ACCOUNTS
of chemical research

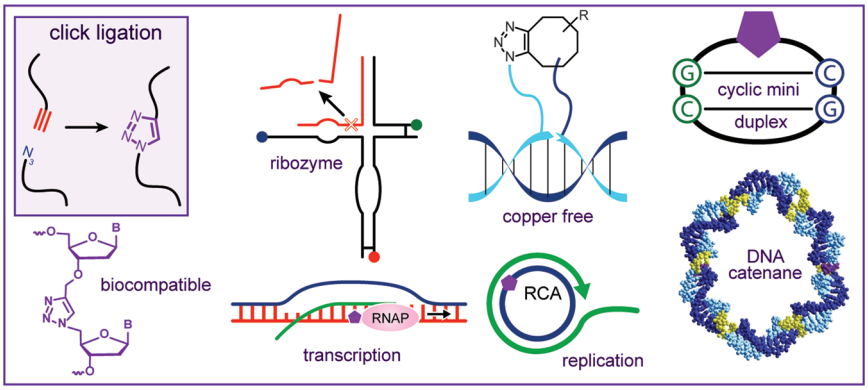
Click Nucleic Acid Ligation: Applications in Biology and Nanotechnology

AFAF H. EL-SAGHEER[†,‡] AND TOM BROWN[*,†]

[†]*School of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom, and* [‡]*Chemistry Branch, Department of Science and Mathematics, Faculty of Petroleum and Mining Engineering, Suez Canal University, Suez 43721, Egypt*

RECEIVED ON DECEMBER 12, 2011

CONSPECTUS



Biochemical strategies that use a combination of synthetic oligonucleotides, thermostable DNA polymerases, and DNA ligases can produce large DNA constructs up to 1 megabase in length. Although these ambitious targets are feasible biochemically, comparable technologies for the chemical synthesis of long DNA strands lag far behind. The best available chemical approach is the solid-phase phosphoramidite method, which can be used to assemble DNA strands up to 150 bases in length. Beyond this point, deficiencies in the chemistry make it impossible to produce pure DNA. A possible alternative approach to the chemical synthesis of large DNA strands is to join together carefully purified synthetic oligonucleotides by chemical methods. Click ligation by the copper-catalyzed azide–alkyne (CuAAC) reaction could facilitate this process. In this Account, we describe the synthesis, characterization, and applications of oligonucleotides prepared by click ligation.

The alkyne and azide oligonucleotide strands can be prepared by standard protocols, and the ligation reaction is compatible with a wide range of chemical modifications to DNA and RNA. We have employed click ligation to synthesize DNA constructs up to 300 bases in length and much longer sequences are feasible. When the resulting triazole linkage is placed in a PCR template, various DNA polymerases correctly copy the entire base sequence. We have also successfully demonstrated both in vitro transcription and rolling circle amplification through the modified linkage. This linkage has shown in vivo biocompatibility: an antibiotic resistance gene containing triazole linkages functions in *E. coli*. Using click ligation, we have synthesized hairpin ribozymes up to 100 nucleotides in length and a hammerhead ribozyme with the triazole linkage located at the substrate cleavage site. At the opposite end of the length scale, click-ligated, cyclic mini-DNA duplexes have been used as models to study base pairing. Cyclic duplexes have potential therapeutic applications. They have extremely high thermodynamic stability, have increased resistance to enzymatic degradation, and have been investigated as decoys for regulatory proteins. For potential nanotechnology applications, we have synthesized double stranded DNA catenanes by click ligation. Other researchers have studied covalently fixed multistranded DNA constructs including triplexes and quadruplexes.

Introduction

Current biochemical strategies to produce large DNA constructs utilize a combination of synthetic oligonucleotides, thermostable DNA polymerases, and DNA ligases. Entire bacterial and mitochondrial genomes and sections of eukaryotic chromosomes up to 1 megabase in length can be assembled using such approaches.[1−3] Such ambitious targets are feasible because the error rates of proofreading

Published on the Web 03/22/2012 www.pubs.acs.org/accounts
10.1021/ar200321n

DNA polymerases are lower than 10^{-6}.[4] In comparison, technologies for the chemical synthesis of long DNA strands lag far behind. The best available chemical approach is the solid-phase phosphoramidite method[5] which can be used to assemble DNA strands up to 150 bases in length. Beyond this point, deficiencies in the chemistry make it impossible to produce pure DNA. This is in part due to the physical properties of the resins on which the synthesis is conducted, but it is also a consequence of imperfect coupling and undesirable side-reactions that create mutagenic modifications to the nucleobases. For very long oligonucleotides, such modifications cannot be removed by purification, and error rates as high as 10^{-2} have been reported.[6] A possible alternative approach to the chemical synthesis of large DNA strands is to join together carefully purified synthetic oligonucleotides by chemical methods. However, it has proved challenging to achieve clean and efficient chemical ligation of oligonucleotides, although progress has been made using cyanogen bromide as a coupling agent.[7,8] An alternative strategy is to design a chemical linkage that mimics the natural phosphodiester bond and which can be formed in high yield in aqueous media from functional groups that are orthogonal to those present in DNA. This goal has been partly achieved by the reaction between oligonucleotides with 3′-phosphorothioate and 5′-tosylate or iodide.[9,10] Unfortunately, this chemistry does not satisfy the key requirements of a robust chemical method of gene synthesis, namely, the use of functional groups that are stable in aqueous media, the ability to synthesize bifunctional oligonucleotides with reactive groups at each end for use in multiple ligation reactions, a mechanism to initiate the ligation reaction when participating oligonucleotides have been hybridized to complementary splints (to arrange the DNA strands in the desired order by templated preassembly), and the creation of a very stable backbone linkage.

If an efficient chemical ligation method was available, it would increase the maximum attainable size of chemically synthesized DNA, and might even be applicable to the total chemical synthesis of genes. Such a method could gain popularity in biotechnology and nanotechnology where the availability of large quantities of very long chemically modified oligonucleotides would facilitate experiments which have hitherto been impossible to perform. However, its utility in biotechnology will only be fully realized if DNA strands containing the artificial linkages are able to act as templates in DNA replication. This is a tall order; the phosphodiester linkage is ubiquitous in the DNA of living systems, and there is no reason a priori to expect DNA polymerases to tolerate an artificial backbone. However, polymerases do have the capacity to read through damaged DNA, and engineered enzymes can perform this feat with surprising efficiency.[11] These developments provided an impetus for the design of a biocompatible DNA backbone linkage.

