pGATA: A Positive Selection Vector Based on the Toxicity of the Transcription Factor GATA-1 to Bacteria


P. Trudel¹, S. Provost¹, B. Massie¹, P. Chartrand¹,²,³ and L. Wall²,³

¹Montreal Center for the Canadian Red Cross Society,
²Institut du cancer de Montréal/ Centre de Recherche Louis-Charles Simard,
³University of Montreal and ¹National Research Council of Canada, Montreal, QC, Canada

ABSTRACT

The transcription factor GATA-1 is a zinc finger DNA-binding protein essential for the development of red blood cells. When we expressed different regions of the zinc finger domain in bacteria using an isopropyl-β-D-thiogalactoside (IPTG) inducible system, growth of bacteria harboring the active DNA-binding domain of GATA-1 was rapidly inhibited upon IPTG induction. The growth inhibition pattern suggested it may be occurring at the level of the initiation of replication, and GATA-1 was found to bind to three of the four DNA A protein-binding sites in the origin of replication. This toxicity was used to develop a positive selection vector system in which cloned DNA fragments interfered with the production of the GATA-1 DNA-binding domain. Thus, vector molecules containing the insert of interest are selected for when bacteria are grown in the presence of IPTG. With this system, the vector does not need to be dephosphorylated, purified or completely digested with a restriction enzyme for the efficient cloning of DNA fragments even when the vector-to-insert DNA molar ratio in ligation reactions is 10 to 1. Moreover, no special strain of Escherichia coli is required, and the selection might also be applicable to other species of bacteria if the toxicity of GATA-1 relates to inhibition of the DNA A protein.

INTRODUCTION

Most prokaryotic cloning vectors that are in common use today provide a marker gene (ampicillin, kanamycin, tetracycline, etc.) to select against the growth of cells that do not contain the vector, but cells will grow if they harbor the vector whether or not it contains the DNA fragment one is attempting to clone. If a DNA insert is being cloned into a single restriction site of such a vector, to ensure that a good proportion of colonies able to grow on the selection medium will contain the insert in the vector, it is necessary to dephosphorylate the digested ends of the vector to prevent religation to itself, and it is necessary to ensure that the vector is completely digested. This is done by using a large excess of the restriction enzyme and/or by purifying the digested from the undigested vector. These steps are time-consuming, and dephosphorylation of the vector reduces the cloning efficiency considerably (14) so that it can be difficult to clone DNA fragments which are available in limited quantities. An efficient positive selection system is one in which dephosphorylation and purification of the vector are unnecessary because cells containing the original vector will not grow under a simple selection regime. Moreover, the positive system should be usable in any common E. coli strain, the vector should contain a good number of unique restriction sites and the selection protocol should be simple enough to be easily reproduced.

Previously, several positive selection vector systems have been developed (for example, see References 4, 6 and 16), but none has seen wide use. We believe this is because, in most cases, special strains of bacteria are required to maintain the vector or for the selection regime, the selection system requires specialty reagents not commonly found in most laboratories, or usefulness has been limited by the lack of unique restriction enzyme sites and the use of low copy number vectors or a combination of these factors. A type of positive selection system that has seen wide use is the blue-white color selection based on the β-galactosidase enzyme that is available in pUC vectors and its derivatives (20). However, because there is no growth selection against vector without insert in this case, the vector must still be completely digested and dephosphorylated for cloning purposes.

GATA-1 is a eukaryotic transcription factor that plays a critical role in the development of the red cell lineage (12). GATA-1 binds to DNA sequences containing the trinucleotide GAT and has a preference for particular nucleotides flanking either side of the GAT core (5,11). The DNA-binding domain of GATA-1 is signified by two repeated zinc finger domains with the general structure CysX₂CysX₁₇CysX₂₋Cys (X denotes any amino acid), each of which is followed by a highly basic region (Figure 1). Only the second zinc finger and basic region of the protein are required for site-specific interaction with DNA. The first zinc finger region does not bind to DNA on its own, although it appears to increase the DNA-binding affinity and/or alter the DNA sequence specificity of the protein (9,19). It has not been determined whether it is both the first zinc finger and first basic domain or whether it is either region alone that influences both the affinity and binding specificity of GATA-1. With this question in mind, we attempted to express the different regions of the GATA-1 zinc finger...
Table 1. Synthetic Oligonucleotides Used in Gel-Mobility Shift Assays

