The EMBO Journal vol.5 no.9 pp.2115-2121, 1986

Structure of DNA formed in the first step of CAD gene amplification

Elena Giulotto[1], Izumu Saito and George R.Stark

Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK

[1]Present address: Dipartimento di Genetica e Microbiologia, Via S.Epifanio 14 — 27100 Pavia, Italy

Communicated by G.Stark

Thirty-three independent mutant cell lines were selected in single steps for resistance to low concentrations of N-(phosphonacetyl)-L-aspartate and the structure of their amplified DNA was probed, using a set of recombinant phage and cosmids containing a total of 380 kb of amplified DNA. In all 33 cell lines, the selected CAD gene and at least 65 kb of flanking DNA were amplified, an average of 2.6-fold. Six other regions of DNA were co-amplified in all 33 mutants, but sometimes to a different extent than CAD. Novel joints, marking recombinations which link amplified regions to each other, were found surprisingly rarely. There were only three within the 380 kb of DNA sequence examined in the total of 33 cell lines. Each novel joint was present in only one copy per cell, was found in a different cell line and was homologous to a different probe. The low frequency of novel joints is consistent either with very large amplified regions in the single-step mutants, possibly 10 000 kb of co-amplified DNA for each copy of the CAD gene, or with a strong bias against recombination in the cloned sequences used as probes. Our previous finding that CAD probes hybridize *in situ* to unusually large chromosome arms in several single-step mutants is most consistent with the first possibility.

Key words: CAD/gene amplification/novel joints/co-amplified DNA/first-step mutants

Introduction

Gene amplification has a role in the normal development of lower eukaryotes and in the abnormal expression of genes related to drug resistance and malignancy in many organisms (Cowell, 1982; Hamlin *et al.*, 1984; Schimke, 1984; Stark and Wahl, 1984). Mammalian cells become resistant to N-(phosphonacetyl)-L-aspartate (PALA), a specific inhibitor of aspartate transcarbamylase, through amplification of the gene for the trifunctional protein CAD, which catalyzes the first three reactions of pyrimidine nucleotide biosynthesis (Kempe *et al.*, 1976; Wahl *et al.*, 1979a). In BHK cells selected in several steps for resistance to high concentrations of PALA, about 1000 kb of DNA is co-amplified along with the CAD gene, which is itself about 25 kb long (Wahl *et al.*, 1982). By *in situ* hybridization, the amplified DNA is usually found in a single expanded region within only one of the two homologous chromosomes, often near the normal site of the CAD gene (Wahl *et al.*, 1982). In mouse cells resistant to methotrexate, the amount of amplified DNA is also much larger than the dihydrofolate reductase (DHFR) gene, and is often located on double-minute chromosomes (Kaufman *et al.*, 1979).

The structure and organization of amplified DNA in highly resistant cells, selected in several discrete steps, has been analyzed in several laboratories. Caizzi and Bostock (1982) cloned amplified DNA from methotrexate-resistant mouse cells and showed that different sequences are co-amplified with the DHFR gene in independently selected cell lines. Using the differential method of Brison *et al.* (1982), Ardeshir *et al.* (1983) cloned 230 kb of amplified DNA from two BHK cell lines resistant to high concentrations of PALA. Some of the cloned probes hybridized to novel restriction fragments present in only one of the several mutant cell lines analyzed and not present in wild-type DNA. These novel fragments probably include junctions between different amplified regions arranged linearly within the chromosomes. Several major conclusions concerning the structure of highly amplified DNA were drawn from this work; amplified regions are not comprised of identical units, different sequences can be co-amplified with CAD in different independent events, novel joints are unique to each cell and 'hot spots' for rearrangement within amplified DNA are not evident. Federspiel *et al.* (1984) reached similar conclusions, using chromosome 'walking' to clone and analyze 240 kb of DNA contiguous with the DHFR genes from mouse cells resistant to methotrexate.

A high degree of amplification is achieved only in multiple-step mutants which have experienced several independent amplification events and which have been grown in culture for many months. Thus, the structure of amplified DNA in these mutants may be complex because the consequences of several independent amplification events have been superimposed, and possibly also because of secondary rearrangements of the DNA during growth of the cells. To learn about the structure of DNA generated during the primary event of amplification, in which one copy of the CAD gene is increased several-fold within a single chromosome, Zieg *et al.* (1983) isolated newly-formed, independent colonies of PALA-resistant cells ('single-step mutants'), using low concentrations of PALA. Analysis indicated that the organization of the amplified DNA was complex even in single-step mutants, and no novel joints were found in the DNA of five such mutants, within a total of about 150 kb of amplified DNA. We have now extended this analysis by examining about 380 kb of amplified DNA in 33 single-step mutants and have found three novel joints. Our results provide new evidence concerning the copy number and distribution of novel joints in the single-step mutants.

Results

Single-step mutants and probes for co-amplified DNA

Thirty-three mutants were isolated at PALA concentrations ranging from 15 to 60 μM. Eight of these mutants had been isolated previously by Zieg *et al.* (1983) and 25 were newly isolated. The PALA concentrations used were slightly higher than the minimum amount necessary to give selection ($\sim$6 μM). The selection scheme employed ensured that each mutant was independent and that it arose just before selection was applied. Separate groups of 10^3 WT-BHK-2 cells were expanded to about 5×10^6 cells each in non-selective medium, re-dispersed and selected with PALA. Mutants which arise independently during the last few

Table I. Probes for investigating the structure of DNA co-amplified with the CAD gene

Probe	Size of insert in the probe (kb)	Number of *Eco*RI fragments in the probe	Size of homologous *Eco*RI fragments in normal genomic DNA (kb)	Number of *Eco*RI fragments in normal genomic DNA	Size of homologous novel fragments in single-step mutants (kb)
cosC31[a]	36	10	50	10	
λ11-200[b]	18	2	18	2	10 (in mutant 15.3)
λ11-600[b]	11	1	11	1	
λ11-1700[b]	18	5	28	6	
λ11-1800[b]	14	2	14	2	15 (in mutant 30.7)
λ11-1900[b]	17	1	33	2	
Total	114		154		

