Evidence that populations of Dictyostelium single-copy mRNA transcripts carry common repeat sequences

Karen L. Kindle and Richard A. Firtel

Department of Biology (B-022), University of California, San Diego, La Jolla, CA 92093, USA

Received 19 March 1979

ABSTRACT

Two recombinant plasmids, M4 and KH10, carrying *Dictyostelium* DNA inserted into the Eco RI restriction endonuclease site of pMB9 by poly(dA)-poly(dT) tailing, were selected for study because they are complementary to abundant mRNA populations from *Dictyostelium*. Both plasmids have been shown to hybridize a heterogeneous size class of mRNAs which, in the case of KH10, comprise 5-10% of the pulse-labeled poly(A)⁺ RNA from vegetative cells. Analysis of the sequence organization of the two pieces of *Dictyostelium* DNA shows that they consist mostly of single-copy sequences with a short DNA sequence which is repeated in the genome and interspersed with single-copy DNA. These and other results suggest that the majority of the hybridization of pulse-labeled mRNA to M4 and KH10 is to the short "repeated" DNA sequences. In the genome, members of these repeat families appear to be transcribed onto a population of different single-copy mRNAs. Additional results show that M4 DNA contains a sequence which is entirely complementary to a discrete mRNA.

INTRODUCTION

Dictyostelium discoideum is a favorable system in which to study gene control during a simple developmental pathway. Removal of the food source induces vegetative amoebae to aggregate and form a multicellular organism which synchronously undergoes a developmental program of morphological, biochemical, and transcriptional changes (1). The organism has a small genome, with a haploid size of 3×10^{10} daltons (2,3).

Previously, we have shown that the general pattern of interspersion of short, moderately repeated sequences and unique sequences is qualitatively similar to that observed in a majority of metazoans (4). In addition, we presented data on the hybridization of poly(A)⁺ RNA to total genomic repeat DNA which suggested that a fraction of *Dictyostelium* mRNA and hnRNA consists of short sequences, transcribed from reiterated DNA which are covalently associated with longer single-copy sequence transcripts (5). In this communication, we describe the isolation and properties of recombinant plasmids

carrying randomly sheared Dictyostelium nuclear DNA sequences which are complementary to a large population of mRNAs. Results suggest that two of these plasmids carry short sequences which are reiterated in the Dictyostelium genome and are transcribed and present on a subpopulation of mRNAs. Isolation and analysis of the mRNAs carrying these common repeat sequences suggests that the protein coding regions are transcribed from single-copy DNA.

EXPERIMENTAL SECTION

The materials and methods employed in the work described in this paper have been previously described (10). Certain details are given in the Legends to Figures in this communication.

RESULTS

Construction and selection of recombinant plasmids complementary to poly(A)⁺ RNA. Recombinant plasmids carrying randomly sheared Dictyostelium DNA inserted into pMB9 (6) were constructed using the poly(dA)-poly(dT) homopolymer tailing procedure (7,8). Details of the methods have been described previously (9,10). Plasmids complementary to pulse-labeled poly(A)⁺ RNA were selected using the colony hybridization technique (11). Three plasmids, M4, M6 and KH10, were selected for further examination because, in colony filter hybridization experiments, they showed strong hybridization to labeled poly(A) containing RNA. Plasmid M6 was subsequently shown to be complementary to actin mRNA, and the complete analysis of plasmid M6 is presented elsewhere (10,12,13). The data on M6 in this manuscript are presented as a contrast to the results we describe with plasmids M4 and KH10, the foci of this communication.

To examine the relative abundance in the pulse-labeled poly(A)⁺ RNA of molecules containing sequences complementary to the plasmid DNAs, DNA from individual clones was prepared and attached to nitrocellulose filters. Duplicate filters were hybridized to in vivo labeled whole cell poly(A)⁺ RNA. The results of this experiment are shown in Table 1. Plasmids M4, M6 and KH10 hybridize a relatively high percentage of the poly(A)⁺ RNA. These experiments have been repeated with polysomal poly(A)⁺ RNA and the results are virtually identical to those obtained with whole cell or cytoplasmic poly(A)⁺ RNA.

TABLE 1 Hybridization Between Recombinant Plasmid DNA and Dictyostelium mRNA.

Recomb. Plasmid	Input cpm	% Bound	% Bound (corrected) ^a	% RNase _b resist.	% Bound cpm that are RNase resist. ^c	Relative RNase _d resistance
M4	200,000	1.4	2.3	0.22	15.9	22%
M6	200,000	0.5	0.8	0.21	42.4	61%
KH10	200,000	5.5	9.2	0.55	10.7	15%

Plasmid DNA was prepared and attached to nitrocellulose filters as described (10) and hybridized for 24 hrs to vegetative [³²P]-poly(A)⁺ whole cell RNA prepared by adsorption to poly(U) sepharose. Hybridization buffer contained 5x SSC, 45% formamide, 0.2% SDS, 2mg/ml poly(A) and washes were done in this buffer lacking poly(A). After two days at 37°C, the filters were washed extensively, counted and treated with 10µg/ml RNase A in 2x SSC for 30 min at room temperature. Filters were washed several times in 2x SSC and counted in scintillation fluid.

^aControl experiments in which purified 17S + 26S rRNA was hybridized to filter-bound cloned rDNA indicated that when the DNA is present in excess, hybridizations go 60% to completion under our conditions. Thus, the percent bound was normalized to 100% reaction by multiplying by 5/3.

