Isolation of repeated and self-complementary sequences from *E. coli* DNA

Ann C. Kato, Lorraine Borstad, Murray J. Fraser and David T. Denhardt

Department of Biochemistry, McGill University, Montreal, Quebec,
Canada H3G 1Y6

Received 28 August 1974

ABSTRACT

We have used the single-strand specific nuclease from *Neurospora crassa* and chromatography on methylated albumin-kieselguhr to purify and characterize repeated and self-complementary sequences from *Escherichia coli* DNA. Approximately 0.5% of the genome renatures spontaneously at zero time and another 2% renatures somewhat more rapidly than the total DNA. The early renaturing DNA has a base composition and a T_m similar to the total DNA and contains on the average 100 base pairs; the self-complementary DNA also has a base composition like *E. coli* but contains a mean of 170 base pairs. No evidence was obtained for the presence of a highly redundant sequence.

INTRODUCTION

Prokaryotes do not contain repetitious DNA, at least to the degree that eukaryotes do, and searches for repeated sequences in *E. coli* have been negative (1) or have revealed a fraction of extrachromosomal DNA (2). Nevertheless, the *E. coli* genome is known to contain 5-8 copies of the ribosomal RNA cistrons (3,4) and "insertion DNA sequences" (5) and it is reasonable to believe that there may be more than one copy of certain operator and promotor sequences. Also it has been suggested that multiple copies of a DNA sequence exist which serve as a signal for the initiation of synthesis of the nascent DNA fragments that are intermediates in DNA replication (6). Because previous work had not ruled out the possible existence of a small fraction of highly repeated DNA in *E. coli*, we decided to initiate this more detailed investigation.

The proportion of those sequences in the DNA that are present more than once, and their degree of repetition, can be deduced from the kinetics of reassociation of sheared, denatured DNA. Most commonly, chromatography on HAP¹ is used to quantitate the amount of reassociated DNA (7). We describe here a

method for the quantitation of reassociated DNA using the N. crassa nuclease (8,9). The enzyme is highly specific for single-stranded nucleic acids and was used previously by Brahic and Fraser (10) to isolate a rapidly-renaturing fraction from mouse Ehrlich Ascites DNA. Similar procedures using the single-strand specific nuclease S1 have been employed to isolate repetitive and self-complementary sequences from eukaryote sources (11-14); in this study we report the isolation of "low-repetitive" and self-complementary DNA from a prokaryote source.

MATERIALS AND METHODS

³H- and ³²P-labelled E. coli C and HF 4704 (a C derivative) DNAs were prepared by conventional procedures using lysozyme, EDTA and dodecyl sulfate to lyse the cells; Pronase and phenol to deproteinize; and finally isopropanol precipitation and isopycnic CsCl centrifugation (15). The specific activities ranged from 0.05-1 μ Ci/ μ g for the [³H]thymidine labelled DNA and 0.01-0.5 μ Ci/ μ g for the [³²P]DNA. N. crassa nuclease was prepared from conidia by the method of Rabin and Fraser (8). The predominant activity in the preparation was exonucleolytic, although there was some endonucleolytic activity specific for single-stranded DNA (16). The enzyme preparation showed very little activity towards duplex DNA (9). The E. coli DNA (2-200 μ g/ml in 0.05 M Tris-HCl, pH 8.0) was sheared by sonication in an Artek sonicator using the microprobe for 30 sec intervals for a total of 3 min. The DNA was denatured at 100°C for 10 min and renatured at 60°C or quick-cooled. All reassociation reactions were performed in sealed 1 ml test tubes at 60°C in 0.09 M Tris-HCl, 0.09 M NaCl, pH 7.5, at a DNA concentration ranging from 0.2-200 μ g/ml. The annealed, quick-cooled DNA was treated for 1-4 hours with sufficient N. crassa nuclease (1-5 units/ml) at 37°C in 0.01 M MgCl₂, 0.09 M Tris-HCl, 0.09 M NaCl, pH 7.5, to degrade all single-stranded DNA (10). Determination of nuclease-resistant, acid-precipitable material was made either by filtration through Millipore (or glass fiber) filters