Design of a Biocompatible Click DNA Linkage

In order to produce a biocompatible chemical linkage, we turned to click chemistry, a concept that was developed to join together organic molecules under mild conditions in the presence of a diverse range of functional groups.[12] The best example of click chemistry is the Cu^{I} catalyzed version of Huisgen's [3 + 2] azide−alkyne cycloaddition reaction,[13] recently discovered independently by Sharpless and Meldal (the CuAAC reaction).[14,15] This highly selective and efficient reaction was chosen for nucleic acid ligation for several reasons; alkynes and azides can be attached to nucleic acids by simple methods, the CuAAC reaction proceeds efficiently in aqueous media, the reaction can be switched on at will by addition of Cu^{I}, and the resultant triazole unit is extremely stable. Studies in the fields of reaction discovery,[16] templated synthesis,[17] and nanotechnology[18] had shown that this chemistry is applicable to DNA, and the new challenge was to design a triazole linkage with a structure that is biocompatible. Initial investigations involved the synthesis of DNA containing the unnatural triazole linkage A (Figure 1d). This is formed by the reaction between oligonucleotides functionalized with 3′-azido-dT and 5′-propargylamido-dT in the presence of a complementary template oligonucleotide (splint) which holds the reactants in close proximity. The click reaction also occurs in the absence of the splint, but it is much slower and requires higher oligonucleotide concentrations. Splints confer another important advantage; they allow the assembly of multiple alkyne/azide DNA fragments in the correct order prior to Cu^{I}-catalysis. A click-ligated DNA strand containing such a triazole linkage was used as a template in PCR (Figure 1a). Amplification was successful but DNA sequencing of the progeny strands revealed only a single thymidine at the ligation site instead of the two thymidines of the original template.[19] Although this was disappointing the overall result was encouraging. Importantly, it provided the first example of efficient chemical DNA strand ligation combined with PCR amplification through an unnatural DNA backbone linkage. It also gave hope that a truly biocompatible DNA linkage could be designed if the limitations of the first

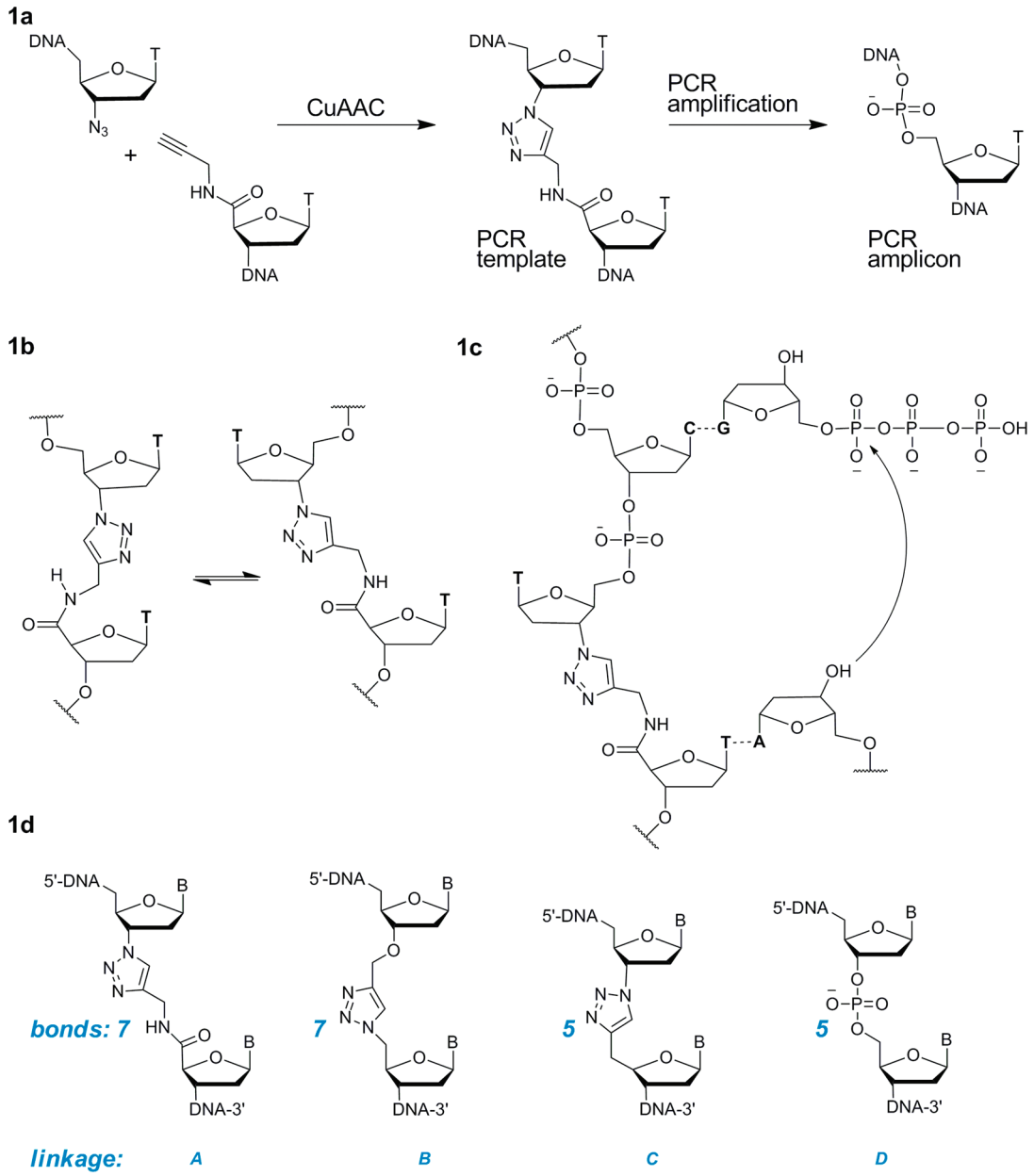


FIGURE 1. (a) Synthesis and PCR amplification from the first generation triazole DNA backbone linkage. (b) cis−trans isomerization of amide bond. (c) Replication bypasses one of the thymine bases around the triazole linkage. (d) First generation triazole linkage A, second generation triazole linkage B, triazole linkage C in ^{TL}-DNA, and canonical DNA linkage D. The number of chemical bonds between the C3′ and C4′ atoms of adjacent sugar rings is written in blue.

generation linker could be rationalized. A possible reason for the failure of DNA polymerase to correctly read the base sequence around triazole linkage A is the presence of the rigid amide bond. This will prefer to adopt the planar *trans*-configuration (Figure 1b), and perhaps cause the thymidine adjacent to the triazole to turn away from the incoming deoxynucleotide triphosphate (dNTP) during replication (Figure 1c). The lack of a 3′-oxygen atom and the absence of a 5′-methylene group might also have caused polymerases to misread the triazole DNA template.

