NON-ENZYMATIC TEMPLATE-DIRECTED SYNTHESIS OF RNA COPOLYMERS[1]

Gerald F. Joyce

The Salk Institute for Biological Studies, Post Office Box 85800, San Diego, California 92138, USA

Abstract. Cytosine-rich RNA copolymers facilitate the template-directed synthesis of complementary oligomers from mononucleotide 5'-phospho-2-methylimidazolides (Fig. 1). The efficiency of this reaction falls sharply as the ratio of cytosine to non-cytosine in the template is decreased. This is a severe limitation for self-replication because it means that any cytosine-rich polynucleotide that can serve as a good template will produce a cytosine-poor complementary strand that is unable to direct further rounds of synthesis. Studies with low-ratio random copolymer templates have shown that the efficiency can be increased by adjusting initial monomer concentrations and by providing additional activated monomers during later stages of the reaction. The oligomeric reaction products can be studied in detail using high performance liquid chromatography. It is possible to separate oligonucleotides on the basis of chain length and base composition. Thus a wealth of information is available to characterize the distribution of products over the course of the reaction and under a variety of reaction conditions.

INTRODUCTION

At the present time it is not possible to construct a self-replicating system based on polynucleotides and activated mononucleotides. Polyguanylate forms a stable self-structure that prevents guanine-cytosine base pairing, polyuridylate forms a

triple-helix with adenosine derivatives, and polyadenylate will
not bind uridine mononucleotides (1).

Inoue and Orgel (2) have shown that a random polynucleotide
containing a substantial excess of cytosine will facilitate the
template-directed synthesis of complementary oligomers from
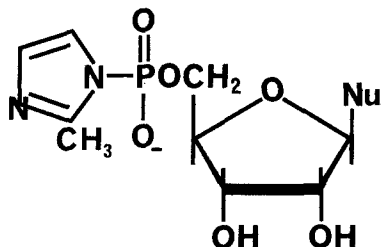mononucleotide 5'-phospho-2-methylimidazolides (Fig. 1).

Figure 1.   Structure of 2-MeImp Nu, where Nu is
cytosine (C), uracil (U), adenine (A), or guanine (G)

Monomers are incorporated into newly-synthesized oligomers
if and only if their complement is present in the template. How-
ever, the efficiency of this reaction falls sharply as the ratio of
cytosine to non-cytosine in the template is decreased. This is
a severe limitation for self-replication because it means that
any cytosine-rich polynucleotide that can serve as a good tem-
plate will produce a cytosine-poor complementary strand that
is unable to direct further rounds of synthesis.

REACTION CONDITIONS

The problem can be illustrated by the reaction of $poly(C_3U_1)$
template with activated guanine and/or adenine mononucleotides
in the presence of Mg^{2+}. With guanine alone the reaction pro-
ceeds rapidly and efficiently, while with adenine alone oligomer
formation does not exceed template-free controls. When both
guanine and adenine are present, both are incorporated into
newly-synthesized oligomers, though adenine incorporation is
only about half as efficient as guanine. The highest yields are
obtained with guanine alone, producing oligomers up to about the
30-mer. Since the probability of encountering 30 consecutive
cytosine residues on the template is only about 10^{-4}, it seems
that guanine must be forming either intrahelical (wobble) or

extrahelical (loop-out) non-complementary pairs (3). Thus the relatively poor incorporation of adenine is a result of its own inefficiency as well as the overefficiency of guanine.

It has been possible to increase the efficiency of adenine incorporation both by increasing the concentration of adenine and by decreasing the concentration of guanine in the reaction mixture. Compared to the standard condition of 0.1 M 2-MeImpG and 0.1 M 2-MeImpA, doubling the concentration of 2-MeImpA to 0.2 M results in a 16% increase in the efficiency of adenine incorporation and a 6% increase in the efficiency of guanine incorporation. Presumably the greater concentration of 2-MeImpA allows it to compete more effectively at wobble sites (e. g. $3'CCU5'$), and increases the likelihood that newly-synthesized oligomers will be able to continue beyond "difficult" sites (e. g. $3'CUU5'$). Subsequently halving the concentration of 2-MeImpG to 0.05 M results in an additional 14% increase in the efficiency of adenine incorporation and a 6% decrease in the efficiency of guanine incorporation. The initiation of oligomer synthesis, which requires the formation of a 2-MeImpGpG dimer, is delayed under these conditions (4). Once oligomer synthesis is underway, 2-MeImpA enjoys a 4:1 concentration at wobble sites, while the amount of surviving 2-MeImpG is close to the threshold concentration necessary to sustain the reaction. If the concentration of 2-MeImpG is further reduced to 0.025 M, the efficiency of both adenine and guanine incorporation falls by over 50%.

