Computer simulation of DNA ligation: determination of initial DNA concentrations favouring the formation of recombinant molecules

Frédéric Dardel

Laboratoire de Biochimie, Ecole Polytechnique, UA no. 240 du CNRS, 91128 Palaiseau Cedex, France

Received August 17, 1987; Revised and Accepted November 10, 1987

ABSTRACT
A computer program was used to simulate the dynamic process of a ligation of DNA fragments. More specifically, the influence of the initial DNA fragments lengths and concentrations on the relative abundance of the various end-products was systematically investigated. Depending on the nature of the DNA extremities (asymmetric or symmetric, dephosphorylated or not), sets of initial conditions could be found that optimized the yield of active recombinant molecules. These results can be directly used to increase the efficiency of the ligation step, in particular for the construction of cDNA or genomic libraries.

INTRODUCTION
In vitro construction of recombinant DNA molecules is one of the most widely used techniques of modern molecular biology. This kind of experiments always involve a ligation step, in which DNA fragments and a linearized vector are incubated with a DNA ligating enzyme, such as T4 DNA ligase. The events that occur in the test tube during this step are quite complicated and not fully understood. In particular, the random joining of compatible DNA extremities can theoretically produce infinite combinations of different molecules, of which only a small number are wanted. It is therefore of great interest to scrutinize the process of ligation in order to define experimental conditions under which formation of the desired end-products is favored. However, the parameters defining the reaction mixture are quite numerous (sizes and concentrations of the various DNA molecules, type of the extremities, concentration of DNA-ligase and buffer composition...), and the exhaustive analysis of the end-products is experimentally difficult. Optimal buffer conditions have been quite well studied (1), and the enhancements obtained by the use of various polymers, such as polyethylene glycol, is well documented (2). On the other hand, little is known about the optimal concentrations and ratios of the various DNA molecules.
Recently, Levene and Crothers (3,4) have proposed a model for cyclization of DNA molecules, which fits well with experimental data. Taking advantage of their results, this work describes a computer simulation of the ligation process and proposes optimal reaction conditions for various types of experiments. Experimental evidences that support the model are also provided. This should prove particularly valuable when the number of recombinant clones is to be maximized, i.e. when constructing cDNA or genomic libraries, or when only traces of the DNA fragment to be cloned are available.

MATERIALS AND METHODS

Molecular biology
T4-DNA Ligase and restriction endonucleases were from either Pharmacia or Boehringer Mannheim. Incubation conditions were those recommended by the suppliers. Ligations were allowed to proceed overnight at 14°C. For scoring of recombinant clones, strain IBPC111 (F<sup>−</sup> Δ(lac-pro), gyrA, rpoB, metB, argEam, supF, ara, recA1, (5)) was transformed as described by Hanahan (6). The same batch of competent cells was used for all transformation experiments. Plasmid DNA was purified by banding twice in CsCl equilibrium gradient. DNA restriction fragments were purified by size exclusion chromatography (7). General microbiological techniques were as described (8,9).

System
The simulation was performed with an IBM-PC AT2, with no special extensions. Languages used were MS-BASIC 3.0, IBM Basic compiler and Turbo Pascal (Borland). Result files were transferred to an Apple MacIntosh+ using MacLink and analyzed with Excel spreadsheet (Microsoft).