or by spotting on 1 in² pieces of Whatman #5 filter paper and washing with 5% trichloroacetic acid. The filters were dried and the radioactivity determined in a scintillation counter. Chromatography on MAK was performed according to the method described by Brahic and Fraser (10). Chromatography on HAP (Bio-Rad) columns was performed at 60°C using stepwise elutions with 0.14 M Na-phosphate (pH 6.8) and 0.4 M Na-phosphate (pH 6.8) (17).

RESULTS

The Cot curve for total E. coli DNA shows no indication of a repetitive fraction (1). Thus, in order to demonstrate its existence, the following procedure was used to enrich for these sequences: The DNA was denatured and renatured at a low Cot (0.5- 1) such that only a portion (roughly 10%) of the DNA could reassociate. This was repeated one or more times, each time removing unrenatured DNA with the N. crassa single-strand specific nuclease and

TABLE 1

Percent Recovery of DNA After Successive Cot 0.5 Isolations.

DNA Reassociated	Concentration μg/ml	Percent Recovered	Percent of Starting DNA
1 x Cot 0.5	170	3.7	3.7
2 x Cot 0.5	34	12	0.45
3 x Cot 0.5	25	24	0.11

E. coli C DNA labelled with [³H]methyl-thymidine at a specific activity of 25 μCi/mg was sonicated, denatured, reannealed to a Cot of 0.5, digested with the N. crassa nuclease, and purified on MAK. The cycle was repeated 3 times. The Table gives the percent recovered after each cycle, and also the percent of the original starting DNA.

purifying the remaining duplex DNA by chromatography on MAK. Table 1 shows the results of a typical 3-cycle purification. By the end of the third cycle half of the DNA remaining appears to be self-complementary in that it renatures in a concentration-independent manner.

Figure 1. Reassociation Rate Curves of *E. coli* DNA.

Panel A: [³H]thymidine or [³²P]phosphate-labelled *E. coli* DNA was sonicated, denatured, reassociated to various Cot values, treated with the *N. crassa* nuclease, and the acid-precipitable radioactivity determined as described in Materials and Methods. (●—●), total *E. coli* DNA; (□—□), DNA isolated after one cycle of Cot 1 purification; (○—○), DNA isolated after two cycles of Cot 0.5 purification.

Insert: Comparison of the nuclease and HAP procedures. (○—○), reassociation rate curve for total *E. coli* DNA determined by the nuclease method; (●—●), reassociation rate curve replotted from the data of Britten and Kohne (1), determined using HAP chromatography.

Panel B: Reassociation rate curve of "fast-reassociating" [³H]DNA (Cot 0.65) and "slow-reassociating" [³²P]DNA (Cot >2) annealed and nuclease treated together. The total DNA concentration was 1.3 µg/ml. (●—●), [³H]DNA. (○—○), [³²P]DNA. Further details are described in the text.

Fig. 1A shows the kinetics of reassociation of the total *E. coli* DNA and the early-renaturing fraction of the *E. coli* genome that has been purified by one or two cycles of denaturing and reannealing at a Cot of 0.5 -1. The Cot 1 DNA constitutes about 8% of the genome and approximately one-quarter of it appears repetitive to some degree; thus roughly 2% of the *E. coli* genome comprises repeated sequences. The principal difference between the Cot 1 DNA and the 2 x Cot 0.5 DNA is that the latter shows a substantial enrichment for a component that renatures spontaneously at zero time. The insert in Fig. 1A compares the use of HAP chromatography and the *N. crassa* nuclease to quantify the amount of duplex DNA formed. The curves are very similar, although

the $Cot_{1/2}$ is slightly higher for the nuclease as compared with HAP; this is to be expected since single-stranded tails on duplex segments would be considered as duplex by the HAP method but not by the nuclease procedure.

