Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 8664–8668, November 1986
Genetics

Charomids: Cosmid vectors for efficient cloning and mapping of large or small restriction fragments

(tandem repeats/*recA*/restriction mapping/*CAD* gene/inverted duplication)

IZUMU SAITO* AND GEORGE R. STARK

Department of Molecular Biology, The Imperial Cancer Research Fund, P.O. Box 123, Lincoln's Inn Fields, London, WC2A 3PX, United Kingdom

Communicated by Walter Bodmer, July 14, 1986

ABSTRACT **Charomids are cosmid vectors up to 52 kilobases (kb) long, bearing 1–23 copies of a 2-kb spacer fragment linked in head-to-tail tandem arrays. Like cosmids and λ phage, charomids can be packaged *in vitro* for efficient introduction into bacteria. Charomids contain a polylinker with nine unique restriction sites for cloning and can be used without preparing vector arms. Using a charomid of appropriate size, one can clone inserts of any size up to 45 kb. For example, charomid 9-36 (9 cloning sites, 36 kb long) is too small to be packaged efficiently without an insert and can be used to clone fragments of 2–16 kb. The structure of charomids facilitates restriction mapping of the insert DNA and, after cloning, all the spacer fragments can be removed easily. After enrichment by size fractionation in an agarose gel, a specific single-copy genomic sequence can be cloned rapidly from ≈3 μg of DNA. Using charomid 9-36, we have cloned and mapped an amplified novel DNA fragment from a cell line resistant to *N*-(phosphonoacetyl)-L-aspartate and carrying about 100 copies of the *CAD* (carbamoyl-phosphate synthetase/aspartate carbamoyltransferase/dihydroorotase) gene. The fragment lies at the center of an inverted duplication of this gene.**

In studying DNA rearrangements such as amplification, translocations, and integration of oncogenic viruses, one often wishes to clone a specific restriction fragment for more detailed investigation. Although such fragments can be enriched by electrophoretic fractionation of restriction digests in an agarose gel, libraries made from partial digests of total genomic DNA are usually used for their cloning, because neither λ phage nor cosmid vectors are ideal for isolating restriction fragments of all sizes from complete digests. Cosmids are bacterial plasmids to which the cohesive-end site (*cos*) of bacteriophage λ DNA has been attached to allow packaging by extracts of lysogenic bacteria. Charomid vectors are cosmids modified especially for cloning restriction fragments of any size up to ≈45 kilobases (kb). Charomids carry both a polylinker and a spacer region of head-to-tail repeats of a 2-kb fragment. The number of spacer fragments is variable and determines the lengths of inserts that can be accommodated within the total length of 38–52 kb necessary to form infectious cosmid particles after packaging *in vitro*. As shown previously (1, 2), individual cosmid DNAs can be used to probe Southern transfers after a simple competition with total genomic DNA to remove repetitive sequences. This method provides a rapid and efficient way to identify rearranged fragments, which can be cloned easily by using the charomid vector system and can be mapped conveniently by using labeling sites located in the vector. Taken together, these techniques provide an effective way to study DNA rearrangements in detail.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

We grew charomid vectors and plasmids containing repeats in *Escherichia coli* ED8767 (3), a *recA*$^-$ strain often used for cosmid cloning. Because ED8767 is F$^+$, an F$^-$ strain, such as NM554, is probably better. All other plasmids were grown in *E. coli* HB101. *E. coli* LE392 was used as a *recA*$^+$ strain. Plasmid and charomid DNAs were prepared from 200-ml cultures by the alkaline-lysis method (4).

RESULTS

Construction of Charomid 9 Vectors. Charomid 9-36 (Fig. 1*C*) was derived from cosmid vector pTBE-PL9 by adding multiple copies of a 2-kb *Ava* I spacer fragment, constructed as shown in Fig. 1*A* and contained within the plasmid pBRd[RS]sp(RCHB). pTBE-PL9 was linearized with *Ava* I and linked to the spacer fragment in tandem head-to-tail arrays (Fig. 1*B*, lines a and b). Because the *Ava* I site in pBR322 (C/TCGGG, the slash shows the position of cleavage) is an incomplete palindrome, the cleaved ends are not rotationally equivalent and head-to-head ligation does not occur (8). This property is important because the head-to-head structure destabilizes the resultant plasmid in bacteria. Upon ligation in high concentration of the *Ava* I-linearized cosmid vector and a 50-fold molar excess of the spacer fragment, long head-to-tail repeats are formed (Fig. 1*B*). These serve as good substrates for *in vitro* packaging whenever two *cos* sites are present at a distance of 38–52 kb (4). The efficiency of this procedure was very high; more than 10^4 colonies were obtained from 50 ng of linearized vector and 1 μg of the spacer fragment. A mixed culture containing 10^3 independent colonies yielded a plasmid population of 40–50 kb (Fig. 2, lane p). Ten colonies were picked randomly and the plasmid DNAs were examined. All were 40–52 kb long, seven contained one cosmid unit (Fig. 1*B*, line a) and three contained two cosmid units (Fig. 1*B*, line b). These charomid vectors were obtained in yields of ≈1 μg/ml and appeared quite stable in ED8767 cells because all 10 DNA preparations (after ≈30 generations) yielded apparently pure plasmids (Fig. 2, lanes 42–52). However, smaller plasmids that have fewer spacer fragments were sometimes evident in large-scale DNA preparations after ≈50 generations of growth (for example, Fig. 3, first lane).

