W. J. J. van den Tweel, *ibid.*, p. 677; E. Dorn, M. Hellwing, W. Reineke, H.-J. Knackmuss, *Arch. Microbiol.* **99**, 61 (1974); J. Latorre, W. Reineke, H.-J. Knackmuss, *ibid.* **140**, 159 (1984); M. A. Rubio, K.-H. Engesser, H.-J. Knackmuss, *ibid.* **145**, 116 (1986); *ibid.*, p. 123.
71. S. W. Pirages, L. M. Curran, J. S. Hirschhorn, in *Impact of Applied Genetics in Pollution Control*, C. F. Kulpa, Jr., R. L. Irvine, S. A. Sojka, Eds., Symposium, University of Notre Dame, 24 to 26 May 1982 (Univ. of Notre Dame Press, Notre Dame, IN); G. Schraa *et al.*, *Appl. Environ. Microbiol.* **52**, 1374 (1986); J. C. Spain and S. F. Nishino, *ibid.* **53**, 1919 (1987); E. Schmidt, G. Remberg, H.-J. Knackmuss, *ibid.* **46**, 1038 (1983).
72. A. K. Koch, J. Reiser, O. Kappeli, A. Fiechter, *Bio/Technology* 6, 1335 (1988); L. Wender, Ed., *Bioprocess. Technol.* **10**, 3 (1988).
73. H. J. Gilbert *et al.*, *J. Gen. Microbiol.* **134**, 3239 (1988).
74. M. Emori, T. Tojo, B. Maruo, *Agric. Biol. Chem.* **52**, 399 (1988).
75. V. L. Seligy *et al.*, *Biotechnol. Adv.* **2**, 201 (1984).
76. C. Ganter, A. Boeck, P. Buckel, R. Mattes, *J. Biotechnol.* **8**, 301 (1988).
77. J. D. Windass *et al.*, *Nature* **287**, 396 (1980): Imperial Chemical Industries, Ltd., European Product Number 35831 (1981); European Product Number 66994 (1982).
78. D. J. Drahos, B. C. Hennig, S. McPherson, *Bio/Technology* **4**, 439 (1986); D. J. Drahos *et al.*, in *The Release of Genetically-Engineered Micro-Organisms*, M. Sussman *et al.*, Eds. (Academic Press, London, 1988), pp. 181–191.
79. We thank our colleagues M. N. Schroth, B. J. Staskawicz, J. Loper, and E. Clark for comments, suggestions, and criticism during manuscript preparation. We also thank H. J. Evans and F. Cannon for useful comments and personal communications. Cited work in the authors' laboratories has been supported by grants DBM-8706129, DBM-8409723 PCM-8313052 from the NSF and grant DE-FG03-86ER13518 from the Department of Energy.

Construction of Large DNA Segments in *Escherichia coli*

MICHAEL O'CONNOR,* MARK PEIFER, WELCOME BENDER

Recombinant DNA clones containing large pieces of DNA are useful in the study of large genetic units, but these are difficult to make in most bacterial cloning vectors. A strategy is described that uses general and site-specific recombination to construct large pieces of eukaryotic DNA from smaller cloned segments. The large clones are propagated on F factor–based plasmids in *Escherichia coli*. They can be easily modified to introduce mutations or rearrangements. These techniques were applied to the construction of large DNA segments from the bithorax complex of *Drosophila*.

GENES IN HIGHER EUKARYOTES ARE SURPRISINGLY LARGE and complex. The bithorax and Antennapedia complexes of *Drosophila* contain transcription units of 75 and 100 kb (*1*), respectively. Certain mammalian genes are even larger. The current record holder is dystrophin (*2*), whose transcribed region may exceed 2000 kb. The noncoding DNA of some large loci contains regulatory sequences critical for appropriate spatial and temporal regulation. For example, the *Ultrabithorax* (*Ubx*) gene is regulated by sequences up to 50 kb away from the messenger RNA start site, both upstream and within the introns (*3*). Whereas the human β-globin gene has a transcription unit spanning only a few kilobases, sequences as far as 50 kb upstream appear to be needed to obtain full expression in transgenic mice (*4*).

For such large genes, a method is needed that allows rapid cloning and manipulation of large functional units, so that the whole unit can be tested by transformation into the appropriate organism. Cosmid vectors are limited to a narrow size range, typically 40 to 45 kb of insert DNA. More recently, a new vector system has been developed (*5*), based on the creation of yeast artificial chromosomes (YAC clones). This allows the cloning of much larger pieces of DNA (up to several hundred kilobases), but the yeast chromosomes are more difficult to work with than bacterial plasmids.

We have developed a method, called "chromosomal building," that allows rapid construction in bacteria of large pieces of defined DNA, in F factor–based vectors. It relies on a combination of general and site-specific recombination to join large pieces of DNA from smaller, overlapping cloned segments in vivo. The replication and partition systems of the F factor (*6*) ensure stable maintenance of the resulting large plasmids. The vectors permit the large clones to be further modified. They can be rearranged, mutations can be introduced, or selectable markers can be added for transformation into various organisms. The product is a supercoiled, circular molecule that is resistant to shearing.

Basic Building Strategy

The basic building strategy is a process in which the F plasmid serves as the recipient of DNA transferred from the shuttle plasmid by recombination. This transfer is repeated a number of times; with each repetition the F plasmid increases in size by an amount roughly equivalent to the size of the insert in the shuttle plasmid. The process begins with the cloning of a DNA fragment (designated A-B-C in Fig. 1) into a polylinker cloning site in the F plasmid vector, pMBO132 (*7*). The F factor origin maintains the plasmid at one to two copies per cell, which enhances its stability. The plasmid also carries the resolution site for the F factor's site-specific recombination system, *rfsF* (*8*).

A second DNA segment (designated C-D-E in Fig. 1) is then cloned into the polylinker cloning site of the shuttle plasmid, pMBO96 (*7*). The second DNA segment is chosen to overlap the first by 500 bp or more, depending on the convenience of restriction sites. The origin of replication for the shuttle plasmid is temperature-sensitive. At 30°C the plasmid can replicate, conferring tetracy-

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115.

