The EMBO Journal vol.7 no.2 pp.407–417, 1988

The multicopy appearance of a large inverted duplication and the sequence at the inversion joint suggest a new model for gene amplification

Olivier Hyrien, Michelle Debatisse, Gérard Buttin and Bruno Robert de Saint Vincent

Unité de Génétique Somatique (UA CNRS 361), Institut Pasteur, 25, rue du Dr Roux, 75724 Paris Cedex 15, France

Communicated by G.Buttin

The amplified DNA of $HC_{50}474$, a Chinese hamster fibroblast cell line selected in three steps for high resistance to coformycin, consists chiefly of 150 copies of a large inverted duplication including the adenylate deaminase gene. Most if not all of these units are more than 2 × 120 kb long. The inverted duplication was first detected in the cells recovered from the second selection step, at the same chromosomal location as the first step amplified units. Its formation and amplification appear to be coupled since the second step cell line already contained 40 copies of this novel structure. Reamplification of the inverted duplication occurred at the third step of selection concomitant with the loss of amplified DNA acquired during the first step. The head-to-head junction has been formed by recombination within a recombinational hotspot described previously [Hyrien,O., Debatisse,M., Buttin,G. and Robert de Saint Vincent,B. (1987) *EMBO J.*, 6, 2401–2408]. Sequences at the joint and in the corresponding wild-type region reveal that the crossover sites, one of which occurs in the putative promoter region of B2 repeat, are located at the top of significant stem-loop structures and that patchy homologies between the parental molecules on one side of the breakpoints allow alignment of these crossover sites. We present a model which explains the formation and amplification of this and other large inverted duplications by errors in DNA replication.

Key words: adenylate deaminase/DNA replication/gene amplification/illegitimate recombination/inverted duplication

Introduction

Gene amplification is a widespread phenomenon of considerable biological importance. In some organisms, it occurs as a developmentally regulated process (e.g. the chorion genes of *Drosophila*; Spradling and Mahowald, 1980) or as an accidental process revealed by selection (e.g. genes encoding detoxification enzymes in pesticide-resistant insects; Mouchès *et al.*, 1986). In mammals, amplification of oncogenes is found in numerous tumour cells and amplification of genes encoding target enzymes or membrane glycoproteins is a major mechanism of drug resistance in cultured cell lines and in tumours (for reviews see Schimke, 1982, 1984; Stark and Wahl, 1984; Hamlin *et al.*, 1984; Kafatos *et al.*, 1985; Stark, 1986).

How the amplification DNA appears and how it is organized in amplified cell lines has been the subject of numerous studies. Several reports on the extent and amount of DNA co-amplified with the selected gene concluded that amplified regions are not composed of identical units and that the 'novel joints' which link together adjacent amplified units are apparently randomly distributed (Ardeshir *et al.*, 1983; Federspiel *et al.*, 1984). Other observations (Looney and Hamlin, 1987; Borst *et al.*, 1987) indicate homogeneity among amplified units within other cell lines.

Using Chinese hamster cell lines selected for overproduction of adenylate deaminase (AMPD) and containing several linked co-amplified genes (W,X,Y_1,Y_2), we found that independent cell lines can co-amplify different DNA sequences (Debatisse *et al.*, 1986). However novel joints are distributed non-randomly along the amplified DNA: many cluster within a 2.6-kb hotspot which is a mosaic of *Alu*-equivalent repeats and long A+T-rich DNA segments enriched in long pseudopalindromes (Hyrien *et al.*, 1987). The breakpoint corresponding to one of these novel joints was found at the centre of the best pseudopalindrome.

Here we report the analysis of another novel joint (NJ474) which appeared at the second selection step of a different cell line. This joint links two amplified units in a 'head-to-head' configuration, forming a large inverted duplication which is itself highly amplified. Large inverted duplications have been found to be associated with gene amplification in an increasing number of cases (Ford *et al.*, 1985; Ford and Fried, 1986; Saito and Stark, 1986; Nalbantoglu and Meuth, 1986; Looney and Hamlin, 1987), suggesting that generation of these structures might relate to the mechanism of DNA sequence amplification in eukaryotic cells (Passananti *et al.*, 1987). A detailed study of inverted duplications might therefore help us to understand these mechanisms. Nucleotide sequences at the inversion joint were determined to investigate the mechanism of formation of NJ474 and the basis for the recombinogenicity of the region in which it took place. The chromosomal location of the amplified copies was also determined and was found to be unchanged during the amplification process. We propose a model which accounts for intrachromosomal DNA amplification as a direct consequence of the formation of an inverted duplication.

Results

Cloning of the NJ474 junction

Chinese hamster fibroblasts which resist increasing doses of coformycin in the presence of adenine and azaserine overproduce AMPD, the target for coformycin, as a result of DNA sequence amplification (Debatisse *et al.*, 1984, 1986). HC4, $HC_{10}47$ and $HC_{50}474$ are first, second and third step mutants respectively which were selected stepwise from the wild-type GMA32 line.

The NJ474 junction, which is absent from the DNA of HC4, appeared in the second step mutant $HC_{10}47$ where it is amplified 40-fold. It is highly reamplified (up to 150-fold) in the third step mutant $HC_{50}474$ which concomitantly

©IRL Press Limited, Oxford, England

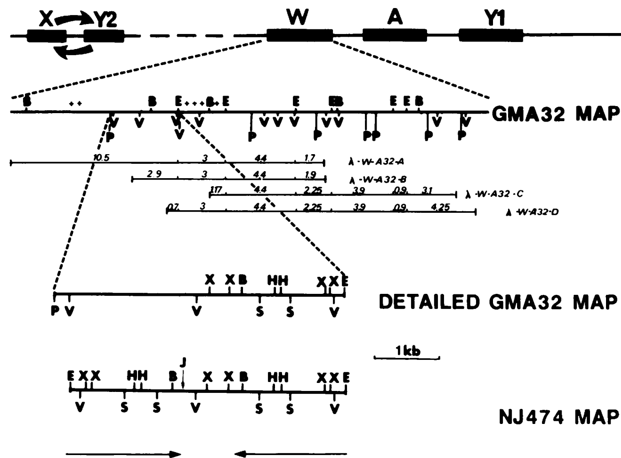


Fig. 1. Molecular clones used in this study and their position in the amplified domain. **Upper line**: a genetic map of the amplified domain. A: AMPD gene; W,X,Y_1,Y_2: co-amplified genes. The W–A–Y_1 linkage has been confirmed by a cosmid walk (Debatisse *et al.*, 1988). **Middle line:** restriction map of the wild-type (GMA32) W region (Hyrien *et al.*, 1987). E: *Eco*RI; B: *Bam*HI; P: *Pst*I; V: *Pvu*II. The positions of the overlapping phages and the sizes of their *Eco*RI fragments are shown under this line. **Bottom lines**: an enlargement of the relevant portion of the GMA32 map is aligned together with the NJ474 map. E, B, P, V as above. X: *Xba*I; S: *Sst*I; H: *Sph*I. The position of the junction is indicated (J).

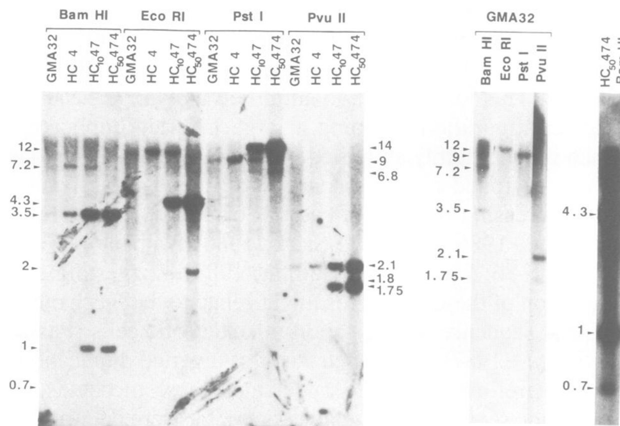


Fig. 2. Southern blot analysis of wild-type (GMA32) DNA and DNA of mutants HC4, $HC_{10}47$ and $HC_{50}474$. Genomic DNAs were digested as indicated, separated on a 0.6% agarose gel, transferred to Hybond-N and probed with pBSJ474 (the insert of which is the 4.3-kb *Eco*RI fragment containing the NJ474 junction) after repeated sequences were competed out as described in Materials and methods. Note that lane $HC_{50}474$/*Bam*HI was loaded with about one-third (3 μg) the amount of DNA loaded in other lanes (10 μg). The middle panel is a composite figure from various exposures of other blots where the GMA32 fragments were better seen. The rightmost lane is an overexposure of an $HC_{50}474$/*Bam*HI digest electrophoresed through 1% agarose. A *Bst*eII digest of λ DNA was used as a size marker.

