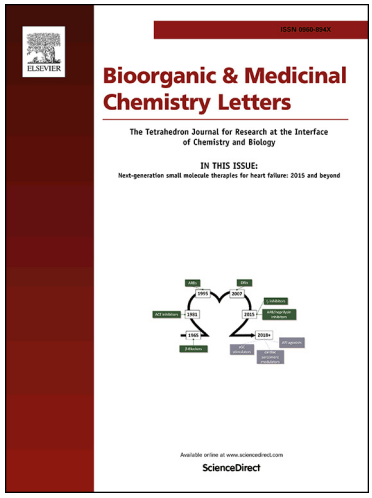
Accepted Manuscript



A double-headed nucleotide with two cytosines: DNA with condensed information and improved duplex stability

Kasper Beck, Charlotte Reslow-Jacobsen, Mick Hornum, Christian Henriksen, Poul Nielsen

PII: S0960-894X(19)30002-2
DOI: https://doi.org/10.1016/j.bmcl.2019.01.002
Reference: BMCL 26239

To appear in: *Bioorganic & Medicinal Chemistry Letters*

Received Date: 5 December 2018
Accepted Date: 2 January 2019

Please cite this article as: Beck, K., Reslow-Jacobsen, C., Hornum, M., Henriksen, C., Nielsen, P., A double-headed nucleotide with two cytosines: DNA with condensed information and improved duplex stability, *Bioorganic & Medicinal Chemistry Letters* (2019), doi: https://doi.org/10.1016/j.bmcl.2019.01.002

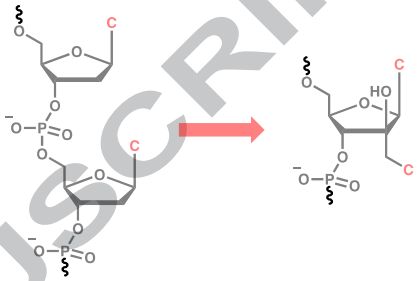
Graphical Abstract

A double-headed nucleotide with two cytosines: DNA with condensed information and improved duplex stability

Kasper Beck, Charlotte Reslow-Jacobsen, Mick Hornum, Christian Henriksen and Poul Nielsen*

A double-headed nucleotide monomer with an additional cytosine base in the 2´-position is prepared and found to behave as a compressed dinucleotide thereby condensing the information in DNA while also stabilizing the resulting duplex.

Leave this area blank for abstract info.



Bioorganic & Medicinal Chemistry Letters
journal homepage: www.elsevier.com

A double-headed nucleotide with two cytosines: DNA with condensed information and improved duplex stability

Kasper Beck, Charlotte Reslow-Jacobsen, Mick Hornum, Christian Henriksen and Poul Nielsen*

[a]*Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, 5230 Odense M, Denmark*

ARTICLE INFO

Article history:
Received
Revised
Accepted
Available online

Keywords:
Nucleic acids;
Double-headed nucleotides;
Oligonucleotides;
Duplex;
Triplex

ABSTRACT

Double-headed nucleotide monomers are capable of condensing the genetic information of DNA. Herein, a double-headed nucleotide with two cytosine bases ($\mathbf{C_C}$) is constructed. The additional cytosine is connected through a methylene linker to the 2´-position of arabinocytidine. The nucleotide is incorporated into oligonucleotides and its effect on duplex stability is studied. For single incorporations, a thermal stabilization of 4.0 °C is found as compared to the unmodified duplex and it is shown that both nucleobases of $\mathbf{C_C}$ participate in Watson-Crick base pairing. In combination with the previously published $\mathbf{U_T}$ monomer, it is also shown that multiple incorporations are tolerated. For instance, a 16-mer sequence is targeted by a 13-mer oligonucleotide by using one $\mathbf{C_C}$ and two $\mathbf{U_T}$ monomers without compromising the overall duplex stability. Finally, the potential of double-headed nucleotides in triplex forming oligonucleotides is studied, however, with the conclusion that the present design is not well-suited for this function.