Charomid vectors, carrying many copies of the same spacer fragment, were expected to be unstable in *recA*$^+$ cells because of homologous recombination. The largest charomid (52 kb) was packaged *in vitro* and transferred into the *recA*$^+$ strain LE392. As shown in Fig. 2 (lanes m), a set of charomids differing by one spacer fragment was formed, ranging in size from 52 kb (23 spacers) to 5.3 kb (no spacers, from recombination with homologous sequences in the vector). To

Abbreviations: kb, kilobase(s).

*Present address: Department of Bacteriology, Faculty of Medicine, University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113, Japan.

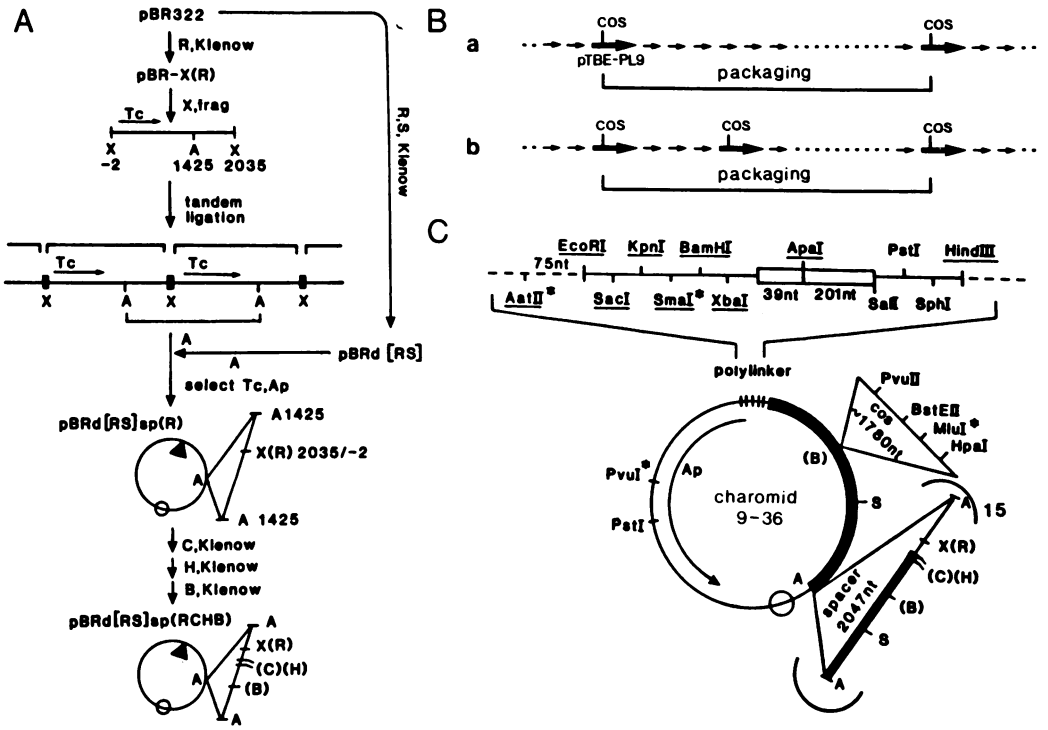

FIG. 1. Construction of charomids. Tc, tetracycline; Ap, ampicillin; open circle, replication origin of pBR322; X, *Xmn* I; R, *Eco*RI; A, *Ava* I; C, *Cla* I; H, *Hin*dIII; B, *Bam*HI; S, *Sal* I; the restriction site destroyed is shown by the name of the enzyme in parentheses. (*A*) Construction of the spacer fragment. To prepare pBR-X(R), the *Eco*RI site of pBR322, G/AATTC (the slash shows the position of cleavage), was converted to an *Xmn* I site, GAATT/AATTC, by filling in. pBRd[RS] was made from pBR322 by deleting the *Eco*RI–*Sal* I fragment; the deletion is indicated as a filled triangle. "frag," isolation of a restriction fragment; "Klenow," use of the Klenow fragment of DNA polymerase I to fill in a restriction site. Numbers show nucleotide positions in pBR322. (*B*) Cloning multiple copies of the spacer fragment. Short arrow, spacer *Ava* I fragment from pBRd[RS]sp(RCHB); long arrow, *Ava* I-linearized cosmid vector pTBE-PL9. Two cases are shown in which charomids bearing one (line a) and two (line b) copies of pTBE-PL9 are generated. Cosmid vector pTBE-PL9 was constructed from pUC18Apa and pTBE. pUC18Apa, donor of the polylinker, was made by inserting the 240-nucleotide (nt) *Hin*cII fragment of simian virus 40 (nt 2057-2297) into the *Hin*cII (*Sal* I) site of pUC18 (5). This procedure introduced an *Apa* I site to the polylinker of pUC18 and regenerated a single *Sal* I site. Cosmid vector pTBE (6) bears the *cos* (cohesive end)-containing 1780-nt *Bgl* II fragment from λ phage Charon 4A at the *Bam*HI site of pBR327 and a *Bam*HI linker at the *Sal* I site. pTBE-PL9 was made from pTBE by removing the *Bam*HI linker and then inserting the polylinker from pUC18Apa between the *Eco*RI and *Hin*dIII sites. (*C*) Charomid 9-36. In the polylinker region, the restriction sites underlined are useful for cloning; broken lines, solid lines, and open boxes denote the sequences from pBR327, the pUC18 polylinker, and the simian virus 40 late region, respectively. The four restriction sites useful for terminal labeling in connection with restriction mapping are shown with an asterisk. Thick lines indicate the sequences homologous between the spacer fragment and the main body of the vector. To construct charomids of 42–52 kb, pTBE-PL9 was methylated with M.