^b%RNase resistant = $\frac{\text{cpm bound after RNase}}{\text{total cpm}} \times 100\%$

^c% bound cpm that are RNase resistant = $\frac{\% \text{ RNase resistant}}{\% \text{ Bound (uncorr.)}} \times 100\%$

^dWhen rRNA:rDNA hybrids are immobilized on nitrocellulose and treated with RNase as described ~ 70% of the bound cpm are resistant to digestion; so relative RNase resistance = % bound cpm that are RNase resistant $\times 10/7$.

When the RNA:DNA hybrids formed with M4 and KH10 DNA are treated with ribonuclease, a large fraction of the RNA is digested relative to rRNA:rDNA hybrids (see Table 1). The large fraction of RNA which hybridizes to M4 and KH10 plasmid DNA and the sensitivity of the hybrids to RNase suggest that M4 and KH10 hybridize only a short sequence on molecules from a large population of different mRNAs. This phenomenon will be discussed in more detail in the following sections.

Nature of RNA hybridizing to M4, M6 and KH10. Because of the large fraction of RNA hybridizing to these clones and the differing degrees of ribonuclease resistance of the hybrids, we were interested to determine the number and size of the RNA species complementary to plasmids M4, M6 and KH10. To do this, [³²P] pulse-labeled poly(A)⁺ RNA was size fractionated on 3.6% polyacrylamide tube gels containing 99% formamide. The gels were sliced, and the RNA from each slice was hybridized separately to DNA filters carrying plasmid DNA from one of the three clones. The filters were then washed extensively and counted before and after treatment with RNase A. The results of this experiment were shown in Fig. 1. The figure shows that M6 hybridizes to a relatively discrete peak of RNA and that most of the RNA hybridized is resistant to RNase treatment, as indicated in Table 1.

Figure 1. Hybridization of Size-Fractionated RNA to M4, M6, and KH10 DNA Filters. Whole cell $[^{32}\text{P}]$ *in vivo* labeled poly(A)⁺ RNA was size fractionated on 3.6% polyacrylamide tube gels in 99% formamide (32). The gels were frozen, sliced, and total counts determined by Cerenkov radiation. Hybridization buffer lacking poly(A) and SDS was added to the fractions and the gel was disrupted by two 5 sec. bursts of sonication. The buffer was brought to 50% formamide, 5x SSC, 10 mM EDTA, pH 7.2, 0.7 mg/ml poly(A), 0.2% SDS, and the samples were hybridized for two days at 37°C with nitrocellulose filters to which 1.7 μg plasmid DNA had been adhered (10). All fractions were extensively washed, and the separate DNA filters counted by Cerenkov radiation. The DNA filters were then treated with 20 $\mu\text{g}/\text{ml}$ RNase A in 2 x SSC for 30 min. at room temperature, rinsed several times in 2x SSC and counted in aquasol. $\Delta\text{--}\Delta$, total $[^{32}\text{P}]$ Cerenkov cpm; $\circ\text{--}\circ$, $[^{32}\text{P}]$ Cerenkov cpm hybridized to DNA filters before ribonuclease digestion; $\bullet\text{--}\bullet$, $[^{32}\text{P}]$ Aquasol cpm hybridized after ribonuclease digestion.

N.B. Cerenkov counting is approximately 1/3 as efficient as counting in Aquasol. Since the data are plotted as total assayed cpm, peaks appear higher after RNase treatment.

In contrast, plasmid KH10 DNA, which hybridizes approximately 10% of poly(A)⁺ RNA, shows hybridization to a very broad size range of RNA. RNase treatment results in removal of 90% of the hybridized label in agreement with the results shown in Table 1, and the fraction of RNA which is nuclease resistant is relatively constant along the gel. To account for the fact that it binds such a large fraction of heterogeneous sized RNA, and that only 10-20% of the hybrids are RNase resistant, we suggest that KH10 contains a short sequence which is complementary to many different mRNA molecules. We would anticipate that such a sequence would be reiterated in the genome in order to be transcribed onto a population of messages, and we present evidence below that KH10 does contain a sequence in its Dictyostelium DNA insert that is repeated.

M4 shows properties similar to both M6 and KH10. As is seen with KH10, there is a broad region of hybridization which is weighted towards the higher molecular weight RNA molecules. As with KH10, approximately 80-90% of the RNA in these hybrids is RNase sensitive, though it should be noted that the fraction of total mRNA which hybridizes to M4 DNA is smaller than that complementary to KH10 DNA. In the region of the gel corresponding to lower molecular weight RNA, there is a sharp peak of hybridization which is relatively resistant to ribonuclease attack, properties similar to that of RNA hybridized to M6 DNA. This second, low molecular weight peak represents a specific mRNA whose sequence is encoded by clone M4, (see below) and we suggest that the heterogeneous hybridization is due to a sequence common to many messages as explained above for KH10. In all cases, the amount and size of the RNA hybridizing could not be competed by the addition of unlabeled poly(A), indicating that the hybridization is not to the poly(dT) linkers used to construct the plasmids.