* Prof. Poul Nielsen. Tel.: +45-6550-2565; e-mail: pouln@sdu.dk

The double-helical structure of DNA self-assembles through the formation of predictable and highly specific Watson-Crick base pairs. The DNA duplex, as well as the more complex DNA triplex, provides a brilliant platform for the design of therapeutics based on supramolecular chemistry, e.g. anti-sense oligonucleotides, nucleic acid aptamers or triplex forming oligonucleotides.[1-3] With the convenient automated synthesis of DNA, it is easy to introduce and study the effect of chemically modified nucleotides on these systems, and a focus on modifications that improve or introduce new properties to the oligonucleotides has emerged.[4,5]

In this context, an interesting modification is the addition of a second nucleobase to the nucleotide skeleton – the so-called double-headed nucleotide.[6] We have previously reported the synthesis and recognition properties of several double-headed nucleotides with the additional nucleobase attached to the 5´-, 2´- or 5-position through various linkers,[7-17] while others have used the 4´-position or the 2´-*N* of amino-LNA as attachment points.[18,19] For the **dU$_B$** design (Fig. 1) with the additional nucleobase attached via a methylene linker to the 2´-position of a 2´-deoxynucleotide, it has been shown that both the natural and the additional nucleobase participate in Watson-Crick base pairing.[10,11] More recently, we have examined a similar design with an arabino-configuration (**U$_B$**, Fig. 1), with a significantly easier synthesis and little or no reduction in the recognition potential as compared to **dU$_B$**.[13] For this design, the full set of uridine-based double-headed nucleotides have been reported comprising **U$_T$**, **U$_C$**, **U$_G$** and **U$_A$**.[14] In practice, these double-headed nucleotides behave like compressed dinucleotides, and consequently, the information of two nucleotides can be condensed to a single double-headed nucleotide, thereby reducing the molecular size and the number of phosphates in a given oligonucleotide. We expect this to form the basis for improved nucleic acid based therapeutics.

In this study we introduce the first cytidine-based double-headed nucleotide **C$_C$** (Fig. 1) with two cytosines. With this expansion, we are for the first time able to examine the recognition properties of double-headed nucleotides binding only through G:C base pairs. Together with **U$_T$**, this new **C$_C$** monomer provides an excellent basis for evaluating the potential of double-headed nucleotides in triplex-forming oligonucleotides (TFOs). These two monomers were therefore incorporated into a standard homo-pyrimidine TFO sequence[20] at various positions resulting in a selection of modified oligonucleotides with both single and multiple incorporations of double-headed nucleotides. The recognition potential of these strands in both duplexes and triplexes is presented.

The synthesis of the double-headed nucleoside is shown in Scheme 1. Starting from uridine **1**, the protected ketone **2** was easily prepared using the optimized and chromatography-free procedure from Lemaire et. al.[21] Stereoselective conversion of **2** to the 2´(*S*)-spiroepoxide **3** was achieved using trimethylsulfoxonium iodide.[13,22] The epoxide was opened by uracil in a N^1-regioselective manner to afford the protected double-headed nucleoside **4**. This regioselectivity was confirmed by the presence of a $^3J_{CH}$ coupling between the protons of the 2´-methylene linker and both C2 and C6 of the newly introduced uracil in the ^{1}H,^{13}C-HMBC NMR spectrum of **4**. Hereafter, the two uracils were converted to protected cytosines in a three-step one-pot synthesis of **5** with a decent yield of 55%. First, O4 of uracil was activated by tosylation, then replaced by a nucleophilic attack from ammonia at the C4-position, and finally the newly introduced amino group was protected as a benzoyl amide. This procedure was originally reported by Du et. al.[23] for a single uracil-to-cytosine conversion. Removal of the silyl protecting group was effectively achieved using triethylamine trihydrofluoride to give the double-headed nucleoside **6**, which was readily reprotected at the 5´-*O*-position by treatment with 4,4´-dimethoxytrityl chloride, affording **7**. Finally, phosphitylation at the 3´-*O*-position was accomplished to give the fully protected and activated phosphoramidite **8**.