With these considerations in mind, a more flexible second generation triazole linkage was synthesized (linkage B, Figure 1d).[20] Importantly the DNA strands required for multiple click ligation (3′-alkyne oligonucleotide, 5′-azide oligonucleotide, and bifunctional 3′-alkyne-5′-azide oligonucleotide) can be made by the solid-phase phosphoramidite method. Three 81-mer DNA templates were made to test the biocompatibility of the new linkage, each containing a single triazole. In all three cases, the PCR reaction produced amplicons in which the bases around the triazole were read

FIGURE 2. Click DNA cyclization and rolling circle amplification (RCA) through triazole linkage B.

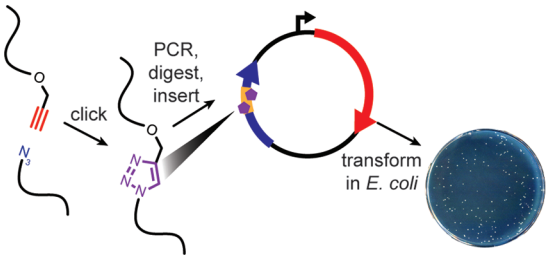


FIGURE 3. Insertion of triazole linkage B into each strand of the BLA gene (blue) of plasmid DNA followed by transformation and growth of *E. coli*. The insert is yellow and the triazole linkages are purple.

through accurately. Successful linear enzymatic copying of one of the 81-mers was then carried out to prove that PCR amplification of the chemically modified DNA was not a rare event.[20] To demonstrate the utility of click chemistry for the assembly of large DNA molecules with more than one click ligation site, three 100-mer oligonucleotides were ligated to give a 300-mer PCR template with two triazole linkages. This template was subjected to PCR and sequenced, and the amplicon was shown to be a faithful copy of the original template. The synthesis of cyclic DNA was also achieved; a 100-mer oligonucleotide with 5′-azide and 3′-alkyne functionalities was cyclized by addition of Cu^{I} to produce a construct containing a single triazole (linkage B, Figure 1). Rolling circle amplification (RCA)[21,22] was then carried out to produce a linear construct containing multiple copies of the original cyclic template (Figure 2).[20]

In Vivo Biocompatibility of Triazole Linkage B

The biocompatibility of the modified DNA was investigated in *E. coli* using a plasmid containing a triazole linkage in each strand of its antibiotic marker gene (Figure 3).[20] The number of colonies obtained from *E. coli* transfected with the modified plasmid was similar to the native plasmid, and DNA sequencing confirmed that the bases around the triazole linkage were copied correctly. The experiments were repeated in a strain of *E. coli* deficient in UvrB,[23] an essential component of DNA repair.[24] Transformation with the triazole plasmid gave 93% of the number of colonies from the native plasmid, and the DNA sequence around the triazole was replicated accurately. This supports the hypothesis that DNA repair does not make a significant contribution

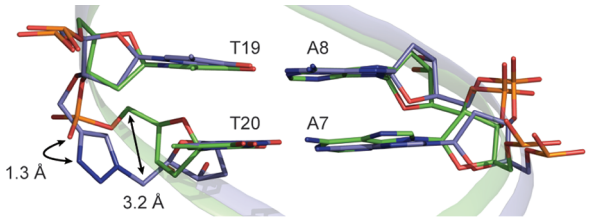


FIGURE 4. Structure of triazole linkage B (Figure 1) in DNA determined by NMR spectroscopy. The backbone of the triazole DNA strand is distorted to accommodate its extra length and to allow efficient base stacking. This requires the 5′-carbon to point down, shifted by 3.2 Å relative to its position in the native duplex. The N3 nitrogen of the triazole duplex and phosphate oxygen of the control duplex are only 1.3 Å apart.

to the biocompatibility of the artificial DNA backbone, and that replication occurs through the triazole linkage in vivo. A rationale for the surprising discovery that polymerases can copy a DNA template containing a completely unnatural backbone linkage comes from the X-ray structure of the complex between Taq polymerase and DNA.[25] In this structure, several nucleotides in the DNA template interact with the enzyme. In the click-ligated DNA template, only one of the internucleotide linkages can be a triazole because the template is synthesized by click ligation of long oligonucleotides, and the triazole linkages only occur at the ligation points. Hence, very few hydrogen bonding interactions between the DNA backbone and the enzyme can be destroyed by the triazole substitution.[20] This partly explains why triazole linkage B is read through correctly during PCR amplification. Further support comes from an NMR study of a DNA duplex containing a single triazole linkage.[26] The triazole introduces local structural changes but leaves the duplex and Watson–Crick base pairs intact. Moreover, the N3 atom of the triazole ring can mimic the hydrogen bond acceptor properties of the phosphodiester group in a complex with DNA polymerase. It lies close to the position of a phosphate branching oxygen (Figure 4), and it has a large dipole moment and proven hydrogen bond acceptor capacity.[27]

Biophysical Studies on Triazole DNA

The effect of the triazole linkage on the kinetics and thermodynamics of duplex formation has been explored by thermal denaturation studies on a 13-mer DNA duplex containing a single central triazole linkage.[26] Both triazole and normal control duplexes gave sharp cooperative melting transitions (at 55 and 63 °C, respectively) with a similar degree of hyperchromicity, indicating that the triazole duplex is fully base paired. The effect of a mismatched base pair at the triazole site is particularly interesting. It lowers the

duplex melting temperature by 11 °C on average, similar to a mismatch in the native duplex. The average decrease in hyperchromicity caused by a mismatch was 18.8% for the triazole duplex but only 7.5% for the native duplex, indicating that the combination of a triazole linkage and mismatched base pair causes a much greater loss of base stacking than is observed in native mismatched DNA. A detailed NMR study[26] revealed that the opening rates of the TA base pairs on either side of the triazole are greatly increased in the modified duplex, and that the destabilization is spread over the four base pairs immediately surrounding the triazole linkage. The Gibbs energy required to dissociate each of the two base pairs on either side of the modified backbone is reduced by 11 kJ/mol compared to the native duplex, and the base pair on the 5′-side of the triazole modification is perturbed rather more than the base pair on the 3′-side (7.2 and 3.4 kJ/mol, respectively). Despite this destabilization, the duplex accommodates the unnatural linkage in a normal B-DNA helical structure with little distortion of the major and minor grooves. The above biophysical data allow us to suggest a model of replication through the triazole linkage. When the correct dNTP is selected by the template–polymerase complex, a normal Watson–Crick base pair forms as the incoming nucleotide is incorporated. In contrast, on the rare occasion that a mismatched dNTP is inserted, severe duplex instability and local melting triggers removal of the rogue nucleotide by the polymerase enzyme. This "double hit" of triazole linkage and mismatched base pair is an easy target for the proofreading functionality of DNA polymerase.