A second approach to the characterization of messenger RNA complementary to the cloned DNAs yields results which are consistent with the above observations. [³²P] pulse-labeled poly (A)⁺ RNA was hybridized to filters separately affixed with M4, M6 and KH10 DNA. After extensive washing, the RNA was eluted and size fractionated on high resolution urea polyacrylamide gels (see Legend to Fig. 2) or agarose gels containing methyl mercury (14; data not shown). RNA complementary to M4 contains a specific band with a mobility indicating that the message is 800-900 nucleotides long. In addition, there is a broad heterogeneous sized population of RNA eluted from M4 DNA filters which represents the mRNA molecules carrying a common sequence complementary to M4 DNA. KH10

Figure 2. Size Fractionation of Plasmid Complementary RNA. [^{32}P] *in vivo* labeled poly(A)⁺ RNA was hybridized to M4, M6 and KH10 filters in 50% formamide, 0.72 M NaCl, 10 mM EDTA, pH 7.2, 0.2% SDS, and 1 mg/ml poly(A) for 2 days at 37°C. After extensive washing with hybridization buffer lacking poly(A), the filters were washed twice in 0.045 M NaCl at 55°C, and the RNA was eluted with 70% formamide, 5 mM EDTA at 55°C. Following ethanol precipitation the RNA fractions were sized by electrophoresis on 3.5% polyacrylamide gels in 7 M urea (33) with rRNA as markers on parallel slots (not shown).

shows no specific bands in most experiments, but there is a broad background of message throughout the gel which represents mRNAs with a common sequence complementary to a sequence in KH10 DNA. We would suggest that if KH10 does contain a coding region for a specific message, the message is not abundant, since a species representing over 0.1% of the poly(A)⁺ RNA of any given size should be easily visible over the heterogeneous "background" hybridization. M6 complementary RNA shows two bands in urea polyacrylamide or methyl mercury containing agarose gels not resolved on the 3.6% polyacrylamide-formamide gels shown in Fig. 1, which migrate with a mobility indicating molecular weights of approximately 1250 and 1350 nucleotides (13). There is negligible background on the M6 RNA slot, which indicates that the heterogeneous hybridization phenomenon which is observed for M4 and KH10 does not occur with M6 DNA. Recombinant plasmids containing cDNA complementary to single-copy genes hybridize only a single RNA species when analyzed in this manner, indicating again that the heterogeneous pattern of hybridization is not simply a function of the nature of this experiment (Rowekamp and

Firtel, unpublished observations).

Stability of RNA: Plasmids DNA hybrids. To determine the stability of RNA: plasmid DNA hybrids, thermal melting profiles of RNA hybridized to cloned DNA were performed. DNA filters carrying M4, M6 and KH10 DNA were hybridized to RNA which had been size fractionated on formamide-polacrylamide gels as described in Fig. 1. After extensive washing, the filters corresponding to the peaks of hybridization were pooled, and the melting profiles were performed. The results are shown in Fig. 3. For the analysis, M4 hybrids were divided in two fractions: M4-I filters, which hybridized to the higher molecular RNA heterogeneous population of RNA, and M4-II filters, which hybridized to the more discrete, lower molecular weight peak of RNA. Both M4-I and KH10 show a very broad melting profile with a low T_m , while M4-II and M6 show a much higher T_m ,

Figure 3. Melting Profile of poly(A)⁺ RNA Hybridized to M4, M6 and KH10 DNA [32P] poly(A)⁺ RNA was size-fractionated as described in the Legend to Fig. 1 and hybridized to M4, M6 and KH10 DNA filters. After extensive washing, the filters showing hybridization were combined and incubated in 0.03 M phosphate buffer (PB) at increasing temperatures. The counts eluted at each temperature were measured, and the results are plotted as the fraction of RNA melted from the hybrid at each temperature. For plasmid M4 two fractions were taken: fraction I contained filters hybridizing to the higher molecular weight heterogeneous population of RNA; fraction II contained filters hybridizing to the discrete, lower molecular weight peak of RNA plus to the lower molecular weight heterogeneous fraction.

consistent with the notion that these two clones code for a specific message. Since length of the hybrid region over 100 bp has a relatively minor effect on T_m (15), it is likely that the 20° difference in T_m between mRNA hybrids with M6 and M4-II and those with KH10 and M4-I is due to significant base mismatch in the latter, on the order of 15-20% (16). M4-II also shows a component that melts at a low T_m , which suggests that some of the counts hybridized from this region of the gel are probably due to the heterogeneous type of hybridization observed in fraction M4-I.

Restriction endonuclease maps. The location of restriction enzyme specific sites in M4 and KH10 was determined so that detailed sequence organization of the cloned fragments could be studied. Products of single and multiple restriction enzyme digestions of plasmid DNA were size fractionated on agarose and polyacrylamide gels and compared to known standards for molecular weight determination. The restriction maps of M4 and KH10 are shown in Fig. 4.

To locate the region of M4 complementary to the 0.8-0.9kb mRNA, the DNA blots were hybridized to [32 P] labeled poly(A)⁺ RNA, washed, and then autoradiographed before and after treatment with RNase. Hybrid-

Figure 4. Restriction Endonuclease Map of Plasmids M4 and KH10.