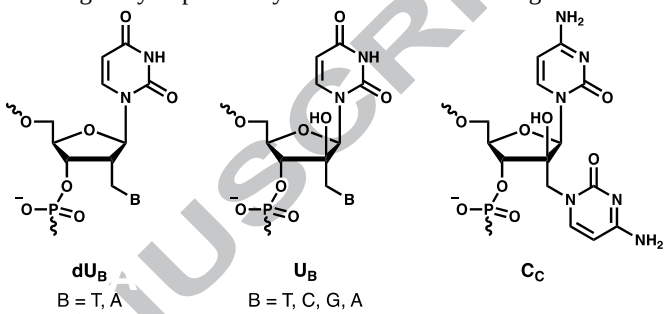


Figure 1. Double-headed nucleotides. T = thymin-1-yl, C = cytosin-1-yl, G = guanin-9-yl, A = adenin-9-yl.

Phosphoramidite **8** was then used in standard solid-phase DNA-synthesis in order to introduce the modified nucleotide **C$_C$** into oligonucleotides. In parallel, the corresponding phosphoramidite of **U$_T$** was synthesized as previously reported.[13] Standard conditions were used in the DNA-synthesis, but with 4,5-dicyanoimidazole as activator and extended coupling times (15 min) for both modified phosphoramidites. Treatment with concentrated ammonia at room temperature (24 h) was used for deprotection and cleavage from the solid-support. All oligonucleotides were purified using reversed phase HPLC, and MALDI-TOF MS and ion-exchange chromatography analysis were used to confirm the identity and determine purity, respectively.

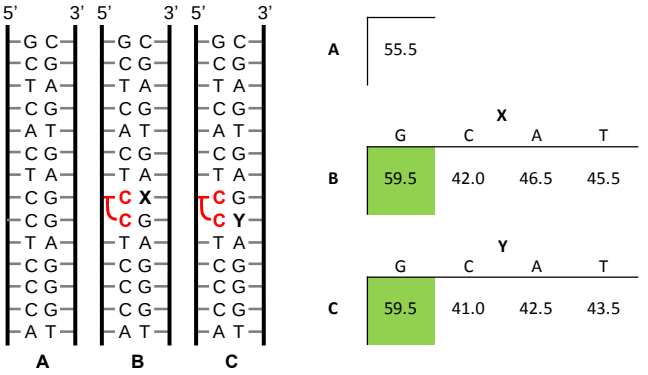
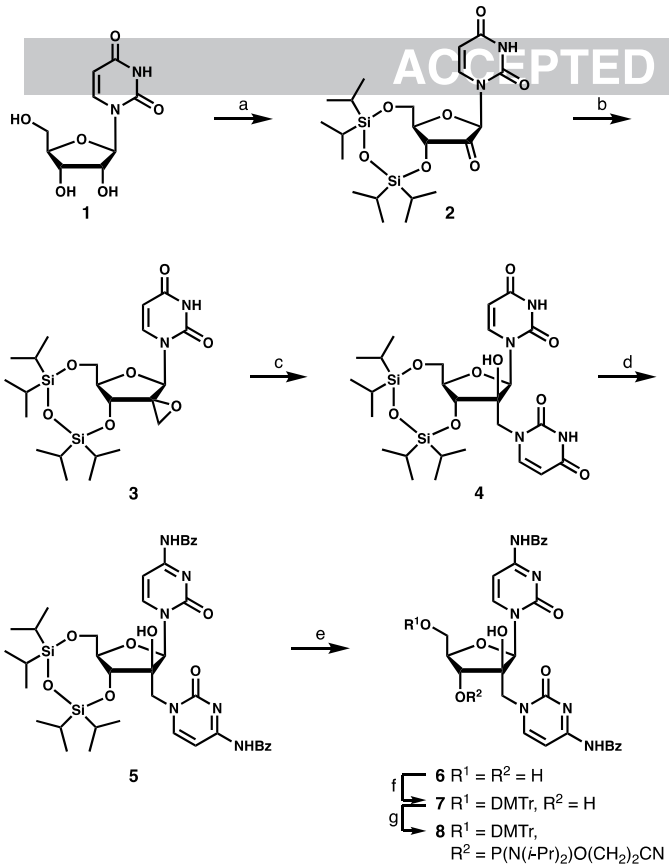


Figure 2. Structures of studied DNA duplexes and their melting temperatures[a] (T_m,°C). (A) 14-mer reference duplex. (B + C) Duplexes used to study match and mismatch properties of the two cytosines in **C$_C$** (marked in red). [a] Melting temperatures were determined in a buffer containing 2.5 mM Na_2HPO_4, 5.0 mM NaH_2PO_4, 100 mM NaCl and 0.1 mM EDTA at pH 7.0 using 1.5 µM concentrations of both strands.



