Journal of Experimental Microbiology and Immunology (JEMI) Vol. 18: 91 – 95

Nucleotide Composition Immediately Flanking the DNA Overhangs Generated by Restriction Endonucleases was not Demonstrated to Contribute to the Ligation Efficiency of T4 DNA Ligase

ChiaYing Chen, SeungHyun Choi, Yue Hu, Mei Mu Zi Zheng
Department Microbiology & Immunology, University of British Columbia

The apparent efficiency of T4 DNA ligase has been observed to vary for fragments cut by different restriction enzymes. In this study, restriction enzymes that leave the same overhang were used to investigate whether nucleotide content immediately flanking the overhangs generated by the restriction endonuclease digestion affect the ligation efficiency of T4 DNA ligase on pBR322 and pACYC177. Two sets of enzymes were used to digest the plasmids, one set resulted in G,C nucleotides immediately flanking the staggered ends, while the other set had A,T nucleotides. Our results suggest that on average T4 ligase showed higher calculated number for ligation efficiencies of fragments cut by G,C-enzymes than A,T- enzymes. However, the small sample size limited our ability to obtain statistically significant results.

In molecular biology, restriction endonucleases (RE) are widely employed for the construction of recombinant DNA molecules and cloning. These enzymes recognize specific palindromic DNA sequences and sever the phosphate-sugar backbone of both DNA strands (1). Some products of these restriction enzymes are blunt-ended, while others are staggered-ended, resulting in cohesive or sticky ends. Fragments cut by the same restriction enzyme have complementary sticky ends, which enable them to anneal together (1). The ends can then be annealed by another type of enzyme, a DNA ligase, which catalyzes the formation of a covalent phosphodiester bond between 5'-phosphate group of one strand and 3'-OH of the other (2). Restriction endonucleases and DNA ligases are both widely used in molecular biology and therefore understanding the factors under which they function would facilitate construction of recombinant DNA molecules and cloning.

DNA samples can be assayed by agarose gel electrophoresis. The speed at which DNA travels through these gels is affected by their size and shape (3). Electrophoresis of circular plasmids may result in the formation of several bands, each representing plasmids of a different topology: relaxed circular, concatenanes and supercoiled (4). Digestion of plasmids with a RE results in linear fragments that travel at different rates in gel electrophoresis (3).

There are restriction enzymes that recognize and cut the same palindromic sequence in nucleic acids, known as isoschizomers. For example, *Ase*I (5'-AT↓TAAT-3') and *Vsp*I (5'-AT↓TAAT-3') are isoschizomers. Because they are isolated from different bacterial strains, the subsequent ligation may behave differently (2). In this way, we are using isoschizomers as different enzymes to increase the effective size of our samples.

It has been observed that cohesive ends containing more G,C nucleotides than A,T nucleotides have a higher ligation efficiency (5). Moreover, lambda DNA fragments cut by *Nde*I resulting in an A,T flanking the overhangs have a lower ligation efficiency compared to *Hind*III cut fragments, which are flanked by G,C nucleotides (6, 7). Because A,T base pairs have two hydrogen bonds keeping them together, while G,C pairs have three, A,T base pairing is thought to be weaker than that of G,C. Therefore, the A,T nucleotide base pair flanking the staggered ends should provide a higher probability of unzipping the DNA oligonucleotide, thus hindering ligation by T4 ligase. In this experiment, we compared the ligation efficiencies of T4 DNA ligase on plasmids cut by restriction enzymes that result in either A,T or G,C nucleotides immediately flanking the overhangs.

MATERIALS AND METHODS

Growth of DH5α *Escherichia coli* cells. One colony of DH5α cells containing pACYC177 (MICB 421 culture collection, Department of Microbiology & Immunology) was streaked on Luria agar plates (1% NaCl, 1% tryptone, 0.5% yeast extract and 0.75% Bacto agar) supplemented with 100 µg/ml filter-sterilized ampicillin. The plates were incubated at 37°C for 16 hours. One colony was chosen to be inoculated to Luria broth (1% NaCl, 1% tryptone, and 0.5% yeast extract). Growth occurred in a 37°C shaking incubator at 200 rpm for 19.5 hours.