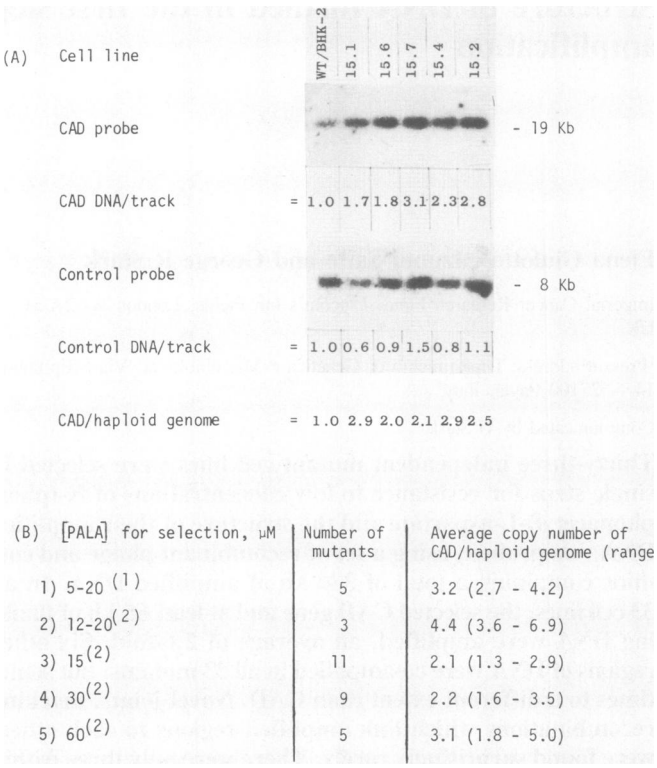
[a]CosC31 contains about 10 kb of DNA from the transcribed region of CAD and 25 kb adjacent to the 5′ end of the gene. The size of the insert in cosC31 is different from the size of the genomic fragment because it was isolated from a library containing *Mbo*I inserts and is homologous to *Eco*RI genomic fragments outside the borders of the 36 kb *Mbo*I insert. The total size of *Eco*RI genomic fragments sharing homology with cosC31 is 50 kb.
[b]The five λ probes were obtained from a Charon 4 library with *Eco*RI inserts, derived from the highly amplified cell line B5-4 (Ardeshir *et al.*, 1983). Probes 11-1700 and 11-1900 contain amplified novel joints from B5-4. Thus each is homologous to two novel fragments in the other cell lines. This explains why the total size of the two homologous normal fragments is larger than the size of the novel fragment in the phage.

generations of growth will give only one or few colonies, whereas any mutant present in the original group of 10^3 cells will give many resistant colonies. Mutants picked from plates with <10 colonies were grown to populations of about 10^7 cells in selective medium and analyzed after two generations of growth without PALA.

As shown in Table I, five phage-containing DNA fragments co-amplified with CAD in the multiple-step mutant B5-4 (Ardeshir *et al.*, 1983) were used to probe the genomic DNAs of the 33 single-step mutants. Two of the fragments used as probes contained amplified novel joints from B5-4. Thus, each of these novel fragments is homologous to two DNA fragments in wild-type DNA, so that the five probes detect seven different fragments in wild-type DNA. These seven fragments are not linked to each other or to the CAD gene (see below). All seven fragments were co-amplified with CAD in the single-step mutants studied previously by Zieg *et al.* (1983), and also in the 25 new mutants (see below). We also used as a probe a cosmid containing part of the CAD gene, plus additional 5′-proximal DNA. As shown in Table I, a total of about 130 kb of wild-type DNA sequences external to the CAD gene were homologous to this initial set of probes.

Copy number of the CAD gene and co-amplified DNA in single-step mutants

Southern transfers of genomic DNAs from the 33 single-step mutants were probed with pCAD 142, a 6.5 kb cDNA (Shigesada *et al.*, 1985), and with a single-copy DNA probe (C6-1) to normalize the amount of DNA in each track. Five examples are shown in Figure 1, together with the results of densitometric analysis and a summary of all the data. More than one copy of the CAD gene per haploid genome was detected in every mutant studied. Thus, gene amplification remains the only mechanism

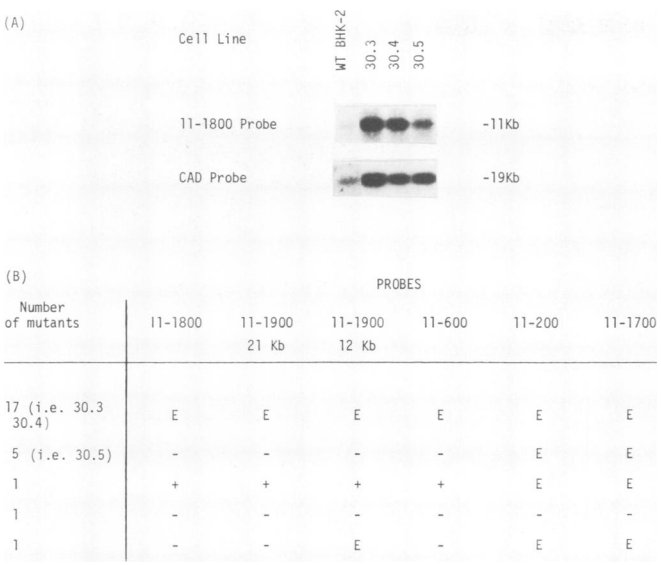


(B) [PALA] for selection, μM	Number of mutants	Average copy number of CAD/haploid genome (range)
1) 5-20 (1)	5	3.2 (2.7 - 4.2)
2) 12-20 (2)	3	4.4 (3.6 - 6.9)
3) 15 (2)	11	2.1 (1.3 - 2.9)
4) 30 (2)	9	2.2 (1.6 - 3.5)
5) 60 (2)	5	3.1 (1.8 - 5.0)