*To whom reprint requests should be addressed at Department of Molecular Biology and Biochemistry, University of California, Irvine, CA 92717.

cline resistance (tet^r). At 42°C, however, the plasmid cannot replicate and is diluted out during growth, leaving the progeny cells tetracycline sensitive (tet^s). The shuttle plasmid also carries a copy of the *rfsF* site, as shown (Fig. 1).

At the start of each cycle, the F plasmid and the shuttle plasmid are introduced into the same *rec*$^+$ cell by transformation at 30°C (Fig. 1, step I). At this temperature, both plasmids replicate independently. Next, we streak out these cells at the restrictive temperature, 42°C, selecting for both chloramphenicol resistance (cam^r) and tet^r (step II). At 42°C the shuttle plasmid cannot replicate and most cells become tet^s. Tetracycline resistance is only maintained if the F plasmid and the shuttle plasmid join by homologous recombination at the overlap (region C), forming a cointegrate with a tandem duplication. This cointegrate replicates by using the F plasmid origin. Cells carrying cointegrates are both tet^r and cam^r and form colonies that grow out of the streak. The rate at which cointegrates arise depends on the size of the overlapping homology between the two plasmids. A single colony streaked out at 42°C will yield three to five cointegrate colonies when the overlap between the plasmids (region C) is 500 bp, and more than 50 colonies with a several kilobase overlap.

Plasmid DNA minipreps are made from several cointegrate colonies and are used to transform *recA*$^-$ cells to cam^r and tet^r at the restrictive temperature (Fig. 1, step III). This serves two purposes. First, it purifies cointegrate DNA from the parental plasmids, since only cointegrates can give cam^r and tet^r at 42°C. Second, a *rec*$^-$ host is necessary for the next steps, because in a *rec*$^+$ cell cointegrates are unstable, forming and resolving via homologous recombination. Cointegrates resolved by homologous recombination mask those resolved by site-specific recombination. In a *rec*$^-$ host, cointegrates are stabilized. DNA minipreps are prepared at this stage to ensure that the correct cointegrate is present. (Since both plasmids carry the *rfsF* site, homologous recombination can also occur between these sequences, but such events can be easily distinguished from the correct cointegrates by a restriction digest.)

Cells carrying the cointegrate are next transformed with the cleanup plasmid (DCM111) (7) (Fig. 1). The cleanup plasmid carries the *resD* gene of the F factor; *resD* encodes a resolvase that mediates a site-specific recombination event between two copies of the *rfsF* site, like those carried on the F plasmid and the shuttle plasmid (*8*). The F, shuttle, and cleanup plasmids carry different selectable markers and compatible origins of replication, allowing all three to be stably maintained in the same cell. Cells with the cointegrate plasmid and the cleanup plasmid are grown at 42°C with selection for cam^r and ampicillin resistance (amp^r) (Fig. 1, step IV). During growth, the F factor resolvase mediates site-specific recombination between the two *rfsF* sites on the cointegrate, removing the shuttle plasmid and the duplicated region (C′) from the cointegrate. The resolution reaction yields an F plasmid that is unchanged, except for the addition of the DNA segment D-E to A-C. The process can be repeated by using a second shuttle plasmid carrying a DNA segment overlapping E. Successive rounds of building allow one to construct a large defined DNA molecule. A single round of building, as diagrammed in Fig. 1, takes about 5 days.

We have used the basic building procedure, along with the variations outlined below, to construct plasmids carrying DNA from the bithorax complex of *Drosophila* (*9*). The inserts of these plasmids are illustrated in Fig. 2; they fall into two classes. The first class consists of large segments of the bithorax complex carrying the *Ubx* and *abdA* genes, respectively, along with associated regulatory DNA. The insert in the largest of these constructs measured 125 kb (making the whole F plasmid 136 kb). This and the other F plasmids we have generated appeared to be completely stable when propagated in *Escherichia coli*. DNA blots of restriction digests of the 125-kb plasmid hybridized with probes covering the entire *Drosophila* insert are shown in Fig. 3, and all the expected fragments were present. The second class includes gene fusions with a reporter gene, β-galactosidase, adding either the natural 5′ flanking region of the *Ubx* promoter (the *bxd/pbx* regulatory region) or a regulatory region normally found in the *Ubx* gene's intron (the *abx/bx* regula-