There are five bonds between the C3′ and C4′ atoms of adjacent sugar rings in a normal DNA backbone. In contrast, triazole linkage B is seven bonds in length, possibly explaining why it destabilizes the DNA duplex. Reducing the length of the triazole linkage has been shown to produce more stable duplexes; triazole linkage C (Figure 1) with five bonds has been used to construct ^{TL}DNA in which every phosphate linkage is replaced by a triazole. It is important to state that although oligo dT ^{TL}DNA forms a stable duplex with oligo dA,[28] the properties of *a single* triazole linkage C in a normal DNA duplex have not yet been explored; its effects on duplex stability and its compatibility with DNA and RNA polymerases should be investigated.

Transcription through the Triazole Linkage

The biocompatibility of triazole-linked DNA in *E. coli* cannot be directly ascribed to efficient transcription through the modified backbone.[20] This is because replication might have occurred prior to transcription, in which case the cellular RNA polymerase enzymes would not have encountered the artificial linkage. Therefore, to evaluate the suitability of triazole DNA for in vitro synthesis of RNA, it was necessary to carry out experiments in a controlled environment. This was done using T7 RNA polymerase (T7-RNAP), an enzyme which is used in biotechnology to direct the synthesis of small RNAs and proteins. The RNA transcript selected for this study contained the 54-mer DicF sequence[29] which inhibits the growth of *E. coli*. Two triazole-containing DNA template strands were studied, one with the triazole linkage inside the coding sequence and a second with the modification within the essential T7 RNA promoter region. When the triazole linkage was placed in the coding region (Figure 5), the transcription reaction produced ~80% of the quantity of RNA obtained from the native control DNA template, and mass spectrometry showed that the two transcripts were identical.[30] This demonstrates that T7-RNAP can accurately transcribe through triazole-modified DNA to synthesize fully complementary RNA in good yield. This is the first example of transcription through a purely synthetic analogue of a DNA backbone, and it suggests that DNA constructs made by solid-phase synthesis and click ligation can be used directly for protein expression. When the transcription reactions were repeated with the triazole in the promoter region of the DNA template there was no RNA product. This could be due to a lack of binding between the T7-RNAP and the triazole DNA. Further studies are required to confirm this.

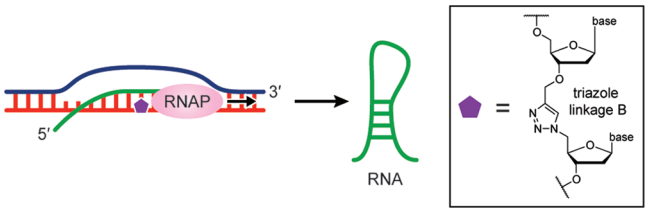


FIGURE 5. Transcription to produce RNA from a DNA template containing triazole linkage B.

Click Ligation for the Synthesis of RNA

In vitro transcription is commonly used to make long RNA strands, but this enzymatic process does not permit the site-specific incorporation of chemical modifications at sugars, bases, or phosphates. In contrast, the chemical synthesis of RNA by solid-phase methods offers advantages over biochemical approaches. It is compatible with the insertion of fluorescent tags, isotopic NMR labels, and modifications

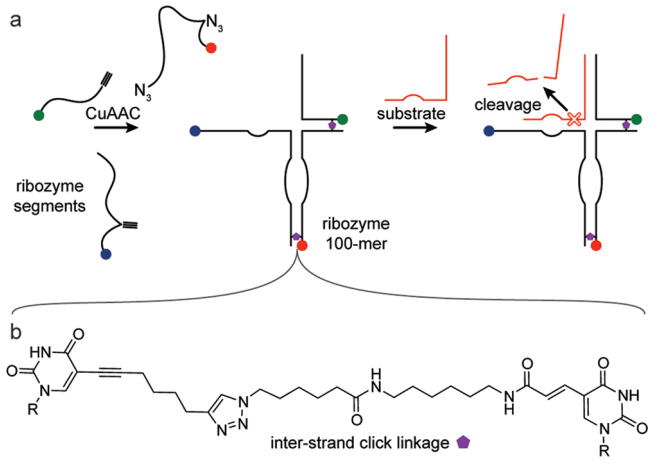


FIGURE 6. (a) Assembly of the hairpin ribozyme from three prepurified fluorescently labeled RNA strands followed by cleavage of the RNA substrate. (b) Chemical structure of the click linkage.

designed to improve the biological activity and stability of RNA. In addition, the scale of chemical synthesis is essentially unlimited. An efficient methodology for chemical ligation would be of great value as it would bring synthetic RNA into the size range required for many biological studies. Enzymatic ligation is an established method of synthesizing large RNA constructs from smaller fragments,[31] but it cannot be used in certain situations, for example, to link together RNA strands across the bases or between modified sugars. Click ligation can overcome these limitations, and it has recently been used to synthesize the chemically modified hairpin and hammerhead ribozymes described below.[32]

Synthesis and Properties of the Click Hairpin Ribozyme and Analogues

The hairpin ribozyme belongs to a family of catalytic RNAs that cleave their RNA substrates to generate 2′,3′-cyclic phosphate and 5′-hydroxyl termini.[33] It was chosen for click ligation studies because its direct chemical synthesis is very difficult, its biochemical properties are well understood,[34] and its functionality can be extended by the incorporation of modifications such as fluorescent tags.[35] Hairpin ribozyme analogues were synthesized by cross-linking the side-chains of modified uracil bases in opposite strands of three separate RNA segments (*trans*-ligation). The individual oligonucleotides were prepared by standard solid-phase synthesis, and they were functionalized with an alkyne or azide group at the 5-position of selected uracil bases (Figure 6). To construct the ribozymes, the individual segments were linked in two simultaneous self-templated CuAAC reactions across the major groove[36] of the ribozyme stems. Cross-strand linking

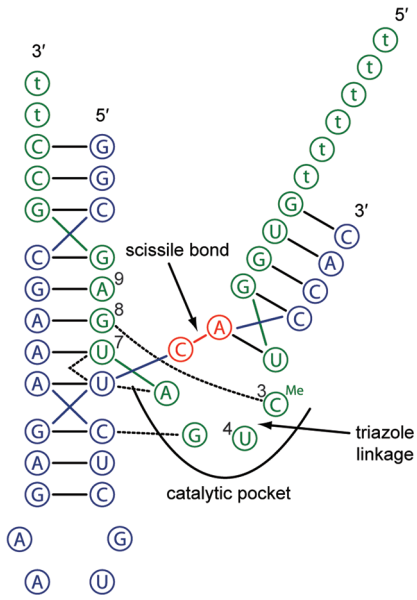


FIGURE 7. Hammerhead ribozyme in green and RNA substrate in blue with cleavage site in red.

between nucleobases in RNA is unique to chemical ligation and cannot be achieved by enzymatic methods. The catalytic activity of the click hairpin analogues was shown to be similar to the native ribozyme. The click ribozymes could be isolated from the gels of the cleavage reactions and reused with no significant loss of activity,[32] demonstrating their high stability.