Fig. 1. Copy number of CAD in single-step mutants. **(A)** Hybridization of ^{32}P-labelled CAD cDNA and a control probe to the same transfer of *Eco*RI-digested genomic DNA from a PALA-sensitive parental line (WT-BHK-2) and from five lines resistant to 15 μM PALA. The concentration of DNA was determined by ethidium bromide staining and 5 μg of digested DNA was loaded per track. The amount of DNA available for hybridization was determined by densitometric analysis of the 8 kb band homologous to the control probe. The amount of CAD DNA per track, determined by densitometric analysis of the 19 kb band homologous to the cDNA probe, was normalized to the amount of control DNA to give the final value of CAD sequences per haploid genome. **(B)** Average copy number of CAD per haploid genome in 33 independent single-step mutants selected from WT-BHK-1 (group 1) or WT-BHK-2 (groups 2–5). The mutants in groups 1 and 2 were obtained previously by Zieg *et al.* (1983).

we have found for resistance to PALA. The average copy number of CAD was 2.6 per haploid genome, or about 5 per cell. Since single-step amplification appears to occur within only one of the two homologous chromosomes (Zieg *et al.*, 1983), one copy of the CAD gene is increased to an average of about four copies within this chromosome. The copy number of CAD was determined again for the eight single-step mutants previously analyzed by Zieg *et al.* (1983). The new values were lower, probably because a more precise method has now been used to analyze the intensity of each band and to correct for differences in the amount of DNA in each track.

The Southern transfers used for the experiments of Figure 1 were also probed with the DNA fragments listed in Table I. These fragments were also hybridized to new Southern transfers from 14 of the mutants after digestion of the DNAs with *Bam*HI or *Hin*dIII. The 25 kb of DNA contiguous with the 5′ end of the CAD gene were always co-amplified to the same extent as the gene itself. However, as shown in Figure 2 for the example of the 11-1800 fragment, the copy number of CAD was sometimes different from the copy number of the co-amplified unlinked fragments. The results for the 25 new mutants are summarized in Figure 2. For 17 of these, all the sequences tested were amplified



(B)

Number of mutants	PROBES 11-1800	11-1900 21 Kb	11-1900 12 Kb	11-600	11-200	11-1700
17 (i.e. 30.3 30.4)	E	E	E	E	E	E
5 (i.e. 30.5)	-	-	-	-	E	E
1	+	+	+	+	E	E
1	-	-	-	-	-	-
1	-	-	E	-	E	E

Fig. 2. Copy number of amplified fragments in single-step mutants. **(A)** The same genomic transfers of *Eco*RI-digested DNA were hybridized with probes pCAD 142 and 11-1800. In mutant 30.5, the 11 kb fragment homologous to the 11-1800 probe was amplified less than the 19 kb fragment homologous to part of the CAD gene. **(B)** The degree of amplification relative to CAD of the fragments homologous to the six probes was estimated by eye, after hybridization of the same set of Southern transfers with all the probes. (E), amplification approximately equal to CAD; (−), amplification less than CAD; (+), amplification more than CAD. Only the 25 mutants isolated in the present work were analyzed. Both the 11-1700 and 11-1900 phage contain novel joints but the amount of genomic DNA on one side of the 11-1700 joint is only about 200 bases, too small to determine the copy number conveniently.

to the same extent as CAD. For seven mutants, some of the sequences were amplified less than CAD. In only one mutant were some of the sequences amplified more than CAD.

Novel joints

The amplified novel fragments from the multiple-step mutant B5-4 (11-1700 and 11-1900 in Table I) did not hybridize to a novel fragment in any of the 33 single-step mutants. With the remaining probes, we found only two novel fragments in all the single-step mutants. Probe 11-1800 hybridized to a 15 kb *Eco*RI fragment only in mutant 30.7 and probe 11-200 hybridized to a 9 kb *Eco*RI fragment only in mutant 15.3. Because these novel fragments were not cloned, we did not know how much sequence they had in common with the probes and, therefore, we could not estimate their copy number from the intensities of bands in the autoradiograms. We determined the copy number of the novel fragment from mutant 30.7 by using as a probe a sub-fragment contained wholly within both the phage and the novel fragments, as illustrated and explained in Figure 3. From these data, we can conclude that the novel fragment is present in only one copy per diploid genome. A similar experiment was used to show that the novel fragment homologous to the 11-200 probe in mutant 15.3 is also present in only one copy per diploid genome (data not shown).

Isolation and use of cosmids containing amplified DNA

We have attempted to map the seven different regions of co-amplified wild-type DNA represented by the five phage clones of Table I with respect to each other and with respect to the region of the CAD gene. To do this, we prepared a cosmid library of genomic DNA from the single-step mutant 10.1 and screened