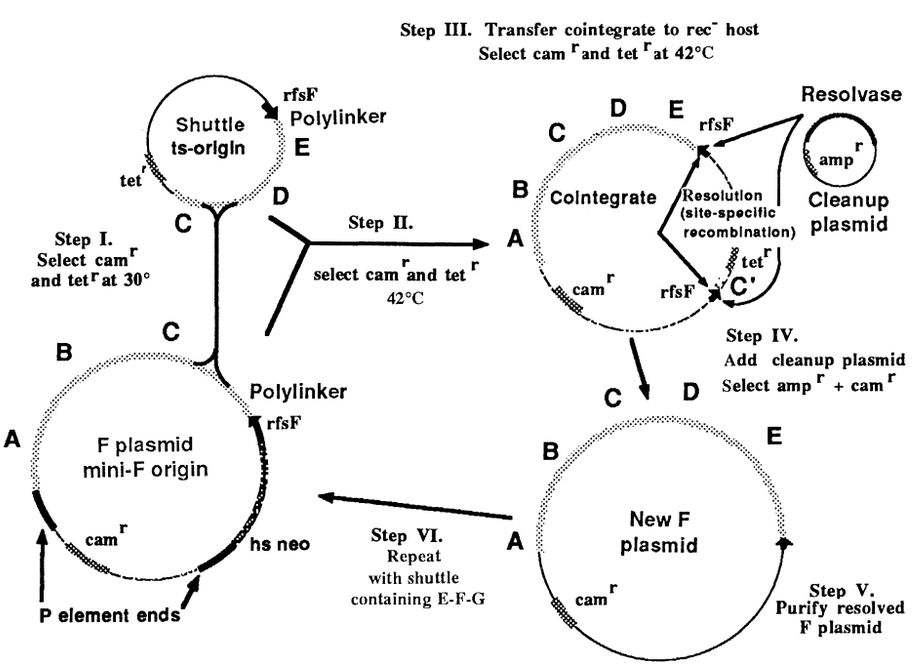

Fig. 1. The basic building strategy. (Step I) Introduce by transformation (*22*) the F and the shuttle plasmid into the same *rec*$^+$ cell at 30°C, selecting for cam^r (30 μg/ml) and tet^r (10 μg/ml). The F plasmid contains a cloned fragment of eukaryotic DNA (A-B-C). The shuttle plasmid contains an overlapping segment of DNA (C-D-E) (C ≥ 0.5 kb). (Step II) Streak out cells carrying both plasmids, selecting for cam^r and tet^r at 42°C. Replication of the shuttle plasmid stops. The only cells surviving are those in which a cointegrate has formed between the two plasmids, via homologous recombination between the overlap regions (C). These cells form colonies papillating out of the streak. The cointegrate carries a duplication of C (designated C′). (Step III) Restreak the cointegrate under the same conditions, and isolate cointegrate DNA by alkaline miniprep procedure (*11*). Transform a *rec*$^-$ cell (SCS1) (Stratagene) with this DNA, selecting for cam^r and tet^r at 42°C. Grow these cells at 30°C and check cointegrates by restriction enzyme digests of plasmid DNA obtained by the boiling miniprep procedure (*10*). (Step IV) Transform cells containing the cointegrate with the cleanup plasmid, selecting for cam^r and amp^r. When a cointegrate and the cleanup plasmid are maintained in the same cell, the F factor site-specific recombinase causes recombination between the two *rfsF* sites on the cointegrate, resolving it into the shuttle plasmid (which is lost) and the F plasmid, now carrying the enlarged DNA insert (A-B-C-D-E). (Step V) Recover the F plasmid by transformation back into the *rec*$^+$ cell, selecting for cam^r. Screen for tet^s and amp^s to ensure that the shuttle plasmid and the cleanup plasmid have been lost during transformation. (Step VI) Repeat with a new shuttle carrying the overlapping segment E-F-G.

tory region). One could, in a similar fashion, fuse regulatory sequences to complementary DNAs or juxtapose regions that are normally widely separated.

DNA Preparation and Analysis

Most procedures for isolation of plasmid DNA from *E. coli* were designed for small high-copy plasmids. The plasmids produced by the building strategy can be quite large and are maintained at a low copy number. Despite these differences, only slight modifications of standard protocols are necessary to prepare plasmid DNA. Cointegrates, when maintained in a *rec*$^-$ cell, can be grown at the permissive temperature for the replication via the temperature-sensitive origin. Cointegrates as large as 150 kb have been propagated stably at a copy number of six to ten. These plasmids are prepared by the boiling miniprep procedures (*10*) and visualized on an ethidium bromide-stained agarose gel. (We use boiling miniprep DNA for restriction analysis and alkaline miniprep DNA (*10*) for transformation of bacteria.) After resolution of the cointegrate, F plasmid derivatives can be prepared by the same miniprep procedures. Because of the large number of restriction fragments in the larger plasmids, it is often difficult to detect all the bands by ethidium bromide staining, but DNA blotting and hybridization permit easy evaluation. To prepare large amounts of plasmid, we have used the alkaline lysis procedure for making plasmid DNA (*10*), followed by two cesium chloride gradients to remove contaminating bacterial chromosomal DNA. We have obtained yields of 30 μg of plasmid DNA per liter of cells with plasmids as large as 136 kb.

One initial concern was that large plasmids would be easily broken by shear forces, but we have not found these plasmids to be overly fragile. There is not significant breakage or nicking of closed circles in the standard manipulations to prepare the plasmids. We were particularly concerned about possible shearing of plasmid DNA in the 1-μm tip of the microcapillaries used for injection of *Drosophila* embryos. An 80-kb plasmid was examined by electron microscopy before and after passage through such an injection needle. No significant increase in either broken or relaxed circle molecules was observed. We assume that the supercoiled nature of these plasmids is responsible for their resistance to shearing.

Mapping large DNA molecules can be difficult. We have constructed F plasmids carrying a single Not I site, which can be linearized to help in restriction analysis. In addition, probes corresponding to one or the other of the flanking vector sequences can be used to identify end fragments. In our experience, however, for constructs as large as 80 to 100 kb, such additional information is not necessary. The pattern of DNA fragments generated by digests with restriction enzymes that have six–base pair recognition sequences, as analyzed on an ethidium bromide–stained gel, is sufficient to verify the identity of clones.

The basic building strategy, as outlined above, was sufficient for much of the work illustrated in Fig. 2. We have developed a few procedures that speed the process and that permit further manipulation of constructed DNA. These are outlined in the sections below.

Reverse Building

Since the process of building requires little labor relative to time, it is efficient to construct a large segment by making several portions simultaneously and then to join them together. However, the intermediate-size segments end up on the F plasmid. Two F plasmids cannot be joined together by the building procedure; one of the two segments must be on the shuttle plasmid. A procedure called reverse building can be used to transfer DNA from the F plasmid to the shuttle.