The Hammerhead Ribozyme

In an alternative approach to click RNA ligation, the hammerhead ribozyme,[37] a small naturally occurring catalytic RNA, was made from two RNA strands by splint-mediated intrastrand click ligation (*cis*-ligation) of a 3′-alkyne and a 5′-azide oligoribonucleotide (Figure 7). This strategy produced a ribozyme containing triazole linkage B (Figure 1). The click hammerhead ribozyme cleaved its substrate with similar efficiency and specificity to the native ribozyme, confirming the biocompatibility of the triazole linkage, even though the unnatural triazole linkage is located at the active site between residues C3 and U4 of the ribozyme (Figure 7).[37−39]

In a separate study, a combination of chemical and enzymatic synthesis was used to incorporate an azide group into RNA. Templated and nontemplated CuAAC ligation reactions were carried out between azide and alkyne-functionalized RNA.[40] This provided active ribozymes containing a single triazole, confirming that the triazole backbone is not detrimental to RNA function. This is despite the fact that this artificial linkage destabilizes RNA duplexes, as shown in a UV melting and CD study on several short self-complementary

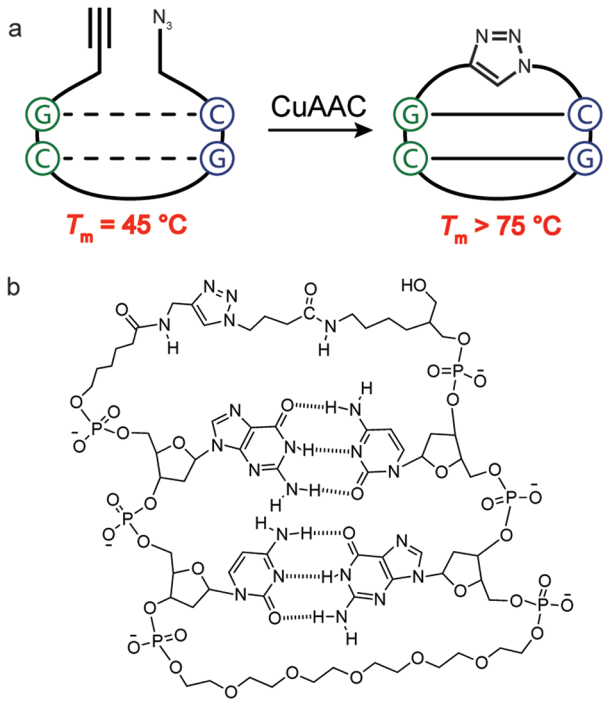


FIGURE 8. (a) Schematic representation of a stable cyclic GC/GC dinucleotide mini-duplex. (b) chemical structure.

RNA duplexes containing triazole linkages. NMR studies on an 8-mer duplex in which the triazoles were present in both strands and directly opposite to each other indicated that all base pairs were formed despite instability around the modified linkages.[41]

Cyclic Mini-DNA Duplexes

The CuAAC reaction has been used to construct very stable cyclic mini-DNA duplexes from hairpin oligonucleotides functionalized with a 5′-alkyne, a 3′-azide, and a hexaethylene glycol loop.[42] They have similar properties to normal double-stranded DNA, a notable exception being their very high thermodynamic stability, which makes them excellent model systems for detailed studies on base pairing. ^{1}H NMR studies on a cyclic GC/GC dinucleotide duplex (Figure 8) showed stable hydrogen bonding between the GC base pairs, and UV melting indicated that the caging effect holds the structure together, making the duplexes highly resistant to thermal denaturation. In contrast to the GC/GC dinucleotide, a cyclic construct with a single G in one strand and C in the other (a single base pair), showed no evidence of base pairing, presumably because there is no possibility of base stacking interactions. This illustrates the crucial influence of base stacking on DNA duplex stability. Cyclic duplexes are resistant to enzymatic degradation in biological media, and have potential uses as decoys for

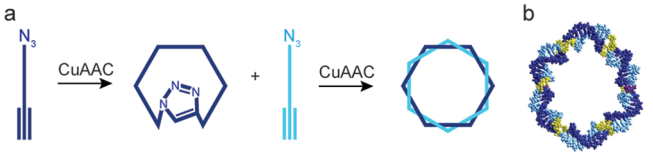


FIGURE 9. (a) Formation of double-stranded DNA catenane from single-stranded cyclic template oligonucleotide and linear complementary strand. (b) Double stranded DNA catenane. Tandem TT hinges are shown in yellow and DNA strands in light and dark blue.

transcription factors.[43] Incubation of a fluorescein-labeled cyclic oligonucleotide and its hairpin counterpart in fetal calf serum showed that the cyclic construct has significantly greater stability to enzymatic degradation. Cell uptake studies in the presence of a transfection agent showed that the cyclic oligonucleotide localized to the nucleus after only 1.5 h and persisted for 24 h. Other cyclic dumbbell oligonucleotides synthesized by click chemistry have also been shown to have high duplex melting temperatures and strong resistance to degradation by snake venom phosphodiesterase. These oligonucleotides were designed to bind to the NF-κB p50 homodimer as in vivo decoys.[44] In a study on DNA drug binding, the mode of action of a novel threading intercalator was elucidated using a click-ligated cyclic duplex.[45]

Synthesis of a DNA Catenane

Click chemistry has been applied to the assembly of a double-stranded DNA catenane consisting of six turns of B-DNA.[18] To construct the catenane, a stepwise strategy was adopted as shown in Figure 9. Two complementary oligonucleotides were used, each labeled with a 3′-azide and a 5′-alkyne. One of them was cyclized in a nontemplated CuAAC reaction and used as a template for the second click cyclization. Tandem TT mismatches were placed at intervals of 10 base pairs in the cyclic duplex to produce points of flexibility, without which the duplex would be too rigid to bend into a circle.