The sites of several restriction endonuclease digestion are shown; fragment lengths are indicated in kilobases (1 kb = 1000 nucleotide pairs). In both the maps of M4 and KH10, a deletion of approximately 200-300 nucleotides was observed between the Eco RI site of insertion of the *Dictyostelium* and the Hind III site in pMB9. This deletion was confirmed by heteroduplex mapping in M4 (Bender, unpublished observations) and is also seen in M6 (12). ■ represents poly(dA):poly(dT) regions used to join pMB9 to *Dictyostelium* DNA.

zation to both the 1.6 and 1.7kb Hind III-Hinc II-Hae III fragments showed only a small change in intensity of hybridization after RNase treatment and were taken as hybridizing to the 0.8-0.9kb RNA. These two fragments are adjacent and one includes the poly(dA):poly(dT) connecting region. Since the distance between the Hinc II site which bisects the mRNA complementary region and the poly(dA):poly(dT) tail is larger than the message which is complementary to M4, we suggest that the entire RNA-complementary region is included in the cloned DNA. There exists the possibility that M4 contains a message whose coding sequence is interrupted; however, the RNase resistant mRNA from an M4 DNA-RNA hybrid co-migrates with the RNA eluted prior to RNase treatment (data not shown). Thus, M4 DNA can protect the entire mRNA from ribonuclease digestion, suggesting again that the entire mRNA coding region is included in the plasmid. Other results (17) show that the sequence which hybridizes the complex mRNA population is located elsewhere in the M4 insert in the 0.45kb Hind III-Eco RI fragment as indicated in Fig. 4.

Presence of repeat and single-copy sequences in M4 and KH10 DNA. To determine whether plasmid M4 and KH10 DNA carry sequences which are repeated in the Dictyostelium genome, [^{32}P]-nick translated DNA from M4 and KH10 was hybridized to Southern DNA blot filters (18) carrying Dictyostelium DNA which had been digested with various restriction endonucleases and size fractionated on agarose gels. The filters were washed at a stringent criterion and then autoradiographed. The results are shown in Fig. 5. Both plasmids hybridize to a large number of discrete restriction fragments and therefore contain sequences which are reiterated in the Dictyostelium genome. As we will show below, the majority of the insert consists of single-copy DNA regions; thus, the repeated sequence must be relatively short. In the case of M4, 1-2 darker bands are present in each lane; these probably represent the single-copy sequences homologous to this clone. When cRNA complementary to M4 was used as the hybridization probe and the unhybridized tails were removed with RNase A digestion, only the one or two discrete fragments which hybridized over the heterogenous background using a nick-translated probe are observed (data not shown). This result suggests that the majority of the M4 insert is complementary to single-copy sequences.

In order to quantitate the amount of single-copy and repeated DNA in the Dictyostelium DNA insert of the recombinant plasmids, DNA excess hybridization kinetics experiments were performed. Nick-translated M4 and KH10 plasmid DNAs were hybridized to a vast excess of Dictyostelium DNA. As is shown

Figure 5. Hybridization of Nick Translated DNA from M4 and KH10 to *Dictyostelium* DNA Southern Filters. *Dictyostelium* DNA was digested with restriction endonucleases, size fractionated on agarose gels and transferred to nitrocellulose filters by the method of Southern (18). These were prehybridized in 0.24 M PB, 2x Denhardt's solution (34) at 60°C overnight and then hybridized with [³²P] nick translated DNA made from total M4 or KH10 DNA or specific restriction fragments in 0.24 M PB, 2x Denhardt's solution, 0.2% SDS at 62°C for 2 days. Washing was done in 0.12 M PB at 62°C for a number of hours with approximately six changes of buffer. Autoradiograms were exposed using preflashed Kodak RP x-omat film with Kodak intensifying screens at -70°C (35).

A. Total M4 plasmid DNA; B. Total KH10 plasmid DNA; C. 650 bp Mbo II fragment from KH10 (complementary to mRNA); D. 350 bp Mbo II fragment from KH10 (complementary to mRNA); E. 1.5 kb Hap II fragment from KH10 (not complementary to mRNA); F. 3.4 kb Hap II fragment from KH10 (complementary to mRNA).

in Fig. 6, the majority of the M4 insert hybridizes with kinetics showing a transition which coincides with *Dictyostelium* single-copy DNA (observed $Cot_{1/2}$ of 120 M-sec; 3,4). In addition, a small fraction (~10%) of the insert hybridizes with kinetics indicating that it represents a short sequence

Figure 6. Hybridization of Nick Translated DNA from Plasmid M4 and KH10 to a Vast Excess of *Dictyostelium* DNA.

A. *Dictyostelium* DNA and M4 plasmid DNA were nick translated separately using [³H] and [³²P] deoxyribonucleotide triphosphates, respectively. These were hybridized with a vast excess of *Dictyostelium* DNA in solution under conditions such that the nick translated probe did not renature in a parallel reaction without added *Dictyostelium* DNA. The driver DNA was present in a 100-fold single-copy sequence excess over the M4 tracer. The reaction was stopped and samples were assayed by S1 nuclease digestion as described in Kindle and Firtel (10). All data were corrected as described (36, 37) for the S1 nuclease sensitivity of single-stranded tails on hybrids, which are the inevitable result of renaturing randomly broken nucleic acids. The correction takes the form: $(C/C_0)^{2.27} = (1 + kC_0t)^{-1}$. Since approximately 50% of the plasmid DNA is *Dictyostelium* insert, the maximum level of hybridization of the M4 probe to *Dictyostelium* DNA will be approximately 50%.

B. The large *Hae* III fragment from plasmid KH10 DNA, which includes the entire *Dictyostelium* insert, was isolated, labeled with [³²P] by nick translation and hybridized and assayed as described above.

which is reiterated in the genome. Results using a KH10 plasmid fragment as probe show that the majority of this insert is single-copy, with a small fraction of the DNA (~15%) hybridizing to reiterated sequences. Since the rapidly hybridizing portion represents a small fraction of the probe, it is difficult to determine the exact repetition number of the repeats. However, from the general location of the transition, it is estimated that the repeat sequences are reiterated approximately 50-100 times per genome.