To verify that the early-renaturing DNA seen in the Cot 1 and Cot 0.5 DNA was not produced by exposure to *N. crassa* nuclease, fractions of "repetitive" [3H]DNA and "non-repetitive" [^{32}P]DNA were obtained using HAP chromatography rather than the nuclease method. The [3H]DNA represented DNA renaturing at a Cot of 0.6, the [^{32}P]DNA represented DNA that had not renatured at a Cot of 2. Fig. 1B shows the renaturation kinetics of the two preparations combined as a function of C_0 and t where C_0 represents the sum of the two DNAs. This experiment indicates that the repeated and self-complementary sequences detected in the experiment described in Fig. 1A are not artifacts of the nuclease digestion.

Figure 2. Melting profiles of *E. coli* DNA. Melting profiles were determined according to the method of Marmur and Doty (18) with a Gilford 220 modification of a Beckman DU spectrophotometer equipped with a wide-range digital readout thermosensor (19). Temperature was controlled with a Haake circulating bath. The absorbance readings were not corrected for thermal expansion of the solution. (■—■) melting profile of Cot 1 DNA (45 $\mu g/ml$) in 0.15 M NaCl, 15 mM sodium-citrate (pH 7); (●—●), melting profile of Cot 0.5 DNA (10 $\mu g/ml$) in 7.2 M NaClO₄ - 0.01 M EDTA; (○—○), melting profile of total *E. coli* DNA (18 $\mu g/ml$) in 7.2 M NaClO₄, 0.01 M EDTA.

Some properties of the early-renaturing DNA are illustrated in Figs. 2 and 3. Fig. 2 shows melting profiles in SSC of the DNA isolated at a Cot of 1 and in 7.2 M NaClO₄ of the DNA isolated at a Cot of 0.5. In SSC, the Cot 1 DNA melts with a T_m of 89°C whereas native *E. coli* DNA melts at 90.5°C (18); in 7.2 M NaClO₄ both the native and reassociated Cot 0.5 DNA melt at approximately

50°C. This result, together with other experiments showing the repetitive and self-complementary DNAs have base compositions similar to the total *E. coli* DNA (data not shown), indicates that this DNA comprises a well-matched duplex with an average base composition close to that of *E. coli* DNA. The hyperchromic shift was approximately 30% in both solvents, also indicative of a well-matched duplex DNA.

Figure 3. Sedimentation-Velocity Analyses of the Repetitive and Self-Complementary DNA. Repetitive [^{32}P]DNA was isolated from *E. coli* C by two cycles of denaturation, renaturation to a Cot of 0.5, nuclease digestion, and purification on MAK. From this DNA the self-complementary DNA was isolated by omitting the renaturation step. Portions of the initial DNA (circles), the $2 \times \text{Cot}$ 0.5 DNA (squares), and the self-complementary DNA (triangles), were centrifuged on linear neutral (open symbols) and alkaline (solid symbols) sucrose density gradients. Centrifugation was for 13 hours at 50,000 rpm in the SW 56 rotor at 15°C in a Beckman L2-65B ultracentrifuge. The neutral gradients contained 1 M NaCl, 50 mM Tris-HCl (pH 8), 1 mM EDTA, and 5-20% sucrose. The alkaline sucrose gradients contained 0.8 M NaCl, 0.2 M NaOH, 1 mM EDTA, 0.1% lauroyl sarcosinate, and 5-20% sucrose. The arrows indicate the positions of fragment IV and VII produced from the ^3H -labelled ϕX RF I by digestion with the restriction endonuclease from *Hemophilus aegyptius*; the fragments comprise approximately 630 and 205 base pairs respectively (K. Bartok and F. Grosveld, personal communication). Sedimentation is from right to left.

Panel A: Neutral gradients. The percent of the recovered DNA found in each fraction is plotted. (○—○) initial sonicated DNA; (□—□) $2 \times \text{Cot}$ 0.5 DNA, (Δ — Δ), self-complementary DNA.

