Two Enabling Architectures for DNA-Templated Organic Synthesis**

Zev J. Gartner, Rozalina Grubina, Christopher T. Calderone, and David R. Liu*

DNA templates can direct a wide variety of chemical reactions[1–24] without obvious structural requirements[23–25] by sequence specifically recruiting reactants linked to complementary oligonucleotides. The generality of DNA-templated synthesis raises the possibility of applying powerful processes such as in vitro selection and PCR (polymerase chain reaction) amplification to the discovery of synthetic molecules with desired properties.[23] In addition, the scope of DNA-templated synthesis provides an experimental basis for models of prebiotic chemistry [26–32] in which nucleic acid templates may have mediated the earliest translation of replicable information into a variety of functional structures. The DNA-templated format also enables modes of reaction that are not possible using conventional synthetic approaches.[23]

A DNA-templated chemical reaction is mediated by a specific architecture in which a portion of a DNA template anneals to a complementary DNA-linked reagent. Annealing places functional groups on the template and reagent in reactive proximity. The structure of this architecture could profoundly affect the nature of the resulting reaction, and it may therefore be possible to manipulate reaction conditions by rationally designing template–reagent complexes with different secondary structures. We previously used the end-of-helix (E) and hairpin (H) architectures (Figure 1) to mediate a variety of DNA-templated reactions.[23,25,33] During the course of these studies, two challenges emerged. First, some DNA-templated reactions do not proceed efficiently when the annealed reactive groups on the template and reagent are separated by even a small number of bases.[24] When the E or H architectures are used, these “distance-dependent” reactions can only be encoded by template bases at the reactive end of the template. Second, the presence of double-stranded DNA between annealed reactive groups can greatly reduce the efficiency of templated reactions because the flexibility of a single-stranded template is required.[23] This precludes the possibility of performing two or more reactions in a single DNA-templated step with the E or H architectures, even though the template oligonucleotide may contain enough bases to encode multiple reactions. Herein we describe two new architectures for DNA-templated synthesis that overcome each of these challenges.

We hypothesized that the distance dependence of certain DNA-templated reactions such as 1,3-dipolar cycloaddition and reductive amination could be overcome through the design of a new architecture that enables a reagent to anneal simultaneously to two widely separated regions of the template. In this omega (Ω) architecture (Figure 1), the template oligonucleotide contains a small number of constant bases at the reactive 5' end of the template in addition to distal coding regions. A reagent oligonucleotide for the Ω architecture contains at its reactive 3' end the bases that complement the constant region of the template followed by bases that complement a coding region anywhere on the template. The constant regions are designed to be of insufficient length to anneal in the absence of a complementary coding region. When the coding region of the template and reagent are complementary and anneal, however, the elevated effective molarity of the constant regions induces their annealing. Annealing of the constant regions forms a bulge (resembling an Ω) in the otherwise double-stranded template–reagent complex and places groups at the ends of the template and of the reagent in reactive proximity. In theory, this design would enable distance-dependent DNA-templated reactions to be encoded for the first time by bases distal from the reactive end of the template (Figure 1).

We compared the efficiency of the Ω architecture in DNA-templated synthesis with that of the standard E and H architectures. The Ω architectures studied consisted of a) three to five constant bases at the 5' end of the template, followed by b) a 5- to 17-base loop, and c) a ten-base coding region. As a basis for comparison, we performed four different DNA-templated reactions that collectively span the range of distance dependence that we have observed to date.

Amine acylation reactions are representative of distance-independent reactions that proceed efficiently even when considerable distances (e.g. 30 bases) separate the amine and carboxylate groups.[24,25] As expected, amine acylation (20 mM DMT-MM,[34] 30°C, pH 7.0, 12 h) proceeded efficiently (46–96% yield) in all architectures regardless of the distance between reactive groups on the reagent and template (Figure 2, lanes 1–5, and Figure 3a). The Ω architecture mediated efficient amine acylation with three, four, or five constant bases at the reactive ends of the template and reagent and 10 or 20 bases between annealed reactants (n = 10 or 20). Importantly, control reactions in which the distal coding region contained three sequence mismatches failed to generate significant product despite the presence of the complementary three- to five-base constant regions at the ends of the template and reagent (see Figure 2, lane 5 for a representative example). The Ω architecture therefore did not impede the efficiency or sequence specificity of the distance-independent amine acylation reaction.