Reverse building differs from the basic building process in that DNA moves from the F plasmid to the shuttle, rather than the opposite (Fig. 4a). Cointegrates are made as in the basic building procedure (step I). However, in this case the resolution is done at 30°C with selection for amp^r and tet^r, to allow recovery of the shuttle plasmid (step II). Resolution is less efficient at this temperature and many cointegrates remain. A DNA miniprep is made from these cells and used to transform a *rec*$^-$ cell carrying the pushout plasmid. The pushout plasmid is incompatible with the F origin

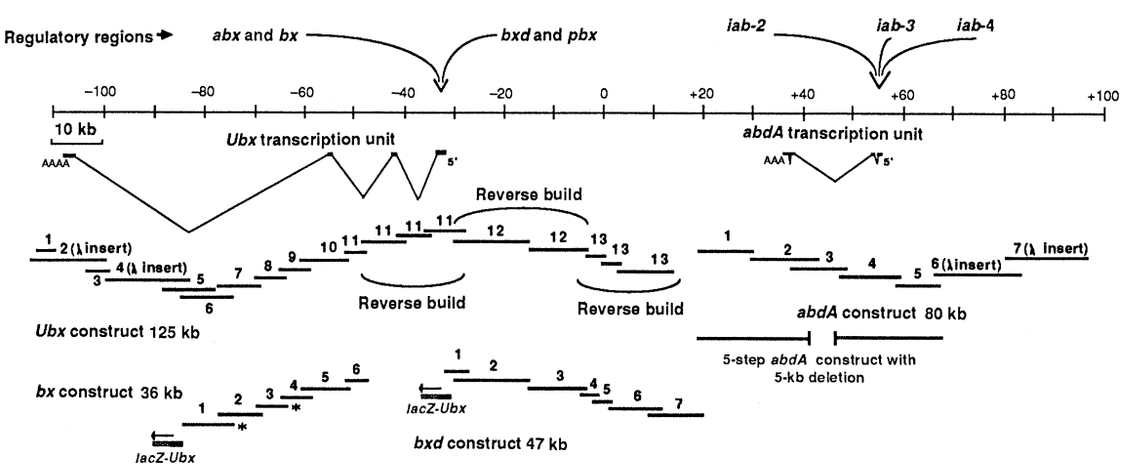

Fig. 2. Assembled DNA segments from the bithorax complex of *Drosophila*. The hatched line at the center represents the molecular coordinates of the DNA of the bithorax complex (*9*). Distance is in kilobases from an arbitrary zero point. Immediately below the coordinate axis are diagrammed the structure of the *Ubx* and *abdA* transcription units. Above the coordinate axis are the *abx, bx, bxd*, and *pbx* regulatory regions, as defined by mutations affecting spatial regulation of the *Ubx* product. Likewise, the *iab* regulatory regions are defined by mutations disrupting segmental regulation of the *abdA* product. The bottom half of the figure shows the plasmids constructed, which fall into two classes. The first class (the *Ubx* and *abdA* constructs) contain transcription units and associated regulatory DNA. The *Ubx* construct, at left, was made by cloning fragment 1 into the F plasmid and by adding successively the other fragments (each carried on the shuttle plasmid). The largest *Ubx* construct contains 125 kb and spans the *Ubx* transcription unit, including the *abx/bx* regulatory region, and part of the *bxd/pbx* regulatory region. On the right is the 80-kb *abdA* construct, containing the *abdA* transcription unit, and the *iab-2*, *iab-3*, and *iab-4* regulatory regions. Also shown is a derivative of the *abdA* construct, in which a 5-kb deletion has been introduced by recombination. Fragments recombined directly from Charon 4 phage are indicated with a λ. Three parts of the *Ubx* construct were obtained by reverse building, as indicated. In the second class of constructs, regulatory sequences from the *abx/bx* region or from the *bxd/pbx* region were connected to a *Ubx-lacZ* fusion gene. Successive steps in these constructions may be analyzed like a series of nested deletions.

plasmids. Selection for tet^r, carried by the shuttle, and kanamycin resistance (kan^r), carried by the pushout plasmid pML21 (7), results in the loss of both cointegrates and resolved F plasmids, allowing purification of the shuttle (step III). We have used the reverse building process in our work with DNA from the bithorax complex. The segments combined in this way are indicated in Fig. 2.

Building from Phage

Often DNA from the locus under study has been collected as a set of overlapping clones in a λ phage vector. If so, DNA can be transferred by recombination from λ phage to the shuttle plasmid, without in vitro biochemistry. We have implemented this strategy for the λ vector Charon 4 (*11*), but it should be applicable to other phage vectors. The strategy relies on a modified shuttle vector, carrying the *E. coli lacZ* gene flanked by portions of the left and right arms of the λ vector Charon 4 (Fig. 4b). These λ sequences are normally found immediately adjacent to the DNA insert in the recombinant λ phage. These two segments of homology allow a pair of recombination events between the modified shuttle and Charon 4, exchanging *lacZ* for the phage insert. Two λ pickup shuttle vectors, pMBO99 and pMBO100 (7), differ in the orientation of the right and left arm fragments, allowing us to recover the phage insert in either orientation on the shuttle.

The transfers are done in three steps. First, the recombinant phage is integrated into the bacterial chromosome, via recombination with a preexisting wild-type lysogen (Fig. 4b). Double lysogens arise in 10^{-2} to 10^{-3} of the infected cells and they are selected by colony hybridization. Next, the pickup shuttle is introduced into the cell. Transformants are streaked out at 42°C in the presence of tetracycline. The shuttle is lost unless it is integrated into the chromosome by recombination with one of the segments derived from the λ arms. To resolve this chromosomal cointegrate, the cells are streaked on a lactose indicator (MacConkey) plate at the 30°C, still selecting for tet^r. The host cells are *lac*$^-$, but the shuttle plasmid carrying *lacZ* makes them *lac*$^+$. At 30°C, the plasmid begins replication and is excised from the chromosome by a second homologous recombination event. Half of the time the recombination occurs at the other arm from the initial event, resulting in loss from the shuttle plasmid of *lacZ* and its replacement by the insert of the recombinant phage. In this process, the chromosome picks up *lacZ*. This gene can be lost, however, through further recombination, giving rise to *lac*$^-$ (white) colonies. One-half of the white colonies carry the desired plasmid. The other half result from recombination between the shuttle and the wild-type λ lysogen. These can be distinguished by restriction digests. We have used this process during the incorporation into the shuttle plasmid of those DNA segments marked with a λ in Fig. 2.