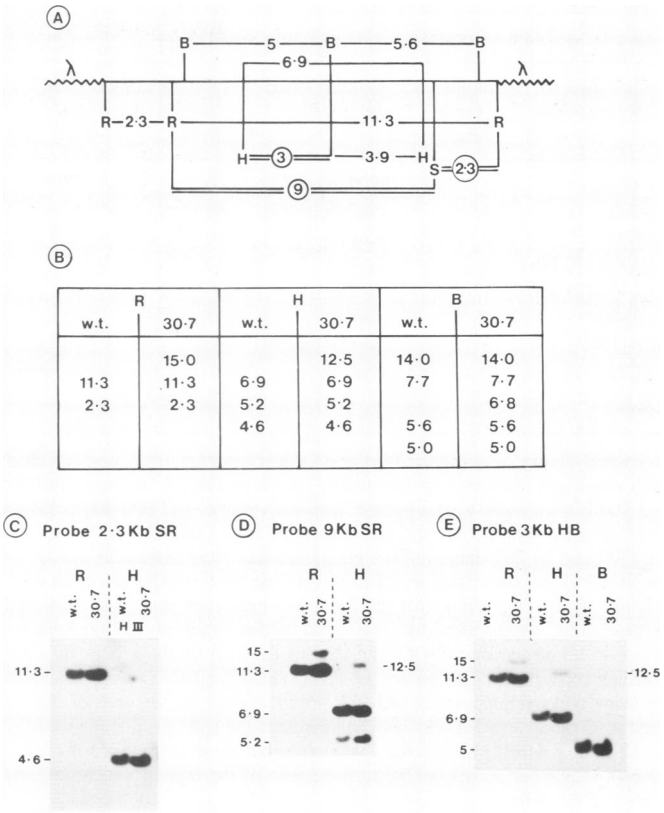


Fig. 3. Copy number of the novel joints homologous to probe 11-1800 in the mutant 30.7. **(A)** Restriction map of the insert in phage 11-1800. The horizontal line represents the insert DNA flanked by λ arms. The vertical lines mark the sites of restriction enzyme cleavage; B, *Bam*HI; R, *Eco*RI; H, *Hin*dIII; S, *Sal*I. The sizes of the restriction fragments are shown in kb. The circled numbers show the fragments used for the hybridizations below. The entire 11-1800 phage was used to probe *Hin*dIII and *Bam*HI digests of genomic DNAs from WT-BHK-2 and mutant 30.7. **(B)** Sizes (kb) of the genomic restriction fragments hybridizing with the entire 11-1800 phage. In each digestion there were fragments in common between WT-BHK-2 and 30.7 which were always amplified in 30.7, and an extra fragment was always found in 30.7. The transfers were then hybridized with subfragments from 11-1800. The 2.3 kb *Sal*I−*Eco*RI fragment was not homologous to the novel *Eco*RI or *Hin*dIII fragments **(C)** whereas the 9 kb *Sal*I−*Eco*RI fragment was **(D)**. Therefore, the novel joint must lie within the 9 kb *Sal*I−*Eco*RI fragment. The 3 kb *Bam*HI−*Hin*dIII fragment was homologous to the 15 kb *Eco*RI novel fragment and the 12.5 kb *Hin*dIII novel fragment but not to the 6.8 kb *Bam*HI novel fragment **(E)**. Therefore, the novel joint must lie within the 3.9 kb *Hin*dIII−*Bam*HI fragment. Consequently, the 3 kb *Bam*HI−*Hin*dIII fragment is fully contained within both the 11.3 kb normal *Eco*RI fragment and the 15 kb novel *Eco*RI fragment.

it with each of the five phage probes. Each of the three probes which contain a wild-type DNA sequence (11-200, 11-600 and 11-1800, average insert size 14 kb) should detect a set of overlapping cosmids of average size ~40 kb, and each of the two probes which contains an amplified novel joint (11-1700 and 11-1900) should detect two sets of cosmids, representing the two wild-type sequences which recombined to form the amplified joint. Cosmids from each of the seven sets could then be hybridized with each other and with cosmids containing the CAD gene, to obtain overlaps and linkage maps. As shown in Table II, 22 cosmids were obtained, corresponding to seven independent regions. The total amount of cloned genomic DNA was about 380 kb. Two independent sets of cosmids were isolated with probes from the novel-joint phage 11-1900, but only one set was isolated with probes from the novel-joint phage 11-1700 since

Table II. Isolation of cosmids containing amplified DNA

Probes used to detect cosmids	Number of cosmids isolated	Cosmids used as probes	Approximate length of each genomic region (kb)
11-1700	7	cos77a,cos773	80
11-200	4	cos276,cos271	60
11-600	2	cos671	40
11-1900 (12 kb)	2	cos171	40
11-1900 (21 kb)	1	cos971	35
11-1800	1	cos872	35
pCAD142	3	cosC31,cosC81	90
cosC31	2	cosC64	
Totals	22		380

A cosmid library, constructed from the DNA of mutant 10.1, was screened with sequences derived from the probes shown in Table I. Between 100 000 and 300 000 colonies were screened with each probe, and 500 000 colonies were screened with pCAD142. More cosmids were isolated with probes 11-1700 and 11-200 than with the others because these sequences are amplified to the same level as CAD in cell line 10.1, while the others are amplified less. Organization of the cosmids within each group was analyzed by restriction mapping and by cross-hybridization. Each cosmid contains about 40 kb of genomic DNA. The regions detected by the groups of cosmids corresponding to probes 11-1700, 11-200 and CAD were analyzed by using two overlapping cosmids per group, containing the ends of the region. Cosmids cosC81 and cosC64 represent the ends of a contiguous 90 kb region which includes the CAD gene, but these cosmids do not contain sequences corresponding to the centre of the region, which contains a part of the gene.

the overlap between the probes and the second region of wild-type DNA proved to be too small in this case.

The DNAs of the 22 cosmids shown in Table II, plus cosmid cCAD6 (Robert de Saint Vincent *et al.*, 1981) were digested with *Eco*RI and *Sal*I, and Southern transfers of each were probed with cosmids from each set, using competition with wild-type and cosmid DNAs (Materials and methods). There was no detectable overlap among the seven groups of cosmids (data not shown). The results could be used to organize the cosmids within each group. Representative cosmids from each group were then used to probe Southern transfers of the entire set of 33 single-step mutants. Three new sets of transfers were prepared from each mutant, using different enzymes (*Eco*RI, *Apa*I and *Kpn*I) to minimize the possibility that we would fail to detect novel joints. The amount of amplified DNA surveyed in each mutant was increased from about 150 kb to about 380 kb in this new experiment. The novel joints in mutants 30.7 (Figure 4) and 15.3 (not shown) previously detected with the phage probes, were detected again with the corresponding cosmid probes. Only one new novel joint was found, in mutant 15.6 with cosC64 as a probe. The new joint is located about 25 kb upstream of the 5′ end of the CAD gene and is also present in one copy per cell (data not shown). Except for these three novel joints, no rearrangement of DNA was detected in 33 mutants with a total of 380 kb of probes.