Scheme 1. Reagents and conditions: (a) i. $TIPDSCl_2$, imidazole, CH_2Cl_2, ii. TEMPO, $PhI(OAc)_2$, AcOH, CH_2Cl_2, 71% (ref. 21); (b) NaH, $(CH_3)_3SOI$, THF, DMSO, 84% (ref. 22); (c) NaH, uracil, DMF, 71%; (d) i. TsCl, DMAP, Et_3N, CH_3CN, ii. NH_4OH, iii. BzCl, pyridine, 55%; (e) $Et_3N{\cdot}3HF$, THF, 88%; (f) DMTrCl, pyridine, 86%; (g) $NC(CH_2)_2OP(N(i\text{-}Pr)_2)Cl$, $(i\text{-}Pr)_2NEt$, DCE, 66%. DMTr = 4,4´-dimethoxytrityl.

At first, we studied the effect of a single incorporation of the $\mathbf{C_C}$ monomer by exchanging the central 5´-CC dinucleotide of our standard mixed 14-mer sequence with the $\mathbf{C_C}$ monomer, giving a modified 13-mer (Fig. 2). This strand was hybridized with the complementary 14-mer sequence and the melting temperature (T_m) was obtained from the maximum of the first derivative of the UV ($A_{260\ nm}$) melting curve. It was found that the introduction of $\mathbf{C_C}$ caused a substantial increase in thermal stability of 4.0 °C as compared to the unmodified duplex (Fig. 2, 59.5 °C compared to 55.5 °C, respectively). Hence, the ability of $\mathbf{C_C}$ to function as a dinucleotide is well-founded, and the $\mathbf{C_C}$ monomer is even better tolerated in the DNA duplex than its uridine-based counterparts ($\mathbf{U_T}$, $\mathbf{U_C}$, $\mathbf{U_G}$ and $\mathbf{U_A}$) showing ΔT_m's ranging from +1.0 °C to –3.5 °C in similar 14-mer duplexes.[14]

A mismatch study was made, targeting not only the additional 2´-attached nucleobase of $\mathbf{C_C}$, but also the "natural" 1´-attached nucleobase. When placing a mismatch nucleotide opposite to the latter nucleobase, a considerable decrease of 13.0–17.5 °C was observed in the melting temperature of the duplex (Fig. 2B). For the additional cytosine, the mismatch discrimination is even slightly better in each of the three mismatches with decreases in melting temperatures of 16.0–18.5 °C (Fig. 2C). Altogether, this shows that the mismatch discrimination of $\mathbf{C_C}$ is excellent, and it confirms that also the additional cytosine base actually takes part in Watson-Crick base pairing with the opposite guanine.

To get further insight into the effect of double-headed nucleotides on both duplex and triplex stability, a series of six modified oligonucleotides were synthesized. All of these were based on the same 16-mer homo-pyrimidine TFO sequence, which has been studied before,[15,17,20] having single, double or triple incorporations of either $\mathbf{C_C}$, $\mathbf{U_T}$ or both in exchange for the corresponding dinucleotides (Fig. 3). First, each strand was hybridised with the 16-mer complimentary homo-purine strand, giving the unmodified duplex (Fig. 3, entry D1) and the six modified duplexes (entries D2–D7). Looking first at the single-modified duplexes, there is a clear correlation between the found stabilizing effect of $\mathbf{C_C}$ of 4.5 °C (entry D4) and the previously found 4.0 °C in the aforementioned mixed sequence (Fig. 2). Likewise, the thermal destabilization of –3.0 °C and –4.0 °C (entries D2 and D3) found for a single incorporation of $\mathbf{U_T}$ is in accordance with the published –3.5 °C for its incorporation into a 13-mer mixed sequence.[13] From these results it is apparent that the thermal stability change provided by the double-headed nucleotides has only a minor dependence on the local environment in the form of nucleotide sequence. It should be noticed, however, that the A:U base pair of $\mathbf{U_T}$ is compared directly with an unmodified A:T base pair, and the lack of a methyl group might add to the decrease in stability.[24]