Crossing on Mutations and Marker Switching

Another variation of the building strategy makes it easy to mutate completed constructs. The introduction of a deletion into an already assembled plasmid is illustrated in Fig. 4c. First, a deletion is made in a small piece of DNA. This fragment (designated B-ΔC-D) is cloned onto the shuttle; it must be homologous to the DNA on the F plasmid (A-B-C-D-E) on both sides of the deletion, and the regions of homology should be close in size (B ≈ D). Both plasmids are joined as in the standard building strategy. Cointegrates are selected and stabilized in a *rec*$^-$ cell. Cointegrates are analyzed by restriction digests to identify the ones that have recombined on the side where the homology is the smallest. Those cointegrates are then transformed back into a *rec*$^+$ background, with selection for cam^r at 42°C. The resulting transformants are restreaked for single colonies

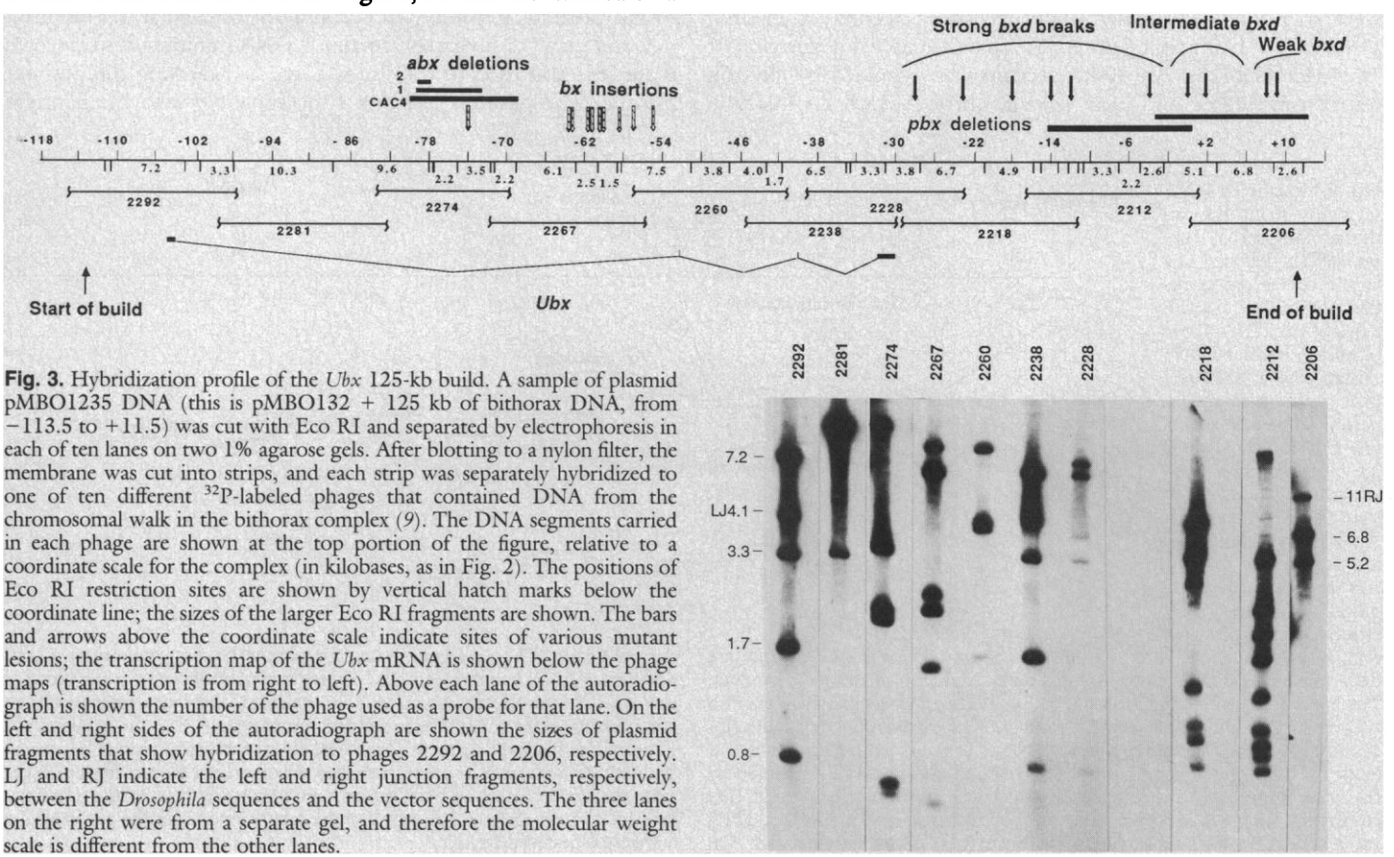

Fig. 3. Hybridization profile of the *Ubx* 125-kb build. A sample of plasmid pMBO1235 DNA (this is pMBO132 + 125 kb of bithorax DNA, from −113.5 to +11.5) was cut with Eco RI and separated by electrophoresis in each of ten lanes on two 1% agarose gels. After blotting to a nylon filter, the membrane was cut into strips, and each strip was separately hybridized to one of ten different ^{32}P-labeled phages that contained DNA from the chromosomal walk in the bithorax complex (*9*). The DNA segments carried in each phage are shown at the top portion of the figure, relative to a coordinate scale for the complex (in kilobases, as in Fig. 2). The positions of Eco RI restriction sites are shown by vertical hatch marks below the coordinate line; the sizes of the larger Eco RI fragments are shown. The bars and arrows above the coordinate scale indicate sites of various mutant lesions; the transcription map of the *Ubx* mRNA is shown below the phage maps (transcription is from right to left). Above each lane of the autoradiograph is shown the number of the phage used as a probe for that lane. On the left and right sides of the autoradiograph are shown the sizes of plasmid fragments that show hybridization to phages 2292 and 2206, respectively. LJ and RJ indicate the left and right junction fragments, respectively, between the *Drosophila* sequences and the vector sequences. The three lanes on the right were from a separate gel, and therefore the molecular weight scale is different from the other lanes.

at 42°C. At a frequency dependent on the size of the overlap, the cointegrates will resolve via a second homologous recombination, and the shuttle plasmid, unable to replicate, will be lost. Resolved F plasmids can be identified by screening for cam^r tet^s colonies (this occurred at about 5% for an overlap of 3 kb). If recombination occurs at the original site, it restores the parental wild-type plasmid. If recombination occurs on the opposite side of the deletion, the deletion is transferred to the F plasmid. The two modes of resolution happen at equal frequency if the flanking homologies are of equal size. The resolved plasmids are tested by restriction enzyme digests to determine whether the mutation has been acquired.