Discussion

Co-amplified DNA and novel joints

Different regions are amplified to different extents within the DNA of an individual single-step mutant. Furthermore, when two different mutants are compared, the relative extent of amplification of any particular sequence may be different (Figure 2). Previously, Zieg *et al.* (1983) reached similar conclusions for single-step mutants on the basis of a more limited analysis. Qualitatively, these results for single-step mutants are not different

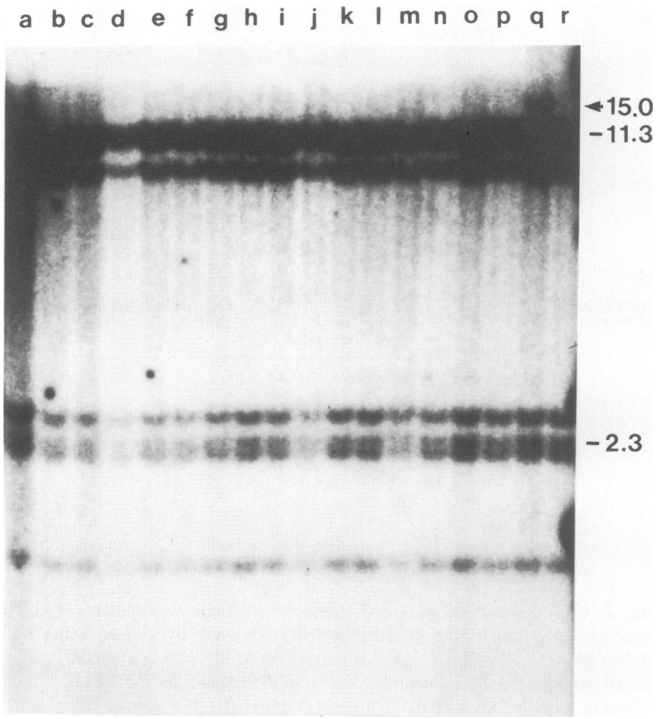


Fig. 4. A representative Southern transfer. *Eco*RI-digested genomic DNAs from several single-step mutants were probed with cos 872. In some of the WT lanes, 4, 2 and 0.5 times the standard amount of DNA was loaded, as indicated. An arrow shows the novel fragment in mutant line 30.7 **(lane q)**. The lengths of fragments common to λ11-1800 (Figure 3) are shown in kb.

from those obtained for multiple-step mutants (Ardeshir *et al.*, 1983). We conclude that the DNA co-amplified with CAD in the first step of amplification is not arranged in identical units and that the sort of complexity previously observed in multiple-step mutants can arise even during the first step.

It is reasonable to assume that the further a DNA sequence is from CAD in wild-type DNA, the greater will be its chance to be amplified to a different extent. The regions detected by probes 11-200 and 11-1700 may be relatively close to CAD, since they are amplified differently in only one mutant (Figure 2). However, the region detected by the 12.2 kb fragment of the 11-1900 probe may be further away and relatively close to the regions detected by probes 11-600, 11-1800 and the 21 kb fragment of probe 11-1900, since the abundance of these fragments is usually similar to each other in different mutants and is often different from the abundance of CAD.

In multiple-step mutants, novel joints unique to each mutant are amplified, i.e. present in many copies per cell (Ardeshir *et al.*, 1983). Two quite different events might lead to this situation: at each step of amplification, a single site might be used for recombination several times, so that amplified joints might be formed during the primary step of amplification; alternatively, each recombination and each novel joint might be unique in the primary step. However, the single-copy joints formed initially might be amplified in subsequent steps. In the present work, we have found only three novel joints. Each is present in a single copy in a different mutant cell line. As shown by I.Saito, E.Giulotto and G.R.Stark (in preparation), these joints may either be amplified or lost in second or third steps of amplification. Thus, it seems likely that some of the amplified joints observed in multiple-step mutants began as single-copy joints. If amplified joints are formed in the primary event, their frequency must be

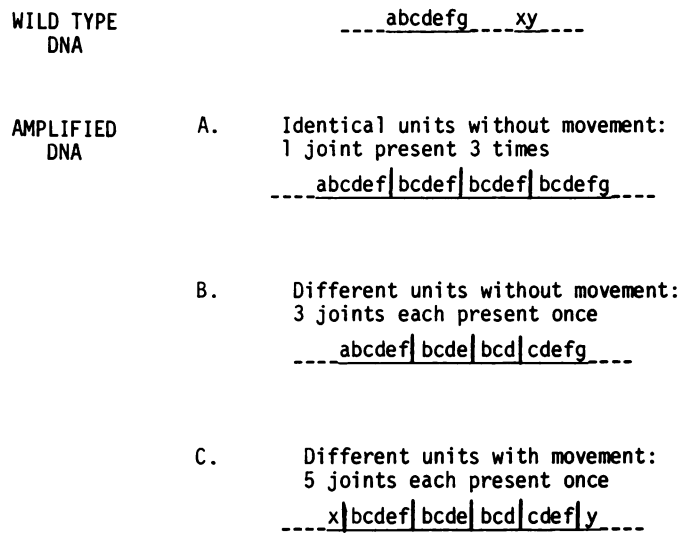


Fig. 5. Models for amplified DNA in single-step mutants. The letters a to g represent a DNA sequence which includes the selected gene (letter c) and the dashed line represents flanking DNA. Note that sequence xy in wild-type DNA might or might not be linked to the amplified DNA. Vertical lines signify joints, linking two sequences that are not adjacent in wild-type DNA.

extremely low, since none were detected in our survey of 33 mutants with about 380 kb of probes. Only one novel fragment was detected in each of the three cell lines whereas, in a reciprocal recombination, two novel fragments homologous to a single normal fragment would be generated. The recombinations we have detected may not be reciprocal, or the second fragments may have been lost before analysis.

Why so few novel joints?

We can estimate the minimum number of joints to be expected. As shown in Figure 5, at least one copy of a joint must be present for each new copy of the CAD gene, and some of the possible chromosomal arrangements of these extra copies are illustrated. In the cases shown, amplification has occurred on only one of the two homologous chromosomes, as actually found by *in situ* hybridization in the single-step mutants analyzed by Zieg *et al.* (1983). Note, however, that the relationship between the number of CAD genes and the number of novel joints does not depend on whether amplification takes place in one or more than one chromosome. We define an amplified unit as the region of DNA between two novel joints. If there are four exact repeats of a unique amplified unit, we should find three copies of the same novel joint (Figure 5A). If each unit is different, there will be three different single-copy joints (Figure 5B). If the DNA moves from its original position during or after amplification, additional single-copy joints will be generated (Figure 5C). Note that movement of DNA has already been observed to accompany amplification in several cases (Andrulis *et al.*, 1983; Flintoff *et al.*, 1984).