<table>
<thead>
<tr>
<th>Synthetic Oligonucleotides Used in Gel-Mobility Shift Assays</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>α-Globin:</strong></td>
</tr>
<tr>
<td><strong>C-oligonucleotide:</strong></td>
</tr>
<tr>
<td><strong>D mutant:</strong></td>
</tr>
<tr>
<td><strong>OCT:</strong></td>
</tr>
<tr>
<td><strong>R1:</strong></td>
</tr>
<tr>
<td><strong>R2:</strong></td>
</tr>
<tr>
<td><strong>R3:</strong></td>
</tr>
<tr>
<td><strong>R4:</strong></td>
</tr>
<tr>
<td><strong>β-CACC:</strong></td>
</tr>
</tbody>
</table>

- α-Globin: GGATCTCCGGGCAACTGATAAGGATTCCTGT
- C-oligonucleotide: TCAGGATGTTTAAGATTAGCATTCAGGAAG
- D mutant: GTCTGCTGGTCCTATTAAATGTCCTCTA
- OCT: CGATTCTCAATGCAATATCTGTCTGAAAT
- R1: CGCACTGCCCCTGTGGATAACAGGGATCCGG
- R2: TCAGAATGAGGGGTATACCACTCAAAA
- R3: ACATGCTTTTGATTATACGAGCTCACTGCA
- R4: TCCTGACAGAGTTATCCACAGTAGATCGCA
- β-CACC: CACTAGGTGTCGTCCACAGGGTGAGGTC

domain in *E. coli*. However, the active DNA-binding domain was found to be highly toxic to the bacteria cells. This phenomenon was used to develop an efficient and simple positive selection vector that is described.

**MATERIALS AND METHODS**

**Construction of Expression Vectors**

DNA fragments were amplified by the polymerase chain reaction using a mouse GATA-1 cDNA as template and DNA primers specific for different regions of GATA-1 that introduce an in-frame *Bam*HI restriction site 5' and an *Eco*RI restriction site 3'. The amplified DNA fragments were cloned into the corresponding sites of the pGEX-2T expression vector (Pharmacia Biotech, Uppsala, Sweden).

*pGATA* was generated by cloning a blunted *Bss*HII (127 bp) fragment containing the full-length polylinker of Bluescript® (Stratagene, La Jolla, CA, USA) into the *Swa*I site of GST-Z2B2.

**Production and Purification of GST Fusion Proteins**

Glutathione S-transferase (GST) fusion proteins were produced in Electromax DH10B™ (Life Technologies, Gaithersburg, MD, USA) and extracted as described (7) except that buffer c was replaced by 20 mM HEPES (pH 7.6) containing 20% glycerol, 100 mM KCl, 0.2 mM EDTA, 1 mM dithiothreitol, 1 mM paramethylsulfonylfluoride, 1 μg/mL leupeptin, 1 μg/mL pepstatin, 2 μg/mL aprotinin and 1.5% Triton® X-100. Proteins were purified by affinity chromatography on glutathione-agarose beads (Pharmacia Biotech) as described (15).

**Gel-Mobility Shift Assays**

Gel-mobility shift assays, preparation of nuclear extracts, etc., were performed as previously described (1,17). The synthetic oligonucleotides used in these studies are shown in Table 1 (only one strand is given).

---

Figure 1. Structure of the GATA-1 protein. GATA-1 is a 413 amino acid protein which contains 3 transcriptional activation domains (denoted R1, R2 and R3) and a DNA-binding domain which is described in the text. The position of lysine and arginine residues in the DNA-binding domain is indicated by dots.
Pulse Labeling with [3H]Thymidine and [35S]Methionine

Two cultures each of the GST and GST-Z1B1Z2B2 recombinants were grown in M9 medium (14) supplemented with 50 μg/mL of all amino acids. Methionine and cysteine were omitted from cultures to be [35S]methionine-la
dated. When the absorbance at 600 nm (A600) reached 20%, 0.1 mM isopropl-
β-D-thiogalactoside (IPTG) was added to one of the cultures. One mL of each culture was taken at various times and
incubated for 5 min at 37°C with 10 μCi of [3H]thymidine or 10 μCi of [35S]methionine, after which 100 μL of
100 mM EDTA and 6% sodium dode-
cyl sulfate were added to the samples, and
then samples (100 μL) were collected on Whatman GF/C filter disks (What-
man, Clifton, NJ, USA) and submersed
in ice-cold 10% trichloroacetic acid
(TCA). The filters were washed with
10% TCA and dried, and the radioac-
tivity was measured by liquid scintilla-
tion counting.