We have used this strategy to introduce a 5-kb deletion into the large intron of the *abdA* transcription unit (Fig. 2). The strategy should work equally well to introduce insertions, substitutions, or point mutations. Plasmids that have acquired point mutations can be identified by differential hybridization of an oligonucleotide probe (*12*) or by denaturing gradient gels (*13*).

We have also used a similar strategy to substitute alternate selectable markers. In this case, the *rfsF* site is used as one of the segments of homology flanking the "mutated" marker gene. Site-specific recombination catalyzed by the F resolvase is used to accomplish the second recombination event. We have used this procedure to replace the neomycin resistance gene in many of our contructs with either the *Alcohol dehydrogenase* (*14*) or the *rosy* (*15*) genes. We have also modified a large plasmid to carry the yeast *his3* and *ura5* genes, a yeast *ars*, a centromere, and an inverted repeat of a yeast telomere. We have transformed this plasmid into yeast and have linearized the molecule to form a linear yeast mini-chromosome. The marker substitution strategy should facilitate the transformation of large plasmids into other organisms besides yeast and *Drosophila*.

Discussion

The procedures described here can be used to construct large, defined clones of DNA in *E. coli*. Despite their size, the large plasmids are stably maintained and can be easily isolated in a form suitable for reintroduction into many eukaryotic organisms. We have used this building procedure to make numerous large clones of DNA from the bithorax complex of *Drosophila*, carrying what genetic and molecular analyses suggest are complete functional units. We are beginning to use these constructs to examine the regulation and function of the bithorax complex gene products. If our model of the bithorax complex (*3*) is correct, it will be difficult to dissect its regulatory regions with small pieces of DNA. Some aspects of the regulation seem to require the correct arrangement of very large DNA domains. It is for the study of such large domains that the building strategy was designed.

Whereas we have applied the building technique to the bithorax complex, this approach is applicable to any gene with a large transcription unit or regulatory regions. The 125-kb plasmid inser-