If novel joints fall randomly in amplified DNA, the probability of finding a joint in any particular mutant will depend only on the degree of amplification and on the fraction of total amplified DNA used as a probe. We use 1000 kb as an initial estimate of the average size of the amplified unit in single-step mutants, since this is the average amount of DNA co-amplified with each copy of CAD in multiple-step mutants, as determined by *in situ* hybridization. There must be at least three novel joints per single-step mutant, and if we have probed with 380/1000 of the total length of amplified DNA, we expect to find at least 3 × 380/1000 novel joints per mutant, or about 38 novel joints in the 33 mutants examined. If there is any movement of the amplified DNA (Figure 5C) or other rearrangement prior to analysis, more than three joints per mutant will be present. Furthermore, since each novel restriction fragment is composed of parts of two normal fragments, it will be complementary to both normal fragments, doubling our chance of detecting it. Although we might occasionally fail to detect a novel joint due to poor overlap with a particular probe, the probability of poor overlap is low, because we used three sets of Southern transfers in which the DNA had been digested with different enzymes.

Table III. Probability of overlaps among seven different 50 kb amplified sequences as a function of the average size of an amplified unit

Size of amplified unit (kb)	Probability of overlaps (%)		
	No overlaps	One overlap	Two or more overlaps
500	1	5	94
1000	11	26	63
2000	34	38	28
4000	59	32	9

An overlap between two 50 kb sequences occurs when the centre of one is <50 kb from the centre of the other, on either side. For example, if the length of an amplified unit is taken as 500 kb, the probability of an overlap (P) is 20% (2 × 50/500). The probabilities were calculated using the formula $(21P)^n(1-P)^{21-n}$, where n is the number of overlaps and 21 is the number of different pairwise combinations of seven sequences.

One possible explanation for finding novel joints at about one-tenth of the expected frequency is that each amplified unit is about 10 000 kb long in the single-step mutants, rather than 1000 kb as estimated from the data for multiple-step mutants. This possibility is discussed further below.

It is also possible that novel joints (recombinations) do not fall randomly in the amplified DNA, and that the probes we have used are relatively poor at detecting such events. The differential screening method originally used to obtain the probes preferentially selects highly amplified fragments and the usual sites for recombination might fall outside such fragments. For example, they might be preferred within long stretches of repetitive DNA, such as ribosomal DNA, as suggested previously by Wahl *et al.* (1983). The possibility of non-random recombination is not favoured by the fact that we have used so much amplified DNA as probes (38% if the average amplified region is 1000 kb) and by the fact that two of the probes used (11-1700 and 11-1900 in Table I), did contain novel joints present in the multiple-step mutant B5-4.

Is a very large amount of DNA co-amplified with CAD in single-step mutants?

In addition to the low frequency of novel joints, two other types of observation are relevant to this possibility. First, seven different regions of DNA, average size ~50 kb, that are amplified in single-step PALA-resistant mutants do not overlap with one another (Table II). The probability of finding overlaps is a function of the size of each amplified unit, as shown in Table III. It would be unlikely (11%) to find no overlaps if the average amplified unit were 1000 kb long, and no overlaps would be found about half the time if the average amplified unit were 4000 kb long. This evidence is consistent with a large amplified unit, but is not compelling. A second type of observation was made by Zieg *et al.* (1983), using *in situ* hybridization of a CAD

cDNA probe to metaphase spreads of nine single-step mutants. Autoradiographic grains usually appeared over extremely long, abnormal chromosomes and not over regions similar to the short arm of chromosome B9, the site of the CAD gene in wild-type cells (Wahl *et al.*, 1982). Three copies of a region 10 000 kb long would represent about 0.5% of the diploid genome of a Syrian hamster, or about 20% of an average Syrian hamster chromosome. Because the single-step mutant cell lines tested were heteroploid and because we did not perform chromosome banding and *in situ* hybridizations on the same metaphase spreads, we did not distinguish between expansion of the short arm of B9 and translocations which might have moved the amplified region to the long arm of another chromosome. In this regard, it is relevant to note that in three of four multiple-step mutants studied by Wahl *et al.* (1982), marker chromosomes clearly derived from B9 carried the amplified DNA in an expansion of the short arm, with no apparent translocation at this site.

More critical experiments to distinguish between non-random recombination and very large amplified regions as the explanation for the low frequency of novel joints are currently in progress. Although we favour the latter explanation, the evidence is not conclusive. If the amount of co-amplified DNA is much larger than expected previously our ideas concerning the mechanism of amplification must be revised. A popular mechanism discussed in several recent reviews (Cowell, 1982; Hamlin *et al.*, 1984; Schimke, 1984; Stark and Wahl, 1984), involves multiple initiation of replication with a single cell cycle, followed by recombination to resolve the resulting 'onion skins' into amplified chromosomal or extrachromosomal DNA. The region of DNA involved has been thought to be small relative to the size of a chromosome, perhaps involving only one or a few origins of replication. We can now imagine a quantitatively different event in which a very large region, perhaps even all of a chromosome, is over-replicated in the primary event of amplification, followed by more or less random recombination to give a large amplified region. In later steps of amplification, we can imagine that there may be two independent and parallel selections, one to increase the copy number of the gene which is the basis of drug resistance and the second to reduce the total amount of DNA. It would probably be difficult for a cell to maintain 100 copies of a 10 000 kb amplified unit, which would be equivalent in size to approximately seven normal Syrian hamster chromosomes. DNA sequences co-amplified with CAD during the first step may be lost, retained or amplified more or less than CAD in further steps of amplification, so that finally the average amount of co-amplified DNA is reduced to ~1000 kb per CAD gene in highly resistant cells. An analysis of a second- and third-step amplification events consistent with this model will be presented by Saito *et al.* (in preparation).