*Hpa* II (New England Biolabs) to prevent *Ava* I from cleaving at the *Sma* I site and then was digested with *Ava* I (7). The resulting linear molecule (50 ng) was ligated with the spacer fragment (1 μg) in 7 μl of buffer at 15°C for 16 hr. The mixture was packaged *in vitro* and adsorbed to ED8767 cells. To isolate charomids of 20–40 kb, charomid 9-52 was packaged *in vitro* and adsorbed to LE392 cells. A pool of $\approx 10^3$ independent transformants was grown directly in 5 ml of Luria–Bertani (LB) broth with ampicillin. A sample of DNA from the culture was separated by electrophoresis directly in a 0.4% agarose gel. Closed circular charomid DNAs of different size were recovered by electroelution from gel slices and used to transform ED8767 cells by the calcium chloride method (4). For large scale preparations, a fresh single colony was transferred into 200 ml of medium and was grown without chloramphenicol amplification. This procedure minimizes the number of cell divisions and hence contamination by smaller and larger charomids.

isolate charomid vectors smaller than 40 kb, DNAs of various size were purified from a single gel and the vectors were obtained after transformation of ED8767 cells (Fig. 2). The charomids isolated were identical to the original 52-kb vector except for the number of repeats (data not shown). As an example, the structure of charomid 9-36 is shown in Fig. 1*C*. The restriction digest shown in Fig. 3 (first and second lanes) confirmed this structure.

Use of Charomid 9-36 as a Cloning Vector. To measure the cloning efficiency, we digested Syrian hamster genomic DNA with *Eco*RI or *Apa* I and isolated the size fractions 7.5–8.5 kb and 9.2–9.7 kb. Each fraction (0.1 μg) was ligated with charomid 9-36 DNA, linearized with either *Eco*RI or *Apa* I (2.5 μg, 6-fold molar excess). In three experiments, 1.5–5.5 $\times 10^6$ colonies per μg of the fractionated genomic DNA were obtained. Each library should contain two to seven recombinants carrying a single-copy genomic fragment of the selected size, assuming that the fragment is enriched 10-fold by size fractionation and that four-fifths of the colonies contain genomic inserts. There were about one-fifth as many colonies in control experiments without inserts. Approximately 10% of these colonies contained the 36-kb charomid itself, and the rest contained larger charomids (42–52 kb). Both types were also recovered from an uncut DNA preparation of charomid 9-36 by *in vitro* packaging.