Panel B: Alkaline sucrose gradients. The percent of the total recovered DNA found in each fraction is plotted. (●—●), initial sonicated DNA; (■—■), $2 \times \text{Cot}$ 0.5 DNA; ($\blacktriangle$ — $\blacktriangle$), self-complementary DNA.

Panel C: A cumulative plot of the percent of the total recovered radioactivity found sedimenting faster than a certain fraction in the gradient versus that fraction. The symbols are the same as in Panels A and B.

The repeated and self-complementary DNAs were centrifuged on neutral and alkaline sucrose gradients to determine their size. These gradients are shown in Panels A and B of Fig. 3. For purposes of clarity, the data are plotted as the cumulative amount of DNA sedimenting faster than a certain fraction in the gradient as a function of that fraction (Panel C). It can be seen that: (i) The self-complementary DNA is slightly larger than the DNA prepared after two cycles of reannealing at a Cot of 0.5. This observation verifies that the enzyme is not nicking duplex DNA since the self-complementary DNA was isolated after an additional denaturation step and enzyme digestion from the 2 x Cot 0.5 DNA. (ii) The reduction in sedimentation rate under alkaline conditions as compared with neutral conditions is less for the self-complementary than for the initial or the 2 x Cot 0.5 DNA. The $\overline{M}_{W_s}$ (determined from the 50% positions in the gradients) for the initial sonicated DNA, the 2 x Cot 0.5 DNA, and the self-complementary DNA in the neutral gradients are 3.8×10^5 , 5.5×10^4 , and 1.0×10^5 daltons respectively; in alkaline gradients these values are 1.9×10^5 , 3.5×10^4 , and 8.1×10^4 respectively.

DISCUSSION

Chiscon and Kohne (2) reported the existence of a small, highly-repetitive sequence in E. coli BB. They argued that this repeated DNA was derived from a plasmid because the number of copies of the sequence varied with the growth phase of the cell. We observed no difference in the renaturation kinetics of the repeated DNA characterized here in the presence of DNA from exponential or saturated cultures (data not shown). Also, the absence of closed circular DNA in our E. coli C strains was verified by isopycnic ethidium bromide-CsCl centrifugation.

We find (Fig. 1) that approximately 2% of the E. coli genome is present in more than one copy. The base composition resembled that of E. coli DNA and there was very little base-pair mismatching in this DNA as judged by

the melting behaviour (Fig. 2). Because the early-renaturing DNA is heterogeneous in its degree of repetition and represents such a small fraction of the genome, it is not surprising that it does not manifest itself in the reassociation rate curve of the total DNA. The conventional HAP procedures (1) are not sensitive enough to detect such a small fraction of the DNA.

Approximately 0.5% (estimated from Figure 1 and Table 1) of the DNA appears to be composed of self-complementary sequences since it renatures in a concentration independent manner at very low Cot values. This possibility is supported by the sedimentation data (Fig. 3). The molecular weight of "good" duplex DNA measured by sedimentation in neutral sucrose gradients should be $2x$ that measured in alkaline gradients, and this is the case for the initial sonicated DNA. In comparison, the self-complementary DNA appears to be only $1.3x$ larger, as would be expected if this DNA contained mostly cross-linked fragments (20) or hairpin structures such as those constituting the tRNA genes (21). If the self-complementary DNA does arise from sequences capable of forming hairpins, then the hairpin loops in these structures must be small since they are not sensitive to the *N. crassa* nuclease. The presence of hairpins with loops resistant to the single-strand specific nucleases in eukaryote DNA has been elegantly demonstrated by Wilson and Thomas (14). Although they concluded that self-complementary DNA was absent from *E. coli*, it is likely that their procedures (chromatography on HAP) would not have revealed the small amount present.