DNA-templated Wittig olefination proceeds at a significantly lower rate when the aldehyde and phosphorane are separated by larger numbers of template bases, even though product yields are typically excellent after > 12 h of reaction.

[**] Funding was generously provided by the Searle Scholars Program (00-C-101), an Office of Naval Research Young Investigator Award (N00014-00-1-0596), an Arnold and Mabel Beckman Foundation Young Investigator Award, and an Alfred P. Sloan Foundation Research Fellowship (BR-4141). Z.J.G. is an NSF Graduate Research Fellow. C.T.C. is an NDSEG Graduate Research Fellow.
regardless of the intervening distance. After only 2 h of reaction (pH 7.5, 30°C) in the E or H architectures, yields of olefin products are three- to sixfold lower when the reactants are separated by ten or more bases (n = 10 or 20) than when reactants are separated by only one base (n = 1) (Figure 2, lanes 6 and 7, and Figure 3b). In contrast, the W architecture with four or five constant bases at the reactive end results in efficient and sequence-specific Wittig product formation after 2 h even when 10 or 20 bases separate the coding region and the reactive end of the template (Figure 2, lanes 8 and 9, and Figure 3b). These results suggest that the constant regions at the reactive ends of the template and reagent in the W architecture enable the aldehyde and phosphorane to react at an effective concentration comparable to that in the E architecture when n = 1 (Figure 1).

Of the many DNA-templated reactions that we have studied to date, the 1,3-dipolar cycloaddition and reductive amination reactions[24] demonstrate the most pronounced distance dependence. Both reactions proceed with low to modest efficiency (7–44% yield) in the E or H architectures when 10 or 20 bases separate the annealed reactive groups (Figure 2, lanes 10, 11, 14, and 15, and Figure 3c–d). This distance dependence limits the positions on a DNA template that can encode these or other distance-dependent reactions. In contrast, both 1,3-dipolar cycloaddition and reductive amination proceed efficiently (up to 97% yield) and sequence

---

**Figure 1.** Four architectures for DNA-templated synthesis. The Ω and T architectures are presented in this work. Templates are colored black and reagents are colored red. The number of single-stranded-template bases that separate the reactive groups of the template and the reagent is designated “n”. For the Ω architectures, the number of fixed bases at the reactive end of the template is shown following the Ω symbol.

**Figure 2.** Denaturing polyacrylamide-gel electrophoresis analysis of representative DNA-templated amine acylation, Wittig olefination, 1,3-dipolar cycloaddition, and reductive amination reactions with the end-of-helix (E) and Ω architectures. For lanes 5, 9, 13, and 17, reagents with coding region sequences containing three mismatches were used. R = template oligonucleotide, R′ = reagent oligonucleotide.
T able 1: Melting temperatures of selected template–reagent combinations using the Ω and end-of-helix (E) architectures.[a]

<table>
<thead>
<tr>
<th>Architecture</th>
<th>Buffer</th>
<th>Tm [°C]</th>
</tr>
</thead>
<tbody>
<tr>
<td>E (n = 10)</td>
<td>PBS</td>
<td>45</td>
</tr>
<tr>
<td>Ω (n = 10)</td>
<td>PBS</td>
<td>46</td>
</tr>
<tr>
<td>E (n = 10)</td>
<td>HSP</td>
<td>55</td>
</tr>
<tr>
<td>Ω (n = 10)</td>
<td>HSP</td>
<td>54</td>
</tr>
<tr>
<td>E (n = 20)</td>
<td>PBS</td>
<td>40</td>
</tr>
<tr>
<td>Ω (n = 20)</td>
<td>PBS</td>
<td>39</td>
</tr>
</tbody>
</table>

[a] For conditions, see Experimental Section. PBS = phosphate-buffered saline, HSP = high salt phosphate buffer.