ALGORITHM

Computational methods
The simulation is straightforward, the dynamical process of ligation is approximated through a discretization of time. For each iteration of the calculation, firstly extremities of the DNA molecules are considered pairwise: if fragment i and fragment j have compatible termini, the program increases the amount of the i-j linear concatemer by ∆C = K.([i]).([j]).8t and decreases the amount of i and j by the same value ∆C, were [i] and [j] are the concentrations of linear species i and j, respectively, 8t the time increment, and K some kinetic constant depending on the ligase, buffer conditions and the nature of the DNA termini. This process is repeated for all the compatible combinations of fragment i and fragment j extremities, thus yielding up to four different species of i-j linear dimers with different respective orientations of i and j. Secondly, each linear fragment i is examined for cyclization: if it has compatible termini, the amount of cyclic i molecules is incremented by ∆Cc = K.[i].[ia].8t, while linear i amount is decreased by the same amount ∆Cc, were K, [i] and 8t are as above and [ia] is the apparent concentration of one extremity of fragment i seen by the other extremity (10). The value of [ia] is a function of the length L of fragment i only and is calculated from the empirical formula of Levene and Crothers (3), for fragment lengths between 230bp and 1,900bp. For DNA fragments smaller than 230 bp, [ia] was approximated to be zero due to physical and topological constraints. For large DNA fragments (L>1,900bp), the following formula was used: [ia] = C.(L/(L<sup>2</sup> + B<sup>2</sup>))<sup>3/2</sup>. This function has an asymptotic behaviour O(L<sup>−3/2</sup>), corresponding to that proposed for large DNA molecules in Maniatis et al. (9, pp286-287). The values of parameters B and C were fitted to ensure (i) continuity of [ia] (L) and (ii) a value of [ia] for phage lambda genome (L=48,502 bp) corresponding to experimental data.
Under these conditions, parameter B, which is homogeneous to a length, is equal to 524 bp. This means that the maximum deviation between the function \([ia] = C \cdot \left( \frac{L}{(L^2 + B^2)^{3/2}} \right)\) and the classical \([ia] = C \cdot \left( \frac{1}{L} \right)^{3/2}\) is at most 7% over the valid range (i.e. \(L > 1,900\) bp). The value of the kinetic constant \(K\) is not known and is probably dependent on the nature of the extremities to be ligated (i.e. protruding or blunt end). However, since our main interest is to determine the final ratios of the end-products, the velocity of the ligation is not really important, at least if there is only one type of termini, in which case \(K\) is the same for all ligation events. The program therefore use \(K \cdot \delta t\) as parametrable step increment instead of \(\delta t\), removing the need to know its value explicitly.

The number of different DNA species generated by the ligation process is theoretically infinite (dimers, trimers, tetramers... with all possible combinations and orientations). In practice, it is not possible to handle a simulation with an infinity of different molecules. Since our main interest is to focus on dimers (i.e. 1 vector + 1 insert) or eventually trimers, we shall limit ourselves in considering lower order oligomers. The higher order polymers appear late in the reaction mixture, and though they may account for a significant percentage of the total DNA mass, they represent a small number of free DNA extremities and therefore have little influence on the results obtained for dimers, trimers and tetramers. The number of \(N\) different DNA species treated by the simulation program varied from 30 to 100 linear species, plus their cyclic counterparts. During the first iteration steps, the table of the \(N\) considered fragments is dynamically constructed by the program.

### TABLE 1

<table>
<thead>
<tr>
<th>Molecule</th>
<th>concentration after 60 steps (K \cdot \delta t = 1/(20 \cdot [i]^{0}_{\text{max}}))</th>
<th>concentration after 300 steps (K \cdot \delta t = 1/(100 \cdot [i]^{0}_{\text{max}}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>vector</td>
<td>2624</td>
<td>2632</td>
</tr>
<tr>
<td>insert</td>
<td>266</td>
<td>284</td>
</tr>
<tr>
<td>cyclic insert</td>
<td>1778</td>
<td>1788</td>
</tr>
<tr>
<td>insert-vector dimer</td>
<td>446</td>
<td>450</td>
</tr>
<tr>
<td>cyclic insert -vector dimer</td>
<td>346</td>
<td>344</td>
</tr>
<tr>
<td>insert-insert dimer</td>
<td>175</td>
<td>176</td>
</tr>
<tr>
<td>cyclic insert-insert dimer</td>
<td>259</td>
<td>254</td>
</tr>
<tr>
<td>trimers</td>
<td>434</td>
<td>428</td>
</tr>
<tr>
<td>cyclic trimers</td>
<td>94</td>
<td>93</td>
</tr>
<tr>
<td>higher order oligomers</td>
<td>53</td>
<td>51</td>
</tr>
<tr>
<td>cyclic oligomers</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

**Table 1:** Effect of the value of the iteration step \(K \cdot \delta t\) on the simulation result. The shown example illustrates the symetric ligation of a 2000 bp insert into a dephosphorylated pUC plasmid (=2500bp). Initial conditions were 1\(\mu\)g of insert and 1\(\mu\)g of pUC in 80\(\mu\)l ligation buffer (0.8/1 molar ratio). 45 linear species were considered in each case. Concentrations, taken at the same time in both simulations, are expressed in billions of molecules/ml. As seen, decreasing \(K \cdot \delta t\) by a factor of 5 does not significantly affect the results.
Figure 1: Concentrations of various DNA species during the ligation process. The figure shows the simulation of the ligation of a 2000 bp insert to a 2500 bp dephosphorylated vector. The iteration step was $K \cdot \Delta t = 6.10^{-16} \text{ml} = 1/(500.\text{[i]}_\text{max})$. Along the y axis are the concentrations expressed in billions of molecules/ml. Various species were grouped according to their relative abundance in the reaction mixture. The oligomers are described as follows, i indicates one molecule of insert and v one molecule of vector: i-i stands for an insert dimer, i-v is an insert ligated to a vector, v-i-v is a trimer composed of one insert sandwiched between two vectors...