suffered a mass deletion of pre-existing amplified units (Debatisse *et al.*, 1986). The junction was obtained on a 4.3-kb *Eco*RI fragment cloned in λgt10. Figure 1 shows a restriction map of this fragment aligned along the wild-type map. It can be seen that the 4.3-kb *Eco*RI fragment from $HC_{50}474$ contains an inverted duplication organized around a non-palindromic centre: the map of the $HC_{50}474$ DNA region from the central *Pvu*II site up to the right side *Eco*RI site is identical to the wild-type map; a DNA region including

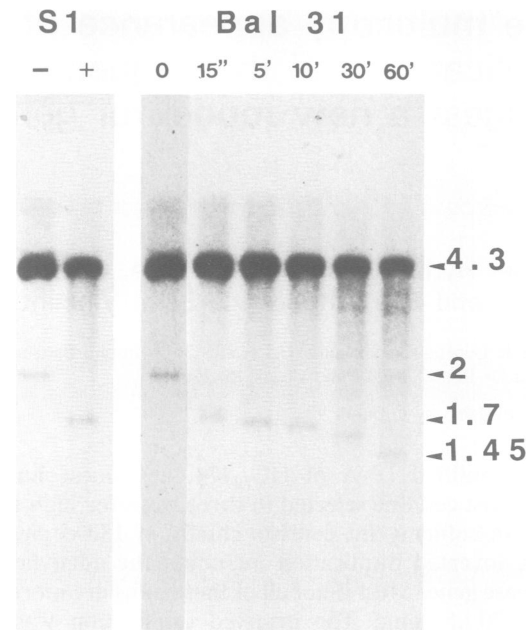


Figure 3. Digestion of $HC_{50}474$ inverted duplicated DNA by S1 and by *Bal*31. **Left**: $HC_{50}474$ DNA was digested with *Eco*RI, precipitated, resuspended in S1 buffer with (+) or without (−) S1 nuclease for 1 h. **Right**: $HC_{50}474$ DNA was digested wtih *Bal*31 for the times indicated, purified, and cut with *Eco*RI. The samples (5 μg) were fractionated in a 1% agarose gel, blotted and hybridized with pBSJ474 after repeated sequences were competed out as described in Materials and methods. A *Bst*eII digest of λ DNA was used as a size marker.

the wild-type 1.6-kb *Bam*HI–*Eco*RI fragment has been duplicated and recombined in inverted orientation upstream of the central *Pvu*II site, thus creating a novel junction between the left *Bam*HI site and the central *Pvu*II site. The central segment which was not duplicated during the process is now flanked by both arms of the duplication.

To rule out the possibility that this structure arose as a cloning artefact and to demonstrate its presence in cellular DNA, DNAs from GMA32 and its derivatives HC4, $HC_{10}47$ and $HC_{50}474$ were digested with various restriction endonucleases and probed with the 4.3-kb *Eco*RI fragment (pBSJ474) for the presence of the fragments predicted from the restriction map (Figure 2). The predicted wild-type fragments (3.5-kb and 7.2-kb *Bam*HI, 12-kb *Eco*RI, 8.8-kb *Pst*I, 2.1-kb and 1.75-kb *Pvu*II) are present in GMA32 DNA. These include fragments which are incorporated either in the first step units or in the inverted duplication without being rearranged (3.5-kb *Bam*HI, 2.1-kb *Pvu*II) and fragments which encompass the site of inversion and hence are incorporated without rearrangement only in the first step units (7.2-kb *Bam*HI, 12-kb *Eco*RI, 8.8-kb *Pst*I and 1.75-kb *Pvu*II). As expected, the former are amplified ~10-fold in HC4, 50-fold in $HC_{10}47$ and 150-fold in $HC_{50}474$, while the latter are amplified 10-fold in HC4 and $HC_{10}47$ DNAs and return to a near wild-type level in $HC_{50}474$ DNA. The fragments bearing the NJ474 junction are absent from GMA32 and HC4 DNAs, and display the expected size (1-kb *Bam*HI, 4.3-kb *Eco*RI, 14-kb *Pst*I, 1.75-kb *Pvu*II) in $HC_{10}47$ and $HC_{50}474$ DNAs where they are amplified 20- and 75-fold respectively. The novel *Pvu*II fragment co-migrates with one of the wild-type *Pvu*II fragments (1.75 kb); this explains why the 1.75-kb *Pvu*II band is stronger than the 2.1-kb *Pvu*II band in the $HC_{10}47$ and $HC_{50}474$

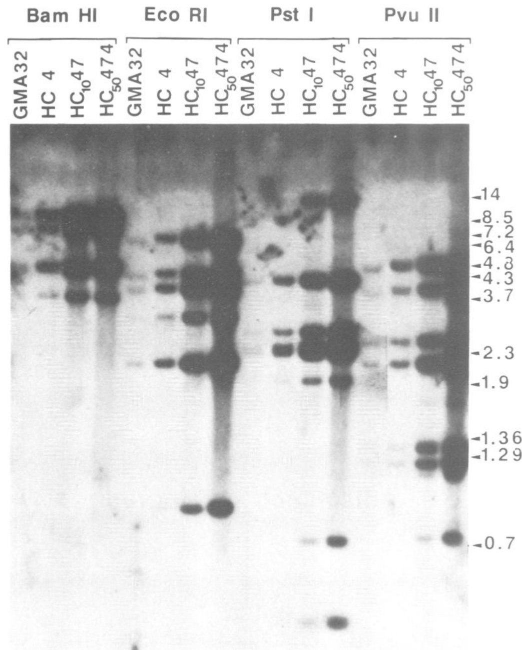


Fig. 4. Analysis of the amplified inverted duplicated DNA. The same blot as in Figure 2 (left) was probed with DNA of the entire λ-W-A32-D phage clone after repeated sequences were competed out as described in Materials and methods. A *Bst*eII digest of λ DNA was used as a size marker.

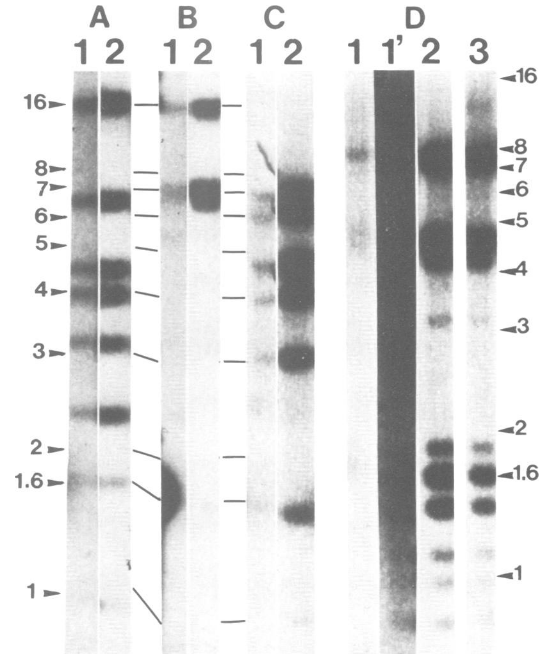


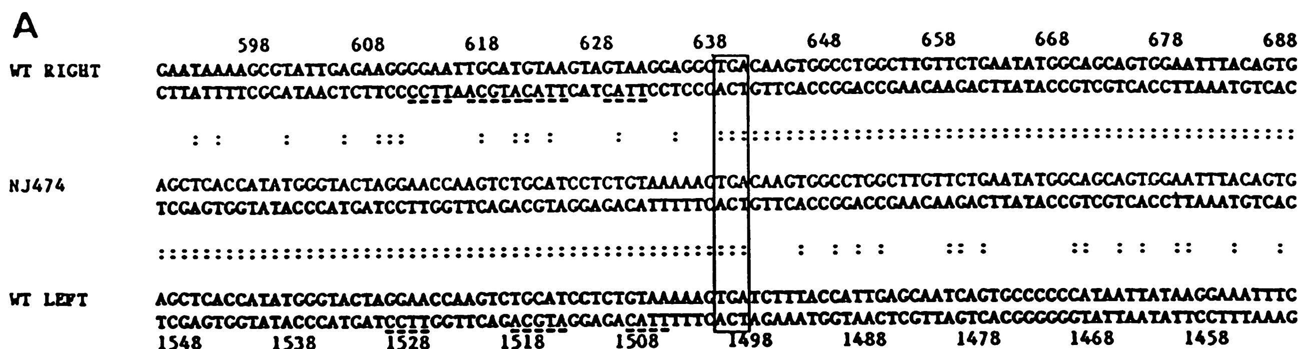
Fig. 5. Long range analysis of the amplified domain. *Eco*RI digests (10 μg) of GMA32 (**lanes 1**), $HC_{50}474$ (**lanes 2**) and $HC_{50}611$ (**lane 3**) were separated through 0.7% agarose gels, transferred to aminothiophenol paper and hybridized with entire cosmids covering 120 kb of amplified DNA downstream of the NJ474 joint (Debatisse *et al.*, 1988) after repeated sequences were competed out as described in Materials and methods. **A**: cosmid 61/W/1. **B**: cosmid 61/A/3. **C**: cosmid $56/Y_1/5$. **D**: cosmid $56/Y_1/2$. Lane D 1′ is an overexposure of lane D1. The BRL 1 kb ladder was used as a size marker.