Isolation of pACYC177: The cultures were centrifuged at 6000 x g for 15 min at 4°C in the Beckman Model J2-21 Centrifuge (Beckman Rotor JA-14). The supernatant was removed and the pellet was used for isolation of pAYCY177 using the Qiagen QIAfilter Plasmid Maxi Kit (#12262). Procedure for low copy plasmid was followed as outlined in the QIAfilter Plasmid Purification Handbook, 3rd Edition. The sample was stored in the -20°C freezer until further processing.

Quantitation of DNA: The concentration and purity of the isolated pACYC177 was determined by spectrophotometry using a Nanodrop (NanoDrop 2000c spectrophotometer, Thermo Scientific)

Digestion of pACYC177 and pBR322 by restriction enzymes: The restriction enzymes were grouped into two sets based on the base pair (bp) immediately flanking the overhangs. The properties of the six restriction enzymes used are presented in

Table 1. Recommended plasmid digestion protocols for enzymes *Pvu*I, *Ase*I, *Nde*I, *Vsp*I, *AsiS*I, *SfaA*I, and *Hind*III were used (Thermo Scientific catalogues). Restriction sites of each RE on pACYC177 and pBR322 are shown in Figure 1.

*Hind*III was used as a control for digestion of both pACYC177 and pBR322. Each reaction was made up to a final volume of 100 ul with DNA concentration of 1 ug/ul. *Pvu*I, *Nde*I, *Vsp*I, *SfaA*I, and *Hind*III digestion reactions were incubated at 37°C for 6 hours. *Ase*I and *Asis*I were digested in FastDigest Buffer (Thermo Scientific) for 5 minutes at 37°C, and subsequently heat inactivated at 84°C for 5 minutes.

TABLE 1 Plasmids with endonuclease recognition site.

Restriction Enzymes	Plasmids with Restriction Recognition	Digestion Site (one cut site per plasmid)	Nucleotides Flanking Overhangs
*Nde*I	pBR322	5′...CATATG...3′ 3′...GTATAC...5′	A/T
*Vsp*I	pBR322	5′...ATTAATT...3′ 3′...TAATTAA...5′	A/T
*Pvu*I	pBR322	5′...CGATCG...3′ 3′...GCTAGC...5′	G/C
*Ase*I	pBR322	5′...ATTAATT...3′ 3′...TAATTAA...5′	A/T
*SfaA*I	pACYC177	5′...GCGATCGC...3′ 3′...CGCTAGCG...5′	G/C
*AsiS*I	pACYC177	5′...GCGATCGC...3′ 3′...CGCTAGCG...5′	G/C
*Hind*III	pBR322, pACYC177	5′...AAGCTT...3′ 3′...TTCGAA...5′	A/T (used as control)

DNA extraction/purification using phenol/chloroform: The digested samples were purified to remove unwanted proteins and salts using phenol-chloroform extraction and ethanol precipitation (8).

Partial T4 DNA ligation: 200 ng of each digested sample were used to perform ligation reactions using 0.5 units of T4 DNA ligase (1 Weiss U/ul) according to the recommended protocols in the Thermo Scientific catalogue for T4 ligase. Each reaction was made up to a final volume of 20 ul. The ligation reactions were performed for all of the samples at 24°C for 1 hour in the thermal cycler (Techne Thermal Cycler PHC-3, Mandel Scientific Company Ltd). The samples were then immediately stored in the -20°C freezer until further processed.

Agarose gel electrophoresis: 20 ul (10 ng/ul of DNA) of each ligated sample and each sample that was restriction enzyme digested but unlighted combined with 6X loading buffer and electrophoresed. Undigested pACYC177 and pBR322 were used as standards and *Hind*III digested pACYC177 and pBR322 were used as positive controls for the digestions. Agarose (Invitrogen) gels (1.0 % w/v) in 1X TAE buffer (40 mM Tris pH 8.0, 0.114% v/v glacial acetic acid, 1 mM EDTA) were used to analyze products of digestion and ligation reactions. The speed at which DNA travels through these gels is affected by their size and shape (3). High DNA Mass Ladder (Invitrogen) was used as a size standard to locate the fragments containing the linear bands. Electrophoresis was performed at 120 volts for 1 hour. The gel was stained in 0.5 ug/ml ethidium bromide bath for 20 minutes prior to visualization using a MultiImage Light Cabinet (Alpha Innotech Corporation). A digital image of the gel was analyzed using the AlphaImager program (Figure 2).