The construction of this table takes into account the type and eventual dephosphorylation of the termini. When this table is full, the higher order multimers produced are stored as an $N+1^{th}$ undifferentiated species called "n-mers". The calculation is stopped when the precursors of the molecule of interest (the cyclic recombinant dimer) have almost completely disappeared, so that its final concentration has been reached. This required from 200 to 500 iteration steps under the chosen parameters.

Convergence of the model
The value of the iteration step $K \cdot \Delta t$ used in most cases was equal to $1/(20.\text{[i]}_\text{max})$, where $\text{[i]}_\text{max}$ is the initial concentration of the most abundant precursor. This insured that at most 5% of each molecule was consumed in each step. In fact this value of 5% is only reached during the few first steps since the number of free extremities in solution keeps decreasing and in the expression $K \cdot \text{[i]} \cdot \text{[j]} \cdot \Delta t$, $\text{[i]}$ and $\text{[j]}$ are soon smaller than $\text{[i]}_\text{max}$. If the value $K \cdot \Delta t$ is decreased to $1/(100.\text{[i]}_\text{max})$, the final results are not significantly different (see Table 1). The other important factor that might affect the final result is the number of different DNA species $N$ considered in the simulation. Usually 45 linear species were considered, together with their cyclic counterparts, when they existed. In a typical simulation where a linearized vector, dephosphorylated to prevent its closure without insert, is ligated to a single insert with symmetric termini, 45 species are sufficient...
Figure 2: Symmetric ligation: characterization of an optimum DNA concentration. For a given vector/insert ratio, the efficiency of the ligation (i.e. the percentage of the insert fragment recovered as circular recombinants at the end of the ligation) was calculated for various initial DNA concentrations. This figure illustrates the results obtained with a 2000 bp insert and a dephosphorylated pUC vector. As seen on the figure, an optimal concentration is found. The position of this optimum expressed in terms of total DNA concentrations is roughly independent of the initial vector/insert ratio, although it is sharper for large vector/insert ratios.

to include monomers, dimers, trimers and most of the tetramers. If the number of species is increased to 100, the final ratios of monomers, dimers and trimers is affected by less than 1% (data not shown), significant changes being only observed for higher order multimers.

IMPLEMENTATION
The original simulation program was written in MS-BASIC 3.0. Since each iteration step required N.(N+1) combinations of molecules, the overall calculation is somewhat lengthy. Compilation of the program did significantly improve things, but a complete simulation required about 20 minutes. A Turbo Pascal version was also written, but the speed was not significantly enhanced. Two version of the program were developed, an interactive one, which allows the user to test a given set of conditions, and a batch version, which was used to systematically analyze various initial concentrations and DNA fragments lengths. This latter version was used for repeated overnight runs.

RESULTS AND DISCUSSION
Symetric ligation
In a first step, insertion of a 2,000 bp fragment into a pUC12 (11) sized plasmid (=2,500 bp) was
Figure 3: Symetric ligation: effect of the initial vector/insert ratio on the maximum efficiency percentage. For various vector/insert ratios (vector is pUC12, insert is 2000 bp long), the maximum percentage of insert recovered as recombinant plasmid was calculated: for each curve in Figure 2, the value at the maximum was plotted as a function of the initial vector/insert ratio (white boxes). Stacked onto this value are shown the respective amounts of recombinant plasmids carrying two tandem insert which are also able to yield transformant colonies (hatched boxes).

Examined. The extremities of the vector plasmid were considered to be dephosphorylated, which prevents its circularization without insert.