Cloning of DNA Fragments by
GATA-1 DNA-Binding Selection

Vector DNA was isolated using the
simple miniprep procedure (14). In all
cases, vector DNA was cleaved with the
appropriate restriction enzyme(s), the
digestion reaction was phen-
ol:CHCl₃-extracted, and then the DNA
was precipitated with ethanol, redis-
solved in H₂O and used directly in liga-
tion reactions without further purifica-
tion. The 1.6-kb BamHI fragment of
the kanamycin resistance gene in
pUC4-K1XX (Pharmacia Biotech) and
the 800-bp Clal-NotI fragment of the
β-globin gene promoter in the μ-locus
vector (3) were purified from agarose
gels. The mixture of λ DNA fragments
was obtained by restriction digestion of
purified λ phage DNA (obtained from
Amersham International, Little Chal-
font, Bucks, UK) with BamHI.

Vector DNA and DNA fragments
were ligated together at 16°C over-
night. One hundred nanograms of vec-
tor DNA (estimated by EtBr staining
on agarose gels) were used per 10-μL
reaction, while the amount of DNA
fragment to be cloned was varied.

RESULTS AND DISCUSSION

Expression of the GATA-1 DNA-
Binding Domain in Bacteria

The zinc finger region of GATA-1
(Figure 1) was divided into 4 domains:
the first zinc finger (Z1), the first basic
region (B1), the second finger (Z2) and
the second basic region (B2). DNA
fragments containing various combina-
tions of these domains (see Figure 2)
were cloned into the pGEX-2T vector, an
ampicillin-selectable vector in which proteins of interest are synthe-
sized as fusions with the GST enzyme
and can be purified using glutathione
affinity chromatography (15). Expression
is controlled by the IPTG-inducible
tac promoter so that recombinant
protein is only synthesized in the
presence of IPTG. Recombinants are
referred to as GST followed by the
GATA-1 domains (Z1, B1, Z2 and/or
B2) contained in the fusion protein.

To determine if any of the recombi-
nants were able to synthesize active
DNA-binding protein, cultures were
grown in the presence of 0.1 mM IPTG
for several hours, the proteins were
extracted and then purified by glu-
thione-Sepharose® (Pharmacia Biote-
tch) affinity chromatography, and then
the purified preparations were assayed
by gel-mobility shift (Figure 2). The
three recombinants that contained the
complete Z2B2 subregion (GST-
Z1B1Z2B2, GST-B1Z2B2 and GST-
Z2B2) all gave a strong signal with the
GATA-1-specific α-globin probe from
the mouse α-globin gene promoter
(13), and this signal was partially com-
petitively inhibited by the strong
GATA-1 binding C-oligonucleotide
from the β-globin gene 3’ enhancer
(17) in each case. The signals were
not inhibited by non-GATA-1 binding se-
quences at the concentrations of com-
petitors used in the figure (not shown),
and none of the other recombinants
gave any detectable signal (Figure 2).
Thus, only fusion proteins that con-
tained the Z2B2 region bound to
GATA-1 sites; this was in agreement
with previous results showing that the
second zinc finger domain (Z2B2) of
the GATA-1 protein is necessary and is
sufficient alone to bind to GATA-1.

Figure 2. Gel-mobility shift assays of recombinant proteins. An equal volume of the glutathione-affinity purified-recombinant protein indicated on the top of each set of lanes was assayed by gel-mobility shift using a GATA-1 binding sequence from the α-globin gene promoter as the DNA probe, either alone or in the presence of unlabeled competitor DNA. In lane 1, no protein was added, while in lane 2, nuclear extracts from murine erythroleukemia cells were used as a positive control.
cognate sequences (9,19).