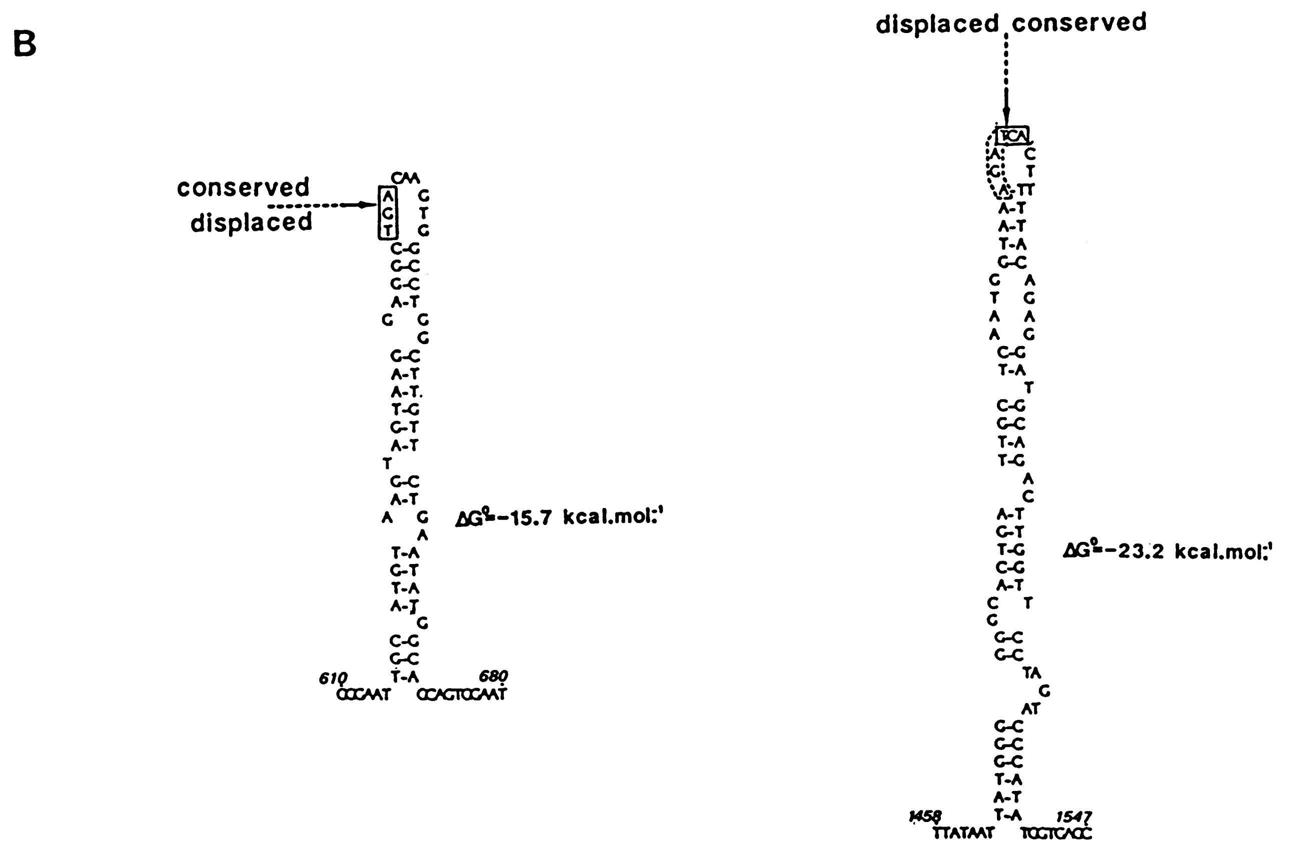
lanes—where the amplified novel joint is present—while it is fainter than the 2.1-kb *Pvu*II band in the GMA32 and HC4 lanes. These blotting data all confirm that the NJ474 novel joint is at the centre of the inverted duplication shown in Figure 1.

The probe detects in the $HC_{10}47$ and $HC_{50}474$ DNAs an additional minor band. The size of this band (0.7-kb *Bam*HI, 2-kb *Eco*RI, 6.8-kb *Pst*I, 1.8-kb *Pvu*II) corresponds for each restriction endonuclease to the size that would display a fragment ending approximately at the centre of the duplication (the 0.7-kb *Bam*HI band is better reproduced on the rightmost lane of Figure 2). This suggests that this band is not a different novel joint but arose from the presence of a free end mapping at the centre of a few copies of the inverted duplication. This is supported by experiments presented in Figure 3. When $HC_{50}474$ DNA is first cut with *Eco*RI and then digested with S1 nuclease, the 2-kb *Eco*RI band is reduced to 1.7-kb while the 4.3-kb band does not change in size. The distance between the end of the inverted duplicated sequences and the first encountered duplicated *Eco*RI site is 1.7 kb; this suggests that the 1.7-kb fragment released by S1 digestion of the 2-kb band corresponds to the stem of a stem-loop structure formed by intrastrand pairing of the inverted duplicated DNA. Further support for this interpretation is provided by *Bal*31 digestion experiments. *Bal*31 is a double-strand exonuclease which degrades simultaneously both the 3′ and the 5′ termini of duplex DNA; it also has endonuclease and exonuclease activity against single-stranded DNA. Digestion of a stem-loop structure by *Bal*31 is therefore expected to proceed in two steps: first, elimination of the single-stranded loop as with S1 nuclease and second, progressive shortening of the remaining stem from the double-stranded free end thus created. DNA from $HC_{50}474$ was treated for various times with *Bal*31 then analysed by *Eco*RI digestion and Southern hybridization with pBSJ474 as a probe. The expected two-step shortening process is observed specifically for the 2-kb *Eco*RI band while the 4.3-kb band (the interstrand duplex form of the centre of the inverted duplication) remains unchanged in size. The gradual decrease in intensity together with the progressive smearing down of this 4.3-kb band is explained by statistical digestion of the whole cellular DNA from the randomly distributed free ends generated during DNA preparation.

Size and amplification level of the inverted duplication

To estimate a minimum size for the inverted duplication, DNAs from GMA32 and its derivatives HC4, $HC_{10}47$ and $HC_{50}474$ were digested with various restriction endonucleases and analysed by Southern blotting using λ-W-A32-D as a probe (Figure 4). λ-W-A32-D (Figure 1) has a 700-bp overlap with the 4.3-kb insert of pBSJ474 and extends 20 kb downstream toward the centre of the amplified unit. Except for the novel fragment containing NJ474, which is amplified in $HC_{10}47$ and $HC_{50}474$ but absent from HC4, all the fragments detected by this probe in the DNA of the amplified cell lines HC4, $HC_{10}47$ and $HC_{50}474$ are present in the wild-type GMA32 DNA. The smaller ones (e.g. 0.9-kb *Eco*RI), which are not always visible in the GMA32 lanes at this exposure of the blot, are nevertheless present in the restriction map established from the GMA32 genomic clones. A few bands seem not to be in molar yield (e.g. the 3.1-kb *Eco*RI band is less dense than fragments of lower mol. wt). Some of these correspond to the ends of the probe which cover them only partially. Also since competition with total DNA is used to remove the repetitive sequences of the probe the intensity of the signal is not necessarily proportional to the mol. wt of the band.





