ChemPubSoc Europe

DOI: 10.1002/cbic.201600320

CHEMBIOCHEM Communications

Conjugation and Evaluation of Triazole-Linked Single Guide RNA for CRISPR-Cas9 Gene Editing

Kaizhang He, Eldon T. Chou, Shawn Begay, Emily M. Anderson, and Anja van Brabant Smith*[a]

The CRISPR-Cas9 gene editing system requires Cas9 endonuclease and guide RNAs (either the natural dual RNA consisting of crRNA and tracrRNA or a chimeric single guide RNA) that direct site-specific double-stranded DNA cleavage. This communication describes a click ligation approach that uses alkyne–azide cycloaddition to generate a triazole-linked single guide RNA (sgRNA). The conjugated sgRNA shows efficient and comparable genome editing activity to natural dual RNA and unmodified sgRNA constructs.

The CRISPR-Cas9 system has rapidly transitioned from an RNA-guided defense system in bacteria to a remarkably effective and flexible technique for modifying the genome of higher eukaryotes, particularly in mammalian cells.[1–3] In the most commonly used *Streptococcus pyogenes* type II CRISPR-Cas9 system, the Cas9 (CRISPR-associated 9) protein that cleaves the DNA is guided by two required small RNA sequences: the CRISPR RNA (crRNA), which binds the target DNA and guides cleavage, and the trans-activating crRNA (tracrRNA), which base pairs with the crRNA and enables the Cas9–crRNA complex to form. A single RNA chimera, which combines the targeting specificity of the crRNA with the scaffolding properties of the tracrRNA, has been engineered to direct sequence-specific dsDNA cleavage by Cas9.[4] The single guide RNA (sgRNA), typically about 100 nucleotides (nt) long, can be generated by using a DNA vector expressing the sgRNA[5–7] or by in vitro transcription (IVT).[8] Both methods have a number of drawbacks, including time-consuming cloning steps, unintended genome modifications from random integration of DNA plasmids encoding the sgRNA, and the complexity of generating high-quality IVT material.[9]

An ideal alternative to an sgRNA expression plasmid or IVT-generated sgRNA would be solid-phase chemical synthesis. This method has the flexibility to quickly and accurately generate different sequences and lengths of RNA without the need for multiple cloning and sequencing steps. Although long RNAs (length around 100 nt) can be prepared by using automated, solid-phase oligonucleotide synthesis machines with 2′-acetoxyethyl orthoester (2′-ACE)[10,11] or 2′-thionocarbamate (2′-TC) chemistry,[12,13] effective and efficient synthesis of hundreds of high quality long RNAs with lengths of ~100 nt is still a challenge when considering yield, purity, and cost. An alternative approach to the synthesis of larger oligonucleotide strands is to combine solid-phase synthesis with chemical ligation; for example, a click chemistry approach can be applied in which copper(I)-catalyzed alkyne–azide cycloaddition (CuAAC) is used for the synthesis of very long oligodeoxyribonucleotides (DNAs).[14,15] Another benefit of using chemical ligation to conjugate crRNA and tracrRNA is the possibility of incorporating chemical modifications to the conjugation site. The CRISPR-Cas9 gene editing system holds tremendous therapeutic potential to treat unmet medical needs, and chemical modifications of synthetic sgRNA are likely required to improve the stability and pharmacokinetic properties. Modifications including phosphorothioates, 2′-OMe, 2′-F, 2′-constrained ethyl (2′-cEt), and/or 2′-OMe 3′-thioPACE were applied to regions of the crRNA that recognize their DNA targets or to the tracrRNA anti-repeat region and were found to increase the efficiency of gene disruption compared to unmodified crRNA.[13,16] We now report a solid-phase synthesis strategy for ligating two oligoribonucleotides (RNAs) by using the CuAAC reaction to assemble a long sgRNA containing a triazole linkage. This approach for generation of chemically synthesized sgRNAs has several desirable features and potentials: it is easy to carry out, other conjugation chemistries (i.e., amine–NHS, thiol–maleimide, and hydrazine–oxime) can also be used, and it is applicable to the synthesis of libraries of sgRNAs. We also envisage that the conjugation methodology presented here will be invaluable for the synthesis of highly modified sgRNAs with complex structures (i.e., incorporation of an *N*-acetylgalactosamine [GalNAc] moiety for delivery[17] or conjugation of fluorescent dyes for imaging[18]).