Kinetical aspects. Even though the exact time scale is not known, the numerical simulation can be used to examine the ligation process kinetically. A typical example of what takes place in the test tube is shown in Figure 1. The insert is rapidly consumed to form mainly dimers (insert+vector or insert+insert), and then higher order oligomers. Linear oligomeric species concentrations exhibit two different behaviours depending on whether both their extremities are dephosphorylated or not. The concentrations of those with at least one phosphorylated terminus show an increase, then reach a maximum and finally decrease as they are converted to their cyclic counterpart. As expected, the maximum concentration is reached later for higher order oligomers, since firstly their direct precursors appear later in the incubation mixture and secondly they are longer and therefore are more slowly converted into cyclic molecules. Fully dephosphorylated linear species are potential end-products of the reaction, as they cannot circularize. Apart from the vector itself (which is also a dephosphorylated linear species), their concentration keeps increasing in the reaction mixture, until exhaustion of the supply of phosphorylated termini. The cyclic molecules concentrations are directly linked to the those of their respective linear precursors, the amount of a given circle being proportional to the sum of the integrals of its linear counterparts concentrations.
Figure 4: Symmetric ligation: Optimal DNA concentration as a function of insert length. For a given vector size, the optimal initial DNA concentration was calculated for various insert lengths. For each insert, an average of 4 different vector/insert ratios were tested, and the optimum concentration of total initial DNA (vector+insert) was derived as described in Figure 2. The black squares show the values obtained for a 2500 bp dephosphorylated plasmid, while the open squares show the corresponding values for a 2500 bp non-dephosphorylated vector. The crosses indicate the values corresponding to a M13 sized vector (=7500bp).

**Optimal DNA concentration.** In order to investigate the influence of initial conditions on the final state of the reaction mixture, a range of dilutions where tested for a given initial amount of plasmid vector and DNA fragment. For each dilution, the percentage of input DNA fragment that was recovered as recombinant plasmid was calculated (referred to as the efficiency percentage). As expected, there was an optimum dilution (Fig. 2). Above this optimum value, most of the input DNA fragment is recovered as cyclic monomers, while at high concentration, it was found in larger oligomers. This calculation was repeated for various plasmid to insert ratios, ranging from 80/1 to 1/4. In each case, an optimum dilution was found. As expected, higher vector to insert ratios give better efficiency percentages, but there seem to be a saturation, (Fig. 3) a 100-fold excess of plasmid being not significantly more efficient in trapping insert molecules than a 10-fold excess and little more than a 3-fold excess. Surprisingly, if the optimal dilutions for each vector/insert ratio were expressed in terms of initial total termini concentration (vector+insert), the resulting values were roughly identical (see Fig. 2), i.e. independent of the initial vector/insert ratio. In other words, if a ligation is to be performed between a given amount of plasmid and a given amount of insert, the optimal reaction volume is only dependent on the total number of DNA molecules and not on the plasmid/insert ratio. The same analysis was performed, still with a 2,500 bp vector, but with inserts of varying sizes, ranging from 300 to 10,000 bp, with similar results: For each insert...
Figure 5: Symmetric ligation: Influence of helical phase angle for small fragments. Shown is the efficiency percentage (as described in Fig. 2) of the ligation of a 2500 bp vector to fragments of lengths around 250 bp. The 250 bp fragment has a maximal $\left[\text{ia}\right] = 6 \times 10^{-8} \text{ M}$ (4), whereas the 245 bp fragment, which is one-half helix-turn shorter has a minimal $\left[\text{ia}\right] = 2 \times 10^{-9} \text{ M}$ (4).

size, there is an optimum initial concentration of termini (or total DNA) that is independent of the vector/insert ratio. These optimal concentrations are given on Figure 4.