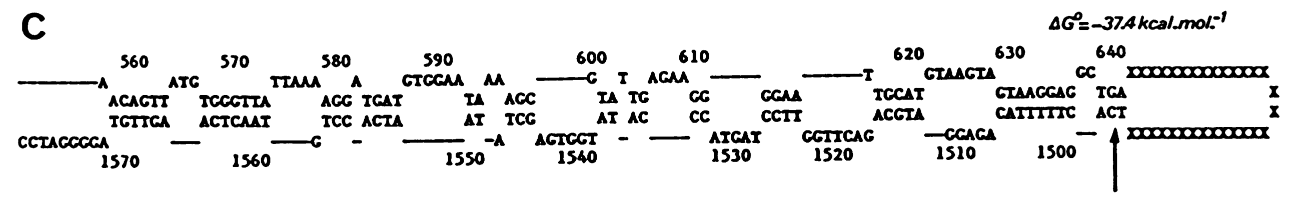


Fig. 6. Nucleotide sequence of the NJ474 joint and its wild-type W counterparts. **A**: alignment of NJ474 and wild-type W counterparts. Numbering represents the position of sequences in the wild-type W region (Hyrien *et al.*, 1987). The sequence to the left of the joint is the complement of the wild-type sequence starting at 1500. Symbol (:) indicates homologous base pairs. The three homologous base pairs at the joint between the recombining molecules are boxed. Additional, patchy homologies are underlined in dotted lines. **B**: possible stem-loop structures at the NJ474 breakpoints. The stem-loop structures shown are those predicted for fragments 610–680 and 1458–1547 of W DNA by the program of Zucker and Stiegler (1981). The homologous inverted trinucleotides at the breakpoints are boxed. One of these coincides with a topo 1 cleavage consensus sequence (boxed in dotted lines). **C**: a possible stable hybrid aligning the parental recombination sites. Numbering as in A. This structure is the most thermodynamically favourable predicted by the folding program of Zucker and Stiegler. The 30 Xs are inert in this program and serve to place the two parental sequences on the same strand to allow folding. The position of the recombination sites is indicated by an arrow.

In addition, *Eco*RI digests of GMA32 and $HC_{50}474$ DNAs were similarly analysed with cosmids extending up to 120 kb from the NJ474 junction and including the AMPD gene itself (Debatisse *et al.*, 1988); we similarly failed to detect any novel joint in the $HC_{50}474$ DNA within the region being probed (Figure 5). We do not believe that the 3.1-kb *Eco*RI fragment revealed by cosmid $56/Y_1/2$ in the DNA of $HC_{50}474$ but not seen in DNA of GMA32 represents a novel joint, because this fragment—like the nine other fragments revealed by cosmid $56/Y_1/2$—is seen in all our six highly amplified cell lines, including for example $HC_{50}611$ (Figure 5D, lane 3) which was derived independently of $HC_{50}474$. If this fragment was a novel joint or any other rearrangement occurred in the $HC_{50}474$ amplified DNA, it could not be detected in any independently derived cell line. More likely this 3.1-kb *Eco*RI fragment is present in the wild-type DNA somewhere in the amplified domain and becomes visible only upon amplification because it cross-hybridizes faintly with a part of cosmid $56/Y_1/2$.

If different units end at different endpoints within the region being probed, then both a novel fragment containing the endpoint and a lower degree of amplification of fragments downstream from this endpoint should be observed. Neither is visible on any of the blots. Of course when only an *Eco*RI digest has been probed, we cannot absolutely rule out that a small subfraction of the amplified units end within the region probed; the novel joint could remain undetected either because of too little homology to the probe or because of co-migration with a highly amplified fragment. The whole set of data however indicates that most, if not all, of the inverted duplications extend over the region probed. This means that the size of the inverted duplication is at least 2×120 kb, and that no other rearrangement is detected in the region covered by this large inverted duplication; since the DNA upstream of NJ474 junction is almost unamplified in $HC_{50}474$ (Debatisse *et al.*, 1986), the amplified DNA of $HC_{50}474$ consists chiefly of 150 copies of this large inverted duplication.

Nucleotide sequences at the junction and corresponding breakpoints

The NJ474 junction was subcloned in M13 and sequenced (Figure 6A). The $HC_{50}474$ sequence to the right of the junction is identical to that of the wild-type to the right of nucleotide 638. The $HC_{50}474$ sequence to the left of the junction is identical to the complementary strand of the wild-type to the right of nucleotide 1497. Sequences from 638 to 1499 are not duplicated but end in a pair of inverted trinucleotides (638–640 and 1497–1499) which are already present in the wild-type DNA and from which the duplication starts. Therefore a 3-bp homology is present at the crossover sites.

The wild-type sequences outside—but not inside—the central fragment 638–1499 present additional, patchy homologies also in inverted orientation (Figure 6A). The folding program of Zuker and Stiegler (1981) was used as described by Bullock *et al.* (1984) to predict whether a heteroduplex could form between these segments arranged in inverted orientation and whether it could align the crossover sites. With fragments starting at the homologous trinucleotides, and extending outward from segment 640–1497, the predicted heteroduplex, at variable window lengths of up to 85 nucleotides, placed the crossover sites in correct alignment. The resulting predicted structure at a window length of 85 nucleotides is shown in Figure 6C. Of course any two random 85-bp sequences can fold in an apparently similar way. However the structure shown in Figure 6C differs from those obtained with random sequences in several regards: (i) when random sequences are used, base pairing occurs as well within each 85-bp sequence as between the two 85-bp sequences. In contrast, with the recombining sequences analysed here, all base pairing occurs between the two 85-bp sequences. (ii) Even if pairing within each random sequence is forbidden to allow only pairing between the two random sequences (thus maximizing the contribution of base pairing towards stable alignments of the two molecules), the average stability of 30 heteroduplexes thus predicted ($\Delta G^o = -28.13 \pm 5.4$ kcal/mol) is significantly ($P < 0.05$) lower than the stability of the structure shown in Figure 6C ($\Delta G^o = -37.4$ kcal/mol). (iii) The structure shown in Figure 6C pairs the homologous trinucleotides at which the recombination event took place. For these reasons we think that the patchy homologies between the parental molecules could have participated in defining the crossover sites during the recombination process.

Particular features are seen at these crossover points. A sequence that matches the consensus for topoisomerase 1 (topo 1) cleavage sites (5′ A/T C/G A/T T; Been *et al.*, 1984) is found at the left crossover site (AGAT in position 1494–1497). This site is also adjacent to a small A+T-rich region (1501–1506). More interesting, this crossover point falls inside the B2 *Alu*-like repeat lying in position 1476–1690. Lehrman *et al.* (1987) have recently reviewed human mutations involving breakage of DNA within *Alu* repeats: eight out of 11 breakpoints fall between two conserved blocks of sequence that are thought to function as promoter elements for RNA polymerase III (Paolella *et al.*, 1983). These two blocks can also be recognized in rodent B2 repeats (Rogers, 1985) and since here they are located in positions 1479–1490 and 1517–1528, the recombination event in NJ474 follows the same pattern as in human *Alu* recombination.

The secondary structures which are predicted at both parental crossover sites are imperfect hairpins in which the crossover points are situated in the terminal loops (Figure 6B). The stability of each of these hairpin loops is significantly ($P < 0.05$) higher than those of predicted secondary structures for 30 randomized sequences of the same length and base composition (compare -15.7 kcal/mol with -9.7 ± 3.6 kcal/mol and -23.2 kcal/mol with -16.8 ± 3.5 kcal/mol). This statistical significance suggests, but does not prove, that these structures are biologically meaningful. In this regard, it is interesting that a secondary structure very similar to that of fragment 1458–1547 (in the B2 repeat) has been found at an equivalent position in the rat B2 element *R.dre.1* (Rogers, 1985).