Nanotechnology and Higher Order Structures

Controlled self-assembly makes DNA an ideal material for building nanostructures. However, an obstacle to its use in constructing intermediates in hierarchic assembly is the thermodynamic instability of the assembled double stranded nanoconstructs, which makes purification difficult. A click-fixation technology has been developed by which robust chemically cross-linked DNA nanoconstructs can be made and purified by denaturing gel-electrophoresis.[46] Such chemically fused building blocks could be used to build larger nanonetworks. Triplex DNA binders have been used in

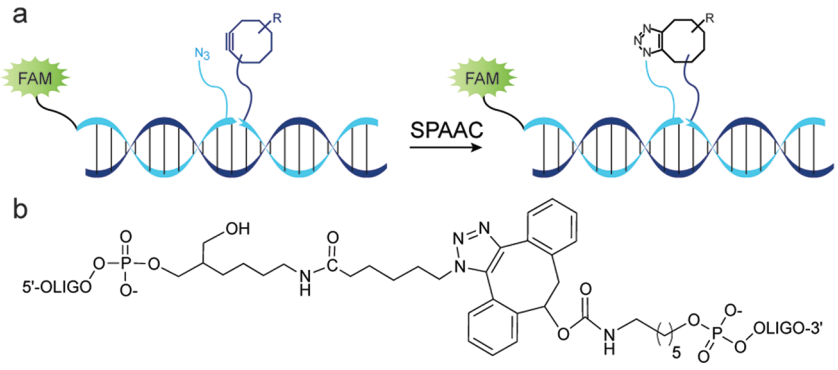


FIGURE 10. (a) SPAAC click DNA ligation between azide and cyclooctyne-labeled oligonucleotides (FAM = fluorescein). (b) Chemical structure of DIBO triazole at the ligation point.

nanotechnology applications involving click chemistry. They can act as external stimuli for controlling chemical reactivity in higher-order DNA-directed reactions.[47] Efficient double click reactions can be controlled simply by adding a strong triplex DNA binder such as naphthylquinoline. This is a high-yielding method for the formation of three-way branched nonsymmetrical DNA sequences that are useful in the construction of DNA nanostructures. Other higher order nucleic acid structures have been isolated in a stable form by click chemistry. Conformational isomers of G-quadruplexes have been trapped by performing click reactions between their termini.[48] This approach led to the discovery of a DNA–RNA hybrid-type G-quadruplex derived from the sequence of the human telomere. The click fixation allows detection and identification of the G-quadruplex reaction products in a complex solution, whereas traditional methods such as NMR spectroscopy and X-ray crystallography were not suitable due to exchange between different quadruplex conformations.

A simple and efficient template-free technique for interstrand cross-linking of DNA has been developed using a "bis-click" CuAAC reaction between alkyne-functionalized oligonucleotides and bifunctional azides.[49,50] The methodology is independent of base composition and is applicable to single stranded DNA, duplexes, and multistranded structures. Four-stranded DNA structures consisting of two cross-linked duplexes were obtained after hybridization of the bis-click constructs. This methodology has applications in nanobiotechnology and material science. The CuAAC reaction has also been used to ligate artificial mimics of DNA. A highly efficient chemical ligation method has been developed for quantitative conjugation of PNA with DNA or PNA with PNA.[51] In this work, PNA oligomers with an alkyne at the C-terminus and an azide at the N-terminus were used in DNA-templated ligation reactions. This enabled the preparation of extended PNA sequences which are difficult to prepare by standard solid-phase methods. Templated PNA click ligation is sequence specific and capable of single nucleotide discrimination.

Copper-Free Click DNA Ligation

The reaction between azides and unactivated terminal alkynes[13] is very slow in the absence of Cu^{I} catalysis, even when templated by DNA.[52] However, the copper-free AAC reaction can be accelerated by using an alkyne which is distorted by incorporation into an 8-membered ring structure. The ring-strain promoted alkyne–azide [3 + 2] cycloaddition reaction (SPAAC reaction) has been pioneered by Bertozzi and Agard,[53] principally for in vivo applications involving carbohydrates. It has recently been applied to DNA strand ligation[54] using dibenzocyclooctyne (DIBO) as the strained alkyne (Figure 10). Unlike electron deficient terminal alkynes, DIBO is stable in aqueous buffers and survives the conditions of oligonucleotide deprotection. The DNA-templated SPAAC ligation reaction is very fast (~1 min at ambient temperature), and the presence of a single mismatched base pair is sufficient to stop the reaction, suggesting applications in genetic analysis.[54] It should be possible to use this chemistry for multiple simultaneous templated DNA ligation reactions if participating oligonucleotides are labeled with either two alkynes or two azides.

The SPAAC reaction on DNA has advantages and disadvantages when compared to the CuAAC reaction. As it does not require catalysis with toxic metals, it has potential uses in vivo. However, undesired ligation reactions can occur before the participating oligonucleotides have had time to assemble against the template. At low reactant concentrations, the reaction does not proceed at an appreciable rate in the absence of DNA templating, so this may not be a major problem. The SPAAC reaction is orthogonal to the Diels–Alder reaction

which has also been used for DNA strand ligation.[55] Unlike the CuAAC reaction, neither the SPAAC or the Diels–Alder reaction is likely to give rise to a DNA linkage that can be read through accurately by polymerase enzymes.

Conclusions and Outlook

Click chemistry on nucleic acids is clean, efficient, tolerant of most other functional groups, and compatible with large scale synthesis. Very long DNA and RNA strands have already been synthesized using this methodology, and the limits have not yet been fully explored. Careful design has given rise to a biocompatible triazole linkage that is treated as normal by DNA and RNA polymerases, and a gene containing triazole linkages has been shown to function in *E. coli*. These surprising results could lead to important future applications in biology and nanotechnology. For example, one could contemplate the use of click DNA ligation for the assembly of synthetic genes on a large scale; for site-specific incorporation of unusual base or sugar modifications (e.g., mutation hot spots, epigenetic modifications) to study DNA replication, repair, and gene regulation, or to incorporate chemical modifications such as fluorescent dyes for in vivo imaging. In the RNA world, segmental labeling followed by the assembly of large chemically modified RNA constructs for functional studies is a distinct possibility. In addition, a number of therapeutic applications could be explored, and copper-free click DNA ligation could be carried out in a cellular context. The future of click nucleic acid ligation holds great promise, and thoughtful and imaginative application of this technology could lead to major advances.

The authors' work that was mentioned in this Account was primarily supported by a grant from the European Community 7th Framework Programme [FP7/2007-2013], Grant Agreement [HEALTH-F4-2008-201418] entitled READNA.

BIOGRAPHICAL INFORMATION

Afaf H. El-Sagheer studied Chemistry and was appointed as a demonstrator at Suez Canal University (Egypt). She completed her Ph.D. in Chemistry at Southampton University in 2000 then returned to Egypt to take up a lectureship at Suez Canal University. She is now Associate Professor, and is currently on a sabbatical leave working with Tom Brown on applications of oligonucleotide chemistry to biology and nanotechnology.

Tom Brown was appointed to a lectureship at Edinburgh University in 1985 where he became Professor of Nucleic Acid Chemistry. He moved to Southampton University in 1995 where he is currently Chair of Chemical Biology. His research interests center on Nucleic acids and the application of oligonucleotide chemistry to biology and medicine. He has received numerous awards including the recent Royal Society of Chemistry prizes for Nucleic Acid Chemistry and Interdisciplinary Research. He is co-founder of three Biotech companies.

FOOTNOTES

*To whom correspondence should be addressed. E-mail: tb2@soton.ac.uk.
The authors declare no competing financial interest.

