J Mol Evol (1994) 38:558–560

Journal of Molecular Evolution

© Springer-Verlag New York Inc. 1994

In Vitro Selection of Optimal DNA Substrates for Ligation by a Water-Soluble Carbodiimide

Kazuo Harada, Leslie E. Orgel

The Salk Institute for Biological Studies, P.O. Box 85800, San Diego, CA 92186-5800, USA

Abstract. We have used in vitro selection to investigate the sequence requirements for efficient template-directed ligation of oligonucleotides at 0°C using a water-soluble carbodiimide as condensing agent. We find that only 2 bp at each side of the ligation junction are needed. We also studied chemical ligation of substrate ensembles that we have previously selected as optimal for ligation by RNA ligase or by DNA ligase. As anticipated, we find that substrates selected with DNA ligase ligate efficiently with a chemical ligating agent, and vice versa. Substrates selected using RNA ligase are not ligated by the chemical condensing agent and vice versa. The implications of these results for prebiotic chemistry are discussed.

Key words: In vitro selection — Chemical ligation — Water-soluble carbodiimide

Introduction

Nonenzymatic template-directed oligomerization of nucleotides has been studied extensively as a model for replication of nucleic acids. It has been shown that some polynucleotides will direct the synthesis of complementary oligomers from activated monomers (Orgel 1992). However, replication has not been achieved and is not likely to be achieved using the systems studied up to now. An alternative model for the replication of DNA or RNA involves the ligation of short, preformed oligonucleotides (Ninio and Orgel 1978; Zielinski 1987; Doudna and Szostak 1989; von Kiedrowski et al. 1989). Sequences that have been brought together using a complementary template can be joined efficiently with chemical condensing agents such as water-soluble carbodiimides (Dolinnaya et al. 1988) or cyanogen bromide (Kanaya and Yanagawa 1986; Dolinnaya et al. 1991). However, it is not known how faithfully an oligonucleotide can be copied, because the dependence of the ligation efficiency on the extent of the complementarity between template and substrates has not been explored. In an attempt to address this question, we have used an in vitro selection technique (Szostak 1992) that we have previously described in detail (Harada and Orgel 1993a,b). Using this method, we were able to characterize an ensemble of DNA substrates that are ligated efficiently by a water-soluble carbodiimide.

Materials and Methods

All materials with the exception of the following were obtained from the same sources as those listed in the previous report (Harada and Orgel 1993a). 1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDAC) was obtained from JBL Scientific Corporation, and 2-(*N*-morpholino)ethanesulfonic acid (MES) came from the Sigma Chemical Company.

Oligodeoxynucleotide sequences (Fig. 1A) are described in the previous report (Harada and Orgel 1993a). The procedure for selection and amplification was identical to that previously described (Harada and Orgel 1993a) except that ligation was performed using EDAC in the place of T4 RNA ligase. Ligation reactions were carried out at 0°C with 0.05–0.5 μg (2–20 pmol) of the 5′-phosphorylated 66-mer Oligo 2, 50 pmol of 18-mer Oligo 1, 50 mM MES-Na (pH 6.0), 20 mM $MgCl_2$, and 0.1 M EDAC (Fig. 1A). The volume of the reaction mixture was 10 μl. Reaction products from 2-h ligation reactions were used to generate 66-mer substrate for the next round

Present address for Kazuo Harada: Gladstone Institute of Virology and Immunology, Department of Biochemistry and Biophysics, University of California, San Francisco, P.O. Box 419100, San Francisco, CA 94141-9100, USA
Correspondence to: Leslie E. Orgel

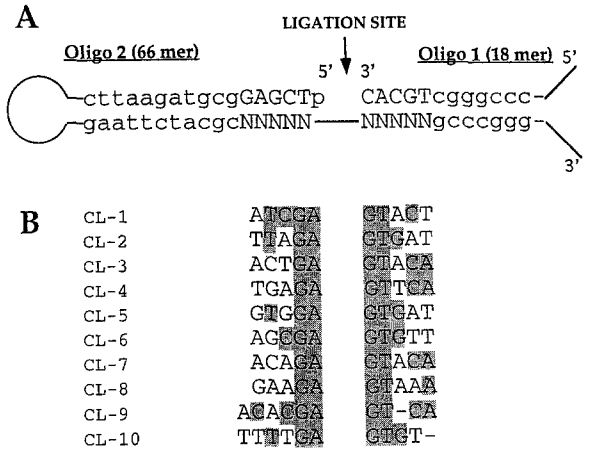

Fig. 1. **A** 18-mer (Oligo 1) and 66-mer (Oligo 2) substrates for ligation. Positions designated "*N*" were randomized with all four bases in approximately equal amounts. **B** The base sequences of the cloned oligomers in the region that was initially randomized (CL-1 to CL-10). Bases that form Watson-Crick base-pairs are *shaded.*

of selection. The mixture of sequences after four rounds of selection and amplification was used to clone individual sequences. Cloning and sequencing were carried out as previously described (Harada and Orgel 1993a).