Qualitatively similar but quantitatively less dramatic results are obtained with guanine & uracil on a poly(CA) template and with guanine & cytosine on a poly(CG) template. Unlike poly(CU), these templates do not contain wobble sites for guanine. However, cytosine and especially uracil have comparatively weak stacking interactions and thus can benefit from a concentration advantage.

HPLC ANALYSIS

The oligomeric products resulting from template-directed condensation of activated mononucleotides can be studied in detail by high performance liquid chromatography (HPLC). Using an RPC-5 column with a linear sodium perchlorate gradient (0 to 0.06 M over 90 minutes), it is possible to separate

oligonucleotides on the basis of chain length (5). The elution
profile is obtained by monitoring UV absorbance at 254 nm,
thus providing a quantitative description of oligomer yields.

Recently, it has also become possible to determine the
base composition of oligonucleotides containing guanine &
adenine, guanine & uracil, and guanine & cytosine. A random
copolymer template containing cytosine and one of the other
three bases is used to direct synthesis of ^{3}H-labeled oligonu-
cleotides. The reaction is carried out in paired tubes, with
[8-^{3}H]2-MeImpG added to the first tube and [8-^{3}H]2-MeImpX
(the other complementary mononucleotide) added to the second.
HPLC elution profiles are obtained for each tube by simulta-
neously monitoring UV absorbance and ^{3}H activity. For each
peak the ratio of UV absorbance to ^{3}H-labeled G activity to
^{3}H-labeled X activity is determined, and taking into account
differences in nucleotide extinction coefficients, integer values
for base composition are assigned. There is an obvious regu-
larity in the separation of subsequent oligomers in a composi-
tional series. In those cases where oligomers with different
base composition have identical retention times (e. g. (pG)$_7$ and
(pG)$_6$pA), this regularity is used to clarify the overlap in posi-
tion.

A detailed quantitative description of the oligomeric product
distribution is necessary for determining the sequence-specific
efficiency of template-directed synthesis. Studies with random
copolymer templates are seen as laying the groundwork for
future work with known sequence templates. Given a cytosine-
rich oligomer template, it is now possible to determine the
yield of its complement and of the various error copies simply
by examining the HPLC profile. This provides a method for
monitoring the time course of the reaction and for character-
izing the product distribution under a variety of reaction condi-
tions.

PULSING

"Pulsing" is a useful technique for studying the efficiency
and regiospecificity of monomer addition to a newly-synthesized
oligomer chain. The reaction of poly(C_3U_1) template with 2-
MeImpG produces a series of 3'-5' linked oligo(G)s, whose mean
chain length is a function of time, temperature, pH, [Mg^{2+}],
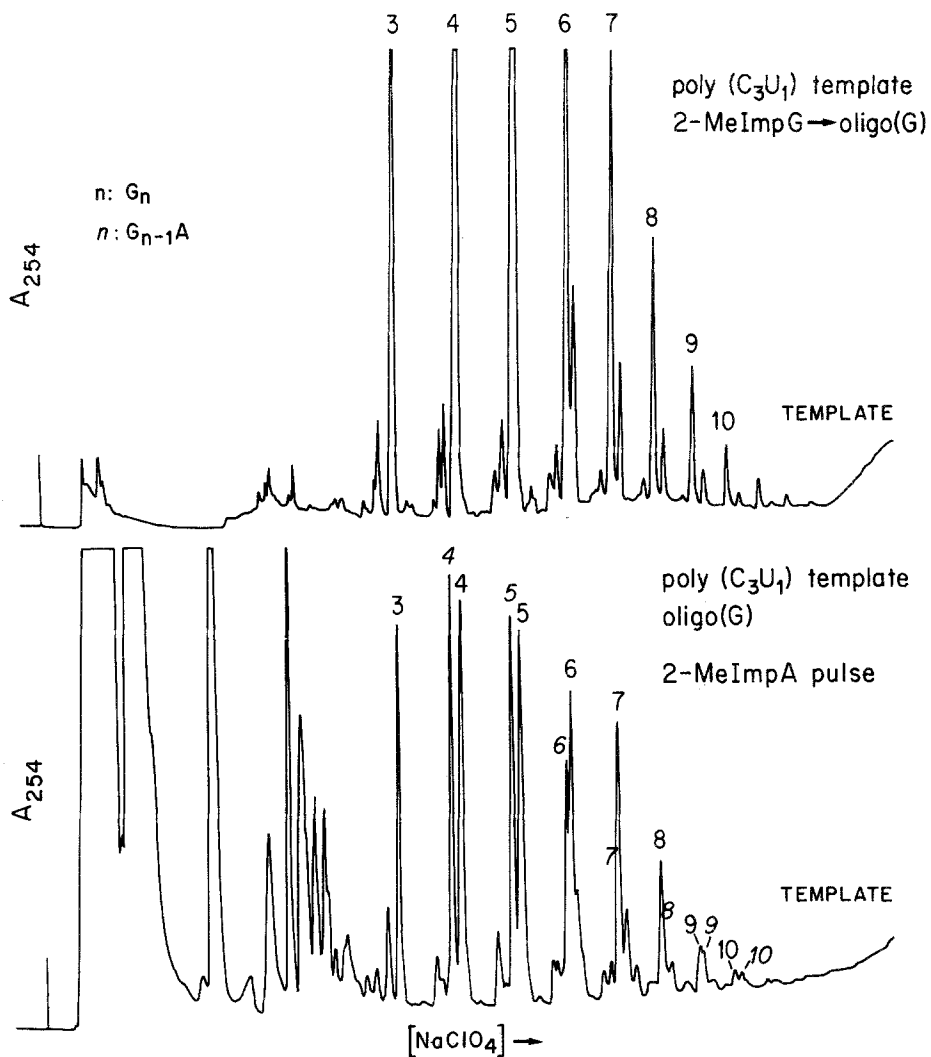
Figure 2.   HPLC elution profile of a pulsing experiment