Recently, Shiloh *et al.* (1985) reported that the amount of DNA co-amplified with each copy of the *N-myc* gene in the human neuroblastoma cell line IMR-32 is ~3000 kb. There are 50–75 copies of *N-myc* per haploid genome in IMR-32 and several other unlinked DNA sequences cloned from IMR-32 are co-amplified to a similar degree. Some of these sequences map to regions of chromosome 2 that are thousands of kb distant from the position of *N-myc* at 2p23-2p24. Although Shiloh *et al.* (1985) propose a different explanation, it seems possible to us that a region of DNA much larger than 3000 kb was co-amplified with *N-myc* in the initial event in IMR-32 cells and then reduced in size during subsequent steps of amplification.

Materials and methods

Cell culture and selection with PALA

Cells were grown as described previously (Swyryd *et al.*, 1974). Dialyzed fetal calf serum was used during drug selection. Forty-five samples of 10^3 WT-BHK-2 cells (Zieg *et al.*, 1983) were grown for about 12 generations to a density of 5×10^6 cells/9 cm plate. The cells were dispersed and, from each independent suspension, 5×10^5 cells were placed onto fresh 9 cm plates. Fifteen plates each were treated immediately with 15, 30 or 60 μm PALA. Resistant colonies, picked after 3 weeks from plates which had <10 colonies, were grown in PALA to populations of about 10^7 cells each. Each population was then divided into five plates without PALA. After growth for two more generations, the cells from two plates were frozen and the remaining cells were used immediately for preparing DNA.

Recombinant cosmids

A cosmid library was constructed as described by Grosveld *et al.*, 1981, 1982). High mol. wt DNA (Brison *et al.*, 1982) was prepared from the single-step mutant 10.1 (Ardeshir *et al.*, 1983), which has about three copies of the CAD gene per haploid genome. Each of the phage probes detects only co-amplified wild-type sequences in this mutant cell line. After partial digestion with *Mbo*I, DNA fragments of 50–30 kb were isolated in a sucrose density gradient and ligated with the cosmid vector pTCF (Grosveld *et al.*, 1982). The genomic DNA was cloned at the unique *Bam*HI site between two *Sal*I sites, so that the insert could be excised with *Sal*I. Cosmids packaged *in vitro* (packaging kit, Amersham) were transduced into *Escherichia coli* strain ED8767. The cloning efficiency was $2-4 \times 10^5$ colonies/μg of fractionated genomic DNA. The library was prepared on nitrocellulose filters (Grosveld *et al.*, 1981) and stored at −70°C (Hanahan and Meselson, 1980). Bacterial DNA was fixed on the filters as usual, except that all cell debris was removed by wiping thoroughly with soft medical tissue in 2 × standard saline-citrate buffer, 0.1% SDS before drying and baking (Grosveld *et al.*, 1982). This procedure greatly reduced the background without apparent loss of specific signal.

The library (6×10^5 colonies) was screened with a CAD cDNA probe (pCAD142, Kempe *et al.*, 1976; Shigesada *et al.*, 1985) and six co-amplified genomic DNA fragments (Ardeshir *et al.*, 1983) (Table II). Genomic fragments were labelled by nick-translation and used as probes after repetitive sequences had been competed out by hybridization with excess unlabelled genomic DNA (Ardeshir *et al.*, 1983). The procedure is the same as the one we have used in Southern analysis, except that 2 μg of sheared pTCF DNA were also added, to compete out any contaminating cos sequences from the phage.

We obtained no cosmid clone containing the middle part of CAD gene, near the largest intron (Padgett *et al.*, 1982a), suggesting that a DNA sequence deleterious to cosmid maintenance (a complete palindrome, for example) may lie in this region. Furthermore, the same region was often deleted in phage clones, and the *Bam*HI fragment bearing this region could not be subcloned into pBR322 in numerous attempts (Padgett *et al.*, 1982a,b). An apparent exception is cosmid cCAD6, which contains a fully functional CAD gene (Robert de Saint Vincent *et al.*, 1981). However, the possibility that this cosmid may have a subtle deletion has not been ruled out. Another example of a sequence difficult to clone in the cosmid–host system that we have used has been reported (Moormann *et al.*, 1984), suggesting that such sequences may not be rare in mammalian DNA.

Screening cosmid libraries with probes containing repetitive sequences

We reported previously (Ardeshir *et al.*, 1983) that large segments of genomic DNA, cloned in a λ phage and including repetitive sequences, could be used directly to probe Southern transfers of genomic DNA, provided that repetitive sequences in the probe were competed first with total cellular DNA solution. We have now extended this competition procedure to the screening of cosmid libraries. The method is very simple and efficient compared with the conventional procedure of identifying and isolating fragments of probe free of repetitive sequences. Competition with the DNA of the cosmid vector as well as with total genomic DNA solves the common problem of contamination of probes with plasmid or λ cos sequences. Moreover, the competition method enables us to use much larger fragments as probes, giving a much more intense signal and dramatically improving the signal-to-noise ratio. Even an entire cosmid (cosC31, Table I) was used successfully as a probe to screen a cosmid library after competition of repetitive and cosmid sequences with genomic and cosmid vector DNAs. In this case, the background was higher than usual but still low enough to isolate cosmids cosC64 and cosC74. Direct probing of cosmid libraries with cosmid DNA might provide a more rapid method of chromosome walking. After two overlapping cosmids have been cloned, one could be used as a probe after competition with the other, giving a result equivalent to that obtained by isolating a terminal fragment from one of the cosmids.

Analysis of amplified DNA

High mol. wt DNA was isolated (Brison *et al.*, 1982) and digested with restriction endonucleases (Davis *et al.*, 1980) for 4 h, using three times more enzyme than required for complete digestion in 1 h. Restriction fragments were separated on 0.8% agarose gels (5 μg DNA/track) and transferred to aminothiophenol paper (Seed, 1982) as described by Wahl *et al.* (1979b). DNA is attached to this transfer medium covalently, and the transfers can be used in many different hybridizations without loss of signal, an important consideration for our work. The degree of digestion and efficiency of transfer were monitored by staining the gels with ethidium bromide before and after transfer.