Results and Discussion

The initial ensemble of 66-mer sequences (Oligo 2) (Fig. 1A) undergoes less than 0.2% ligation with 18 mer (Oligo 1) using EDAC as condensing agent after a 2-h reaction and less than 0.5% reaction after 18 h at 0°C. The ligated product from this 2-h reaction was used to generate the 66-mer substrate for the next round of selection. The ligation yield increased to 1.2% (18-h reaction). The efficiency of ligation increases with successive rounds of selection (Fig. 2). A ligation yield of 5.9% (18-h reaction) is obtained after four rounds of selection. This final selected mixture is ligated more than 10 times as efficiently as the initial ensemble.

Individual sequences in the ensemble obtained after four rounds of selection were cloned. The sequences of the randomized portion of the sequence are shown in Fig. 1B. As expected, we observed strong complementarity opposite the ligation junction. However, none of the 10 cloned sequences formed a perfect complementary stem, and only five to seven of the 10 bases within the randomized region were found to form base-pairs (with the possible exception of CL-9, which may be able to form 8 bp). Two Watson–Crick base-pairs on each side of the ligation junction seemed to be sufficient to ensure ligation by EDAC.

In previous experiments we have selected ensembles of 66-mers that ligate efficiently with T4 DNA ligase (Harada and Orgel 1993b) and T4 RNA ligase (Harada and Orgel 1993a). We used those ensembles as substrates together with the ensemble selected in the present experiments and a 66-mer that formed a completely complementary stem around the ligation junction. We prepared a battery of 12 ligation reactions with these four substrates using DNA ligase, RNA ligase, or the chemical ligating agent EDAC. The results are illustrated in Fig. 3.

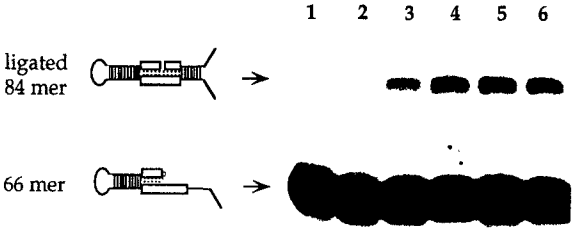

Fig. 2. Selection of substrates for efficient template-directed ligation by a water-soluble carbodiimide. Lane *1,* unselected ensemble (<0.5%); lane *2,* after round 1 (1.2%); lane *3,* after round 2 (2.6%); land *4,* after round 3 (4.9%); lane *5,* after round 4 (5.3%); lane *6,* after round 5 (5.9%). Reaction conditions: 66-mer (5 pmol), 18-mer (50 pmol), 100 mM MES-Na (pH 6), 20 mM magnesium chloride, 100 mM EDAC, 18 h reaction at 0°C.

Selection for chemical ligation gave rise to an ensemble that ligated readily with DNA ligase and vice versa (lanes 2 and 8). This is readily understood since both reactions depend on complementarity around the ligation junction. However, substrates that ligated efficiently with RNA ligase (lane 11) did not ligate efficiently with DNA ligase (lane 7) or the chemical ligation reagent (lane 3) and vice versa. This is again to be anticipated since RNA ligase acts on single-stranded but not double-stranded substrates.

Quantitative comparisons illustrate a number of additional points. RNA ligase has no detectable effect on the completely complementary substrate (lane 9) but gives about 1.7% of ligation with the ensembles selected with DNA ligase (lane 10) or EDAC (lane 12). Presumably, fraying around the ligation junction generates some single-stranded segments that can undergo ligation with RNA ligase if the substrates are not complementary over a sufficiently long distance.

Chemical ligation (lanes 1–4) occurs most efficiently with the fully complementary sequence (lane 1, 8.8%), almost as efficiently with the sequences selected by DNA ligase (lane 2, 7.1%), and significantly less efficiently with the ensemble selected by EDAC (lane 4, 4.6%). The fully complementary sequence must have been present in the initial ensemble. It follows that selection, although extensive, was incomplete after four cycles.

The experiments with DNA ligase (lanes 5–8) are consistent with the above conclusion and with the sequence data. The fully complementary sequence (lane 5) and the ensemble selected with DNA ligase (lane 6) are ligated with roughly equal efficiency (60%), but the ensemble selected by chemical ligation (lane 8) is a significantly poorer substrate (17%). Presumably DNA ligase requires a larger region of complementarity than EDAC.