The finding that the GATA signal seen with the Z2B2-containing recombinants in Figure 2 was only partially competitively inhibited by a 200-fold excess of unlabeled C-oligonucleotide was surprising. However, close inspection of the autoradiograms suggested that the signal actually consists of two closely spaced bands, the upper of which is completely inhibited by the 200-fold excess of C-oligonucleotide and the lower of which did not appear to be strongly affected at this concentration of competitor. For example, in Figure 2 it can be seen that upon

---

**Figure 3. Inhibition of growth of the GST-Z2B2 recombinant upon IPTG induction.** In Panel a, two cultures of the bacteria indicated at the top of the figure were grown in liquid culture to an $A_{600}$ of 0.2 (indicated by the arrow), then 0.1 mM IPTG was added to one of the cultures and the $A_{600}$ was followed. Dotted line and open symbols, cultures grown in the absence of IPTG; closed symbols and solid lines, cultures grown in the presence of IPTG. In Panel b, cultures of the indicated recombinant bacteria at a starting $A_{600}$ of approximately 0.2 were grown in the presence or absence of 0.1 mM IPTG, and at the times indicated, 1 mL of culture was removed, and the cells were pulse-labeled with 10 μCi of $[^3H]$thymidine (open symbols) or with 10 μCi of $[^35S]$methionine (closed symbols). For each time point, the amount of incorporation for cells grown in the presence of IPTG was divided by the amount of incorporation for cells grown in the absence of IPTG, and the results are presented as the % incorporation relative to 0 time taken as 100%.
addition of competitor, the remaining signal actually migrates lower than when no competitor is present in each case. Moreover, it was found that a 1000-fold excess of specific competitors (C-oligonucleotide or α-globin) completely eliminated both signals, whereas a 1000-fold excess of nonspecific competitor (β-CACC) appeared to inhibit the lower band without affecting the upper band (not shown). Thus, the upper band would appear to reflect nonspecific binding, possibly due to the presence of partially denatured GATA protein in the extracts.

Expression of the GATA-1 DNA-Binding Domain Inhibits Growth of Bacteria

The three recombinants, containing the Z2B2 region of GATA-1, that produced active GATA DNA-binding activity (see above) gave very low yields of purified protein (<0.9 mg protein/L of culture), while all other recombinants gave relatively high yields (from 4 to 20 mg protein/L of culture). It was also noticed that cultures of the Z2B2-containing recombinants grown in the presence of IPTG specifically gave rise to much smaller pellets upon centrifugation, suggesting that IPTG induction of protein synthesis might be affecting their growth. To investigate this, cultures were grown to an $A_{600}$ of 0.2, IPTG was added and the $A_{600}$ was followed over several hours (Figure 3a). For the GST, and GST-B2 recombinants, addition of IPTG had a very slight inhibitory effect on growth. But the effect of IPTG on the GST-Z2B2 recombinant was much more dramatic (Figure 3a). The cells doubled their density and then ceased to grow any further. On the other-hand, this clone grew just as well as the other clones in the absence of IPTG. This experiment was repeated several times with both the GST-Z2B2 and the GST-Z1B1Z2B2 recombinants, and adding the IPTG at different culture densities was also tried. In every case, upon addition of IPTG, the cells approximately doubled their density then ceased to grow further.

To check if all cells were affected, overnight cultures grown with ampicillin alone were diluted and plated on medium containing only ampicillin and on medium also containing 0.1 mM IPTG. With cells containing just the GST protein, approximately an equal number of colonies grew on the IPTG-containing plates compared with ampicillin alone. In contrast, the ratio of colonies on the IPTG-containing plates compared with ampicillin alone was 1/50000, 1/20000 and 1/90000 for the GST-Z2B2, GST-B1Z2B2, and GST-Z1B1Z2B2 recombinants, respectively. Moreover, when 20 colonies of the GST-Z1B1Z2B2 clones that grew on IPTG-containing plates were picked and replated on the same medium, only 14 of them grew again, and analysis of the vector DNA in these 14 colonies showed that the vector had rearranged so that they would not be expected to express the fusion protein (results not shown). Therefore, there is a very strong, if not absolute, selection against the Z2B2-expressing vectors in the presence of IPTG.

To address whether the primary growth inhibition effect was at the level of protein or DNA synthesis, protein and DNA synthesis were followed by pulse-labeling cells at various times with $[^{35}S]$methionine and $[^{3}H]$thymidine, respectively (Figure 3b). DNA synthesis was immediately and severely inhibited by addition of IPTG to the GST-Z1B1Z2B2 recombinant so that it was less than 5% of the control after only 90 min, while protein synthesis was affected at a later time and at a slower rate.