Chromosomal location of the amplified inverted duplication

Metaphase spreads of HC4 and $HC_{10}47$ cells were hybridized *in situ* with a W cDNA probe. This probe recognizes the wild-type W region as well as the NJ474 novel joint (Debatisse *et al.*, 1986). In both cell lines the amplified copies are found at a single position on the long arm of the longest metacentric chromosome (Figure 7A). The karyotypes of the two cell lines were prepared by the BrdU–Hoechst 33258–Giemsa (RBG) method (Camargo and Cervenka, 1980) to allow comparison of their chromosomes. The longest chromosome of HC4 is the same as the longest

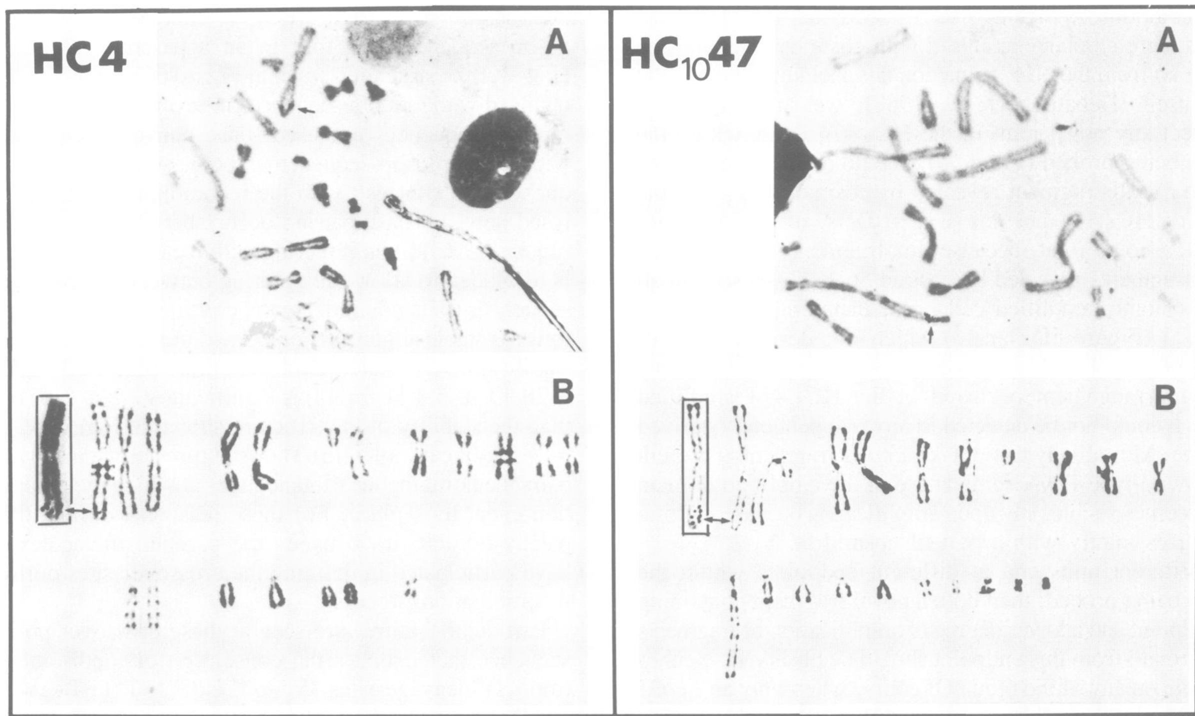


Fig. 7. Localization of the amplified DNA sequences in HC4 (left) and $HC_{10}47$ (right). **A**: metaphase spreads hybridized *in situ* with a W cDNA probe. In both lines the grains are seen on the long arm of the longest metacentric chromosome (arrows): three grains are visible on the HC4 chromosome (6 weeks exposure) and six grains are visible on the $HC_{10}47$ chromosome (3 weeks exposure). In both cell lines, two grains or more were observed at the same position in at least 25 different metaphases, and never at any other location. **B**: the panels show a chromosome (boxed) cut out from one of the hybridized metaphase spreads that contain the cluster of silver grains and a representative RBG-banded karyotype from the same cell line. A double arrow indicates the cluster of grains and the corresponding location on the RBG-banded chromosome.

chromosome of $HC_{10}47$ and the position of the grains is the same in HC4 and $HC_{10}47$ (Figure 7B). No extrachromosomal structures such as double minute chromosomes were seen. We frequently observed abnormal chromosomes or chromosome duplications in both cell lines but no grains were ever seen on them. Therefore most, if not all, of the amplified copies reside in these two cell lines at a unique and identical chromosomal location.

Discussion

We previously reported (Hyrien *et al.*, 1987) the identification and sequencing of a 2.6-kb recombinogenic DNA region within a DNA sequence which is amplified in AMPD-overproducing Chinese hamster fibroblasts. Numerous novel joints were formed in this region during independent amplification events. We also reported the structure of one of these novel joints, NJ551,562, which linked together two unrelated, distant sequences. Here we have analysed another novel joint, NJ474, the formation of which occurred independently of that of NJ551,562 yet in the same short DNA region. NJ474 links two amplified units in a large ($>2 \times 120$ kb) inverted duplication with an 862-bp asymmetric, non-duplicated central fragment at its centre. This conclusion can be reached because (i) the restriction map of the 4.3-kb *Eco*RI fragment containing the novel joint shows that the NJ474 joint is at the centre of an inverted duplication (Figure 1); (ii) all the restriction fragments predicted from this map and characteristic of this structure are indeed found in the cellular DNA of $HC_{10}47$ and $HC_{50}474$ by blot analysis (Figure 2); (iii) no rearrangement has been detected downstream of the joint in 120 kb of the duplicated region (Figures 3 and 4); (iv) the DNA upstream of the NJ474 joint is not amplified in $HC_{50}474$ (Debatisse *et al.*, 1986). We also showed that a small fraction of this inverted duplicated DNA in the DNA preparation appears to be arranged in hairpins formed by intrastrand pairing of the inverted repeats. A similar observation has been made for another amplified inverted duplication in the CAD system (O.Brison, personal communication). Such structures may exist *in vivo* (for example as cruciforms extruded from the intrachromosomal palindromes or as free extrachromosomal hairpins), but they could also represent an artefact of DNA preparation since it has been shown in *Escherichia coli* that DNA extraction can artificially trigger cruciform extrusion from palindromic sequences cloned in pBR322 (Courey and Wang, 1983). We are currently investigating this point but this is not the focus of the present paper. Anyhow, the detection of these hairpin structures directly confirms the structure of the inverted duplication.

This example adds to the growing evidence that the amplified DNA is frequently arranged in inverted duplications in mammalian cells (Ford *et al.*, 1985; Ford and Fried, 1986; Saito and Stark, 1986; Nalbantoglu and Meuth, 1986; Looney and Hamlin, 1987); it also shows that, at least in the system studied here, different kinds of arrangements of amplified DNA can be produced in the same DNA region, by separate events. These different arrangements might be two possible outcomes of the same recombination mechanism. In fact there is a similarity between the crossover sites of both NJ474 and NJ551,562. The NJ551,562 breakpoint lies at the centre of a 70-bp palindrome which could form a significantly stable stem-loop structure. Likewise, both NJ474 breakpoints are situated at the top of potential stem-

loop structures. Also, one of the NJ474 breakpoints is, like the NJ551,562 breakpoint, associated with a putative topo 1 cleavage site. Bullock *et al.* (1985) have proposed a model for illegitimate recombination based on the frequent association of SV40 excision crossover sites with topo 1 cleavage sites. A related mechanism could have been implicated in the formation of both NJ474 and NJ551,562, although it is doubtful that topo 1 could cleave these sites at hairpin loops formed *in vivo*, since topo 1 has been reported to cleave single-stranded DNA *in vitro* only in regions with a potential for intramolecular base pairing (Champoux *et al.*, 1984). Alternatively, the W hotspot could be a preferred site for different types of recombination processes, one of these leading to the formation of inverted duplications. In fact, many of the features described in the first two reports of the sequences at and around an inverted duplication associated with a gene amplification event (Nalbantoglu and Meuth, 1986; Passananti *et al.*, 1987), such as no unusual sequence at the joint, joint formed by a simple illegitimate recombination event, presence of hairpin structures in parental sequences and also a small A+T-rich region at one of the parental sequences, are also present at the NJ474 joint. It is interesting to note that one of the four *Alu*-equivalent repeats present in the W recombinational hotspot (Hyrien *et al.*, 1987) has been involved in the formation of the NJ474 inverted duplication and that the breakpoint in this repeat is at a position equivalent to that which in human *Alu* recombination is frequently involved in genomic rearrangements (Lehrman *et al.*, 1987).