The following experiment was done to determine if the restriction fragments of KH10 which hybridize mRNA are the ones which are reiterated in the *Dictyostelium* genome. To map the location of sequences complementary to mRNA, KH10 plasmid DNA was digested with several restriction endonucleases, and the fragments were size fractionated on agarose gels and then transferred to nitrocellulose paper. The Southern blots were then hybridized to [³²P] poly(A)⁺ RNA to identify RNA complementary regions. As can be seen in Fig. 7, 1-2 restriction fragments from each digest hybridize poly(A)⁺ RNA. When these hybridization data are compared to the restriction map of KH10 (see Fig. 4) the region of hybridization can be localized around the three left hand Hinc II restriction sites in the 3.4 kb Hap II fragment. Although all the Mbo II sites in the insert have not been mapped, the 350 and 650bp fragments which hybridize mRNA have been mapped to the junction of the 2.1 kb Hap II-Hinc II and 1.1 kb Hinc II fragments both by using the procedure of Smith and Birnstiel (19) and by hybridizing the nick-translated probes from both fragments to DNA blot filters of Hap II and Hinc II digests of KH10 DNA (data not shown). Several mRNA complementary fragments (the 2.1 kb Hap II-Hinc II, 1.1 kb Hinc II, 3.4 kb Hap II plus two Mbo II fragments) and several fragments which do not hybridize mRNA (the 1.3 kb Hap II-Hind III, 1.5 kb Hap II plus two Mbo II fragments) were isolated from gels, labeled by nick-translation and hybridized to Southern DNA blots of *Dictyostelium* genomic DNA. The results (see Fig. 5) showed that all fragments which hybridize mRNA are complementary to a large number of genomic restriction fragments, indicating they are repeated in the genome. All insert fragments which do not hybridize mRNA hybridize to 1-2 genomic DNA fragments, indicating they are probably single-copy sequences (only some of data are shown). This result suggests that it is the repeated sequences in KH10 which are complementary to mRNA.

Properties of the KH10 complementary RNA. The data presented suggest that the majority of the RNA molecules complementary to KH10 DNA contain a sequence which is reiterated in the *Dictyostelium* genome. Moreover, 80-90% of the RNA-KH10 DNA hybrids are sensitive to RNase suggesting that only a

Figure 7. Localization of RNA Complementary Region on Plasmid KH10. Plasmid KH10 DNA was digested with various restriction endonucleases and the digests were size fractionated on 0.9 (data not shown) and 1.25% agarose gels. The DNA bands were visualized by staining with ethidium bromide and photographed using a short wave length UV transilluminator (UV Products). Southern DNA blot transfers were made onto nitrocellulose (18), and the filters were hybridized with 2×10^6 cpm of $[^{32}\text{P}]$ poly(A)⁺ RNA. The filters were washed and the restriction fragments which hybridize were identified by autoradiography. Hybridization conditions were performed as described previously (10) except that a 0.01 M Na⁺ wash at 37°C was used after extensive washing under hybridization conditions instead of the 45mM Na⁺ wash at 55°C. Hybridized fragments were visualized by autogradiography.

The DNA fragments were identified as complementary to RNA by comparisons of the photograph (see left side of figure) and autoradiograph (right side of figure). Eco RI-Hind III digests of λ DNA (see lane a), and Hae III digest of ϕ x174 were run on the gels for molecular weight markers. a. Eco RI - Hind III - λ DNA; b. Hae III - KH10 DNA; c. Hae III + Hap II - KH10 DNA; d. Hae III + Hap II + Hinc II + Hind III - KH10 DNA; e. Hap II + Hinc II + Hind III - KH10 DNA; f. Hap II + Hinc II + Eco RI - KH10 DNA; g. Hap II - KH10 DNA; h. Hap II + Eco RI - KH10 DNA; i. Hap II + Hind III - KH10 DNA; j. Mbo II - KH10 DNA; k. Hae III - ϕ x174 DNA.

small fraction of the RNA molecule forms a stable duplex and that the remainder is present as an unhybridized tail. To determine whether KH10 complementary poly(A)⁺ RNA is transcribed from solely reiterated sequences or consists of primarily single-copy sequences coupled to a short repeat sequence transcript, [³²P]-KH10 complementary RNA purified by two cycles of hybridization to KH10 filters was hybridized to sheared *Dictyostelium* genomic DNA and assayed by both hydroxylapatite (HAP) chromatography and RNase sensitivity (20, 5). The RNase assay measures the actual fraction of RNA nucleotides present in a hybrid at various Cots and therefore the fraction transcribed from single-copy and repeated sequences. The HAP assay measures the total fraction of RNA in which some portion of the RNA molecules is hybridized to a DNA molecule at a given Cot. Therefore, any sequence covalently attached to another sequence which forms an RNA:DNA hybrid will bind to HAP. If, as we suggest, the KH10 complementary RNA molecules contain both short repeat and longer single-copy DNA transcripts, they should bind to HAP when driven to the Cot at which repetitive DNA renatures; at this point the bulk of the RNA molecule would be a non-hybridized single strand tail sensitive to ribonuclease. That this is the case can be seen from Fig. 8; only a small fraction of the RNA nucleotides are RNase resistant at a Cot at which the reiterated DNA has renatured, while >70% of the RNA molecules bind to HAP at the same Cot. These results indicate that while >80% of the sequences in KH10 complementary RNA are single-copy transcripts, a majority, and most probably all, of the molecules contain a short sequence transcribed from moderately repeated DNA. When similar experiments are done with total *Dictyostelium* mRNA, it is estimated that ~10% of the sequences are transcribed from repeated DNA and that ~30% of the molecules contain either all repeat as actin mRNA or linked repeat and single-copy sequence transcripts (5). The RNA molecules complementary to KH10 thus make up a substantial portion of the class which contains repeat transcripts.