Taken together, the above results indicated that GATA-1 DNA-binding activity might be acting at the level of the initiation of replication; i.e., it appeared that the cells are able to complete replication cycles in progress, but are not able to initiate new cycles once the Z2B2 polypeptide is synthesized. In fact, the growth inhibition profile that is seen here (Figure 3a) is virtually identical to what has been observed for temperature-sensitive mutants of the $E. coli$ DNA A protein (18). DNA A is involved in the initiation of replication at ori C of the $E. coli$ genome (reviewed in Reference 8). It binds to 4 DNA sequences in ori C that are referred to as R1, R2, R3 and R4.

Three of the DNA A protein binding sites (R1, R3 and R4) contain a potential GATA-1 binding sequence, including the GAT trinucleotide (see Table 1), the central core for GATA-1 binding. Using the C-oligonucleotide, which binds to the GATA-1 and OCT-1 proteins in extracts of murine erythroleukemia cells, as the probe and specific competitors against GATA-1 (α-globin), OCT-1 (D mutant) or both proteins (C-oligonucleotide and γ-globin octamer) and a nonspecific competitor (β-CACC) to control for the specificity of the reactions, competitive gel-mobility shift assays showed that R1, R3 and R4 all bind to GATA-1, while R2 does not (Figure 4). Similar competitive assays using α-globin and R4 as the probes (not shown), both of which bound only to GATA-1 in murine erythroleukemia extracts, suggested that R4 is a very strong GATA-1 binding sequence (about equivalent to α-globin probe), while R1 and R3 are...
somewhat weaker sites than R4 but are still strong (about equivalent to C-oligonucleotide probe). Thus the toxic effect of the GATA-1 DNA-binding domain may be through binding to 3 of the 4 DNA A protein binding sites in the origin of replication, which would be expected to interfere with the DNA A protein function. This hypothesis would be supported by the growth characteristics and rapid cessation of replication seen when the protein is induced. However, since it is expected that there are numerous other GATA binding sites in the bacteria genome, it is possible that the effects on replication are indirect and do not involve the origin of replication. Thus, further studies will be necessary before a definite conclusion can be reached.

Figure 5. Cloning of DNA fragments by GATA DNA-binding selection. In Panel a, DNA from 18 colonies (lanes 1–18) that grew on solid medium containing ampicillin and IPTG from ligation of the GST-Z2B2 BamHI-digested vector with the 1.6-kb kanamycin gene fragment at a 10:1 molar ratio (vector to insert; see Table 2) was cut with BamHI and analyzed by agarose gel electrophoresis. A picture of the EtBr-stained gel is shown. Positions of vector and kanamycin gene fragments are indicated at the left of the figure, while migration positions of some of the molecular size DNA standards (1-kb ladder from Life Technologies) run in lane 19 are indicated to the right of the figure. Note that only clones in lanes 7, 9, 12, 13 and 15 do not contain the kanamycin fragment. In Panel b, GST-Z2B2 vector prepared as in Panel a was mixed with 1/100 of the same amount of uncut vector; then this was ligated in the presence of 1/5 of the molar concentration of λ phage DNA cut with BamHI. DNA from 18 transformed colonies (lanes 2–19) that grew on ampicillin and IPTG-containing medium was digested with BamHI and was resolved by agarose gel electrophoresis. Positions of some of the molecular size markers run in lane 1 are indicated to the left of the figure, while λ DNA digested with BamHI was run in lane 20. Only clones in lanes 4, 8, 11, 13, 15 and 17 did not contain one of the λ DNA fragments. Note that large fragment seen in lane 8 does not co-migrate with either of the large λ BamHI fragments.
Toxicity of GATA DNA Binding to Bacteria Can Be Used as an Efficient Selection System for Routine DNA Cloning