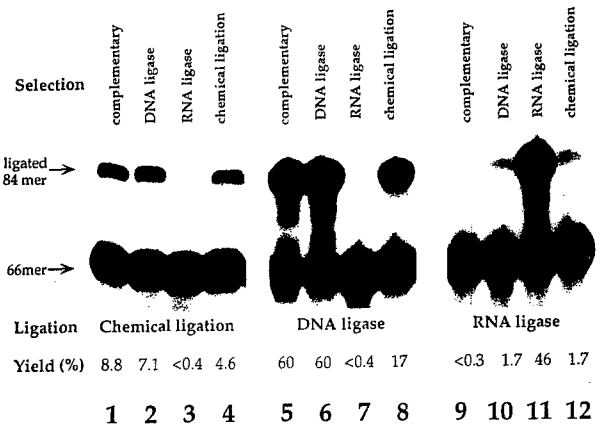

Fig. 3. A matrix of 12 ligation reactions. Ligation of a complementary control sequence and the three selected mixtures from the T4 DNA ligase, T4 RNA ligase, and EDAC selections were ligated with Oligo 1 using EDAC (lanes 1–4; 18-h reaction at 0°C), T4 DNA ligase (lanes 5–8; 1-min reaction at room temperature), and T4 RNA ligase (lanes 9–12; 1-h reaction at 0°C). For detailed reaction conditions, refer to previous reports (Harada and Orgel 1993a,b).

The efficiency of chemical ligation depends on two factors—the efficiency of hybridization and the rate of ligation in a well-formed hybrid. Our experiments imply that at 0°C hybridization is efficient when as few as two complementary bases are present in each oligomer at the ligation junction and that only 2 bp are needed to fix the termini of the substrate oligomers in a conformation productive for ligation. Incomplete hybridization and failure to define the geometry of the ligation junction together reduce ligation by at most a factor of two relative to that achieved with fully complementary substrates.

The above conclusion shows that fidelity in ligation reactions at temperatures higher than 0°C will depend only on the specificity of hybridization, and not on the length of the complementary sequence at the ligation junction. This makes it unlikely that a high degree of discrimination between closely related oligonucleotides can be achieved at any one temperature (or under any one temperature regime) for more than a very narrow range of substrates. This in turn places severe limits on the accuracy of any general scheme of molecular replication that depends upon template-directed ligation of oligonucleotide substrates.

Acknowledgments. This work was supported by NSCORT/EXOBIOLOGY grant No. NAGW-2881 and grant No. NAGW-1600 from the National Aeronautics and Space Administration. We thank Sylvia Bailey for manuscript preparation.

References

Dolinnaya NG, Sokolova NI, Ashirbekova DT, Shabarova ZA (1991) The use of BrCN for assembling modified DNA duplexes and DNA-RNA hybrids; comparison with water-soluble carbodiimide. Nucleic Acids Res 19(11):3067–3072

Dolinnaya NG, Sokolova NI, Gryaznova OI, Shabarova ZA (1988) Site-directed modification of DNA duplexes by chemical ligation. Nucleic Acids Res 16(9):3721–3738

Doudna JA, Szostak JW (1989) RNA-catalyzed synthesis of complementary-strand RNA. Nature 339(6225):519–522

Harada K. Orgel LE (1993a) *In vitro* selection of optimal DNA substrates for T4 RNA ligase. Proc Natl Acad Sci USA 90:1576–1579

Harada K. Orgel LE (1993b) Unexpected substrate specificity of T4 DNA ligase revealed by *in vitro* selection. Nucleic Acids Res 21(10):2287–2291

Kanaya E, Yanagawa H (1986) Template-directed polymerization of oligoadenylates using cyanogen bromide. Biochemistry 25(23): 7423–7430

Ninio J, Orgel LE (1978) Heteropolynucleotides as templates for non-enzymatic polymerizations. J Mol Evol 12:91–99

Orgel LE (1992) Molecular replication. Nature 358:203–209

Szostak JW (1992) *In vitro* genetics. Trends Biochem Sci 17:89–93

von Kiedrowski G, Wlotzka B, Helbing J (1989) Sequence dependence of template-directed syntheses of hexadeoxynucleotide derivatives with 3′-5′ pyrophosphate linkage. Ange Chem Int Ed Engl 28(9):1235–1237

Zielinski W (1987) Autocatalytic synthesis of a tetranucleotide analogue. Nature 327:346–347