DISCUSSION

The working model suggested in this communication was prompted by several surprising observations. Two pieces of cloned *Dictyostelium* DNA immobilized on nitrocellulose filters hybridize a relatively large fraction of [³²P]-labeled poly(A)⁺ RNA, and this is represented by a heterogeneous population of molecules. Moreover, although both cloned DNAs contain repeated sequences according to Southern blot experiments, solution DNA excess hybridization

Figure 8. Linkage of Repeat and Single-Copy Sequences in KH10 Complementary RNA. Vegetative poly(A)⁺ RNA labeled *in vivo* for 3 hours was purified according to Firtel and Lodish (31). The KH10 complementary fraction was purified by two successive hybridization to nitrocellulose filters carrying KH10 plasmid DNA as described in the legend to Figure 1 except that hybridization was for 16 hours in 4x SSC plus 0.12M neutral sodium phosphate buffer (PB). The RNA was hybridized to a vast excess of sheared *Dictyostelium* genomic DNA in 0.4M PB, 0.2% SDS, 61°C at a DNA concentration of 100 µg/ml for low C_0t values and 1.6 mg/ml for higher C_0t values. Under these conditions, the majority of mRNA molecules remain intact (5). The time points were quick cooled, and divided into two parts. The first was diluted to give a final buffer concentration of 0.17M PB, 15% formamide, 0.8% SDS and passed over HAP columns at 42°C (5) to measure the fraction of total RNA molecules bound at a given C_0t . Under these conditions, 1.7% of the RNA bound at $C_0t=0$. The second aliquot was treated with 20 µg/ml RNase A and 10 units/ml RNase T₁ in 0.24M PB at room temperature. Fraction of RNA nucleotides in hybrid was determined by TCA precipitation. All curves are corrected for zero C_0t values. [³H] nick translated *Dictyostelium* nuclear DNA was added as a marker to assay HAP renaturation kinetics of driver DNA.

○—○, RNase assay Δ—Δ, ³H tracer DNA ●—●, HAP assay

experiments using the cloned DNA sequences as probe suggest that the majority of the insert is single-copy. When RNA complementary to one clone (KH10) is hybridized to an excess of *Dictyostelium* DNA, the kinetics of the reaction as assayed by RNase digestion suggest that it is transcribed largely from single-copy DNA. However, when the kinetics of hybrid formation are assayed by HAP, it is evident that a large fraction of the molecules contain a sequence which hybridizes to repeated DNA sequences. These data suggest that the majority

of KH10 complementary RNA molecules contain a short region transcribed from a sequence which is reiterated in the genome. This short repeat transcript is covalently linked to a longer single-copy sequence transcript.

It is important to consider whether the stringency of filter hybridizations done in cloned DNA excess is equivalent to hybridization done in solution in genomic DNA excess. The nature of the reactions is different, and as discussed elsewhere (21), the use of a low complexity non-renaturing driver (fixed to nitrocellulose filters) may allow the formation of poorly matched hybrids which might be displaced if better matched complementary sequences are available. However, two points argue for the specificity of the reaction between the complex class of mRNA and the sequences in M4 and KH10. First, the hybridization is sequence specific; many recombinant plasmids do not exhibit such behavior (M6 is an example) and in M4 and KH10 the hybridization is confined to a discrete sequence (see above; 17). Second, the results are not altered by changing the criterion of RNA-DNA hybridization in solution or on filters. We therefore feel that the hybridization of the complex class of RNA is due to a specific interaction between a repeat transcript on these molecules and one complementary family member on the cloned DNAs.

The data presented show that a relatively large population of poly(A)⁺ RNA (at least 2 and 9% respectively) hybridizes to the *Dictyostelium* DNA inserts in plasmids M4 and KH10. Sizing of the complementary RNA shows that the population is very heterogeneous in molecular weight except for a discrete 0.8-0.9 kb mRNA complementary to and encoded by M4. Analysis of the hybrids by melting profiles suggests that they contain considerable base mismatch. The melting profile combined with the fact that approximately 80-90% of the hybridizing RNA is RNase sensitive suggests that the length of hybridization may be short. An alternative possibility is that the mRNA hybridizes along an extended region but that there is extensive base mismatch. This possibility is difficult to rule out; however, several points argue against it. When KH10 complementary RNA is nicked in alkali (average 4 nicks per molecule), only 20-25% of the RNA hybridizes back to KH10 DNA in comparison to 70% rehybridization by unnicked RNA (Firtel and Kindle, unpublished observations). This molecular weight dependence suggests that a specific, short region of the RNA is complementary to the plasmid DNA. In the case of M4, Kimmel and Firtel (17) have shown that the repeat which hybridizes mRNA is short (<450 base-pairs long). Hybridization of [³²P]-RNA to DNA blot filters of restricted, size-fractionated KH10 DNA indicates that

the mRNA is complementary to a short region of the insert carrying the repeat sequence.