DNA templates of arbitrary length are easy to synthesize and undesired cross-reactivity in a given solution is avoided by using concentrations that are too low to allow noncomplementary reactants to react intermolecularly.[33] These features of DNA-templated synthesis in theory enable more than one DNA-templated reaction to take place on a single template in one solution, eliminating the effort associated with additional DNA-templated steps and product purification. However, multiple DNA-templated reactions per step are difficult to achieve with the E, H, or Ω architectures because the reagent oligonucleotide that remains annealed to the template following the first reaction forms a relatively rigid double helix that can prevent a second reagent, annealed further away along the template, from encountering the reactive end of the template.[23] To overcome this difficulty, we moved the reactive group on the template from the end to the middle of the oligonucleotide, and attached the reactive group to the non-Watson–Crick face of a base. We hypothesized that this T architecture (Figure 1) would enable two DNA-templated reactions, one with a reagent annealed between the reactive group and the 5’ end of the template, and one with a reagent annealed between the reactive group and the 3’ end of the template, to take place sequence specifically in the same solution on a single template.

To test the viability of the T architecture in DNA-templated reactions, we evaluated the efficiency of the amine acylation, Wittig olefination, 1,3-dipolar cycloaddition, and reductive amination reactions in the T architecture with one to five bases between the annealed reactants. The T architecture directed these four reactions sequence specifically with efficiencies comparable to or greater than those of the E or H architectures (Figure 4, 69–100% yield when n = 1 or when n = −1. The observed degree of distance dependence in the T architecture for each of the four reactions is consistent with the findings described above (compare Figure 4 and Figure 3). Together these results demonstrate that the T architecture can mediate sequence-specific and efficient DNA-templated synthesis.

Once the ability of the T architecture to support efficient DNA-templated synthesis had been established, we tested its ability to direct two DNA-templated reactions on one template in one solution. We performed two different two-reaction sequences with the T architecture. In the first of these, a benzaldehyde-linked T template 1 was combined with a phosphine-linked reagent 2 and an α-iodoamide-linked reagent 3 in a single solution (pH 8.5, 1st NaCl, 25°C, 1 h)
Figure 5. Two DNA-templated steps on a single template in one solution, mediated by the T-architecture. a) SN2 reaction–Wittig olefination sequence. b) Analysis of the reactions in (a) by denaturing polyacrylamide-gel electrophoresis. The 30-base T-architecture template 1 with an aldehyde group was present in lanes 1, 2, and 5–10. A template lacking the aldehyde group but otherwise identical to 1 was present in lanes 3 and 4. DNA-linked phosphane reagent 2 was present in lanes 3–6, 9, and 10. DNA-linked α-iodoamide reagent 3 was present in lanes 3, 4, and 7–10. Lanes 1, 3, 5, 7, and 9 show reactions after 30 min; lanes 2, 4, 6, 8, and 10 show reactions after 1 h. c) Amine acylation–Huisgen cycloaddition sequence. For reaction conditions, see Experimental Section.
followed by the addition of 500 μM copper(II) sulfate and sodium ascorbate to induce the recently reported Sharpless-modified Huisgen 1,3-dipolar cycloaddition,\(^{35}\) provided 1,4-disubstituted triazoylalanine adduct 8 in 32% overall yield (Figure 5c). We conclude from these results that the T architecture enables two sequence-specific DNA-templated reactions to take place on one template in one solution. Importantly, the T-architecture templates described above were accepted as efficient templates for both a single cycle of primer extension as well as standard PCR amplification by Tag DNA polymerase, consistent with the known tolerance\(^{36}\) of several DNA polymerases for modifications to the non-Watson–Crick face of DNA templates. In addition to reducing the number of separate DNA-templated steps required for the synthesis of a target structure, this architecture may enable three-component reactions,\(^{37,38}\) commonly used to build structural complexity into synthetic libraries, to be performed in a DNA-templated format.

The Ω and T architectures developed in this work significantly expand the scope of DNA-templated synthesis. Our findings also demonstrate how reactivity can be rationally manipulated in a DNA-templated format by altering the design of template and reagent architectures, rather than by the traditional approach of changing global reaction conditions. By enabling distance-dependent DNA-templated reactions to be encoded by bases located far away from the reactive end of the template, the Ω architecture expands the types of reactions that can be encoded anywhere on a DNA template. The T architecture enables two DNA-templated reactions to take place on a single template in one step. Together these new architectures are proving to be useful in ongoing efforts to synthesize new compounds translated from corresponding DNA sequences.