**Other biologically active molecules.** An interesting question is to determine the respective abundance of the molecule of interest - the recombinant plasmid - and of other biologically active products. In order to yield transformant colonies, such products must be circular and carry a replication origin. Therefore, if one assumes that the insert does not carry a replication origin and that the vector is dephosphorylated, the competitors of the recombinant plasmid will be higher order cyclic oligomers containing at least one copy of the vector. Furthermore, plasmids with extensive palindromic sequences are known to be very unstable, this means that most probably molecules which carry inverted repeats of the insert will not be active. From the numerical simulation, it turns out that the main competitor of the simple recombinant plasmid is the molecule carrying two direct repeats of the insert ligated to one vector. Higher order oligomers are also found but their contribution is negligible in terms of number of molecules if not in terms of mass, unless very low insert/vector ratio are used. The respective amount of recombinant plasmids carrying one or two molecules of insert at the optimal DNA concentration is shown in Figure 3. As expected the low vector/insert ratios favours the oligomerization of the insert and therefore the formation of plasmids carrying two tandem copies of the insert, but in most cases the simple recombinant plasmid will be the most abundant active product.
Figure 6: Influence of vector dephosphorylation. A simulation was performed with a non-dephosphorylated pUC sized vector and a 400 bp insert. As in the example of figure 2, various initial concentrations of DNA were tested, for a fixed ratio of 0.8 vector/insert. The efficiency, measured in terms of percentage of the input fragment recovered as recombinant plasmid (efficiency %), plotted as a curve (black diamonds) similar to those of figure 1. Contrasting with the dephosphorylated case, closed vector with no insert are also obtained in this case. The percentage of recombinants (recombinants %) among total plasmids recovered at the end of the ligation was monitored at the same time (dotted boxes). Conditions that favor the efficiency in terms of number of recombinant clones (i.e. low initial DNA concentrations) also give a strong background of non-recombinant plasmids. The horizontal arrow indicates the maximum efficiency observed for similar conditions, but when the plasmid vector is dephosphorylated prior to ligation.

Effect of the helical phase angle for small inserts. In their report, Levene and Crothers point out the dramatic effect of the helical phase angle on the cyclization of small DNA molecules (4). This is expected to have a significant effect on the ligation of small inserts, especially around 250bp, where the cyclization probability can vary over two order of magnitude. The simulation performed with fragments of 245, 248 or 250 bp and a 2500 bp vector is shown in Figure 5. Ligation with the 245 bp fragment, which has a very low [ia] factor, is very similar to an assymetric ligation (see below), since its cyclization rate is to low to significantly interfere with the rest of the ligation process except at extreme dilutions. On the other hand the 250 bp fragment has a high [ia] factor, which means a high cyclization rate, and this results in significant modifications in the relative amounts of the end products. In the shown example, as much as 40 to 60% of the 250 bp insert will circularize, depending on the initial conditions, whereas in the case of the 245 bp fragment, the cyclic insert recovered does not account for more than 8 to 12% of the initial insert. For the 248 bp fragment an intermediate behaviour is observed. This helical phase angle dependence is not so dramatic for the final concentration of the recombinant molecule, but can however affect the efficiency of the ligation by a factor of 2. For practical purposes, the optimal DNA concentration can be calculated for an average case, such as that of the 248 bp insert, and used with no great loss.
Figure 7: Influence of initial DNA concentration influence in a sample case. A kanamycin
cellarage gene cartridge (1500bp) was ligated to pUC12 linearized by SalI (6/1 vector to insert
ratios). Various DNA concentrations were tested. Following overnight ligation, the
concentrations of the various samples were equalized, and the same batch of competent cells was
used to transform the ligation mixtures. The efficiency is given as a number of recombinant
clones/pmol total input DNA. The optimal DNA concentration predicted by numerical simulation
is 11nM (Fig. 4).

in efficiency. It must also be pointed out that this helical phase effect is dampened when higher
vector/insert ratios are used, since this will reduce the ratio of the insert cyclization rate vs the rate
of oligomerization.

Effect of dephosphorylation
When the same analysis is performed without specifying dephosphorylated termini for the vector
plasmid, the results parallel those obtained with a dephosphorylated vector and the optimal initial
DNA concentration is the same whether the plasmid termini are dephosphorylated or not. However
two significant differences must be pointed out: firstly, the yield measured in terms of number of
recombinant plasmids is lower, from two to four-fold than with a dephosphorylated plasmid, and
secondly, the conditions that yield a maximal number of recombinant plasmids also greatly favor
the cyclization of the vector without insert (Fig. 6). This means that if no positive biological
selection is available for recombinant plasmids, a large number of background colonies carrying
non-recombinants plasmids will be obtained upon transformation of the ligation mixture. On the
other hand, conditions that optimize the recombinant to non-recombinant ratio can be used (i.e.
low vector/insert ratio and high DNA concentration) but they are very inefficient in terms of number
of recombinant clones. The dephosphorylation step is therefore quite useful.