Although the generality of the association of large inverted duplications with mammalian DNA amplification appears now firmly established, the mechanism of their formation and their role in the development of amplified arrays is unclear. The inverted duplication analysed here is absent from the amplified DNA of the first-step mutant line HC4. However, at the following selection step, line $HC_{10}47$ acquired 40 copies of this duplication; no other rearrangement in the DNA amplified during the second step has been found. This suggests that the generation of the inverted duplication was a primary event in amplification. Also, the amplified region resides in the same chromosomal location as before this amplification event. At the third amplification step, no new DNA rearrangement was detected, but reamplification of the same large inverted duplication to give up to 150 copies was observed, together with elimination of amplified DNA formed during the first step. It has been generally accepted that amplified novel joints arise from amplification of novel joints formed as single copy during earlier steps (Ardeshir *et al.*, 1983; Federspiel *et al.*, 1984). We indeed observed this situation for another joint in the AMPD system (Debatisse *et al.*, 1986; Hyrien *et al.*, 1987; Debatisse *et al.*, 1988). But this is clearly not the case for the NJ474 joint; indeed, generation of inverted duplications may frequently be coupled with immediate amplification, as suggested by the properties of another joint under study in the AMPD system and by those of inverted duplications described in other systems (Ford *et al.*, 1985; Saito and Stark, 1986; Nalbantoglu and Meuth, 1986; Looney and Hamlin, 1987).

Different mechanisms of recombination can generate inverted duplications. In lower eukaryotes, during amplification of rRNA genes in *Tetrahymena thermophila* (Yao *et al.*, 1985) and the generation of palindromic repetitions of *rho*$^-$ mitochondrial DNA in yeast (Sor and Fukuhara, 1983), the pairing of inverted repeats present in the wild-type DNA seems to promote intramolecular recombination and formation of a long hairpin molecule which becomes the palindromic molecule after one round of replication. On the other hand the formation of head-to-head dimers of linearized plasmids transfected in yeast cells is a bimolecular process which involves synapsis between the parental molecules and very short patches of homology already present in an inverted orientation in the original linear plasmid molecule (Kunes *et al.*, 1986). The most favoured model of gene amplification in mammalian cells is the 'onion skin' model (Bullock and Botchan, 1982). An onion skin structure generated by muliple, unscheduled initiations of DNA replication at a fixed origin in the chromosome could conceivably be resolved into a head-to-head tandem array of amplified units by a bimolecular recombination process similar to that described in yeast. However, a single amplification step is expected in this case to generate a multiplicity of single copy inversion joints, rather than the highly amplified single joint which has been observed here and in other systems. In two of these cases (Saito and Stark, 1986; Looney and Hamlin, 1987) the amplified novel joint could have arisen in multiple steps because highly resistant cell lines obtained through multiple selection steps were analysed. But here and in the other two cases (Ford *et al.*, 1985; Nalbantoglu and Meuth, 1986) an apparently single step of amplification was analysed. In the onion skin model, an amplified novel joint could be created in a single step only if recombination occurs before replication is complete, and it would be very surprising if it were as highly amplified as 40-fold.

Another possibility would be that the joint was already present as a single copy in a small fraction of the HC4 cells—and therefore undetected in the Southern blot—and was amplified in a second step by, for example, an onion skin replication. Although HC4 and $HC_{10}47$ cells were isolated as clones, the necessary delay between cell line selection and DNA preparation precludes direct exclusion of this hypothesis. However, such a possibility would again lead to heterogeneity of the amplification units with different units ending at different points. We failed to detect any such heterogeneity over 120 kb of amplified DNA (although we could have missed a small amount of amplified units ending at different points). Furthermore, Looney and Hamlin (1987) have found that amplified inverted duplications are organized as a homogeneous tandem array of inverted duplications identical to each other up to both ends in a methotrexate-resistant cell line. Finally, it has to be emphasized that the inverted duplication described by Nalbantoglu and Meuth (1986) was found to be amplified although no selection for amplification was ever applied; this suggests again that the formation and the amplification of inverted duplications are mechanistically coupled and do not occur in separate steps.

A unified model explaining through a single mechanism both the formation of an inverted duplication and its amplification without additional replicon firing and without chromosomal breakage can be proposed on the basis of our sequence data for the NJ474 inversion joint. The potential for hairpin structures and heteroduplex formation at the crossover sites suggests their involvement at some stage in the recombination process. These structures may form when a replication fork progresses through this region, as strongly supported by the recent observations of Shurvinton *et al.* (1987) and Leach

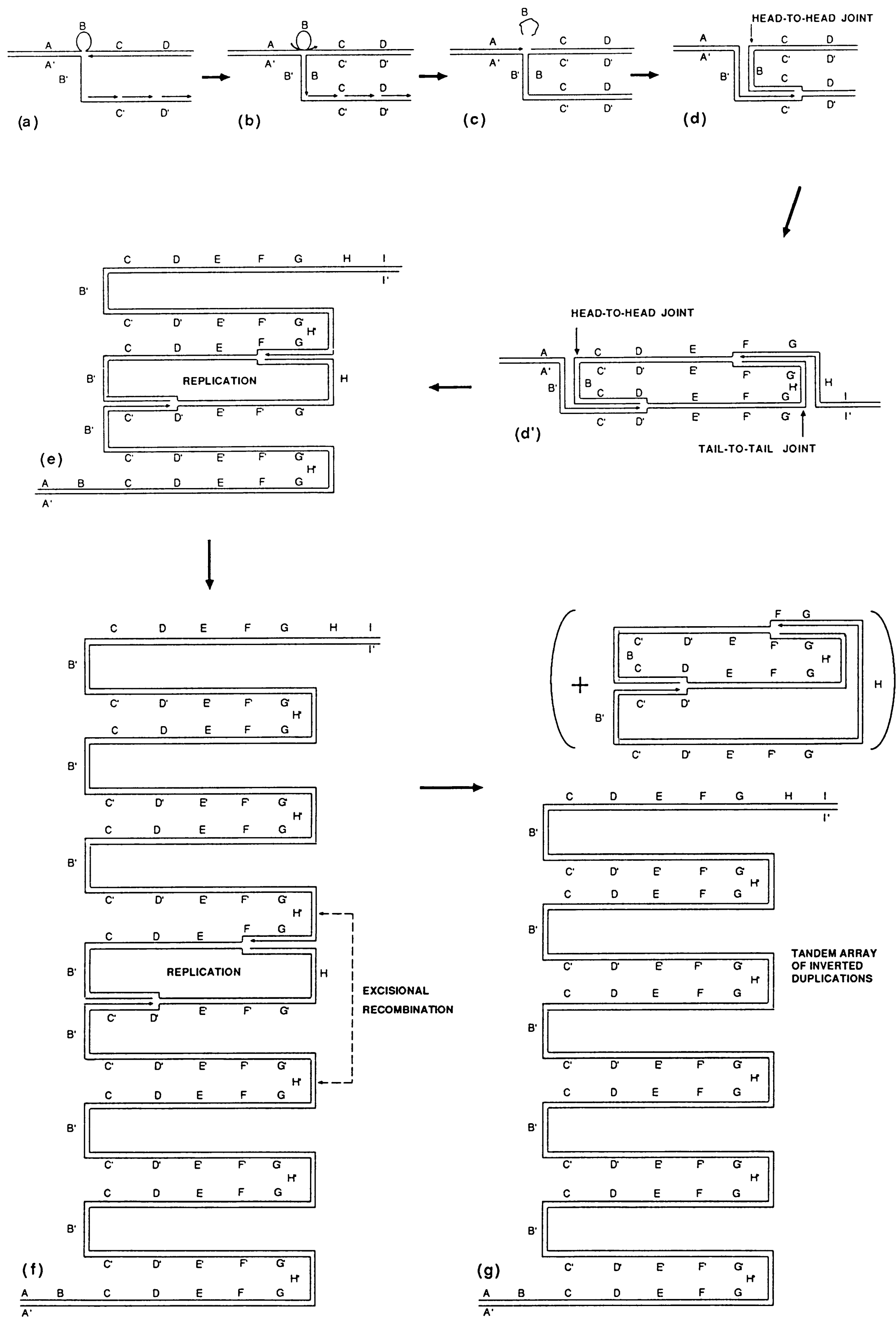


Fig. 8. A model for inverted duplication formation and intrachromosomal amplification. **Steps a–d**: formation of a head-to-head (C′/BC) joint by copy-choice recombination. **Step d′**: structure obtained if a head-to-head (C′/BC) joint and a tail-to-tail (G/H′G′) joint are formed by the same events occurring at both ends of the replication bubble. **Steps d′–f**: amplification of the resulting inverted duplication by intrachromosomal double rolling circle replication. **Step g**: resolution into a chromosomal tandem array of inverted duplications with possible release of an extrachromosomal replicating circle. See text for additional details.