Cloning a Novel Genomic Fragment. Approximately 100 copies per cell of the selected *CAD* (carbamoyl-phosphate synthetase/aspartate carbamoyltransferase/dihydroorotase) gene are present in the *N*-(phosphonoacetyl)-L-aspartate-resistant cell line B5-4 (1). These cells also contain ≈50 copies per cell (see below) of a novel 9.5-kb *Apa* I fragment. A charomid library (1.5×10^5 colonies) was constructed from B5-4 DNA in order to clone this novel fragment. Twelve positive colonies were identified from 16,000 screened, in agreement with the expected frequency of one positive per 1200 colonies, assuming a 10-fold enrichment by size fractionation. Three colonies were picked and all contained the same insert; one, chdJ5B, was studied further.

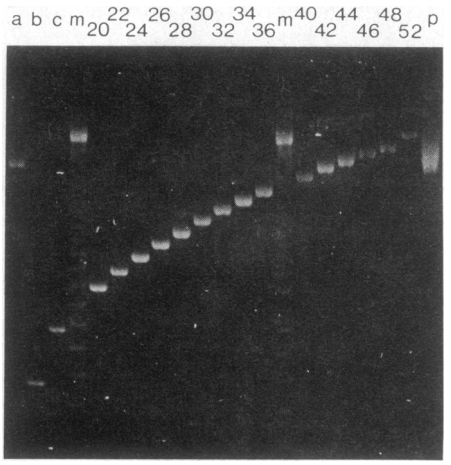

FIG. 2. A set of charomid vectors. Closed circular charomid DNAs prepared by the rapid method (4) (1/50th of a 1.5-ml culture) were separated by electrophoresis in a 0.4% agarose gel in the presence of ethidium bromide. Each number represents the approximate size of a charomid in kb. Lanes m, DNA from charomid 9-52 after growth in the *recA*$^+$ strain LE392; lane p, mixed population of DNAs from the construction experiment (see text). Size markers (100 ng of each): lane a, cosmid cosC81 (42 kb); lane b, pBR2HBV (10.8 kb); lane c, pBRHBV dimer (15.3 kb).

Since the *Sal* I recognition site is relatively rare in mammalian DNA (average spacing ≈50 kb), the charomid vector was constructed to use this enzyme for reducing the vector size while keeping the insert intact. Because *Sal* I cuts both the distal part of the polylinker and the spacer fragment (Fig. 1*C*), recircularization by T4 ligase should yield a plasmid containing the full insert plus 2.9 kb, unless *Sal* I happens to cut the insert. The shortened plasmid, pJ5B, was easily obtained from chdJ5B by using *Sal* I and ligase (Fig. 3).

Restriction Mapping of chdJ5B. As shown in Fig. 3, restriction digests of chdJ5B give very simple patterns after separation in an agarose gel, because most enzymes do not cut the 31-kb spacer region at all (Fig. 3). Nearly all the smaller fragments come from the insert and the insert–vector junctions. To make restriction maps of the insert, we used a simple procedure based on the method of Smith and Birnstiel (11). The *Mlu* I restriction site (A/CGCGT), present once in the charomid vectors, is rare in mammalian DNA (average spacing ≈90 kb; unpublished observation). DNA from chdJ5B was cut with *Mlu* I, labeled at the 3′ ends with T4 polymerase to incorporate [α-^{32}P]dCTP, and used directly for partial digestion with various enzymes that do not cut within the spacer region. Almost every band in the autoradiogram (Fig. 4) represents the distance from a restriction site within the insert to the proximal *Mlu* I end only, because all bands from the distal *Mlu* I end contain the 31-kb spacer region and are retained near the top of the gel. Complete restriction maps were obtained from experiments such as these, together with data from complete digests of chd5B or pJ5B.

Structure of the Novel Joint and Surrounding DNA. The structure of DNA amplified in B5-4 cells is unusual. The DNA up to a point ≈25 kb upstream of the *CAD* gene (in the normal arrangement) is amplified to ≈100 copies per B5-4 cell, but the DNA further upstream of this point (in the normal arrangement) is present at only ≈1 copy per B5-4 cell. For example, when an *Eco*RI digest of B5-4 DNA was probed with cosC64, a cosmid bearing normal DNA sequences (Fig. 5, lane 1), the 8.5-kb fragment was amplified but the 5.8-kb fragment was not; see also the positions of these fragments in the map of normal DNA in Fig. 6. Moreover, amplified novel bands (2.1 kb in an *Eco*RI digest and 9.5 kb in an *Apa* I digest) were detected in B5-4 DNA by using the 2.4-kb *Bam*HI probe