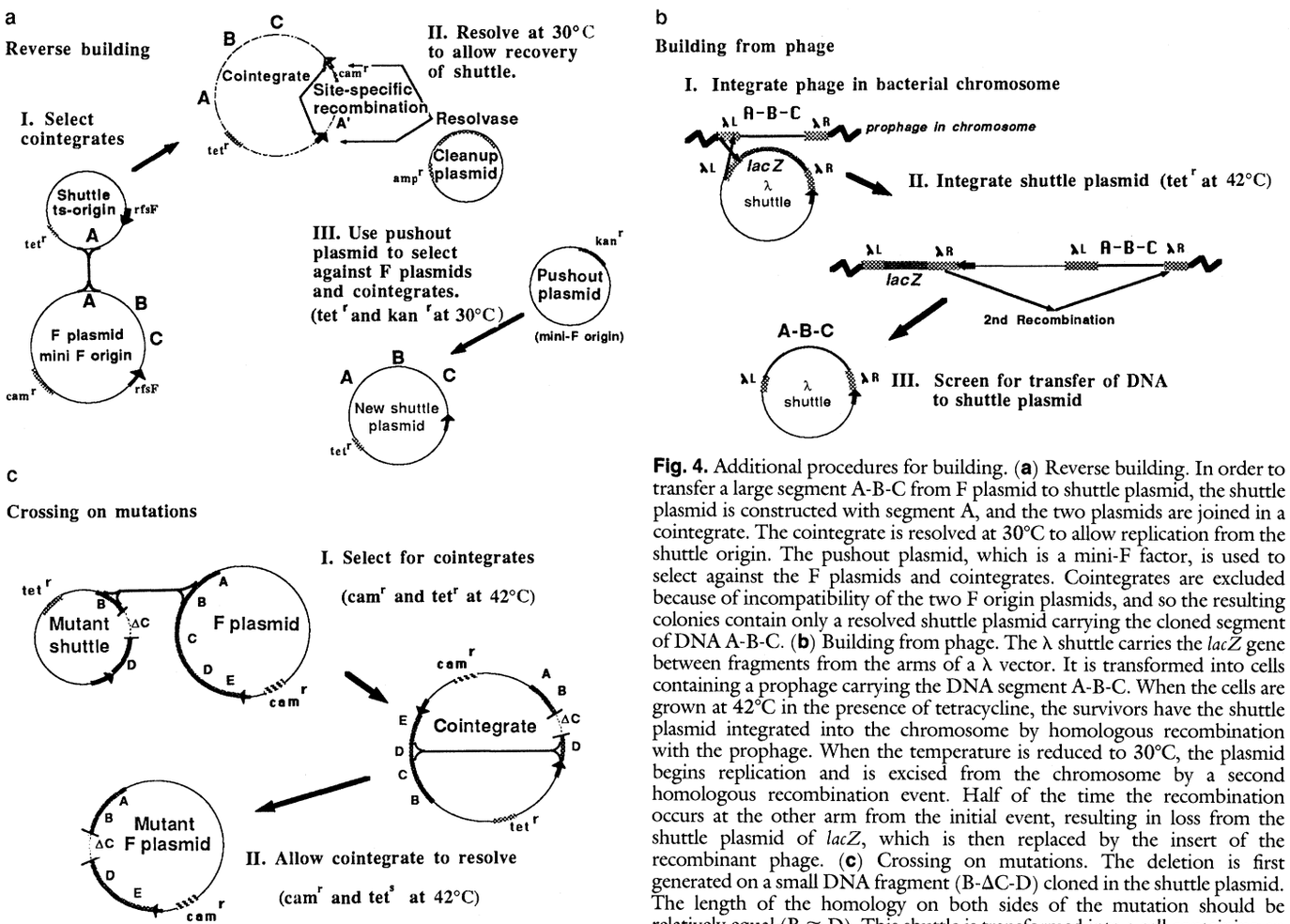

Fig. 4. Additional procedures for building. (**a**) Reverse building. In order to transfer a large segment A-B-C from F plasmid to shuttle plasmid, the shuttle plasmid is constructed with segment A, and the two plasmids are joined in a cointegrate. The cointegrate is resolved at 30°C to allow replication from the shuttle origin. The pushout plasmid, which is a mini-F factor, is used to select against the F plasmids and cointegrates. Cointegrates are excluded because of incompatibility of the two F origin plasmids, and so the resulting colonies contain only a resolved shuttle plasmid carrying the cloned segment of DNA A-B-C. (**b**) Building from phage. The λ shuttle carries the *lacZ* gene between fragments from the arms of a λ vector. It is transformed into cells containing a prophage carrying the DNA segment A-B-C. When the cells are grown at 42°C in the presence of tetracycline, the survivors have the shuttle plasmid integrated into the chromosome by homologous recombination with the prophage. When the temperature is reduced to 30°C, the plasmid begins replication and is excised from the chromosome by a second homologous recombination event. Half of the time the recombination occurs at the other arm from the initial event, resulting in loss from the shuttle plasmid of *lacZ*, which is then replaced by the insert of the recombinant phage. (**c**) Crossing on mutations. The deletion is first generated on a small DNA fragment (B-ΔC-D) cloned in the shuttle plasmid. The length of the homology on both sides of the mutation should be relatively equal (B ≅ D). This shuttle is transformed into a cell containing an F plasmid with a larger DNA segment A-B-C-D-E. Cointegrates are formed and resolved. In about half of the resolved molecules, the F plasmid should carry the deletion.

tion we have assembled spanning the *Ubx* gene would encompass the β-globin cluster (*4*) or half of the smallest yeast chromosome (*5*). A size of 136 kb is probably not near the upper limit of the F plasmid, since much larger F factor derivatives have been identified (up to 1.5 megabases) (*16*). This latter size would cover all but the very largest known genes, as well as many of the yeast chromosomes.

Although our primary goal was to clone large and defined fragments, the mechanics involved in building make it very versatile. The shuttle vector constructs can be modified to allow introduction of mutations in any region of a larger construct. The same fragments used to build the entire gene can be subsequently used to make fusions of large regulatory regions to reporter genes or to make mini-genes carrying part but not all of the locus. These building blocks can also be used to rearrange elements within the gene, by moving regulatory regions normally found upstream to the middle or the end of a large gene, or by reversing their orientation. This versatility makes the effort of making the shuttle clones doubly worthwhile.

A bacterial vector that can carry large DNA inserts could also be used to make a genomic library. Given the size of known F factor derivatives, DNA molecules of 150 to 200 kb and larger should be stable in our vector. We have seen no instability with any of our derivatives. Natural F′ plasmids are relatively stable in *recA*$^+$ cells even when they carry directly repeated insertion sequence (IS) elements (*17*). A random clone library of large pieces of DNA, 50 to 100 kb, in the shuttle plasmid vector or the F plasmid vector, could be rapidly screened by colony hybridization, and overlapping DNA segments could be assembled in F plasmids by the procedures we have outlined.

The building technique complements current technology such as the YAC system (*5*). The YAC system is preferable to F plasmids for isolating, in a single step, DNA segments of more than 100 kb, since there is currently no efficient method for ligating together such large circles and transforming them into bacteria. However, it is difficult to isolate a large gene without much flanking DNA in the YAC system. There are also differences in the ease of manipulation of the DNA products. YAC clones are linear chromosomes in yeast, whereas our strategy yields circular, supercoiled plasmids in *E. coli*. Bacterial colonies are easier to screen than yeast colonies. Yeast chromosomes are more difficult to prepare in bulk than bacterial plasmids, and the chromosomes must be handled more carefully to avoid shearing.

Our F plasmid vector is designed to allow reintroduction into the *Drosophila* genome. DNA added to the F plasmid ends up inside a P element, which can then be used to obtain stable germ line transformants carrying the DNA construct (*18*). We have shown that the P element on the F plasmid transposes successfully into the *Drosophila* germ line and confers G418 resistance (or ethanol resistance, after the switch of selectable markers to *Alcohol dehydrogenase*). Both a 35- and a 20-kb P element in the series in Fig. 2 (indicated by the asterisks) have already been successfully introduced into flies (*19*). We do not yet know how size will affect the rate of P element transposition. The largest P element introduced into flies thus far was a 45-kb transposon carried on a cosmid (*20*). There appears to be an inverse relation between transposon length and transposition ability, but the transposition frequency seems also to be influenced by the DNA sequences involved (*18*). Further advances in P element–mediated transformation are likely to occur as the biology of P element transposition is worked out in greater detail. Other methods of transformation may be necessary, however, to deliver very large (>100 kb) pieces of DNA.

DNA delivery to organisms other than *Drosophila* can be adapted to the biology of those organisms. The products of the building strategy, supercoiled plasmid DNAs, are compatible with most means of DNA delivery. Transfection of cultured mammalian cells requires no special vector. The marker switch strategy described above allows addition of a marker, such as the thymidine kinase gene, for work with mammalian cells or suppressor transfer RNA genes for transformation of nematodes (*21*). Microinjection of large supercoiled plasmids into nematodes or into mammalian eggs should be possible without breakage by shearing. The building technology should allow the testing of large DNA molecules for function in a variety of eukaryotic systems.