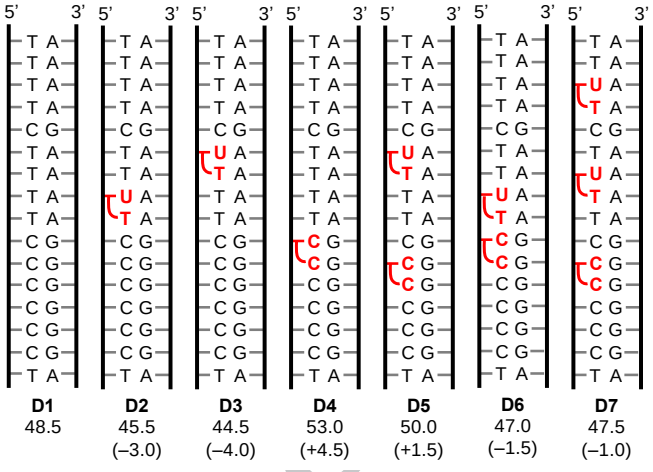


Figure 3. Structures of studied DNA duplexes, their melting temperatures[a] (T_m, °C) and the differences in melting temperatures (ΔT_m, °C) relative to duplex D1. [a] See caption of Fig. 2.

Hereafter, duplexes with multiple incorporations were studied (Fig. 3, entries D5–D7). The melting temperatures of these duplexes can approximately be explained by accumulating the isolated effect of each incorporation. Specifically, the ΔT_m for the doubly modified duplex D5 (+1.5 °C) only differs by 0.5 °C from the sum of the corresponding ΔT_m's for the single $\mathbf{U_T}$ incorporation D3 (–3.0 °C) and the single $\mathbf{C_C}$ incorporation D4 (+4.0 °C), and so the thermal effects of the double-headed nucleotides seem to stack. A limitation to this trend arises when the double-headed nucleotides are incorporated consecutively. In this case a thermal destabilization of –3.0 °C is observed, compared to having a gap of three canonical nucleotides (D6 compared to D5). This might be explained by the geometry of the double-headed nucleotides demanding some degree of adjustment of neighbouring nucleotides. In duplex D7, a total of three double-headed nucleotides were incorporated, and even with gaps of only two canonical nucleotides, the change in melting temperature (–1.0 °C) was still roughly the sum of two single $\mathbf{U_T}$ and a single $\mathbf{C_C}$ incorporation. D7 is essentially comprised of a 13-mer modified strand that recognizes a 16-mer complementary strand.

Finally, the modified oligonucleotides were evaluated for their potential as triplex-forming oligonucleotides (TFOs). Polypyrimidine oligonucleotides already studied in duplexes (Fig. 3, entries D2–D5 and D7) were hybridized with a standard 29-mer target DNA duplex[15,20] in a high salt buffer at pH 6, and the melting temperatures of the resulting triplexes were measured. Triplex formation, however, was only observed for one of the modified sequences having a single $\mathbf{C_C}$ incorporation (Fig. 4). Even this triplex was considerably destabilized by the modification, showing a decrease in melting temperature of 7.5 °C compared to the unmodified triplex (24.5 °C). Consequently, it seems that the design of the presented double-headed nucleotides is not optimal for accommodation in triplexes. Probably neither the constrained double-headed nucleotide structure nor the target duplex allows for the necessary adjustment of the geometry.