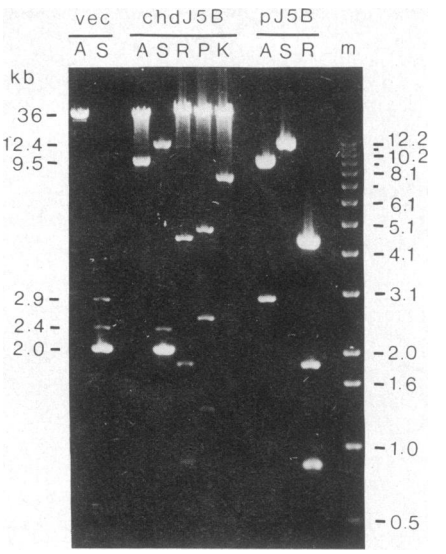

FIG. 3. Charomid containing a novel fragment from cell line B5-4. vec, charomid 9-36; chdJ5B, the recombinant charomid bearing a 9.5-kb *Apa* I fragment from cell line B5-4; pJ5B, shortened plasmid derived from chdJ5B by using *Sal* I and T4 ligase. The DNAs (0.5–1 μg) were digested and separated by electrophoresis in a 1% agarose gel. A, *Apa* I; S, *Sal* I; K, *Kpn* I; P, *Pst* I; R, *Eco*RI; m, 1-kb "ladder" marker (Bethesda Research Laboratories). The sizes of *Apa* I and *Sal* I fragments are shown at left: 36 kb, charomid vector; 12.4 kb, insert plus vector without *cos* or spacer fragments; 9.5 kb, insert; 2.9 kb, vector without *cos* or spacer fragments; 2.4 kb, *Sal* I fragment bearing the *cos* site; 2.0 kb, spacer fragments. Genomic DNA (5 μg) from cell line B5-4 (1) was digested with *Apa* I, and the fraction from 9.2 to 9.7 kb (0.1 μg, purified from an agarose gel by electroelution) was ligated with charomid 9-36 DNA (2.5 μg, linearized with *Apa* I) at 15°C for 16 hr in 7 μl of buffer. The mixture was packaged *in vitro* (Gigapack, Vector Cloning Systems) and adsorbed to ED8767 cells. Some lots of Gigapack yielded 500 times fewer colonies because of contamination with *Eco*K (9). The sequence used for screening was a 2.4-kb *Bam*HI fragment (2.4B in Fig. 6) that reveals a novel joint in B5-4 DNA. The fragment was isolated from the plasmid pV2-104 (a gift of O. Brison, Institut Gustave-Roussy, Paris), labeled using the Klenow fragment of DNA polymerase I (10), and used as a probe after blocking the repetitive sequences by hybridization with total genomic DNA (1, 2). The same methods for labeling and competition were also used for the other Southern experiments. In brief, 130 ng of labeled cosmid DNA or an equivalent amount of a genomic fragment (7 ng for the 2.4-kb fragment), mixed with 80 μg of sonicated DNA from normal BHK cells, was denatured and reannealed in 0.12 M phosphate buffer (pH 6.8) at 65°C for 135 min in a volume of 200 μl.

(2.4B in Fig. 6), which is located 25 kb upstream of the *CAD* gene (data not shown).

To examine the structure further, the novel 9.5-kb *Apa* I fragment was cloned in charomid 9-36 (chdJ5B). The cloned fragment contained a single *Kpn* I site, and double digestion with *Apa* I and *Kpn* I yielded fragments of 7.5 kb (7.5KA in Fig. 6, bottom) and 2.0 kb. Only the 2.0-kb fragment hybridized with the 2.4B probe used for cloning. This fragment contains the novel joint, because it was detected as a novel 2.0-kb band in a double digest of B5-4 DNA with *Apa* I and *Kpn* I (data not shown). To examine the location of the other fragment in normal DNA, the 7.5KA fragment was hybridized to Southern transfers of previously cloned cosmids. The probe hybridized with a 7.5-kb *Kpn* I–*Apa* I fragment in cosC64 (n7.5KA in Fig. 6, center; n denotes normal) located downstream of the position of the novel joint. Restriction maps of 7.5KA and the normal counterpart, n7.5KA, were identical with four enzymes (data not shown). In a Southern experiment with an *Apa* I digest of B5-4 DNA, the 7.5KA probe detected not only the 9.5-kb amplified novel