Various bioconjugation reactions can be used to ligate two oligonucleotides.[19] One widely utilized method is the Huisgen 1,3-dipolar cycloaddition reaction involving an alkyne and azide to yield a highly stable triazole linker. The ligation reaction is fast, efficient, specific, and compatible with aqueous media.[20,21] To synthesize the long conjugated sgRNA containing a triazole linker (Scheme 1), we designed a 5′-hexyne tracrRNA (tracr-1, 65-mer), based on the published *S. pyogenes* tracrRNA sequence,[4] and a 3′-azide crRNA component (cr-1,

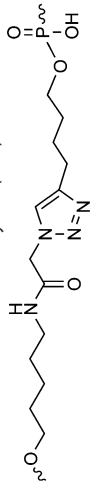
[a] *Dr. K. He, E. T. Chou, S. Begay, Dr. E. M. Anderson, Dr. A. van Brabant Smith*
GE Healthcare Dharmacon Inc.
2650 Crescent Drive, Suite 100, Lafayette, CO 80026 (USA)
E-mail: anja.smith@ge.com

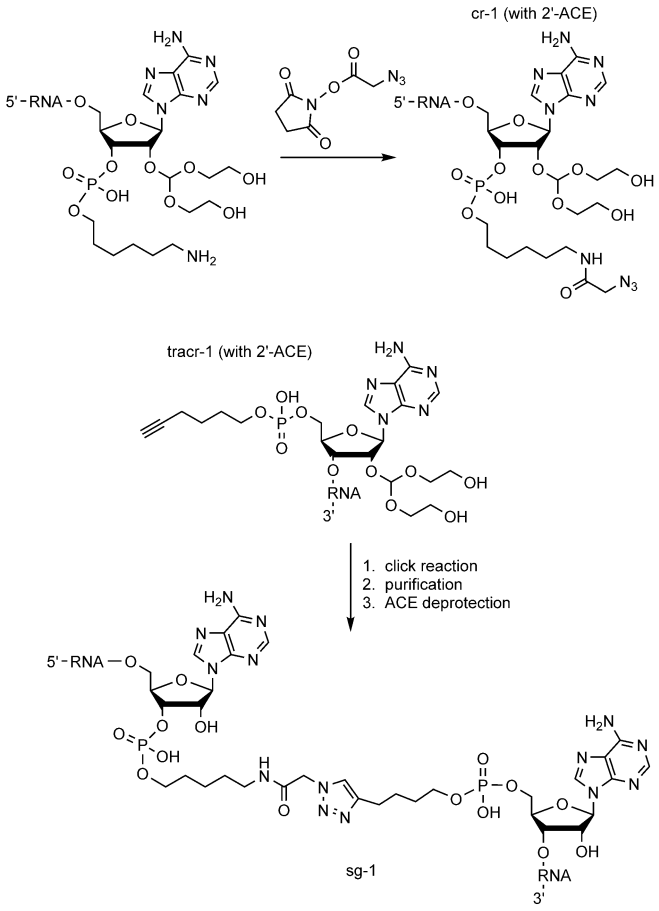
Supporting information for this article can be found under http://dx.doi.org/10.1002/cbic.201600320: experimental procedures (RNA synthesis, CRISPR-Cas9 experiments) and mass spectra of synthesized oligoribonucleotides (RNAs).

Table 1. RNA oligonucleotides used in this study.