REFERENCES AND NOTES

1. A. Laughon, A. M. Boulet, J. A. Bermingham, R. A. Laymon, M. P. Scott, *Mol. Cell. Biol.* **6**, 4676 (1986); V. L. Stroeher, E. M. Jorgenson, R. L. Garber, *ibid.*, p. 4667; M. B. O'Connor, R. Binari, L. A. Perkins, W. Bender, *EMBO J.* **7**, 435 (1988); K. Kornfeld, R. B. Saint, P. A. Harte, D. A. Peattie, D. S. Hogness, *Genes Dev.* **3**, 243 (1989).
2. M. Koenig *et al.*, *Cell* **50**, 509 (1987).
3. M. Peifer, F. Karch, W. Bender, *Genes Dev.* **1**, 891 (1987).
4. F. Grosveld, G. B. van Assendelft, D. R. Greaves, G. Kollias, *Cell* **51**, 975 (1987).
5. D. T. Burke, G. F. Carle, M. V. Olson, *Science* **236**, 806 (1987).
6. N. Willets and R. Skurray, in *E. coli and S. typhimurium Cellular and Molecular Biology*, F. Neidhardt, Ed. (American Society for Microbiology, Washington, DC, 1987), pp. 1110–1133.
7. Construction of the vectors involved in the building strategy. Restriction maps are available upon request. pMBO132: the 9.4-kb f_5 Eco RI fragment of the F factor (*6*) was digested with Sal I and Hind III, and the fragment carrying the *ori* and *par* functions was ligated to a Sal I–Hind III fragment of Tn9 carrying cam^r (pMBO130). A 3.2-kb Hind III–Nco I fragment, carrying the normal *rfsF* site was deleted, filling in and removing both sites, and creating pMBO131. Meanwhile, a 450-bp Mlu I–Bam HI fragment from pMBO80 (*8*), carrying the *rfsF* site, was filled in and ligated into the filled-in Eco RI site in the polylinker within the P element of CosPneo to generate pMBO120. From this, a Sca I–Sal I fragment, carrying the P element and the *rfsF* site, was isolated, filled in, and ligated into the Sal I site of pMBO131, creating pMBO132. pMBO96: pHS1 [T. Hashimoto-Gotoh and M. Sekiguchi, *J. Bacteriol.* **133**, 405 (1977)] was digested with Eco RI and Pvu II, filled in, and a Mlu I–Bam HI fragment carrying *rfsF* was inserted (pMBO93). This can be used as a shuttle for cloning Eco RI fragments. To clone other fragments pMBO96 was created. pMBO93 was cut with Eco RI and the Eco RI–Hind III polylinker of pEMBL18 [L. Dente, G. Cesareni, R. Cortese, *Nucleic Acids Res.* **11**, 1645 (1983)] was filled in and added by blunt-end ligation, creating pMBO94. The polylinker is oriented with the Sac I site closest to the *rfsF* site. pMBO95 has the polylinker in the opposite orientation. pMBO94 was digested with Hind III and Nru I to remove the pSC101 tet^r determinant, and it was replaced with the 2.3-kb Hind III–Bgl II fragment of pRT2 [R. A. Jorgenson and W. S. Reznikoff, *J. Bacteriol.* **138**, 705 (1979)] carrying the tet^r determinant of Tn10. This plasmid was then partially digested with Bam HI, and the site outside the polylinker was filled in, generating pMBO96. pMBO99 and pMBO100: pMBO95 was digested with Pst I and Kpn I and the 6.6-kb Kpn I–Hind III fragment of Charon 4A (*11*) was inserted (pMBO97). This fragment carries 1.0 kb of the left arm and DNA of the insert of Charon 4, including the *lacZ* gene. The 3.4-kb Bgl II fragment of Charon 4, carrying the right arm–insert junction, was inserted into the Bst EII site of pMBO97 (pMBO98). The Kpn I–Xba I fragment carrying the left arm, *lacZ*, and the right arm was cut out and inserted into the polylinker of either pMBO94 or pMBO95 to generate pMBO99 and pMBO100. DCM111: A 3.4-kb Bgl II fragment from the F factor coordinates 44.65 to 48.05 was cloned into the Bam HI site of pBR322. The plasmid was a gift of M. Malamy. pML21: J. J. Manis and B. C. Kline, *Plasmid* **1**, 492 (1978).
8. M. B. O'Connor, J. Kilbane, M. H. Malamy, *J. Mol. Biol.* **189**, 85 (1986).
9. W. Bender *et al.*, *Science* **22**, 23 (1983); F. Karch *et al.*, *Cell* **43**, 81 (1985).
10. T. Maniatis, E. F. Frisch, J. Sambrook, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982).
11. F. R. Blattner *et al.*, *Science* **196**, 161 (1977).
12. J. L. Bos *et al.*, *Nucleic Acids Res.* **12**, 9155 (1984).
13. R. M. Myers, L. S. Lerman, T. Maniatis, *Science* **229**, 242 (1985).
14. D. A. Goldberg, J. W. Posokony, T. Maniatis, *Cell* **34**, 59 (1983).
15. G. M. Rubin and A. C. Spradling, *Science* **218**, 348 (1982).
16. B. Holloway and K. Brooks-Low, in *E. coli and S. typhimurium Cellular and Molecular Biology* F. Neidhardt, Ed. (American Society for Microbiology, Washington, DC, 1987), pp. 1145–1154.
17. R. C. Deonier and L. Mirels, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 3965 (1977).
18. A. C. Spradling, in *Drosophila: A Practical Approach*, D. B. Roberts, Ed. (IRL Press, Oxford, 1986), pp. 175–197.
19. M. O'Connor, J. Simon, M. Peifer, unpublished results.
20. M. Haenlin, H. Steller, V. Pirrotta, E. Mohier, *Cell* **40**, 827 (1985).
21. A. Fire, *EMBO J.* **5**, 2673 (1986).
22. D. Hanahan, *J. Mol. Biol.* **166**, 557 (1983).

23. We are grateful to M. Malamy for strains, plasmids, and advice; V. Pirrotta for the CosPneo; and A. DeLong and F. Winston for comments on the manuscript. Supported by a fellowship from the Damon Runyon–Walter Winchell Cancer Fund to M.B.O. and a grant from the NIH to W.B.