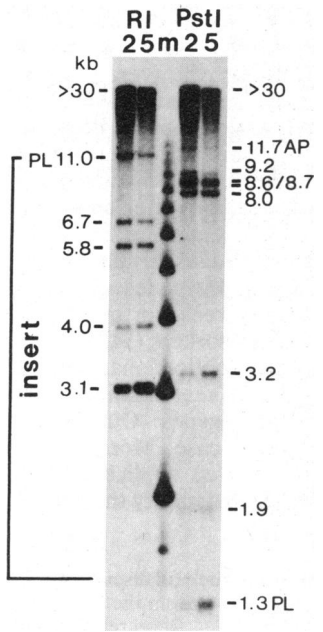

FIG. 4. Restriction mapping of chdJ5B. Recombinant charomid DNA (1 μg, linearized with *Mlu* I) was labeled at the 3′ ends with 50 μCi of [α-^{32}P]dCTP (3000 Ci/mmol; 1 Ci = 37 GBq), using 3 units of T4 DNA polymerase (New England Biolabs), after allowing the exonuclease reaction to proceed for 5–10 min at 37°C. This DNA was diluted and mixed with sheared carrier DNA to 50 μg/ml. For partial digestion, a 20-μl sample was incubated with 6 units of *Eco*RI or *Pst* I for 2 or 5 min (lanes 2 and 5) at 37°C. After electrophoresis, the gel was dried and autoradiographed for 4 hr. The observed sizes of the partially digested fragments are shown. PL, polylinker; AP, ampicillin-resistance gene. The *Apa* I sites at the ends of the insert are 1.5 kb and 11.0 kb from the *Mlu* I site. Lane m: end-labeled 1-kb ladder.

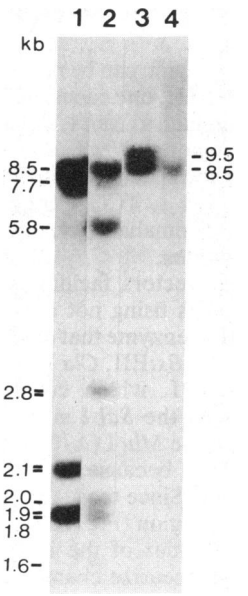

FIG. 5. Southern transfers of B5-4 DNA. DNAs from the G1A96 (30 copies of *CAD* per cell) and SS10.1 (6 copies of *CAD* per cell) cell lines show the same pattern as does wild-type Syrian hamster DNA. *Eco*RI-digested DNAs (3 μg) from cell lines B5-4 (lane 1) and G1A96 (lane 2) were probed with the total DNA of cosmid cosC64, which contains the normal DNA sequence shown in Fig. 6. A novel fragment containing the B5-4 joint comigrated with the 2.1-kb band. *Apa* I-digested DNAs from cell lines B5-4 (lane 3, 1 μg) and SS10.1 (lane 4, 3 μg) were probed with the 7.5-kb *Apa* I–*Kpn* I fragment isolated from chdJ5B (shown in Fig. 6 as 7.5KA).

fragment but also the 8.5-kb *Apa* I fragment with similar intensities (Fig. 5, lane 3). All the results are consistent with the proposed structure shown in Fig. 6, in which most or all units of amplified DNA in B5-4 cells, present in ≈50 copies per cell, contain an inverted duplication involving the *CAD* gene. The structure is also consistent with the result that no rearrangement was detected in the region between the 1.8-kb *Eco*RI fragment and the 3′ end of the *CAD* gene (1). The inverted duplication contains a unique region of about 1.5 kb in its center. This region includes one end of the 9.5-kb novel *Apa* I fragment and is homologous to the 2.4B probe used for cloning. Comparison of the restriction maps of cloned fragments from B5-4 and normal DNA revealed that one *Eco*RI site is missing in the unique central region of the amplified inverted repeat (data not shown).

DISCUSSION

Since charomid vectors can be packaged *in vitro*, the cloning efficiency is comparable to that of λ vectors and is high enough to isolate a restriction fragment present in one copy per cell (6×10^6 kb) from 0.1 μg of size-fractionated DNA, or from 2–3 μg of genomic DNA before fractionation. Charomids have many advantages. (*i*) One can choose the best vector size for a particular purpose. For example, charomid 9-36 is suitable for cloning many restriction fragments because it accepts inserts of 2–16 kb. (*ii*) For charomids of 28–36 kb, there is no need to prepare vector arms, as is necessary for some λ vectors, and partial digestion should not increase the number of clones without inserts. (*iii*) The potential number of cloning sites is very high. Nine are present in the charomid 9 vectors, and any two sites can be used to clone fragments from double digests. (*iv*) Because *recA*$^-$ *E. coli* are used for charomid cloning, a genomic insert should be more stable than in those phage vectors which require a *recA*$^+$ host. (*v*) Restriction patterns of recombinant charomids are very simple (Fig. 3). (*vi*) After isolation of a clone, the vector size can be reduced to 2.9 kb by deletion of the spacer sequences. (*vii*) Restriction maps of the insert can be determined by a simple and rapid procedure. (*viii*) Charomid vectors can be modified easily for special purposes.