Code	Sequence (5′–3′)	nt length	MS calcd	MS found
tracr-1	5′-hexyne-AAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU-3′	65	21 080.7	21 082.3
tracr-2	5′-AAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU-3′	65	20 920.7	20 918.3
cr-1	5′-GUGUAUUUUGACCUACGAAUGUUUUAGAGCUAGA-C6-N_3-3′	34	11 126.0	11 127.0
cr-2	5′-GUGUAUUUUGACCUACGAAUGUUUUAGAGCUAGA-3′	34	10 863.5	10 862.2
cr-3	5′-GUGUAUUUUGACCUACGAAUGUUUUAGAGCUAGA-C6-NH_2-3′	34	10 988.6	10 990.4
sg-1	5′-GUGUAUUUUGACCUACGAAUGUUUUAGAGCUAGA-**L**-AAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUC-AACUUGAAAAAGUGGCACCGAGUCGGUGCUUU-3′	99	32 192.5	32 225.1
sg-2	5′-GUGUAUUUUGACCUACGAAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA-CUUGAAAAAGUGGCACCGAGUCGGUGCUUU-3′	99	31 846.2	31 863.8

Products were characterized by MS (ESI^-). L is:





Scheme 1. Synthesis of triazole-linked single guide RNA (sgRNA) (sg-1).

34-mer; Table 1). The requisite 5′-hexyne-tracrRNA protected with 2′-ACE was prepared by incorporating a 5′-hexyne phosphoramidite monomer in the final coupling step to the solid support-bound RNA, synthesized by 2′-acetoxyethyl orthoester (2′-ACE) chemistry. The 34-nt crRNA was designed by appending part of the *S. pyogenes* crRNA repeat sequence, 5′-GUUUUAGAGCUAGA-3′, to the 20-mer target/spacer sequence (GUGUAUUUUGACCUACGAAU) targeting human peptidylprolyl isomerase B (PPIB). The corresponding 3′-azide-crRNA, protected with 2′-ACE, was synthesized from the reaction of azidoacetic acid NHS ester with HPLC-purified 3′-aminohexyl-modified RNA (cr-3 with 2′-ACE). The 3′-azide-labeled crRNA was desalted and further purified by reverse-phase HPLC.

The Cu^I-catalyzed ligation reaction between 2′-ACE-protected 5′-hexyne-tracrRNA and 3′-azide-crRNA in aqueous solution was conducted on a 50 nmol scale. The resulting conjugated sgRNA was precipitated with acetone. The RNA pellet was washed, dried, and purified by reverse-phase HPLC. The 2′-ACE groups on the conjugated sgRNA were removed according to a standard Dharmacon protocol,[10,11] and the triazole-linked sgRNA (sg-1) was desalted by ethanol precipitation.

To investigate and compare the efficiency of genome editing by the synthetic conjugated sgRNA (sg-1), we also synthesized a natural 65-mer tracrRNA (tracr-2), 34-mer crRNA (cr-2), and

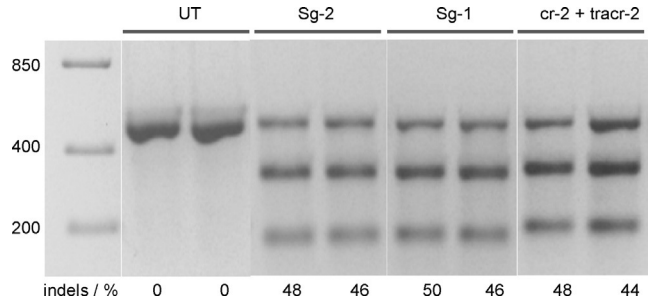
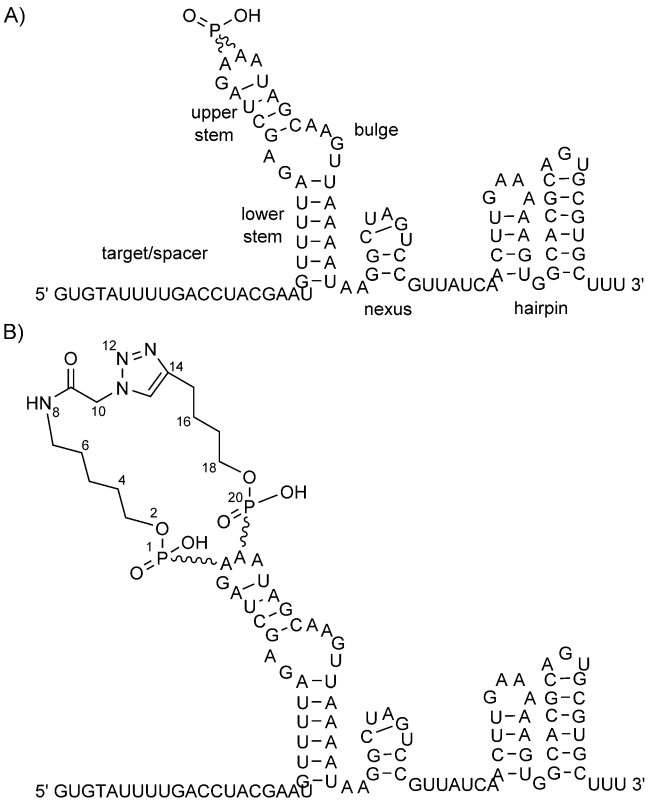