Both plasmids contain short regions, comprising less than 15% of the *Dictyostelium* DNA insert, which are complementary to sequences which are reiterated in the genome (see Fig. 6). A careful analysis of the Southern DNA blot data indicates that the repeat pattern observed for the two plasmids is different and that the repetition frequency is probably approximately 100, in the range estimated from the DNA excess hybridization. The molecular weight range of the restriction fragments which hybridize M4 and KH10 probe is quite large suggesting the repeated sequence is topologically dispersed in the genome.

As described above, the mRNA complementary to the two plasmids is heterogeneous in size. We expect that most of the individual species of RNA molecules complementary to the repeat sequences in M4 and KH10 are present at a relatively low abundance in total mRNA since a population of 10 moderately abundant species (0.1% of mRNA each) would be seen as discrete RNA bands on CH₃Hg-OH agarose gels. It might be expected that the RNA complementary to M4 and KH10 consists of a population of 50-200 different mRNAs each with an abundance of 0.01-0.03% of the total mRNA. From the DNA excess hybridization using KH10 complementary RNA as probe, we conclude that the majority of the RNA is transcribed from single-copy sequences.

From these results, we suggest that individual members of a short interspersed repeat family lie adjacent to or near (if splicing occurs during hnRNA processing) a population of different single-copy genes. The mRNA population transcribed from these genes consists of a short common sequence transcribed from a member of the repeat family covalently associated with a single-copy sequence transcript. Since the bulk of the mRNA complementary to KH10 (and presumably M4) is transcribed from single-copy DNA and since it is represented equally in polysomal and total cytoplasmic poly(A)⁺ RNA (data not shown), we assume that the single-copy sequences code for specific proteins.

Confirmation of this model has come from an in-depth study of M4 (17). Adjacent to the ~450 base pair repeat is a single-copy region which is transcribed at a low level; evidence suggests that these transcripts include sequences complementary to the repeat. We have, at present, no evidence for a single-copy "gene" adjacent to the repeat in KH10, but the use of KH10 as a probe for a specific subpopulation of mRNA has shown that there are KH10 repeat family members which are transcribed adjacent to single-copy sequences.

Of great interest is the prevalence of this type of arrangement for structural genes in Dictyostelium. The original screening of recombinant plasmids was done by hybridizing with in vivo labeled poly(A)⁺ RNA, and the filters were not treated with ribonuclease; such a screen should be efficient in selecting clones with the properties we have described. In fact, three out of four clones initially tested showed this hybridization to a heterogeneous population of RNA. From earlier studies, we estimated that 30% of poly(A)⁺ RNA contains both repeat and single copy transcripts(5). Together M4 and KH10 hybridize some 11-12% of poly(A)⁺ whole cell or polysomal RNA.

It should be pointed out that M4 and KH10 hybridize different populations of RNA. The inserts of the plasmids themselves do not cross-hybridize (data not shown). When RNA complementary to M4 and KH10 is prepared and hybridized to nitrocellulose filters affixed with DNA from KH10 and M4 respectively, the hybridization is slightly above background (21), suggesting that M4 complementary RNA is not a subset of KH10 complementary RNA. There could, however, be a small fraction of molecules which contain sequences complementary to both plasmids. Thus, it seems likely that the observation of a family of transcribed repeat sequences does not represent a single case, but that there may be a number of such families in Dictyostelium.

There have been numerous reports of linked repeat and single-copy transcripts in hnRNA (20, 22, 23, 24), but the case for mRNA is much less clear. In mammalian cells and in sea urchin gastrular mRNA, it appears that the single-copy transcripts do not contain covalently associated repeat transcripts (25, 26, 27); however, evidence has also been reported for such transcripts in Xenopus embryos (28, 29). The observations reported here pertain to nuclear, cytoplasmic and polysomal poly(A)⁺ RNA preparations; any contamination of the latter preparations by nuclear RNA species could not account for the large fraction of poly(A)⁺ RNA which is involved.

That this arrangement of repeat and single-copy sequences is not universal for transcription units of structural genes in Dictyostelium is also clear. Actin mRNA represents a transcript from a repeated "gene" which is flanked by non-reiterated DNA (10). Kimmel and Firtel have demonstrated that the 0.8-0.9 kb mRNA complementary to M4 is transcribed from a single-copy region with no flanking repeat (17). Thus, there seem to be several sequence organization alternatives for transcription units in Dictyostelium.

The function of transcribed repeated sequences is unknown. One possibility would be a common regulatory element for a population of genes transcribed at the same rate or coordinately regulated during development. Such a model has been proposed by Britten and Davidson (30). Direct evidence for such possibilities does not exist at present. It is interesting that KH10 hybridizes significantly more mRNA than does M4. Although accurate repetition frequencies for these two repeats are not known, the data suggest that both are repeated approximately 100 fold in the genome. There are a number of explanations; one intriguing possibility is that the repetitive element might be important in controlling the rate of transcription or otherwise determining the cellular abundance of the associated messages.

Previous results suggested that in hnRNA repeat transcripts are covalently associated with single-copy transcripts and localized toward the 5' end (31, 5). We suggest that the repeat transcript may be similarly localized at the 5' end of a predominantly single-copy transcript in mRNA. We do not know whether the repeated sequences are directly linked to the single-copy sequences in the genome or whether their location in the mRNA results from splicing during hnRNA processing. Since the average size of Dictyostelium hnRNA is only 25% larger than mRNA, a large excised intervening sequence in the hnRNA is unlikely.