In summary, we have shown some systematic effects on duplex stability of the different double-headed nucleotides. The incorporation of $\mathbf{U_T}$ destabilizes the duplex by –3.5 °C,[14] whereas the $\mathbf{C_C}$ monomer shows a stabilizing effect of 4.0 °C. In comparison, the $\mathbf{U_A}$ monomer gave a destabilization of –2.5 °C, whereas $\mathbf{U_C}$ and $\mathbf{U_G}$ showed minor effects on duplex stability of +1.0 °C and 0.0 °C, respectively, in similar sequences.[14] Even though G:C pairs are always stronger than A:T and A:U pairs, the present results demonstrate that G:C base pairs provide a relatively larger stabilizing effect in the condensed DNA provided by the double-headed nucleotides. The reason might be found in the special geometry, and a speculation could be that the third hydrogen bond of the G:C pair thermodynamically drives the neighbouring base-pairs into a more fixed geometry. Further studies are needed to enlighten this observation.

With the present results, we have gained increased knowledge of DNA duplexes with double-headed nucleotides and about the scope in sequence design; (1) Several modifications are allowed in the same sequence without compromising the thermal stability. (2) A small penalty, however, is paid for consecutive incorporations. (3) Stabilization is gained with G:C base pairs, while some destabilization comes with A:T base pairs, relative to native DNA. A larger variety of sequences including double-headed nucleotide monomers with other base-combinations is needed to explore the full scope of modification with double-headed nucleotides in oligonucleotides and, in other words, the maximal level of condensation of the information in DNA.

In conclusion, a double-headed nucleoside with two protected cytosine bases has been synthesized in 8 steps from commercially available uridine, and thereafter protected and activated for oligonucleotide synthesis. A selection of modified oligonucleotides has been prepared containing this $\mathbf{C_C}$ monomer and/or the corresponding uridine-based $\mathbf{U_T}$ monomer. The hybridization studies showed that $\mathbf{C_C}$ is very well-accommodated in double-stranded DNA, and moreover, the mismatch discrimination properties of $\mathbf{C_C}$ were shown to be excellent at both positions. It has also been shown that multiple incorporations of double-headed nucleotides are tolerated, as a triply modified strand (being in fact a 13-mer) recognizes a 16-mer complementary strand with a drop of just 1.0 °C in thermal stability as compared to the unmodified duplex. The potential of $\mathbf{C_C}$ in triplex-forming oligonucleotides has been evaluated as well, but clearly, the present design is not well-suited for this function. From these results, it can be concluded that $\mathbf{C_C}$ is able to condense the information of two deoxycytidines to a single double-headed nucleotide, while it at the same time is improving the thermal stability of the duplex. This is a truly unique set of properties that may have future potential in the development of nucleic acid based therapeutics.

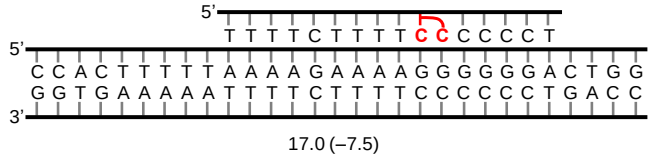


Figure 4. Structure of a selected triplex, its melting temperature[a] (T_m, °C) and the difference in melting temperature (ΔT_m, °C) relative to the unmodified triplex. [a] Melting temperatures were determined in a buffer containing 10 mM sodium cacodylate, 150 mM NaCl and 10 mM $MgCl_2$ at pH 6.0 using 1.5 μM concentrations of the TFOs and 1.0 mM concentrations of each strands of the target DNA-duplex. In all experiments the melting temperature of the duplex was found to be 68 °C.

Acknowledgments

This work was supported by grants from the Danish Council for Independent Research | Natural Sciences.

Supplementary Material

Electronic supplementary material contains experimental details including synthetic procedures, selected NMR spectra, procedures for oligonucleotide synthesis and hybridization studies, as well as MS data for modified oligonucleotides.