Figure 1. Gene editing by using synthetic RNAs. Transfection of a synthetic sgRNA 99-mer (sg-2), conjugated sgRNA 99-mer (sg-1), or synthetic crRNA (34-mer, cr-2) and tracrRNA (65-mer, tracr-2) results in targeted DNA cleavage in Cas9-expressing U2OS cells. Stable Cas9-expressing cells were generated by using Cas9 lentivirus. Cells were transfected with the sgRNA or dual crRNA and tracrRNA and compared to untransfected cells (UT) by using the T7 endonuclease I (T7EI) mismatch detection assay. The percent edited after 72 h is indicated at the bottom of each lane.



Scheme 2. sgRNA sequences and functional modules: A) Natural sgRNA (sg-2) with native phosphodiester bond; B) Conjugated sgRNA (sg-1) with triazole linker.

99-mer sgRNA (sg-2) by using 2′-ACE chemistry (Table 1). A T7E1 mismatch detection assay was used to detect and quantify insertions and deletions (indels) that were created by repair of the CRISPR-Cas9-induced double strand break through the non-homologous end joining repair pathway. As shown in Figure 1, transfection of sg-1 into a U2OS cell line that stably expresses Cas9 results in indels at the PPIB-targeted site. The level of gene disruption (>40% cleavage) is comparable to that of the natural dual RNA system with two RNAs (cr-2 and tracr-2). After replacing the phosphodiester bond with a long linker containing the triazole linkage (Scheme 1), the conjugated sgRNA (sg-1) retained high indel frequency, comparable to the natural 99-mer sgRNA (sg-2). A recent study identified and characterized six conserved modules (spacer, lower stem, upper stem, bulge, nexus, and hairpin) within native crRNA:tracrRNA duplexes and sgRNAs that direct Cas9 endonuclease activity; the bulge and nexus are the most critical features for Cas9 targeting.[21] Compared to the native 99-mer sgRNA, the triazole-linked sgRNA has an additional 20-atom linkage replacing the phosphodiester bond (Scheme 2). The linkage is located at the loop structure of the upper stem; this structure was suggested to be non-essential for CRISPR-Cas9 functionality.[22,23] Our gene editing results confirmed that changing or extending the loop structure with a non-natural linkage did not have a significant effect on knockout efficiency.

In summary, we report the first CRISPR-Cas9 system with a synthetic conjugated sgRNA with efficient gene knockout functionality. Our results indicate that the nucleotides in the loop structure of the upper stem in the crRNA:tracrRNA duplex are not essential for CRISPR-Cas9 targeting and open up the possibility of modifying the loop structure with unnatural linkages by conjugating modified crRNA and tracrRNA. This approach would not necessarily be limited to the triazole linker reported here, and alternative conjugation chemistries[19,21] and the resulting linkers (amide, disulfide, thioester, hydrozone, etc.) might also be suitable for forming new conjugated sgRNA molecules. Furthermore, the ability to synthesize conjugated sgRNAs by using a 5′-end-modified tracrRNA and altered crRNAs (3′-end-modified) with different sequences targeting multiple genes opens new avenues for generating long sgRNA libraries in array or plate formats.

Acknowledgements

We thank Chelsea Cheng and Kristina Larson for assistance with oligonucleotide synthesis.