and Lindsey (1986) on the replication-dependent instability of palindromes in λ DNA. As pointed out by Glickman and Ripley (1984) there is abundant *in vitro* evidence that DNA polymerases behave aberrantly when confronted wtih a template containing hairpins (Sherman and Gefter, 1976; Ikoku and Hearst, 1981; Kaguni and Clayton, 1982). An inverted duplication can be generated (Figure 8a–d) if replication switches strand and proceeds around the replication fork (Nalbantoglu and Meuth, 1986). Such turnaround events have indeed been observed at palindromic sites in bacteria (Backman *et al.*, 1978, and references therein). The presence at position 1501–1506 of a small A+T-rich region [also found at the inversion joint described by Passananti *et al.* (1987) and Nalbantoglu and Meuth (1986)] could have helped the 3′ OH primer terminus to switch template by facilitating its dislocation from its initial template strand. This would have allowed in turn the possibility of inappropriate base pairing between the 3′ end of the primer and the inverted patchy homologies present on the other template strand. An asymmetric, non-duplicated fragment is found at the centre of the duplication: this is the general structure of amplified inverted duplications (Ford *et al.*, 1985; Saito and Stark, 1986; Nalbantoglu and Meuth, 1986; Looney and Hamlin, 1987) and it implies that the other copy of this central sequence must be looped and excised at the site where replication turns around the fork (Figure 8b). The formation of this loop can indeed be understood if the strand switching event does not occur at identical positions on the two strands but is promoted by inverted patchy homologies, located at staggered positions, such as those we observed. It is known that single-stranded loops in heteroduplex DNA transfected into mammalian cells can be repaired efficiently by an excision–repair process which favours the strand without the loop (Weiss and Wilson, 1987). The loop created by the strand switching event could be excised as part of such a process. Whatever the exact mechanism, the site of copy choice at an inversion joint coincides with a break on the parental DNA strand which suffered excision (Figure 8b and c). We propose that the resulting structure (8c), instead of being resolved by a break in the other unreplicated parental DNA strand (Nalbantoglu and Meuth, 1986), can prime a replication step which now proceeds towards the already duplicated DNA (8d). This could represent the outcome of an excision–repair process gone awry due to the abnormal structure of the repair template (8c); or this structure (8c) could be seen as a *bona fide* replication fork by the replication complex which just created it. If a head-to-head and a tail-to-tail joint are formed by template switching at both ends of a replication bubble (8d′), two replication forks at the two triple junctions engage into the inverted duplicated DNA and follow each other in an indefinite spiral (e and f), yielding a multimeric tandem array of inverted duplications. The simplest hypothesis accounting for resolution of this structure is excisional recombination of the replication forks (f), leaving a tandem array of identical inverted duplications in the chromosome, without relocation (g). In fact, the amplified DNA of HC4 and $HC_{10}47$ is found by *in situ* hybridization to be chromosomally integrated and at the same chromosomal position in both cell lines. Excisional recombination will liberate a replicating circle if the recombination event is reciprocal (g). While one can argue that the presence of amplified DNA in HC4 may favour the reintegration of extrachromosomal DNA, we rather think that the identical location of amplified DNA in both cell lines reflects intrachromosomal DNA amplification.

Passananti *et al.* (1987) recently proposed another model for the amplification of inverted duplications. The hypothesis is that, after a single round of local DNA replication, an inverted duplication is excised in a circular form; the inverted dimeric circle could then amplify according to the mechanism proposed by Futcher (1986) and experimentally confirmed by Volkert and Broach (1986) for copy number amplification of the 2μ yeast plasmid. This 6318-bp plasmid is bisected nearly equally by a pair of exact inverted repeats 599 bp long. A plasmid gene (*FLP*) encodes a site-specific recombinase which acts at a pair of short, defined sites in the repeats (*FRT* sites) to invert the unique segments about the repeats. In addition, the single origin of replication of the plasmid is located very close to one of the repeats. Futcher proposes that soon after replication of this repeat, *FLP* recombination takes place between one of the newly replicated *FRT* sites and the yet unreplicated *FRT* site; this inverts the theta-form replication intermediate into a double rolling circle where the two replication forks, which previously converged, now follow each other indefinitely and produce a multimeric array of the original plasmid. Additional *FLP* recombination can resolve this structure and liberate monomer plasmids. This model of amplification requires that exactly one of the diverging replication forks has passed an *FRT* site at the time of recombination. The relative location of the repeats and the replication origin on the 2μ plasmid, together with the action of a site-specific recombination system, maximize the probability that these requirements are satisfied, but such an organization is not expected to be present frequently in an inverted dimeric circle excised from mammalian DNA. Adaptation of this model to mammalian DNA amplification also requires that DNA replication is initiated twice at the same locus: once to generate the duplication and once again—but at only one of the two duplicated origins—to amplify this circle. Finally, additional recombination events have to be postulated to account for intrachromosomal relocation of the amplified units. Nevertheless and despite these restrictions, this model is flexible enough to accommodate our data.

Both models (the '2μ-like' and the 'spiral') predict that the amplified DNA formed at a given step will be organized as a monotonous tandem array of inverted duplications identical to each other up to both ends. This uniform organization is indeed found in the case of a methotrexate-resistant cell line (Looney and Hamlin, 1987) and in the case of the major urinary proteins (MUP) gene family of the mouse which arose by generation and amplification of a 45-kb palindromic unit (Bishop *et al.*, 1985). In our system, this prediction will be checked by pulsed field gel electrophoresis analysis of the amplified units and cloning of the tail-to-tail joint. These models in their simplest form do not predict the existence of single copy novel joints nor the possibility that at least some novel joints are not of the 'head-to-head' type. Yet such structures are also found (Hyrien *et al.*, 1987; Looney and Hamlin, 1987; Debatisse *et al.*, 1988). They may represent secondary rearrangements occurring in the amplified DNA, or reflect integration of separated monomeric or oligomeric circles released by excisional recombination from the multimeric intermediate. Extrachromosomal elements associated with gene amplification could derive from such structures. Also, amplifications found in chromosomal lo-

cations different from the position of the original gene could be accounted for if the whole tandem of inverted duplications is released by the excisional recombination event and further reintegrated at one or several different chromosomal positions. An equally plausible explanation is that more than one mechanism of gene amplification operates in the same cells.

Materials and methods

Cell lines

The GMA32 (wild-type) cell line and its first (HC4), second ($HC_{10}47$) and third step ($HC_{50}474$) derivatives have been described previously (Debatisse *et al.*, 1986). For the isolation of $HC_{10}47$ cells, HC4 cells from the earliest possible transfer were propagated in selective medium containing 0.5 μg coformycin per ml (step I medium), then plated directly at 10^5 cells/dish in step II medium (5 μg of coformycin per ml). The medium was changed every 2–3 days. After 2 weeks an individual colony was recovered and propagated in selective medium. The same general procedure was applied for $HC_{50}474$ cell line selection. DNAs were prepared from the earliest possible transfers.

Phage cloning, plasmid subcloning and restriction mapping

The W full length cDNA probe and the cloning of the wild-type genomic W region in λ EMBL4 have been described (Hyrien *et al.*, 1987). The NJ474 junction is highly amplified (150×) in the cell line $HC_{50}474$. A Southern blot analysis with the 3.5-kb *Bam*HI fragment from λ-W-A32-A (not shown) indicated that the NJ474 junction is contained in a 4.3-kb *Eco*RI fragment. $HC_{50}474$ DNA was digested with *Eco*RI and the size fraction 4–4.5 kb was isolated on an agarose gel. DNA was extracted by the glass-bead method (Davis *et al.*, 1986) and ligated in the *Eco*RI site of λgt10 arms (Vector Cloning Systems). Recombinants were packaged *in vitro* with a Gigapack kit (Vector Cloning Systems) and plated on *E.coli* C600 *Hfl*. Screening with the 3.5-kb *Bam*HI fragment from λ-W-A32-A (Figure 1) yielded positive plaques at the expected frequency (1/500). Six recombinants were further purified and their DNA was prepared as described (Davis *et al.*, 1980); all contained the same 4.3-kb *Eco*RI insert. One was subcloned in Bluescribe $M13^-$ (Vector Cloning Systems) for restriction mapping and use as a probe (plasmid pBSJ474).

Nucleotide sequences

The sequence of the wild-type W hotspot has been reported elsewhere (Hyrien *et al.*, 1987). DNA sequences at and around the NJ474 junction were determined using the method of Sanger *et al.* (1977) with relevant subfragments of pBSJ474 subcloned in M13 vectors tg130 and tg131 (Kieny *et al.*, 1983). Nucleotide sequences were assembled and analysed with the 'Système d'Analyse de Séquences de l'Institut Pasteur' (SASIP) (Claverie, 1984).