Determining T4 ligase efficiency: The AlphaImager (Alpha Innotech, San Leandro, CA, USA) was used to quantify specific DNA bands on the agarose gel. After entering in the concentration of the standard ladder bands, the Spot Density Tool was able to determine the DNA concentration of selected bands, based on the intensity of the bands after background correction. The differences in concentration of the linear fragments before and after ligation reactions were used to determine the DNA T4 ligation efficiency. Data was statistically analyzed using the independent samples Mann-Whitney nonparametric statistical test (IBM SPSS Statistics 20).

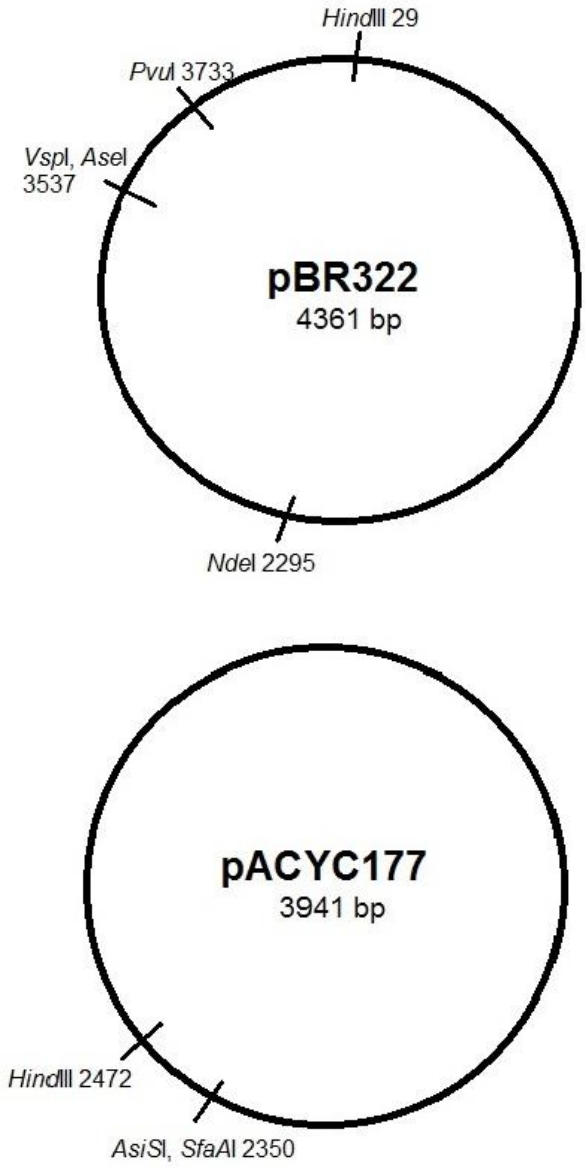

FIG 1. Plasmid map of pBR322 and pACYC177 plasmids and the cut sites by restriction endonucleases used. The numbers indicate the position of the restriction site on the plasmid for each enzyme.

RESULTS

Identification of ligated and unligated plasmid bands of RE digestions. Unligated pBR322 (lane 2, Fig. 2) and pACYC177 plasmids (lane 13) were used as controls to compare the band patterns to those of RE digestions and ligated samples. The linear, double-stranded pBR322 band was expected to appear in between the 6000 bp and 4000