Of the nine cloning sites of charomid 9-36 (Fig. 1*C*), the sites for *Aat* II, *Sma* I, *Kpn* I, and *Apa* I are novel in vectors that can be packaged *in vitro*. The *Bam*HI site accepts fragments cut with *Bam*HI, *Bgl* II, *Bcl* I, *Xho* II, and *Mbo* I, and the *Xba* I site accepts fragments cut with *Nhe* I, *Spe* I, and *Avr* II. *Sma* I produces flush ends that accept flush-ended fragments (24 enzymes commercially available). Altogether, 39 enzymes can be used for charomid 9 cloning, and other fragments can be cloned at the *Sma* I site after the ends are filled in. *Kpn* I and *Apa* I are especially useful because they

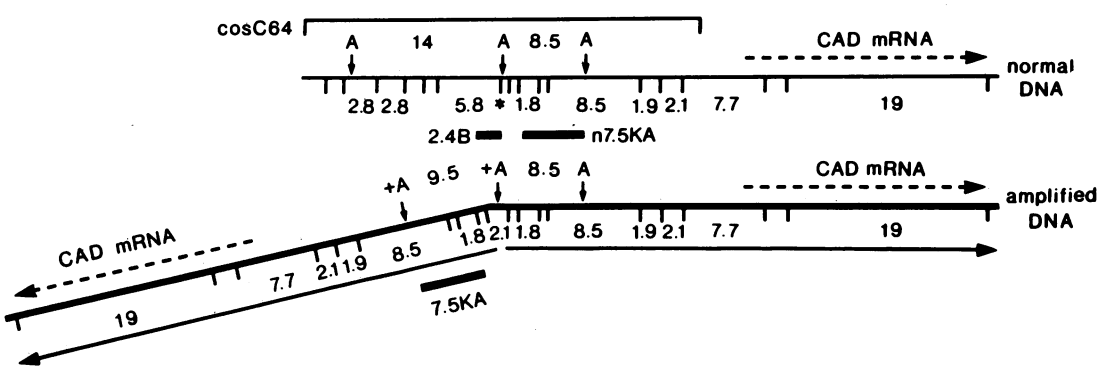

FIG. 6. Structure of amplified DNA near the *CAD* gene in B5-4 cells. Restriction maps of normal DNA (*Upper*) and amplified B5-4 DNA (*Lower*) are shown. The *Eco*RI sites are represented by vertical lines, and the *Apa* I sites near the novel joint are also shown. A, *Apa* I; +A, the end of the novel *Apa* I fragment cloned in chdJ5B; K, *Kpn* I; B, *Bam*HI; asterisk, the *Eco*RI site missing in the amplified DNA. Fragment sizes are shown in kb. Arrows below the map of amplified DNA indicate the inverted duplicated sequences. Probes: cosC64, a cosmid bearing normal DNA; 2.4B, probe used for isolating chdJ5B; 7.5KA, fragment from chdJ5B; n7.5KA, normal counterpart of 7.5KA in cosC64 DNA.

cut mammalian DNAs into fragments of a convenient size, often 6–15 kb (unpublished data).

After isolation of a clone, the spacer region can be removed with *Sal* I. If *Sal* I happens to cut the insert, one can use other enzymes that cut the spacer fragment, such as *Sph* I, *Eco*RV, *Nhe* I (removing the *cos* and spacer regions), *Nru* I, *Bbe* I (= *Nar* I), *Nae* I, and *Xma* III (removing the spacer region only). The last four enzymes (especially *Nru* I, TCG/CGA) are useful because, like *Sal* I, they cut mammalian DNA rarely because of the rarity of CG dinucleotides.

The unique structure of charomid vectors facilitates restriction mapping. One can make maps using not only the nine cloning enzymes but also any other enzyme that does not cut the spacer fragment, such as *Bgl* II, *Bst*EII, *Cla* I, *Dra* I, *Pst* I, and *Pvu* II. *Hin*cII and *Acc* I, which cut more frequently, can be used after blocking the *Sal* I sites with M.*Taq* I methylase. For end-labeling, the *Mlu* I (A/CGCGT) and *Pvu* I (CGAT/CG) sites seem best because both these sites are very rare in mammalian DNA. Since these two sites lie on opposite sides of the polylinker region (Fig. 1*C*), maps can often be constructed from both sides of the insert. λ Terminase (12) could also be used to linearize charomids at the *cos* site before end-labeling.