Nucleic acid hybridization

Southern blots were performed with Hybond-N membranes (Amersham) as described by the supplier or with aminothiophenol paper (Figure 5) as described by Wahl *et al.* (1979). Repetitive sequences were competed out of genomic probes by reannealing with sonicated wild-type DNA to a Cot value of 30 (Figure 2, left) or 10 (other blots) before hybridization as described by Giulotto *et al.* (1986). The copy number of restriction fragments was estimated by comparing their densitometric signal with that of a known single copy fragment after correction for the amount of homology to the probe and normalization to the amount of DNA loaded onto the gel.

Analysis with S1 and Bal31 nucleases

*Eco*RI digested $HC_{50}474$ DNA was precipitated, resuspended at 500 μg/ml in S1 buffer (0.25 M NaCl, 30 mM potassium acetate, pH 4.6, 1 mM $ZnSO_4$, 5% glycerol) with or without Amersham S1 nuclease (1 U/μg), and incubated for 1 h at room temperature. Reactions were stopped in 50 mM Tris pH 8.0, 12 mM EDTA before electrophoresis and blotting.

$HC_{50}474$ DNA (500 μg/ml) was digested with *Bal*31 (5 U/μg) as recommended by the supplier (IBI) at 37°C. Aliquots (10 μl) were removed at the times indicated; the reactions were stopped with 30 mM EDTA and by phenol–chloroform extraction. The time 0 aliquot was removed before addition of the enzyme. The *Bal*31-digested DNA was precipitated and analysed by *Eco*RI digestion and blotting.

Chromosome banding and in situ hybridization

The slides were stained using the BrdU–Hoechst 33258–Giemsa (RBG) technique (Camargo and Cervenka, 1980) in parallel with or after *in situ* hybridization. *In situ* hybridizations were performed according to Wahl *et al.* (1982); the W cDNA probe was labelled by nick translation with $[^3H]$dCTP + $[^3H]$dTTP + $[^3H]$dATP (Amersham) to a specific activity of 1×10^7 c.p.m./μg and used at a concentration of 1 μg/ml.

Acknowledgements

We thank O.Brison for discussions and sharing unpublished results with us, the members of the Unité d'Informatique Scientifique for their help with the computer, F.Rouyer for help with computer graphics, D.Mazel for helpful suggestions, M.Berry for expert technical assistance, G.Lutfalla, F.Rouyer and P.Avner for critical reading of the manuscript and M.Mahérou for typing the manuscript. This work was supported in part by the Université P. et M.Curie, the INSERM (grant no. 861022), the Ligue Nationale Française contre le Cancer and the Fondation pour la Recherche Médicale Française.

References

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

Backman,K., Betlach,M., Boyer,H.W. and Yanofsky,S. (1978) *Cold Spring Harbor Symp. Quant. Biol.*, **48**, 69–76.

Been,M.D., Burgess,R.R. and Champoux,J. (1984) *Nucleic Acids Res.*, **12**, 3097–3113.

Bishop,J.O., Selman,G.G., Hickman,J., Black,L., Saunders,R.D.P. and Clark,A.J. (1985) *Mol. Cell. Biol.*, **5**, 1591–1600.

Borst,P., Van der Bliek,A.M., Van der Velde-Koerts,T. and Hes,E. (1987) *J. Cell. Biochem.*, **34**, 247–258.

Bullock,P. and Botchan,M. (1982) In Schimke,R.T. (ed.), *Gene Amplification.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 215–224.

Bullock,P., Forrester,W. and Botchan,M. (1984) *J. Mol. Biol.*, **174**, 55–84.

Bullock,P., Champoux,J.J. and Botchan,M. (1985) *Science*, **230**, 954–958.

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

Camargo,M. and Cervenka,J. (1980) *Hum. Genet.*, **54**, 47–53.

Champoux,J.J., McCoubrey,W.K. and Been,M.D. (1984) *Cold Spring Harbor Symp. Quant. Biol.*, **49**, 435–442.

Claverie,J.M. (1984) *Nucleic Acids Res.*, **12**, 397–407.

Courey,A.J. and Wang,J.C. (1983) *Cell*, **33**, 817–829.

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

Davis,L.G., Dibner,M.D. and Battey,J.F. (1986) *Basic Methods in Molecular Biology.* Elsevier Science Publishing, Amsterdam.

Debatisse,M., Robert de Saint Vincent,B. and Buttin,G. (1984) *EMBO J.*, **3**, 3123–3127.

Debatisse,M., Hyrien,O., Petit-Koskas,E., Robert de Saint Vincent,B. and Buttin,G. (1986) *Mol. Cell. Biol.*, **6**, 1776–1781.

Debatisse,M., Saito,I., Buttin,G. and Stark,G.R. (1988) *Mol. Cell. Biol.*, **8**, in press.

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

Ford,M., Davies,B., Griffiths,M., Wilson,J. and Fried,M. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3370–3374.

Ford,M. and Fried,M. (1986) *Cell*, **45**, 425–430.

Futcher,A.B. (1986) *J. Theor. Biol.*, **119**, 197–204.

Giulotto,E., Saito,I. and Stark,G.R. (1986) *EMBO J.*, **5**, 2115–2121.

Glickman,B.W. and Ripley,L.S. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 512–516.

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

Hyrien,O., Debatisse,M., Buttin,G. and Robert de Saint Vincent,B. (1987) *EMBO J.*, **6**, 2401–2408.

Ikoku,A.S. and Hearst,E. (1981) *J. Mol. Biol.*, **151**, 245–259.

Kafatos,F.C., Orr,W. and Delidakis,C. (1985) *Trends Genet.*, **1**, 301–306.

Kaguni,L.S. and Clayton,D.A. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 983–987.

Kieny,M.P., Lathe,R. and Lecocq,J.P. (1983) *Gene*, **26**, 91–99.

Kunes,S., Botstein,D. and Fox,M.S. (1986) In Klar,A. and Strathern,J. (eds), *Mechanisms of Yeast Recombination.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 149–156.

Leach,D. and Lindsey,J. (1986) *Mol. Gen. Genet.*, **204**, 322–327.

Lehrman,M.A., Goldstein,J.L., Russel,D.W. and Brown,M.S. (1987) *Cell*, **48**, 827–835.

Looney,J.E. and Hamlin,J.L. (1987) *Mol. Cell. Biol.*, **7**, 569–577.

Mouchès,C., Pasteur,N., Bergé,J.-B., Hyrien,O., Raymond,M., Robert de Saint Vincent,B., de Silvestri,M. and Georghiou,G.P. (1986) *Science*, **233**, 778–780.
Nalbantoglu,J. and Meuth,M. (1986) *Nucleic Acids Res.*, **14**, 8361–8371.
Paolella,G., Lucero,M.A., Murphy,M.H. and Baralle,F.E. (1983) *EMBO J.*, **2**, 691–696.
Passananti,C., Davies,B., Ford,M. and Fried,M. (1987) *EMBO J.*, **6**, 1697–1703.
Rogers,J.H. (1985) *Int. Rev. Cytol.*, **93**, 187–279.
Saito,I. and Stark,G.R. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 8664–8668.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Schimke,R.T. (ed.) (1982) *Gene Amplification*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Schimke,R.T. (1984) *Cell*, **37**, 705–713.
Sherman,L.A. and Gefter,N.L. (1976) *J. Mol. Biol.*, **103**, 61–76.
Shurvinton,C.E., Stahl,M.M. and Stahl,F.W. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 1624–1628.
Sor,F. and Fukuhara,H. (1983) *Cell*, **32**, 391–396.
Spradling,A.C. and Mahowald,A.P. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 1096–2002.
Stark,G.R. (1986) *Cancer Surveys*, **5**, 1–23.
Stark,G.R. and Wahl,G.M. (1984) *Annu. Rev. Biochem.*, **53**, 447–491.
Volkert,F.C. and Broach,J.R. (1986) *Cell*, **46**, 541–550.
Wahl,G.M., Stern,M. and Stark,G.R. (1979) *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.
Weiss,U. and Wilson,J. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 1619–1623.
Yao,M.-C., Zhu,S.-G. and Yao,C.-H. (1985) *Mol. Cell. Biol.*, **5**, 1260–1267.
Zuker,M. and Stiegler,P. (1981) *Nucleic Acids Res.*, **9**, 133–148.

Received on July 6, 1987; revised on November 23, 1987