REFERENCES

1 Gibson, D. G.; Smith, H. O.; Hutchison, C. A.; Venter, J. C.; Merryman, C. Chemical synthesis of the mouse mitochondrial genome. *Nat. Methods* **2010**, *7*, 901–U905.
2 Gibson, D. G.; Glass, J. I.; Lartigue, C.; Noskov, V. N.; Chuang, R. Y.; Algire, M. A.; Benders, G. A.; Montague, M. G.; Ma, L.; Moodie, M. M.; Merryman, C.; Vashee, S.; Krishnakumar, R.; Assad-Garcia, N.; Andrews-Pfannkoch, C.; Denisova, E. A.; Young, L.; Qi, Z. Q.; Segall-Shapiro, T. H.; Calvey, C. H.; Parmar, P. P.; Hutchison, C. A.; Smith, H. O.; Venter, J. C. Creation of a bacterial cell Controlled by a chemically synthesized genome. *Science* **2010**, *329*, 52–56.
3 Dymond, J. S.; Richardson, S. M.; Coombes, C. E.; Babatz, T.; Muller, H.; Annaluru, N.; Blake, W. J.; Schwerzmann, J. W.; Dai, J. B.; Lindstrom, D. L.; Boeke, A. C.; Gottschling, D. E.; Chandrasegaran, S.; Bader, J. S.; Boeke, J. D. Synthetic chromosome arms function in yeast and generate phenotypic diversity by design. *Nature* **2011**, *477*, 471–U124.
4 Kunkel, T. A. DNA replication fidelity. *J. Biol. Chem.* **2004**, *279*, 16895–16898.
5 Caruthers, M. H. Chemical synthesis of DNA and DNA analogs. *Acc. Chem. Res.* **1991**, *24*, 278–284.
6 Hecker, K. H.; Rill, R. L. Error analysis of chemically synthesized polynucleotides. *Biotechniques* **1998**, *24*, 256–260.
7 Sokolova, N. I.; Ashirbekova, D. T.; Dolinnaya, N. G.; Shabarova, Z. A. Chemical reactions in nucleic acid duplexes 0.4. Cyanogen bromide as an efficient reagent in condensation of oligodeoxyribonucleotides. *Bioorg. Khim.* **1987**, *13*, 1286–1288.
8 Luebke, K. J.; Dervan, P. B. Nonenzymatic sequence-specific ligation of double-helical DNA. *J. Am. Chem. Soc.* **1991**, *113*, 7447–7448.
9 Herrlein, M. K.; Nelson, J. S.; Letsinger, R. L. A covalent lock for self-assembled oligonucleotide conjugates. *J. Am. Chem. Soc.* **1995**, *117*, 10151–10152.
10 Xu, Y. Z.; Kool, E. T. Chemical and enzymatic properties of bridging 5′-S-phosphorothioester linkages in DNA. *Nucleic Acids Res.* **1998**, *26*, 3159–3164.
11 d'Abbadie, M.; Hofreiter, M.; Vaisman, A.; Loakes, D.; Gasparutto, D.; Cadet, J.; Woodgate, R.; Paabo, S.; Holliger, P. Molecular breeding of polymerases for amplification of ancient DNA. *Nat. Biotechnol.* **2007**, *25*, 939–943.
12 Kolb, H. C.; Finn, M. G.; Sharpless, K. B. Click chemistry: Diverse chemical function from a few good reactions. *Angew. Chem., Int. Ed.* **2001**, *40*, 2004–2021.
13 Huisgen, R. Kinetics and Mechanism of 1,3-Dipolar Cycloadditions. *Angew. Chem., Int. Ed.* **1963**, *2*, 633–645.
14 Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. A stepwise Huisgen cycloaddition process: Copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angew. Chem., Int. Ed.* **2002**, *41*, 2596–2599.
15 Tornoe, C. W.; Christensen, C.; Meldal, M. Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem.* **2002**, *67*, 3057–3064.
16 Kanan, M. W.; Rozenman, M. M.; Sakurai, K.; Snyder, T. M.; Liu, D. R. Reaction discovery enabled by DNA-templated synthesis and in vitro selection. *Nature* **2004**, *431*, 545–549.
17 Gartner, Z. J.; Grubina, R.; Calderone, C. T.; Liu, D. R. Two enabling architectures for DNA-templated organic synthesis. *Angew. Chem., Int. Ed.* **2003**, *42*, 1370–1375.
18 Kumar, R.; El-Sagheer, A. H.; Tumpane, J.; Lincoln, P.; Wilhelmsson, L. M.; Brown, T. Template-directed oligonucleotide strand ligation, covalent intramolecular DNA circularization and catenation using click chemistry. *J. Am. Chem. Soc.* **2007**, *129*, 6859–6864.
19 El-Sagheer, A. H.; Brown, T. Synthesis and Polymerase Chain Reaction Amplification of DNA Strands Containing an Unnatural Triazole Linkage. *J. Am. Chem. Soc.* **2009**, *131*, 3958–3964.
20 El-Sagheer, A. H.; Sanzone, A. P.; Gao, R.; Tavassoli, A.; Brown, T. Biocompatible artificial DNA linker that is read through by DNA polymerases and is functional in E. coli. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 11338–11343.
21 Dean, F. B.; Nelson, J. R.; Giesler, T. L.; Lasken, R. S. Rapid amplification of plasmid and phage DNA using phi29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res.* **2001**, *11*, 1095–1099.
22 Tsaftaris, A.; Pasentzis, K.; Argiriou, A. Rolling circle amplification of genomic templates for inverse PCR (RCA-GIP): a method for 5′- and 3′-genome walking without anchoring. *Biotechnol. Lett.* **2010**, *32*, 157–161.

23 Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K. A.; Tomita, M.; Wanner, B. L.; Mori, H. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2006**, *2*, 1–11.

24 Truglio, J. J.; Croteau, D. L.; Van Houten, B.; Kisker, C. Prokaryotic nucleotide excision repair: The UvrABC system. *Chem. Rev.* **2006**, *106*, 233–252.

25 Li, Y.; Korolev, S.; Waksman, G. Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of Thermus aquaticus DNA polymerase I: structural basis for nucleotide incorporation. *EMBO J.* **1998**, *17*, 7514–7525.

26 Dallmann, A.; El-Sagheer, A. H.; Dehmel, L.; Mügge, C.; Griesinger, C.; Ernsting, N. P.; Brown, T. Structure and dynamics of triazole-linked DNA - biocompatibility explained. *Chem.—Eur. J.* **2011**, *17*, 14714–14717.

27 Angell, Y. L.; Burgess, K. Peptidomimetics via copper-catalyzed azide-alkyne cycloadditions. *Chem. Soc. Rev.* **2007**, *36*, 1674–1689.