Although charomid 9-36 is particularly useful for cloning specific restriction fragments, vectors of other size may be better for some purposes. For constructing a total genomic library, charomid 9-28 seems the best alternative to Charon 4, Charon 35, or EMBL3/4 phage vectors because it accepts inserts of 10–24 kb efficiently without being packaged itself. After phosphatase treatment, charomid 9-42 could be useful for cloning small restriction fragments and for constructing cDNA libraries. The Okayama–Berg method (13) can be used because the vector has multiple restriction sites.

One drawback of cloning in charomid vectors is an unexpectedly high background (about 20%), though the background is not usually a problem. The background is due partly to generation of charomids bearing more or fewer copies of the spacer during replication of the vector. Overdigestion of the vector also tends to increase the background.

The method for inserting multiple copies of a DNA fragment into a cosmid might be useful for other purposes. For example, a cosmid bearing many copies of an expressed sequence, transfected into mammalian cells, might give cells expressing this sequence at a high level after integration of only one plasmid. As one example, we have constructed a plasmid containing 12 copies of the entire hepatitis B virus genome to test the utility of this approach (I.S., G.R.S., A. Lever, and H. Thomas, unpublished data).

In the mutant cell line B5-4, most or all of the amplified *CAD* genes are found in identical inverted duplications in which two copies of the gene are arranged head-to-head, about 50 kb apart. The distance between *CAD* genes in B5-4 seems surprisingly short because we have shown (2) that, in the initial step of *CAD* gene amplification, the average length of amplified units may be very large, perhaps 10,000 kb or more per gene. Even in the multiple-step mutant B5-4, we estimate an average size of 2000 kb or more (2). One plausible hypothesis is that the inverted duplication in B5-4 arose as a single copy and was retained preferentially after additional amplification events, selected in discrete steps at higher concentrations of *N*-(phosphonoacetyl)-L-aspartate. Amplification of any unit in which two *CAD* genes are close together will give twice the usual increase in gene copy number per event. Ford, Fried, and coworkers (14, 15) have developed a method to detect inverted duplications, and they suggest that such structures are present near amplified *CAD* genes and also near amplified c-*myc* genes. Our present results provide direct evidence in one case. More work is necessary to determine whether inverted duplications are found commonly in amplified DNA and to clarify their role in the development of an amplified array.

We thank E. Giulotto and M. Debatisse for helpful discussions and O. Brison for pV2-104. I.S. is a visiting fellow from the University of Tokyo and was supported in part by a Research Training Fellowship from the International Agency for Research on Cancer.

1. Ardeshir, F., Giulotto, E., Zieg, J., Brison, O., Liao, W. S. L. & Stark, G. R. (1983) *Mol. Cell. Biol.* **3,** 2076–2088.
2. Giulotto, E., Saito, I. & Stark, G. R. (1986) *EMBO J.*, in press.
3. Murray, N. E., Brammar, W. J. & Murray, K. (1977) *Mol. Gen. Genet.* **150,** 53–61.
4. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
5. Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) *Gene* **33,** 103–119.
6. Grosveld, F. G., Lund, T., Murray, E. J., Mellor, A. L., Dahl, H. H. M. & Flavell, R. A. (1982) *Nucleic Acids Res.* **10,** 6715–6732.
7. Nelson, M., Christ, C. & Schildkraut, I. (1984) *Nucleic Acids Res.* **12,** 5165–5173.
8. Hartley, J. L. & Gregori, T. J. (1981) *Gene* **13,** 347–353.
9. Rosenberg, S. M. (1985) *Gene* **39,** 313–315.
10. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132,** 6–13.
11. Smith, H. O. & Birnstiel, M. L. (1976) *Nucleic Acids Res.* **3,** 2387–2398.
12. Rackwitz, H.-R., Zehetner, G., Murialdo, H., Delius, H., Chai, J. H., Poustka, A., Frischauf, A.-M. & Lehrach, H. (1985) *Gene* **40,** 259–266.
13. Okayama, H. & Berg, P. (1982) *Mol. Cell. Biol.* **2,** 161–170.
14. Ford, M., Davies, B., Griffiths, M., Wilson, J. & Fried, M. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 3370–3374.
15. Ford, M. & Fried, M. (1986) *Cell* **45,** 425–430.