Keywords: conjugation · CRISPR-Cas9 · gene technology · oligonucleotides · single guide RNA

[1] J. D. Sander, J. K. Joung, *Nat. Biotechnol.* **2014**, *32*, 347–355.
[2] P. D. Hsu, E. S. Lander, F. Zhang, *Cell* **2014**, *157*, 1262–1278.
[3] A. V. Wright, J. K. Nunez, J. A. Doudna, *Cell* **2016**, *164*, 29–44.
[4] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, E. Charpentier, *Science* **2012**, *337*, 816–821.
[5] L. Cong, F. A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P. D. Hsu, X. Wu, W. Jiang, L. A. Marraffini, F. Zhang, *Science* **2013**, *339*, 819–823.
[6] P. Mali, L. Yang, K. M. Esvelt, J. Aach, M. Guell, J. E. DiCarlo, J. E. Norville, G. M. Church, *Science* **2013**, *339*, 823–826.
[7] M. Jinek, A. East, A. Cheng, S. Lin, E. Ma, *eLife* **2013**, *2*, e00471.
[8] S. W. Cho, S. Kim, J. M. Kim, J. S. Kim, *Nat. Biotechnol.* **2013**, *31*, 230–232.
[9] J. L. Gori, P. D. Hsu, M. L. Maeder, S. Shen, G. G. Welstead, D. Bumcrot, *Human Gene Ther.* **2015**, *26*, 443–451.
[10] S. A. Scaringe, F. C. Wincott, M. H. Caruthers, *J. Am. Chem. Soc.* **1998**, *120*, 11820–11821.

[11] S. A. Scaringe, *Methods Enzymol.* **2000**, *317*, 3 – 18.
[12] D. J. Dellinger, Z. Timár, F. Ferreira, Z. Kupihár, G. Dellinger, *J. Am. Chem. Soc.* **2011**, *133*, 11540 – 11546.
[13] A. Hendel, R. O. Bak, J. T. Clark, A. B. Kennedy, D. E. Ryan, S. Roy, I. Steinfeld, B. D. Lunstad, R. J. Kaiser, A. B. Wilkens, R. Bacchetta, A. Tsalenko, D. Dellinger, L. Bruhn, M. H. Porteus, *Nat. Biotechnol.* **2015**, *33*, 985 – 989.
[14] J. Qiu, A. H. El-Sagheer, T. Brown, *Chem. Commun.* **2013**, *49*, 6959 – 6961.
[15] A. H. El-Sagheer, T. Brown, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 15329 – 15334.
[16] M. Rahdar, M. A. McMahon, T. P. Prakash, E. E. Swayze, F. Bennett, D. W. Clevaland, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E7110 – 7117.
[17] K. G. Rajeev, J. K. Nair, K. Charisse, N. Taneja, J. O'Shea, J. L. S. Willoughby, K. Yucius, T. Nguyen, S. Shulga-Morskaya, S. Milstein, A. Liebow, W. Querbes, A. Borodovsky, K. Fitzgerald, M. A. Maier, M. Manoharan, *ChemBioChem* **2015**, *16*, 903 – 908.
[18] W. Deng, X. shi, R. Tjian, T. Lionnet, R. H. Singer, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 11870 – 11875.
[19] E. M. Sletten, C. R. Bertozzi, *Angew. Chem. Int. Ed.* **2009**, *48*, 6974 – 6998; *Angew. Chem.* **2009**, *121*, 7108 – 7133.
[20] C. R. Becer, R. Hoogenboom, U. S. Schubert, *Angew. Chem. Int. Ed.* **2009**, *48*, 4900 – 4908; *Angew. Chem.* **2009**, *121*, 4998 – 5006.
[21] C. P. Ramil, Q. Lin, *Chem. Commun.* **2013**, *49*, 11007 – 11022.
[22] A. E. Briner, P. D. Donohoue, A. A. Gomaa, K. Selle, E. M. Slorach, C. H. Nye, R. E. Haurwitz, C. L. Beisel, A. P. May, R. Barrangou, *Mol. Cell* **2014**, *56*, 333 – 339.
[23] H. Nishimasu, F. A. Ran, P. D. Hsu, S. Konermann, S. Shehata, N. Dohmae, R. Ishitani, F. Zhang, O. Nureki, *Cell* **2014**, *156*, 935 – 949.

Manuscript received: June 3, 2016
Accepted article published: July 21, 2016
Final article published: August 19, 2016