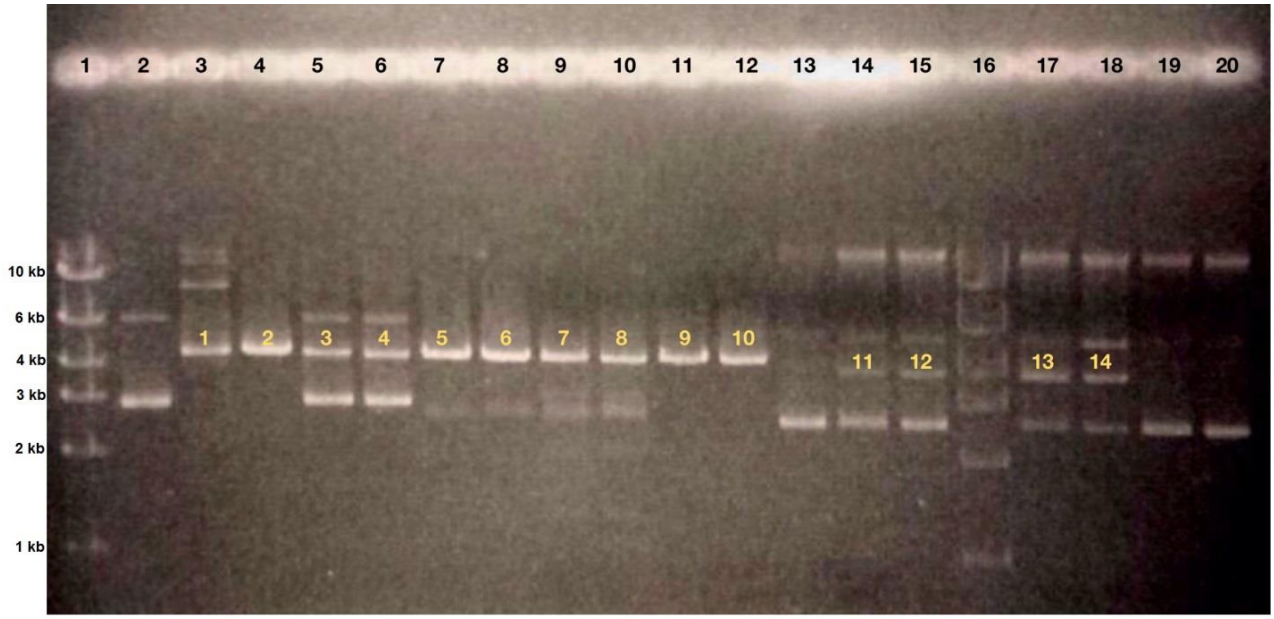

FIG 2. Ligation and digestion reactions of *Hind*III, *Ase*I, *Vsp*I, *Nde*I, *Pvu*I, *Sfa*AI, and *Asi*SI restriction enzyme treated pBR322 and pACYC177. Lane 1: 0.52 ug High DNA Mass Ladder (fragment sizes as indicated); Lane 2: 0.2 ug pBR322 (negative control); Lane 3: 0.2 ug ligated *Hind*III digest; Lane 4: 0.2 ug *Hind*III digest; Land 5: 0.2 ug ligated *Ase*I digest; Lane 6: 0.2 ug *Ase*I digest; Lane 7: 0.2 ug ligated *Vsp*I digest; Lane 8: 0.2 ug *Vsp*I digest; Lane 9: 0.2 ug ligated *Nde*I digest; Lane 10: 0.2 ug *Nde*I digest; Lane 11: 0.2 ug ligated *Pvu*I digest; Lane 12: 0.2 ug *Pvu*I digest; Lane 13: 0.2 ug pACYC 177 (negative control); Lane 14: 0.2 ug ligated *Hind*III digest; Lane 15: 0.2 ug *Hind*III digest; Lane 16: 0.52 ug High DNA Mass Ladder; Lane 17: 0.2 ug ligated *SfaA*I digest; Lane 18: 0.2 ug *SfaA*I digest; Lane 19: 0.2 ug ligated *Asi*SI digest; Lane 20: 0.2 ug *Asi*SI digest. Lanes 3 to 12 are pBR322 reactions, and Lanes 14, 15, 17, 18, 19, and 20 are pACYC 177 reactions. The numbered bands were analyzed in Table 2.

bp bands in lane 1. Lane 2 contained two bands; the upper one is expected to contain plasmids of the relaxed covalently closed circular (CCC) conformation, while the lower one to contain supercoiled plasmids (4). In lane 13, four different bands were observed for the isolated pACYC177 plasmids. The first band from the top likely contained relaxed CCC plasmids, and the bottom one contained supercoiled plasmids. The migration of the second band was equivalent to the distance of a 5,800 bp linear plasmid, indicating a different conformation of the plasmids. The third band had a molecular weight (MW) of 3,850 bp, as calculated from to the linear mass ruler, which was very close to the size of pACYC177 plasmid (3,941 bp), suggesting the band contained linear fragments. However, considering that the plasmids were not treated with any restriction enzymes, the appearance of the linear band might have resulted from improper handling, long storage time, or possibly contamination with nucleases.