Our experiments do not indicate whether the repeated and self-complementary sequences are genes or whether they have some role other than the specification of gene products; for example, operators or promoters common to several genes. If promoters are repeated in the genome, and if they contain pyrimidine runs (22), then the repetitive DNA could have been enriched for particular pyrimidine tracts. However, the self-complementary and repetitive DNA fractions gave base compositions and pyrimidine isostich

distributions similar to those of total DNA (data not shown). No evidence was obtained for the presence of the highly repeated sequence postulated by Denhardt (6) to serve as a start signal for the initiation of synthesis of polydeoxyribonucleotide strands. If such sequences exist either they are too short to be detected by this method or they form hairpins (23).

ACKNOWLEDGEMENTS

We thank Shlomo Eisenberg for the base composition data, Barbara Harbers for the pyrimidine isostich determinations, Katalina Bartok for the Hae endonuclease markers, and David Lane for the plasmid analyses. The research was supported by the Medical Research Council and the National Cancer Institute of Canada.

¹ The abbreviations used are: HAP, hydroxyapatite; MAK, methylated albumin-kieselguhr; $\overline{MW}$, weight average molecular weight; SSC, standard saline (0.15 M NaCl) _w citrate (0.015 M); Cot (moles of nucleotides times seconds per liter) (1).

REFERENCES

1. Britten, R.J. and Kohne, D.E. (1968) *Science* 161, 529-540.
2. Chiscon, J.A. and Kohne, D.E. (1972) *Carnegie Inst. Wash. Yearbook* 68, 388-391.
3. Spadari, S. and Ritossa, F. (1970) *J. Mol. Biol.* 53, 357-367.
4. Yu, M.T., Vermeulen, C.W. and Atwood, K.C. (1970) *Proc. Natl. Acad. Sci. Wash.* 67, 26-31.
5. Saedler, H. and Heiss, B. (1973) *Molec. Gen. Genet.* 122, 267-277.
6. Denhardt, D.T. (1972) *J. Theoret. Biol.* 34, 487-508.
7. Kohne, D.E. (1970) *Quarterly Rev. of Biophysics* 33, 327-375.
8. Rabin, E.Z. and Fraser, M.J. (1971) *Preparative Biochem.* 1, 283-307.
9. Kato, A.C., Bartok, K., Fraser, M.J. and Denhardt, D.T. (1973) *Biochim. Biophys. Acta* 308, 68-78.
10. Brahic, M. and Fraser, M.J. (1971) *Biochim. Biophys. Acta* 240, 23-36.
11. Sutton, W.D. and Walker, P.M.B. (1972) *Biochem. J.* 128, 193-198.
12. Cech, T.R., Rosenfeld, A. and Hearst, J.E. (1973) *J. Mol. Biol.* 81, 299-325.
13. Mace, M.L., Clark, R.W. and Arrighi, F.E. (1974) *Biochem. Biophys. Res. Commun.* 56, 9-13.
14. Wilson, D.A. and Thomas, C.A., Jr. (1974) *J. Mol. Biol.* 84, 115-124.
15. Schekman, R.W., Iwaya, M., Bromstrup, K. and Denhardt, D.T. (1971) *J. Mol. Biol.* 57, 177-199.
16. Rabin, E.Z., Tenenhouse, H. and Fraser, M.J. (1972) *Biochim. Biophys. Acta* 259, 50-68.
17. Kohne, D.E. (1968) *Biophys. J.* 8, 1104-1118.
18. Marmur, J. and Doty, P. (1962) *J. Mol. Biol.* 5, 109-118.
19. Brown, R.H.L., Szer, W. and Spencer, J.H. (1969) *Lab. Practice* 18, 547-548.

20. Alberts, B.M. and Doty, P. (1968) J. Mol. Biol. 32, 379-403.
21. Gierer, A. (1966) Nature 212, 1480-1481.
22. Szybalski, W., Kubinski, H. and Sheldrick, P. (1966) Cold Spring Harbor Symp. on Quant. Biol. 31, 123-124.
23. Sobel, H.M. (1973) Prog. Nucleic Acids Res. and Mol. Biol. 13, 153-190.