References and notes

1. Deleavey, G. F.; Damha, M. J. *Chem. Biol.* **2012**, 19, 937.
2. Wan, W. B.; Seth, P. P. *J. Med. Chem.* **2016**, 59, 9645.
3. Chandrasekaran, A. R.; Rusling, D. A. *Nucleic Acids Res.* **2018**, 46, 1021.
4. Endo, M.; Sugiyama, H. *ChemBioChem* **2009**, 10, 2420.
5. Bandy, T. J.; Brewer, A.; Burns, J. R.; Marth, G.; Nguyen, T.; Stulz, E. *Chem. Soc. Rev.* **2011**, 40, 138.
6. Sharma, P. K.; Kumar, P.; Nielsen, P. *Aus. J. Chem.* **2016**, 69, 1094.
7. Christensen, M. S.; Madsen, C. M.; Nielsen, P. *Org. Biomol. Chem.* **2007**, 5, 1586.
8. Madsen, C. S.; Nielsen, L. J.; Pedersen, N. S.; Lauritsen, A.; Nielsen, P. *RSC Adv.* **2013**, *3*, 10696.
9. Pedersen, S. L.; Nielsen, P. *Org. Biomol. Chem.* **2005**, 3, 3570.
10. Madsen, C. S.; Witzke, S.; Kumar, P.; Negi, K.; Sharma, P. K.; Petersen, M.; Nielsen, P. *Chem. Eur. J.* **2012**, 18, 7434.
11. Kumar, P.; Sharma, P. K.; Madsen, C. S.; Petersen, M.; Nielsen, P. *ChemBioChem* **2013**, 14, 1072.
12. Kumar, P.; Madsen, C. S.; Nielsen, P. *Bioorg. Med. Chem. Lett* **2013**, 23, 6847.
13. Kumar, P.; Sharma, P. K.; Nielsen, P. *J. Org. Chem.* **2014**, 79, 11534.
14. Hornum, M.; Sharma, P. K.; Reslow-Jacobsen, C.; Kumar, P.; Petersen, M.; Nielsen, P. *Chem. Commun.* **2017**, 53, 9717.
15. Kumar, P.; Sorinas, A. F.; Nielsen, L. J.; Slot, M.; Skytte, K.; Nielsen, A. S.; Jensen, M. D.; Sharma, P. K.; Vester, B.; Petersen, M.; Nielsen, P. *J. Org. Chem.* **2014**, 79, 8020.
16. Dalager, M.; Andersen, N. K.; Kumar, P.; Nielsen, P.; Sharma, P. K. *Org. Biomol. Chem.* **2015**, 13, 7040.
17. Kumar, P.; Sharma, P. K.; Hansen, J.; Jedinak, L.; Reslow-Jacobsen, C.; Hornum, M.; Nielsen, P. *Bioorg. Med. Chem.* **2016**, 24, 742.
18. Wu, T.; Nauwelaerts, K.; Van Aerschot, A.; Froeyen, M.; Lescrinier, E.; Herdewijn, P. *J. Org. Chem.* **2006**, 71, 5423.
19. Umemoto, T.; Wengel, J.; Madsen, A. S. *Org. Biomol. Chem.* **2009**, 7, 1793.
20. Højland, T.; Kumar, S.; Babu, B. R.; Umemoto, T.; Albæk, N.; Sharma, P. K.; Nielsen, P.; Wengel, J. *Org. Biomol. Chem.* **2007**, 5, 2375.
21. Lemaire, S.; Houpis, I.; Wechselberger, R.; Langens, J.; Vermeulen, W. A. A.; Smets, N.; Nettekoven, U.; Wang, Y.;

Xiao, T.; Qu, H.; Liu, R.; Jonckers, T. H. M.; Raboisson, P.; Vandyck, K.; Nilsson, K. M.; Farina, V. *J. Org. Chem.* **2011**, 76, 297.

22. Sano, T.; Shuto, S.; Inoue, H.; Ueda, T. *Chem. Pharm. Bull.* **1985**, 33, 3617.

23. Du, J.; Chun, B.-K.; Mosley, R. T.; Bansal, S.; Bao, H.; Espiritu, C.; Lam, A. M.; Murakami, E.; Niu, C.; Micolochick Steuer, H. M.; Furman, P. A.; Sofia, M. J. *J. Med. Chem.* **2014**, 57, 1826.

24. Kamya, P. R. N.; Muchall, H. M. *J. Phys. Chem. A.* **2011**, 115, 12800.