A single linear band (band 2) in *Hind*III-digested pBR322 (positive control, lane 4, Fig. 2) was located at 4370 bp, indicating a complete digestion of *Hind*III. In lane 3, the observation of less intense linear band and higher MW bands indicated a partial ligation. The size of the third band in lane 3 was calculated to be 8,800 bp, which is expected to be a ligated product of two linear fragments. The top two bands migrated slower than the third band and, therefore, we suspected the top band was the open circular form of the linear third band while the second was CCC. The *Hind*III-digested pACYC177 (positive control, lane 15) showed four bands, identical to those of the isolated plasmids (plasmid control, lane 13), but differing in band intensities, indicative of an incomplete digestion. The reduced intensity of the band 11compared to the band 12 inferred a partial ligation.

For each enzyme, we expected the band corresponding to the linear fragments to be more intense in the unligated sample than that of the ligated sample, resulting from successful partial ligation. Bands 1 to 10 (Fig. 2) were calculated to be 4,370 bp using the standard curve of the mass ruler (lane 1), which is close to the size of pBR322. Bands 11 to 14 have a MW of 3,850 bp, also consistent with the size of pACYC177 plasmid. Therefore, these bands contain digested linear pBR322 and pACYC177 fragments. Referring to our plasmid control (lane 2), *Ase*I ligations and digestions in lanes 5 and 6 contained relaxed CCC pBR322 plasmids at the top band and supercoiled ones at the bottom band. *Vsp*I and *Nde*I ligations and digestions in lanes 7 to 10 showed that the bottom bands were slightly lower than the expected supercoiled pBR322 band, suggesting that they were plasmids with different degrees of supercoiling. The *Pvu*I ligation sample (band 9) contained only a single band with the expected MW of linear pBR322, but the intensity was only slightly less than that of the *Pvu*I digestion (band 10), indicative of a low ligation efficiency or a lack of ligase activity. In lane 20, *Asi*SI enzyme digestion activity was not observed since no band was present around the 3,900 bp size. The top bands in lanes 19 and 20 were expected to be the relaxed CCC

and the lower bands are expected to be the supercoiled plasmids. For *Ase*I, *Vsp*I, *Nde*I, and *SfaA*I restriction enzymes, proper digestion and ligation reactions were observed.

TABLE 2 Ligation efficiency of plasmids digested by restriction endonucleases with A/T or G/C nucleotides immediately flanking the restriction site.

Nucleotide Composition Flanking Overhangs	Restriction Enzyme	Ligation Efficiency (%)	Mean Rank
A/T	*Ase*I	4.85	2.0
(n=3)	*Vsp*I	5.78	
	*Nde*I	4.03	
G/C	*Pvu*I	5.93	4.5
(n=2)	*Sfa*I	12.33	
	*Asis*I	No data	

U=0.00, p=0.20, two-tailed

Comparison of T4 ligase efficiency on plasmids cut by RE. The ligation efficiency was determined by taking the difference in DNA concentration of the linear fragments between the ligated and unligated lanes of each enzyme. The difference was then divided by the concentration of the linear fragments in the unligated sample to obtain the percent ligation efficiency. The mean rank of ligation efficiency of T4 ligase on plasmids cut by G,C-enzymes (*Pvu*I and *Sfa*I) was higher than that of A,T-enzymes (*Ase*I, *Vsp*I, and *Nde*I), but the difference was not statistically significant (U=0.00, p=0.20, two-tailed) as determined by the Mann-Whitney test (Table 2). *AsiS*I was not included in the analysis as bands corresponding to the linear fragments did not appear in the lanes with *AsiS*I.