The high toxicity of the GATA-1 DNA-binding domain suggested that it might serve as an efficient positive selection system for cloning DNA fragments. We attempted to clone a BamHI fragment containing the gene for kanamycin resistance into the unique BamHI site located just after GST in the GST-Z1B1Z2B2 and GST-Z2B2 constructs. Insertion of a DNA fragment will prevent the GATA-1 DNA-binding domain from being synthesized if the fragment provides a transcription or translation stop site and/or frame shifts the coding sequences. The vectors were digested with BamHI, but were not dephosphorylated or purified further, and then were ligated in the presence of a 10, 1 or a 0.1 molar ratio of the kanamycin fragment. Transformed bacteria were plated on medium containing either ampicillin alone to measure the total number of cells containing vector, ampicillin with IPTG to test for cells in which the GATA-1 DNA-binding domain is not being expressed, or ampicillin with IPTG and kanamycin to test for the presence of the kanamycin fragment (Table 2, series 1 and 2). The vectors ligated in the absence of insert gave rise to a large number of colonies on the ampicillin plates as expected, while a very small number of colonies were obtained on the ampicillin-IPTG plates and no colonies grew on the plates containing kanamycin. However, when insert and vector were ligated together (Table 2), the proportion of colonies that grew on the IPTG-kanamycin-containing plates was greatly increased. Even when the ratio of insert to vector was only 0.1:1, the results suggested that 42% of the colonies that grew on the ampicillin-IPTG plates could also grow in the presence of kanamycin (average value for the GST-Z2B2 and GST-Z1B1Z2B2 vectors shown in Table 2). When DNA was isolated from 18 colonies (from the ligation of insert to vector at a 0.1:1 ratio) that grew on ampicillin and IPTG and was digested with BamHI to test for the presence of the 1.6-kbp kanamycin fragment, 13
d(72%) of the colonies were found to contain the desired insert (Figure 5a).

For comparative purposes, exactly the same experiment was repeated with the GST-Z1B1Z2B2 vector, except that in this case the vector was dephosphorylated after BamHI digestion, but was otherwise treated identically (Table 2, series 3). Dephosphorylation reduced the cloning efficiency by a factor of 4 to 10 (compare number of colonies on IPTG-containing medium in series 3

### Table 2. Selection Efficiency of the Kanamycin Restriction Fragment by the GATA-1 DNA-Binding Domain

<table>
<thead>
<tr>
<th>Series</th>
<th>Vector</th>
<th>Insert:Vector (ratio)</th>
<th>Ampicillin</th>
<th>Ampicillin and IPTG</th>
<th>Ampicillin, IPTG and Kanamycin</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GST-Z2B2: no insert</td>
<td>&gt;10000</td>
<td>38</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Not dephosphorylated)</td>
<td>10:1</td>
<td>6200</td>
<td>2320</td>
<td>2130</td>
</tr>
<tr>
<td></td>
<td>1:1</td>
<td>3600</td>
<td>340</td>
<td>290</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1:1</td>
<td>3900</td>
<td>67</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>GST-Z1B1Z2B2 no insert</td>
<td>&gt;10000</td>
<td>186</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Not dephosphorylated)</td>
<td>10:1</td>
<td>5500</td>
<td>2010</td>
<td>1820</td>
</tr>
<tr>
<td></td>
<td>1:1</td>
<td>&gt;10000</td>
<td>1470</td>
<td>1130</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1:1</td>
<td>&gt;10000</td>
<td>233</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>GST-Z1B1Z2B2 no insert</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Dephosphorylated)</td>
<td>10:1</td>
<td>311</td>
<td>238</td>
<td>173</td>
</tr>
<tr>
<td></td>
<td>1:1</td>
<td>350</td>
<td>271</td>
<td>250</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1:1</td>
<td>28</td>
<td>19</td>
<td>18</td>
<td></td>
</tr>
</tbody>
</table>

1Vector concentration was kept constant, while insert concentration was varied.

Figure 6. Structure of the pGATA cloning vector. A partial map of the vector is shown. Only unique restriction sites in the polylinker are indicated.
and series 2). Thus, for routine cloning purposes dephosphorylation with this vector system is unnecessary and will substantially reduce the total number of positive clones. However, dephosphorylation of the vector did also appear to reduce the background considerably in that no colonies grew on the ampicillin + IPTG medium when vector was ligated without insert. Moreover, when tested as described for Figure 5, approximately 90% of the colonies from the ligation of insert to vector at a 1:10 ratio that grew on the ampicillin + IPTG medium contained the kanamycin fragment (not shown), compared with only 72% when the vector was not dephosphorylated (see above). Thus, dephosphorylation does increase the positive clone-to-background ratio, although substantially less clones are obtained.