28 Isobe, H.; Fujino, T.; Yamazaki, N.; Guillot-Nieckowski, M.; Nakamura, E. Triazole-linked analogue of deoxyribonucleic acid (^{TL}DNA): Design, synthesis, and double-strand formation with natural DNA. *Org. Lett.* **2008**, *10*, 3729–3732.

29 Faubladier, M.; Cam, K.; Bouche, J. P. Escherichia-Coli cell-division inhibitor DicF-RNA of the DicB operon-evidence for its generation in vivo by transcription termination and by RNase-III and RNase E-dependent processing. *J. Mol. Biol.* **1990**, *212*, 461–471.

30 El-Sagheer, A. H.; Brown, T. Efficient RNA synthesis by in vitro transcription of a triazole-modified DNA template. *Chem. Commun.* **2011**, *47*, 12057–12058.

31 Lang, K.; Micura, R. The preparation of site-specifically modified riboswitch domains as an example for enzymatic ligation of chemically synthesized RNA fragments. *Nat. Protoc.* **2008**, *3*, 1457–1466.

32 El-Sagheer, A. H.; Brown, T. New strategy for the synthesis of chemically modified RNA constructs exemplified by hairpin and hammerhead ribozymes. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 15329–15334.

33 Buzayan, J. M.; Gerlach, W. L.; Bruening, G. Nonenzymatic cleavage and ligation of RNAs complementary to a plant virus satellite RNA. *Nature* **1986**, *323*, 349–353.

34 Fedor, M. J. Structure and function of the hairpin ribozyme. *J. Mol. Biol.* **2000**, *297*, 269–291.

35 Murchie, A. I. H.; Thomson, J. B.; Walter, F.; Lilley, D. M. J. Folding of the hairpin ribozyme in its natural conformation achieves close physical proximity of the loops. *Mol. Cell* **1998**, *1*, 873–881.

36 Kocalka, P.; El-Sagheer, A. H.; Brown, T. Rapid and efficient DNA strand cross-linking by click chemistry. *ChemBioChem* **2008**, *9*, 1280–1285.

37 Scott, W. G.; Murray, J. B.; Arnold, J. R. P.; Stoddard, B. L.; Klug, A. Capturing the structure of a catalytic RNA intermediate: The hammerhead ribozyme. *Science* **1996**, *274*, 2065–2069.

38 Pley, H. W.; Flaherty, K. M.; McKay, D. B. 3-Dimensional structure of a hammerhead ribozyme. *Nature* **1994**, *372*, 68–74.

39 Martick, M.; Scott, W. G. Tertiary contacts distant from the active site prime a ribozyme for catalysis. *Cell* **2006**, *126*, 309–320.

40 Paredes, E.; Das, S. R. Click Chemistry for Rapid Labeling and Ligation of RNA. *ChemBioChem* **2011**, *12*, 125–131.

41 Mutisya, D.; Selvam, C.; Kennedy, S. D.; Rozners, E. Synthesis and properties of triazole-linked RNA. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 3420–3422.

42 El-Sagheer, A. H.; Kumar, R.; Findlow, S.; Werner, J. M.; Lane, A. N.; Brown, T. A very stable cyclic DNA miniduplex with just two base pairs. *ChemBioChem* **2008**, *9*, 50–52.

43 El-Sagheer, A. H.; Brown, T. Synthesis, Serum Stability and Cell Uptake of Cyclic and Hairpin Decoy Oligonucleotides for TCF/LEF and GLI Transcription Factors. *Int. J. Pept. Res. Ther.* **2008**, *14*, 367–372.

44 Nakane, M.; Ichikawa, S.; Matsuda, A. Triazole-linked dumbbell oligodeoxynucleotides with NF-kappa B binding ability as potential decoy molecules. *J. Org. Chem.* **2008**, *73*, 1842–1851.

45 Nordell, P.; Westerlund, F.; Reymer, A.; El-Sagheer, A. H.; Brown, T.; Norden, B.; Lincoln, P. DNA Polymorphism as an Origin of Adenine-Thymine Tract Length-Dependent Threading Intercalation Rate. *J. Am. Chem. Soc.* **2008**, *130*, 14651–14658.

46 Lundberg, E. P.; El-Sagheer, A. H.; Kocalka, P.; Wilhelmsson, L. M.; Brown, T.; Norden, B. A new fixation strategy for addressable nano-network building blocks. *Chem. Commun.* **2010**, *46*, 3714–3716.

47 Jacobsen, M. F.; Ravnsbaek, J. B.; Gothelf, K. V. Small molecule induced control in duplex and triplex DNA-directed chemical reactions. *Org. Biomol. Chem.* **2010**, *8*, 50–52.

48 Xu, Y.; Suzuki, Y.; Komiyama, M. Click Chemistry for the Identification of G-Quadruplex Structures: Discovery of a DNA-RNA G-Quadruplex. *Angew. Chem., Int. Ed.* **2009**, *48*, 3281–3284.

49 Xiong, H.; Seela, F. Stepwise ″Click″ Chemistry for the Template Independent Construction of a Broad Variety of Cross-Linked Oligonucleotides: Influence of Linker Length, Position, and Linking Number on DNA Duplex Stability. *J. Org. Chem.* **2011**, *76*, 5584–5597.

50 Pujari, S. S.; Xiong, H.; Seela, F. Cross-Linked DNA Generated by ″Bis-click″ Reactions with Bis-functional Azides: Site Independent Ligation of Oligonucleotides via Nucleobase Alkynyl Chains. *J. Org. Chem.* **2010**, *75*, 8693–8696.

51 Peng, X. H.; Li, H.; Seidman, M. A Template-Mediated Click-Click Reaction: PNA-DNA, PNA-PNA (or Peptide) Ligation, and Single Nucleotide Discrimination. *Eur. J. Org. Chem.* **2010**, 4194–4197.

52 El-Sagheer, A. H.; Brown, T. Factors influencing hairpin oligonucleotide cyclization by the uncatalyzed alkyne-azide cycloaddition (AAC) reaction. *Pure Appl. Chem.* **2010**, *82*, 1599–1607.

53 Agard, N. J.; Bertozzi, C. R. Chemical Approaches To Perturb, Profile, and Perceive Glycans. *Acc. Chem. Res.* **2009**, *42*, 788–797.

54 Shelbourne, M.; Chen, X.; Brown, T.; El-Sagheer, A. H. Fast copper-free click DNA ligation by the ring-strain promoted alkyne-azide cycloaddition reaction. *Chem. Commun.* **2011**, *47*, 6257–6259.

55 El-Sagheer, A. H.; Cheong, V. V.; Brown, T. Rapid chemical ligation of oligonucleotides by the Diels-Alder reaction. *Org. Biomol. Chem.* **2011**, *9*, 232–235.