Effect of vector size
With a larger vector, such as the modified M13 phages mp8 and mp9 (12) (7500bp), the linear
dimer composed of one vector plus one insert will be also larger, and therefore its cyclization probability will be lower. We therefore expect to find different optimal ligation conditions. As shown in Figure 4, the results of the simulation indicate that this is indeed the case, although the general aspect of the curve is similar to that obtained with a pUC sized vector. The fact that for a similar insert size, the optimal termini concentration is lower is a result of the lowered cyclization probability of the linear dimer. Since the linear dimer vector+insert cyclize more slowly, other linear species compete more efficiently to drive the reaction towards the formation of large concatemers, and a greater dilution is required to prevent this. As a result of this, the yield of recombinant molecules is significantly smaller than with a pUC plasmid (between 1.5 and 2 x). For instance, with a 1000 bp insert and a 8/1 vector to insert ratio the maximum efficiency percentage calculated is 10.3% for M13, while a value of 19.5% was found for pUC sized plasmid. For vectors of intermediate sizes, it seems reasonable to interpolate the optimal initial concentrations from those given in Figure 4.

**Asymmetrical Ligation**

One of the constraints of symmetrical ligation, i.e. avoiding the circularization of the DNA fragment to be cloned, disappears when an asymmetrical ligation is performed, because its termini are no longer compatible. The only remaining difficulty is to avoid the formation of large concatemers. Therefore we expect that low DNA concentrations, which favor cyclization, will enhance the recombination efficiency. This is what is observed in the simulation (Data not shown). At low DNA concentrations, the rate-limiting step is the formation of dimers, cyclization then proceeds with a concentration-independent velocity. There are however practical limitations to work at very low DNA concentrations: the affinity of the DNA ligase for its substrate, the longer incubation times and the large reaction volumes required when performing ligation on significant amounts of DNA. The apparent $K_m$ of T4 DNA ligase for DNA extremities is $=1.5\text{nM}$ (1), thereby fixing the minimal DNA concentration in the range of 1nM (i.e. 2nM extremities).

**Experimental Validity**

A ligation experiment was undertaken in order to test some of the predictions of the model. Since it is not possible to measure the amount of recombinant molecules directly in the ligation mixture, a biological screen was used, i.e. scoring for recombinants plasmids after transformation into a suitable host. In this context, a kanamycin resistance gene cartridge was obtained from plasmid pUC-4K (Pharmacia). 2 pmol of an EcoRI restriction fragment carrying this gene ($=1500 \text{bp}$) were purified by size exclusion chromatography, while 12 pmol of pUC12 were linearized with EcoRI, dephosphorylated and mixed with the insert. This vector+insert mixture (in a 6/1 molar ratio) was divided into several aliquots and ligated overnight in different volumes of the same reaction buffer (i.e. ligation buffer as described in (7), with 10 units/ml of T4 DNA ligase). Tested initial DNA concentrations in the ligation mixture varied from 2 to 100 nM. Prior to transformation, all DNA concentrations were adjusted to the same value of 2nM, by diluting in an appropriate volume of ligation buffer. This was necessary to avoid any bias that could be introduced by the effect on
transformation efficiency of using different volumes with different DNA concentrations. After transformation ampicillin and kanamycin resistant colonies were scored. The transformation efficiency varied from 43 colonies/pmol of total DNA at 2nM to 250 colonies/pmol at 100 nM DNA, passing through a maximum value of 390 colonies/pmol at 10 mM DNA (Fig. 7), in good agreement with the predicted value of \approx 11 nM which can be derived from Figure 4.

Practical rules
The results of this computer modelization of the ligation process can be summarized as a set of practical rules which should prove useful to molecular biologists.
(i) If formation of long linear concatemers is desired, for instance when inserting DNA fragments into lambda vector arms, a high DNA concentration should be used (the higher, the better).
(ii) If cyclic molecules are to be constructed, asymmetrical ligation is more efficient than symmetrical ligation.
(iii) If an asymmetrical ligation is performed, low DNA concentrations enhance efficiency, with the practical limits of DNA ligase affinity and the duration of the ligation, i.e. concentrations below 1 nM DNA should be not be used.
(iv) If a symmetrical ligation is performed, a vector/insert molar ratio of 3 should be used, if possible, and the total initial DNA concentration should be read from Figure 4. Furthermore, the vector should be dephosphorylated.
(v) small vectors give better efficiencies than large ones (because of the increased cyclization probability of the insert+vector dimer).

ACKNOWLEDGEMENTS
The author thanks Dr. J.M. Schmitter for allowing him to use his computer and Pr. Métivier for mathematical advice. This work was supported in part by the following grant: UA240 CNRS. The listing of the simulation program (Basic or Pascal) is available upon request.

REFERENCES