The applicability of GATA-1 toxicity as a positive selection system was further tested. The GST-Z2B2 plasmid was cut with BamHI. The linearized vector was not dephosphorylated or purified further and was mixed with 1/100 of the same amount of undigested plasmid to simulate an incomplete digestion. The vector was ligated in the presence of a 1:5 molar ratio (insert:vector) of λ phage DNA digested with BamHI to simulate a situation when only small amounts of a contaminated DNA fragment insert are recovered from purification on agarose gel. This λ DNA mixture consists of 5 DNA fragments ranging in size from 5600 to 16900 bp. Bacteria transformed with the ligation mixture were plated on ampicillin-IPTG-containing medium, and the plasmid DNA from 18 colonies that grew were analyzed by BamHI digestion (Figure 5b). Twelve of the colonies (66%) contained one of the λ DNA fragments, and 4 of the 5 DNA fragments were represented in these 12 clones. Thus, the selection is so powerful that complete digestion is not mandatory, and different fragments in a mixture can be cloned even when the ratio of total fragment DNA to vector DNA is low in the ligation reaction.

We have also tested that the GATA-1 selection works in a variety of E. coli strains commonly used. We have tested DH10B, DH5α, XL1 blue and HB101. The selection functions well in all cases, although for XL1 blue it was necessary to reduce the IPTG level to 50 μM from 100 μM. Higher IPTG concentrations appear to have nonspecific toxic effects on this strain.

As a more useful cloning vehicle, we have designed the vector shown in Figure 6 that is denoted as pGATA. pGATA contains 13 unique restriction enzyme cloning sites, and in cloning experiments, it was shown to be as efficient as the GST-Z1B1Z2B2 and GST-Z2B2 vectors described above. As a
further example, when pgGATA was digested with Clal and NotI, without dephosphorylation or further purification so that the small Clal-NotI fragment from the polylinker remained, and then was ligated in the presence of an equal molar ratio of an 800-bp Clal-NotI DNA fragment, 100% of the colonies selected with IPTG contained the 800-bp fragment in the vector (not shown).

In conclusion, our results demonstrate that the toxicity of the GATA-1 DNA-binding domain can be used as a very powerful positive selection cloning system. For routine cloning, the vector does not have to be dephosphorylated or purified after restriction enzyme digestion, nor does the vector have to be completely digested. Moreover, the ratio of vector to insert is unimportant because even when the vector-to-insert ratio is 10:1, the majority of clones selected on IPTG-containing medium contain the desired insert (Table 2 and Figure 5a). However, in special situations (e.g., cDNA libraries) where the positive clones-to-background ratio must be maximized, it would be advisable to ensure the ratio of insert to vector is at least 1 to 1, as we found that at this ratio 95% or greater of IPTG-selected clones contain insert (not shown). Moreover, as was described above, dephosphorylation of the vector also reduces the background ratio considerably, but in this case the cloning efficiency is also reduced.

The vector we have developed, pgGATA, being a high copy number vector, routinely gives several micrograms of DNA from only 1 mL of culture. The selection system uses the common reagent IPTG, which is inexpensive. The results indicate that the toxic effect may be at the level of GATA-1 binding to the origin of replication, reflecting an inhibition of DNA A insert:protein. If this is the case, it may work in any strain of E. coli, and we have found that the selection system works in at least 4 different strains of E. coli commonly used. Since the DNA A protein and the DNA sequences it binds to in bacteria originate of replication are highly conserved among different species of bacteria (10), it may also be that the pgGATA system will be applicable to other species of bacteria. Thus, the pgGATA vector fits the criteria for a good positive selection system and should make routine cloning of DNA fragments simpler, especially when limiting quantities of DNA are available.

ACKNOWLEDGMENTS

We thank Chantal Nault for typing this manuscript and Rashmi Kothary for his critical comments. This research was initiated under a grant from the Miles/Canadian Red Cross Society Research and Development Fund to P.T. and L.W. and was concluded under a grant from the National Cancer Institute of Canada (to L.W.). Sylvie Provost was the recipient of a Medical Research Council of Canada studentship, and Lee Wall is the recipient of a scholarship from the Fonds de la Recherche en Santé du Québec.

REFERENCES


Received 19 May 1995; accepted 27 September 1995.

Address correspondence to:
Lee Wall
Institut du Cancer de Montréal
Centre de Recherche L.-C. Simard
1560 Sherbrooke East
Montréal, QC H2L 4M1, Canada
Internet: walll@ere.umontreal.ca