[template], and [2-MeImpG]. The reaction can be interrupted
by dilution and subsequent acid hydrolysis of the guanosine imi-
dazolides. Unreacted monomers and dimers are removed by
gel filtration, and the solution is evaporated to dryness. The
reaction can then be resumed by rehydration with radioactive
2-MeImpA, producing oligo(G)s terminated by labeled adenine
residues.

The production of oligo(G_nA)s and oligo(G_nA_2)s is confirmed
by examining the HPLC elution profile (Fig. 2). The site of 2-
MeImpA addition to an oligo(G) primer can be determined by
exhaustive alkaline hydrolysis (5). Following KOH degradation,
radioactive adenine is distributed into A, Ap, and AppGp in a
ratio of approximately 7:1:2, corresponding to terminal sites at
the 3' end, internal sites at the 3' end, and terminal sites at the
5' end, respectively. The regiospecificity of 2-MeImpA addition
at the 3' end can be determined by ribonuclease T_1 digestion (6).
This enzyme cleaves only 3'-5' internucleotide linkages on the 3'
side of a G residue. Following degradation with T_1, radioactive
adenine is distributed into A and $G^{2'}pA$ in a ratio of 3:1 and into
ApA and $G^{2'}pApA$ in a ratio of about 3:1. Thus the 3'-5' regio-
specificity at both terminal and internal sites is about 75%.

REKINDLING

CCGCC was the first known-sequence RNA template to be
used to demonstrate non-enzymatic template-directed synthesis
of complementary oligomers (7). The reaction of CCGCC with
2-MeImpG and 2-MeImpC produces a host of oligo(GC)s, with
GGCG and GGCGG dominating the distribution of tetramers and
pentamers, respectively (Fig. 3). The identity of GGCGG was
confirmed by HPLC analysis and by its hydrolysis with pancre-
atic ribonuclease to GGC and GG (7).