DISCUSSION

The conformations in which plasmids exist affect the speed at which they travel during electrophoresis (3). *E.coli* cells containing DH5α replicate by theta replication mechanism which may result in various conformations of plasmids: relaxed circular, concatenanes and supercoiled (4). In Figure 2, lane 2 was used as a control to show the different conformations of circular plasmid present in pBR322. Since the relaxed CCC has a greater effective radius than that of the supercoiled plasmid, the plasmids in relaxed CCC were expected to migrate slower in comparison to supercoiled plasmids (3). Therefore, for undigested pBR322 bands in lane 2, we expected that the upper band to be in CCC while the lower to be supercoiled. For undigested pACYC177 (lane 13), four bands were observed. The top band was expected to be the relaxed CCC because CCC plasmid travels the slowest in 1% agarose gel compared to concatenanes, linear and supercoiled conformations (3). The second band from the top was expected to contain concatenanes isolated from cells that are undergoing replication. The supercoiling in the concatenane conformation causes these plasmids to travel faster than relaxed CCC. The third faint band from the top was expected to be the linear form of pACYC177 because the MW of which (3,850 bp), based on its migration distance, was similar to the size of pACYC177 (3941bp). Although the plasmid was not treated with any restriction enzymes, the presence of the linear band suggests that our plasmid sample was likely sheared during pipetting or isolation processes, or breakage due to repetitive freezing and thawing and long storage time.

Figure 1 shows that all restriction enzymes would cut the plasmid at a single site to give a linear band equal to the size of the plasmid. The third band in lane 3 is close to double the size of pBR322, which would indicate a successful ligation between two linear fragments. To compare the band containing the linear fragment in lane 11 (*Pvu*I digestion ligated) with the one in lane 12 (unligated), we used the measured concentration data because the band intensities seemed similar by sight. However, the data indicated a slight decrease in concentration of the linear products in lane 11 when compared to lane 12 (less than 3% change). We expected that this was due to a very low ligation activity which resulted in no observable ligated bands in lane 11 because of the low concentration of ligated fragments generated. But because this difference (3%) is within the uncertainty range, we were not able to make a conclusion on the ligation efficiency.

*Hind*III digested pACYC177 (lanes 14 and 15) showed similar banding patterns as the undigested pACYC177 (lane 13), indicative of incomplete digestion of plasmids. Band 12 has a higher concentration than band 11, and this is expected as a result of T4 DNA ligase activity. We expected bands 13 and 14 to be the linear fragments of pACYC177 after ligation and without ligation, respectively. Neither the concatenane nor the linear form of the plasmids was observed, which was not expected but could be due to insufficient digestion period.

Although the average calculated ligation efficiency for the G,C group was a higher number than that of the A,T-enzymes (Table 2), the small sample size and large standard deviation limits our ability to draw conclusions. Because the sample size was limited, we used a nonparametric statistical test that does not make any assumptions about the distribution of the data. Using the Mann-Whitney test, we found that the distribution of difference in ligation efficiency was not statistically different across the enzymes with A,T and C,G immediately flanking the restriction site.

Furthermore, the ligation efficiency of the A,T-enzymes are below 5%, which is within the range of uncertainty expected from technical errors such as precision of the pipettor or the operator. The low ligation efficiency could be due to insufficient ligation duration for plasmids cut with endonucleases that leave short overhangs. We expected a higher ligation efficiency for *Hind*III digested plasmids compared to that of the other enzymes we used because longer overhangs were shown to have better ligation efficiencies than shorter ones (10). Thus, extending the ligation time for our samples might have increased the ligation efficiencies. Consequently, we were unable to statistically demonstrate a difference in ligation efficiency as attributed by the nucleotide composition flanking the staggered ends.

Our results showed high ligation efficiency in the T4 ligase in both pBR322 and pACYC177 digested with *Hind*III (35.79 ±1.44%), that we used as positive control to demonstrate effective ligation. This was consistent with previous research that showed that *Hind*III-digested plasmid ligated more efficiently than *Nde*I-digested plasmid (9). *Hind*III has G,C nucleotides immediately flanking the staggered ends, which differs from *Nde*I in that *Nde*I results in A/T nucleotides flanking the overhangs. Besides the composition of the nucleotide immediately flanking the overhangs, *Nde*I and *Hind*III also differ in the length of the overhangs, which may play a role in ligation efficiency of T4 ligase (10).