**Experimental Section**

Unless otherwise specified, DNA oligonucleotides were synthesized and functionalized as previously described\(^{39}\) by using 2-[2-(4-

"Experimental Section"}

followed by the addition of 500 μM copper(II) sulfate and sodium ascorbate to induce the recently reported Sharpless-modified Huisgen 1,3-dipolar cycloaddition,\(^{35}\) provided 1,4-disubstituted triazoylalanine adduct 8 in 32% overall yield (Figure 5c). We conclude from these results that the T architecture enables two sequence-specific DNA-templated reactions to take place on one template in one solution. Importantly, the T-architecture templates described above were accepted as efficient templates for both a single cycle of primer extension as well as standard PCR amplification by Tag DNA polymerase, consistent with the known tolerance\(^{36}\) of several DNA polymerases for modifications to the non-Watson–Crick face of DNA templates. In addition to reducing the number of separate DNA-templated steps required for the synthesis of a target structure, this architecture may enable three-component reactions,\(^{37,38}\) commonly used to build structural complexity into synthetic libraries, to be performed in a DNA-templated format.

The Ω and T architectures developed in this work significantly expand the scope of DNA-templated synthesis. Our findings also demonstrate how reactivity can be rationally manipulated in a DNA-templated format by altering the design of template and reagent architectures, rather than by the traditional approach of changing global reaction conditions. By enabling distance-dependent DNA-templated reactions to be encoded by bases located far away from the reactive end of the template, the Ω architecture expands the types of reactions that can be encoded anywhere on a DNA template. The T architecture enables two DNA-templated reactions to take place on a single template in one step. Together these new architectures are proving to be useful in ongoing efforts to synthesize new compounds translated from corresponding DNA sequences.

**Experimental Section**

Unless otherwise specified, DNA oligonucleotides were synthesized and functionalized as previously described\(^{39}\) by using 2-[2-(4-

"Experimental Section"

followed by the addition of 500 μM copper(II) sulfate and sodium ascorbate to induce the recently reported Sharpless-modified Huisgen 1,3-dipolar cycloaddition,\(^{35}\) provided 1,4-disubstituted triazoylalanine adduct 8 in 32% overall yield (Figure 5c). We conclude from these results that the T architecture enables two sequence-specific DNA-templated reactions to take place on one template in one solution. Importantly, the T-architecture templates described above were accepted as efficient templates for both a single cycle of primer extension as well as standard PCR amplification by Tag DNA polymerase, consistent with the known tolerance\(^{36}\) of several DNA polymerases for modifications to the non-Watson–Crick face of DNA templates. In addition to reducing the number of separate DNA-templated steps required for the synthesis of a target structure, this architecture may enable three-component reactions,\(^{37,38}\) commonly used to build structural complexity into synthetic libraries, to be performed in a DNA-templated format.

The Ω and T architectures developed in this work significantly expand the scope of DNA-templated synthesis. Our findings also demonstrate how reactivity can be rationally manipulated in a DNA-templated format by altering the design of template and reagent architectures, rather than by the traditional approach of changing global reaction conditions. By enabling distance-dependent DNA-templated reactions to be encoded by bases located far away from the reactive end of the template, the Ω architecture expands the types of reactions that can be encoded anywhere on a DNA template. The T architecture enables two DNA-templated reactions to take place on a single template in one step. Together these new architectures are proving to be useful in ongoing efforts to synthesize new compounds translated from corresponding DNA sequences.
double-stranded during quantification, changes in staining intensity may result in higher apparent yields. Representative reaction products were characterized by MALDI (matrix-assisted laser desorption/ionization) mass spectrometry in addition to denaturing polyacrylamide gel electrophoresis.

Melting curves were obtained on a Hewlett-Packard 8453 UV/Vis spectrophotometer using a Hewlett-Packard 89090A Peltier temperature controller. Absorptions of template–reagent pairs (1.5 μM each) at 260 nm were measured every 1°C from 20°C to 80°C, holding for 1 min at each temperature, in either phosphate-buffered saline (NaCl (137 mM), KCl (2.7 mM), potassium phosphate (1.4 mM), sodium phosphate (10 mM), pH 7.4) or in high salt phosphate buffer (sodium phosphate (50 mM), NaCl (1 M), pH 7.2).

Received: December 19, 2002 [Z50806]

Keywords: bioorganic chemistry · combinatorial chemistry · DNA · template synthesis