When the reaction is allowed to reach completion (after
about 3 weeks), the yield of GGCG far exceeds that of GGCGG.
It might seem that if the newly-synthesized oligomers could
reach tetramer length, then it would be comparatively easy for
the final guanine residue to be added. However, by the time an
appreciable amount of GGCG has been formed, the supply of
remaining 2-MeImpG has been greatly depleted by ongoing hy-
drolysis. At this point the reaction can be "rekindled" by
removing the hydrolyzed monomers (using gel filtration) and
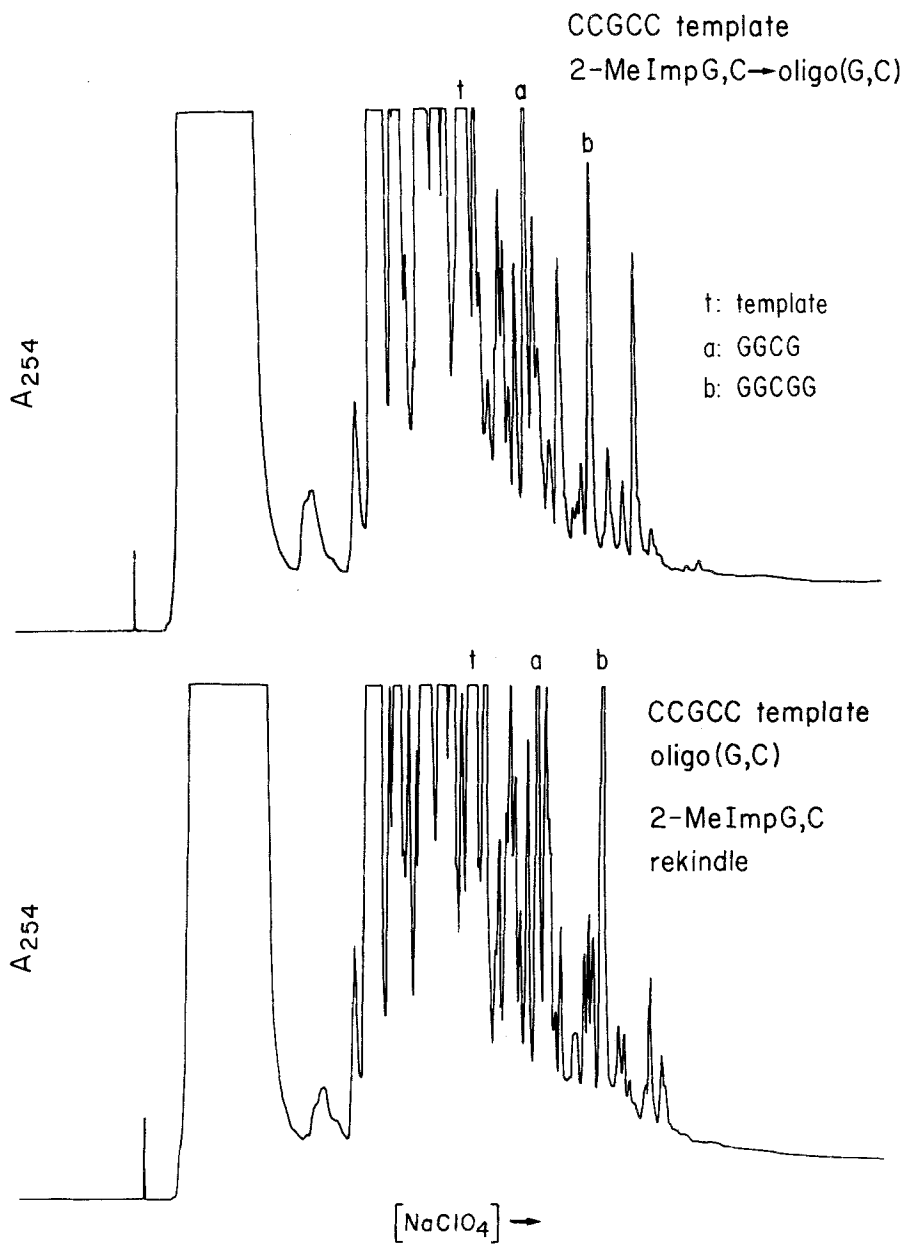
Figure 3.   HPLC elution profile of a recycling experiment

replacing them with activated compounds. The result is a rapid and dramatic increase in the yield of GGCGG.

A process equivalent to rekindling can be envisioned as operating continuously in the context of a flow reactor. The reaction vessel would contain a population of GC-rich oligonucleotides and a supply of activated mononucleotides. Its semipermeable walls would allow continuous removal of unreacted monomers, while fresh monomers are added to maintain the supply of activated materials. At the present time, a major obstacle to the realization of such a system is the requirement for a method to allow periodic disruption of the association between template and complement. This method must permit each strand to begin a new round of synthesis before the strand is made inaccessible by intra- or intermolecular duplex formation.

REFERENCES

1. Ts'o, P. 1974, "Basic Principles in Nucleic Acid Chemistry" (Academic Press, New York) Vol. 1, Chap. 6.

2. Inoue, T. and Orgel, L. E. 1983, Science 219, pp. 859-862.

3. Lomant, A. J. and Fresco, J. R. 1975, Progr. Nucleic Acid Res. Mol. Biol. 15, pp. 185-218.

4. Fakhrai, H., Inoue, T. and Orgel, L. E., unpublished results.

5. Lohrmann, R. and Orgel, L. E. 1980, J. Mol. Biol. 142, pp. 555-567.

6. Sulston, J. et al 1968, PNAS 60, pp. 409-415.

7. Brown, J. et al, unpublished results.