In summary, the nucleotides immediately flanking restriction endonuclease digested overhang were not statistically demonstrated to contribute to the ligation efficiency by T4 DNA ligase. Specifically, the G,C nucleotides flanking short overhangs (2 bp in our experiment) were not demonstrated to contribute to significantly higher ligation efficiency, compared to fragments with A,T. An U value of 0 was obtained in the Mann-Whitney test, which means each ligation efficiency value in the A,T group was lower than that of the G.C group. Normally a low U value would contribute to a low p value. However, even though fragments with G,C showed higher calculated ligation efficiency numbers on average, the small sample size limited our ability to obtain statistically significant results (p=0.20). Further experimentation involving a larger sample size is needed to confirm this correlation.

FUTURE DIRECTIONS

To increase the validity of the results, more replicates should be prepared for each endonuclease. To overcome the problem of low DNA yield from DNA isolation kits, PCR could be employed to provide more starting plasmid DNA for more replicates. In addition, the ligation efficiency in our experiment using enzymes with an overhang length of two nucleotides has shown to be too low to effectively compare between the A,T and G,C groups. Therefore, to obtain results outside the range of uncertainty caused by technical errors, either a longer ligation time can be applied to increase the ligation efficiencies of short overhangs, or endonucleases that leave longer overhangs can be chosen for future studies. Last, a larger sample size is needed to obtain statistically significant results for this study. Keeping the median value of the ligation efficiencies constant for both groups, the statistical software used predicts that a minimum sample size of four (n=4) is required for each group to obtain a statistically significant result ($p<0.05$).

ACKNOWLEDGEMENTS

This project was supported by the Department of Microbiology and Immunology at University of British Columbia. We would like to sincerely thank Dr. William Ramey for his invaluable guidance and patience throughout the course of the project. We would also like to thank Kirstin Brown for her constant guidance and encouragement throughout the term.

REFERENCES

1. **Corkill G, Rapley R**. 2008. Molecular Biomethods Handbook, p. 3–15. *In* Walker, JM, Rapley, R (eds.), 2nd ed. Humana Press, New York, New York.
2. **Lehman IR**. 1974. DNA ligase: structure, mechanism, and function. Science **186**:790–797.
3. **Mickel S, Arena V, Bauer W**. 1977. Physical properties and gel electrophoresis behavior of R12-derived plasmid DNAs. Nucleic Acids Res. **4**:1465–1482.
4. **Del Solar G, Giraldo R, Ruiz-Echevarría MJ, Espinosa M, Díaz-Orejas R**. 1998. Replication and control of circular bacterial plasmids. Microbiol. Mol. Biol. Rev. **62**:434–64.
5. **Harada K, Orgel LE**. 1993. Unexpected substrate specificity of T4 DNA ligase revealed by *in vitro* selection. Nucleic Acids Res. **21**:2287–2291.
6. **Chang E, Ge BIN, Lee M, So M, Wang W**. 2005. Investigation of the ligation efficiency of *Nde*I digested fragments. J. Exp. Microbiol. Immunol **7**:68–72.
7. **Harada K, Orgel LE**. 1993. *In vitro* selection of optimal DNA substrates for T4 RNA ligase. Proc. Natl. Acad. Sci. U. S. A. **90**:1576–1579.
8. **Sambrook JF, Russell DW**. 2001. Molecular Cloning: A Laboratory ManualMolecular Cloning: A Laboratory Manual, 3rd ed. Cold Spring Harbor Laboratory Press, New York, New York.
9. **Chow P**. 2005. Cloning of λ DNA fragments into pUC19 vector to study the ligation efficiency of *Nde*I-digested pUC19 and *Hind*III-digested pUC19 by T4 DNA ligase. J. Exp. Microbiol. Immunol. **8**:8–13.
10. **Bola G**. 2005. Evaluating the role of G, C-nucleotides and length of overhangs in T4 DNA ligase efficiency. J. Exp. Microbiol. Immunol **8**:1–7.