Further analysis of the structure and transcription of these clones during development as well as other plasmids carrying different members of the same repeat families will lead to an understanding of the organization of these genes and may shed light on the function of the repeat sequences.

ACKNOWLEDGEMENTS

We would like to thank A. Kimmel, A. Lee, and C. Lai for useful discussions and for critical readings of the manuscript. We would also like to thank C. Lai for allowing us to use unpublished data on the mapping of some of the restriction sites in plasmid KH10. We would like to thank Robin Timm for technical assistance.

K.L.K. was the recipient of an N.S.F. predoctoral Fellowship and was also supported by an N.I.H. predoctoral Training Grant. R.A.F. was the recipient of a Faculty Research Award from the American Cancer Society. This research was supported by grants from the N.S.F., the N.I.H. and the American Cancer Society to R.A.F.

REFERENCES

1. Loomis, W.F. (1975) D. discoideum: A Developmental System. (New York: Academic Press).
2. Sussman, R.R. and Rayner, E.P. (1971) Arch. Biochem and Biophys. 144, 127.
3. Firtel, R.A. and Bonner, J. (1972) J. Mol. Biol. 66, 339.
4. Firtel, R.A. and Kindle, K.L. (1975) Cell 5, 401.
5. Firtel, R.A., Kindle, K.L. and Huxley, M.P. (1976) Federation Proc. 35, 13.
6. Rodriguez, R.L., Bolivar, F., Goodman, H.M., Boyer, H.W., and Betlach, M.C. (1976) Molec. Mech. in the Control of Gene Expression, D.P. Nierlich, W.J. Rudder, C.F. Fox, eds. (New York: Academic Press), p.471.
7. Lobban, P.E. and Kaiser, A.D. (1973) J. Mol. Biol. 78, 453.
8. Jackson, D., Symons, R. and Berg, P. (1972) Proc. Nat. Acad. Sci. USA 69, 2904.
9. Cockburn, A.F., Newkirk, M.J. and Firtel, R.A. (1976) Cell 9, 605.
10. Kindle, K.L. and Firtel, R.A. (1978) Cell 15, 763.
11. Grunstein, M. and Hogness, D.S. (1975) Proc. Nat. Acad. Sci. USA 72, 3961.
12. Bender, W., Davidson, N., Kindle, K.L., Taylor, W.C., Silverman, M. and Firtel, R.A. (1978) Cell 15, 779.
13. McKeown, M., Taylor, W.C., Kindle, K.L., Firtel, R.A., Bender, W. and Davidson, N. (1978) Cell 15, 789.
14. Bailey, J.M. and Davidson, N. (1976) Anal. Biochem. 70, 75.
15. Hayes, F.N., Lilly, E.H., Ratliff, R.L., Smith, D.A. and Williams, D.L. (1971) Biopolymers 9, 1105.
16. Laird, C.D., McConaughy, B.L. and McCarthy, B.J. (1969) Nature 224, 149.
17. Kimmel, A. and Firtel, R.A. (1979) Cell in press.
18. Southern, E.M. (1975) J. Mol. Biol. 98, 503.
19. Smith, H.O. and Birnstiel, M.L. (1976). Nucleic Acids Res. 3, 2387.
20. Smith, M.J., Hough, B.R., Chamberlin, M.E. and Davidson, E.H. (1974) J. Mol. Biol. 85, 103.
21. Kindle, K.L. (1978) Ph.D. Thesis, Univ. of California at San Diego, La Jolla, California.
22. Darnell, J.E. and Balint, R. (1970) J. Cell Physiol. 76, 349.
23. Spradling, A., Penman, S., Campo, M. and Bishop, J.O. (1974) Cell 3, 23.
24. Ryffel, G.U. and McCarthy, B.J. (1975) Biochem. 14, 1385.
25. Campo, M.S. and Bishop, J.O. (1974) J. Mol. Biol. 90, 649.
26. Goldberg, R.B., Galau, G.A., Britten, R.J., and Davidson, E.H. (1973) Proc. Nat. Acad. Sci. USA 70, 3516.
27. Klein, W.H., Murphy, W., Attardi, G., Britten, R.J., and Davidson, E.H. (1974) Proc. Nat. Acad. Sci. USA 71, 785.
28. Dina, D., Crippa, M., and Beccari, E. (1973) Nature (London) New Biology 242, 101.
29. Dinak D., Meza, I., and Crippa, M. (1974) Nature 248, 486.
30. Britten, R.J. and Davidson, E.H. (1969) Science 165, 349.
31. Firtel, R.A. and Lodish, H.F. (1973) J. Mol. Biol. 79, 295.
32. Duesberg, P.H. and Vogt, P.K. (1973) J. Virol. 12, 594.
33. Spradling, A., Pardue, M.L. and Penman, S. (1977) J. Mol. Biol. 109, 559.
34. Denhardt, D. (1966) Biochem. Biophys. Res. Commun. 23, 640.
35. Laskey, R.A. and Mills, A.D. (1977) FEBS Letters 82, 314.
36. Morrow, J.F. (1974) Ph.D. Thesis, Stanford University, Stanford, California
37. Britten, R.J. and Davidson, E.H. (1976) Proc. Nat. Acad. Sci. USA 73, 415.