr. Nephrol. 12, 721–726
3. Grunfeld, J.-P., Chauveau, D., Joly, D., Fonck, C., and Oualim, Z. (1999) J. Nephrol. 12, 842–846
4. Harris, P. C. (1999) Hum. Mol. Genet. 8, 1861–1866
5. Wu, G., and Song, S. (2000) Mol. Genet. Metabol. 69, 1–5
6. European Polycystic Kidney Disease Consortium (1994) Cell 77, 881–894