Recombinant Charon 4 phage were purified from lysates of *E. coli* LE 392 (Davis *et al.*, 1980). Two hundred nanograms of recombinant phage DNA were labelled with ^{32}P to a specific activity of $2-8 \times 10^8$/c.p.m./μg. Southern transfers were pre-hybridized for at least 2 h at 42°C in standard buffer (Davis *et al.*, 1980) containing 50% formamide. Repeated DNA sequences were removed from the probes before hybridization (Ardeshir *et al.*, 1983): 150 ng of ^{32}P-labelled phage or cosmid DNA (or equivalent amount of genomic fragment) mixed with 80 μg of sheared DNA from wild-type BHK cells, were denatured and re-annealed to a C_0t value of 10. At least 2×10^6 c.p.m. of re-annealed probe were added per ml of pre-hybridization mixture. Hybridization was performed at 42°C for 16–20 h. The paper was washed at 65°C in 0.1% standard saline-citrate buffer and 0.1% SDS before autoradiography.

Analysis of copy number

A random single-copy Syrian hamster sequence free of highly repetitive DNA (C6-1), cloned in Charon 4, was used as an internal standard. C6-1 contains *Eco*RI fragments of 8 and 3.5 kb. After hybridizing with nick-translated C6-1 DNA, the intensity of the 8 kb band in each track was determined by densitometry and normalized to the intensity in the track containing wild-type DNA (Figure 1A). The cDNA clone pCAD 142 corresponds to 6.5 kb of the 7.9 kb CAD mRNA and hybridizes to 19 kb and 1.7 kb *Eco*RI fragments of genomic DNA (Shigesada *et al.*, 1985). pCAD 142 was hybridized to the same Southern transfers and the values were normalized to the amount of DNA in each track, as determined with C6-1. Similar methods were used to determine the copy number of co-amplified DNA fragments.

Acknowledgements

We thank Kit Osborn for her excellent technical assistance. I.S. was supported by a Research Training Fellowship from the International Agency for Research on Cancer.

References

Andrulis,I.L., Duff,C., Evans-Blackler,S., Worton,R. and Siminovitch,L. (1983) *Mol. Cell. Biol.*, **3**, 391–398.

Ardeshir,F., Giulotto,E., Zieg,J., Brison,O., Liao,W.S.L. and Stark,G. (1983) *Mol. Cell. Biol.*, **3**, 2076–2088.

Brison,O., Ardeshir,F. and Stark,G.R. (1982) *Mol. Cell. Biol.*, **2**, 578–587.

Caizzi,R. and Bostock,C.J. (1982) *Nucleic Acids Res.*, **10**, 6597–6618.

Cowell,J.K. (1982) *Annu. Rev. Genet.*, **16**, 21–59.

Davis,R.W., Botstein,D. and Roth,J.R. (1980) *Advanced Bacterial Genetics. A Manual for Genetic Engineering*. Cold Spring Harbor Laboratory Press, NY.

Federspiel,N.A., Beverley,S.M., Schilling,J.W. and Schimke,R.T. (1984) *J. Biol. Chem.*, **259**, 9127–9140.

Flintoff,W.F., Livingston,E., Duff,C. and Worton,R.G. (1984) *Mol. Cell. Biol.*, **4**, 69–76.

Grosveld,F.G., Dahl,H.H.M., de Boer,E. and Flavell,R.A. (1981) *Gene*, **13**, 227–237.

Grosveld,F.G., Lund,T., Murray,E.J., Mellor,A.L., Dahl,H.H.M. and Flavell, R.A. (1982) *Nucleic Acids Res.*, **10**, 6715–6732.

Hamlin,J.L., Milbrandt,J.D., Heintz,N.H. and Azizkhan,J.C. (1984) *Int. Rev. Cytol.*, **90**, 31–82.

Hanahan,D. and Meselson,M. (1980) *Gene*, **10**, 63–67.

Kaufman,R.J., Brown,P.C. and Schimke,R.T. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 5669–5673.

Kempe,T.D., Swyryd,E.A., Bruist,M. and Stark,G.R. (1976) *Cell*, **9**, 541–550.

Moormann,R.J.M., Jonsbloed,R. and Schoenmakers,J.G.G. (1984) *Gene*, **29**, 1–9.

Padgett,R.A., Wahl,G.M. and Stark,G.R. (1982a) *Mol. Cell. Biol.*, **2**, 293–301.

Padgett,R.A., Wahl,G.M. and Stark,G.R. (1982b) *Mol. Cell. Biol.*, **2**, 302–307.

Robert de Saint Vincent,B., Delbruck,S., Eckhalt,W., Meinkoth,J., Vitto,L. and Wahl,G.M. (1981) *Cell*, **27**, 267–277.

Schimke,R.T. (1984) *Cell*, **37**, 705–713.

Seed,B. (1982) *Nucleic Acids Res.*, **10**, 1799–1810.

Shigesada,K., Stark,G.R., Maley,J.A., Niswander,L.A. and Davidson,J.N. (1985) *Mol. Cell. Biol.*, **5**, 1735–1742.

Shiloh,Y., Shipley,J., Brodeur,G.M., Bruns,G., Korf,B., Donlon,T., Schreck, R.R., Seeger,R., Sakai,K. and Latt,S.A. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3761–3765.

Stark,G.R. and Wahl,G.M. (1984) *Annu. Rev. Biochem.*, **53**, 447–491.

Swyryd,E.A., Seaver,S.S. and Stark,G.R. (1974) *J. Biol. Chem.*, **249**, 6945–6950.

Wahl,G.M., Padgett,R.A. and Stark,G.R. (1979a) *J. Biol. Chem.*, **254**, 8679–8689.

Wahl,G.M., Stern,M. and Stark,G.R. (1979b) *Proc. Natl. Acad. Sci. USA*, **76**, 3683–3687.

Wahl,G.M., Vitto,L., Padgett,R.A. and Stark,G.R. (1982) *Mol. Cell. Biol.*, **2**, 308–319.

Wahl,G.M., Vitto,L. and Rubnitz,J. (1983) *Mol. Cell. Biol.*, **3**, 2066–2075.

Zieg,J., Clayton,C.E., Ardeshir,F., Giulotto,E., Swyryd,E.A. and Stark,G.R. (1983) *Mol. Cell. Biol.*, **3**, 2089–2098.

Received on 4 April 1986; revised on 2 June 1986