A strongly pairing fifth base: oligonucleotides with a C-nucleoside replacing thymidine

Tanja J. Walter and Clemens Richert

Institute of Organic Chemistry, University of Stuttgart, 70569 Stuttgart, Germany

Received June 10, 2018; Revised July 06, 2018; Editorial Decision July 11, 2018; Accepted July 13, 2018

ABSTRACT

There are five canonical bases in DNA and RNA. Each base has its particular molecular recognition properties and base pairing strength. Thymine and uracil form only two hydrogen bonds when pairing with adenine, and duplexes rich in A:T base pairs are more labile than duplexes rich in C and G, making some sequences difficult to detect via hybridization in a genomic context. Here we report the synthesis of an ethynylmethylpyridone C-nucleoside, abbreviated 'W', that presents a similar recognition surface as thymidine in the major groove but pairs with A about as strongly as C pairs with G. A phosphoramidite building block was synthesized that allows for incorporation of W residues via automated synthesis in high yield. Melting point increases over duplexes containing T:A pairs of up to 17.5°C, or up to 5.8°C per residue were measured for oligonucleotides containing W. Further, the new base shows excellent fidelity, with a single mismatched G opposite W causing a melting point depression of up to 20.5°C. The strongly pairing replacement for thymidine is only slightly larger than its natural counterpart and performs well in different sequence contexts. It can be used to target weakly pairing A-rich sequences in biological studies.

INTRODUCTION

The most common secondary structure element found in nucleic acid complexes is the double helix. Helix formation is driven by base pairing (1). Base pairs are stabilized by hydrogen bonds and stacking of the aromatic heterocycles. The base pairing rules in duplexes depend on the nucleobases and the sugar/phosphate backbone (2), but in natural DNA, Watson–Crick pairing dominates in duplexes between complementary strands (3). Further, Watson–Crick pairing also underlies replication and the folding of strands into intricate 3D structures (4). The canonical bases found in DNA and RNA (adenine, cytosine, guanine, thymine and uracil) are well adapted to their roles in biology and also allow for rapid formation and dissociation of duplexes, as the charged backbone favors extended structures of well-balanced thermal stability (5), but not ideal for simultaneous detection in sets of diverse sequences.

Among the canonical nucleosides, thymidine is unique, as it is only found in DNA. In the cell, this nucleoside is synthesized enzymatically from deoxyuridine monophosphate in reactions catalyzed by thymidylate synthase, an enzyme with an important role in cancer treatment. The methyl group of thymidine that distinguishes it from deoxynucleosides is then used for molecular recognition by proteins. This includes enzymes critical for genome integrity and repair. When hydrolytic deamination of deoxycytidines leads to deoxyuridines, the most common mutation in DNA (6), the base pairing pattern changes from a base complementary to guanine to a base complementary to adenine. Unless quickly detected and repaired, this can cause mutations during replication. The missing methyl group at the 5-position of deoxyuridines originating from hydrolysis can also lead to a misincorporation of deoxyuridine (7). Thus, the methyl group is an important marker, facilitating recognition by repair enzymes, most notably uracil-DNA glycosylase, which removes uracil from DNA (8).

The base pairs between adenine and its complementary bases have long been known to be weaker than those between guanine and cytosine (9). As a result, duplex stabilities vary greatly between different sequences (10), complicating parallel detection of multiple DNA strands with high fidelity or sequence discrimination (11,12). This has led to a quest for oligonucleotides that give isostable duplexes, i.e. duplexes whose stability is independent of sequence, so that efficient hybridization can be achieved at a fixed, ‘universal’ hybridization temperature (13,14). At this universal temperature, good mismatch discrimination and thus high fidelity would be observed for each probe. The most obvious way to modulate affinity is through modification of the nucleobases. Among the modified nucleobases that bind their target bases with increased affinity are clamp-like base derivatives (15), expanded bases (16), and bases with alkynyl substituents. The latter can improve stacking in duplexes or triplexes (17–19), and are typically introduced via

1To whom correspondence should be addressed. Tel: +49 711 685 64311; Fax: +49 711 685 64321; Email: lehrstuhl-2@oc.uni-stuttgart.de

© The Author(s) 2018. Published by Oxford University Press on behalf of Nucleic Acids Research. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
The pairing of T or U suffers from the fact that adenine displays only two hydrogen bonding sites at its Watson/Crick pairing edge (the amino group at position 6 and the nitrogen at position 1) and that both also pair well with guanine (21, 22), making it difficult to suppress mispairing. This affects experiments, such as Southern or Northern blotting, gene expression profiling with microarrays, or PCR. Given the potential rewards, it is not surprising that derivatives or analogs of thymidine have been synthesized that have the potential to pair more strongly with adenine than T. Figure 1 shows some such structures, together with the UV-melting point difference found for duplexes containing the T analogs, compared to the DNA duplexes featuring a natural T residue. The stabilizing effect of the methyl group at the 5-position of T can be gleaned from the comparison with deoxyuridine. Without the methyl group, the duplex melting point typically drops modestly, as shown for F#, when compared with difluorotoluene (29). This base analog has a strongly destabilizing effect when placed opposite A, but when an ethynyl group replaces the fluorine facing C2 of A, affinity improves, as shown for F8, when compared with difluorotoluene F (30), even though in F8 the methyl group is missing. Reconstitution of the lactame group at positions 3 and 4 of thymine was achieved in the ethynylpyridone found in the nucleoside dubbed "E", significantly improving pairing strength (31).

For the C-nucleoside E, the ethynyl groups replacing the carbonyl oxygen is at the 2-position of thymidine to allow for interactions with the CH fragment at the 2-position of adenine and to strengthen the stacking interactions with neighboring bases in a duplex. But, E lacks the methyl group marker of T and is difficult to incorporate in oligonucleotides (31). Attempts to generate a phosphoramidite for automated DNA synthesis failed. Instead, the nucleoside had to be manually coupled to DNA strands previously phosphitylated on the solid support (31), resulting in modest yields, mostly due to side reactions caused by the poorly shielded lactame group. So, the absence of the methyl group at the 5-position caused both a problematic reactivity and made the base different from T for proteins binding in the major groove. The methyl group was also desirable to increase duplex stability through stacking interactions. Here, we report the synthesis of a C-nucleoside that overcomes these limitations. It features the methyl group at the 5-position of thymine, making it a full replacement for the fifth base. The new nucleobase, dubbed "W", pairs more strongly than T and E and is readily incorporated in oligonucleotides by automated synthesis.

MATERIALS AND METHODS

General

Chemicals and solvents were purchased from Acros Organics (Geel, Belgium), Carbolution (Saarbrücken, Germany), Carbosynth (Compton, Berkshire, Great Britain), Sigma-Aldrich (Deisenhofen, Germany) or TCI (Eschborn, Germany) and were used without further purification. If not otherwise noted, reactions were carried out under argon atmosphere. Thin layer chromatography was performed with Merck Millipore TLC Silica 60 F254 aluminum sheets with visualization by ultraviolet light and staining with phosphormolybdato cerium(IV) sulfate solution (25 g phosphormolybdic acid hydrate, 10 g cerium(IV) sulfate tetrahydrate, 60 mL conc. sulfuric acid, diluted with water to 1 l). Column chromatography was carried out with Merck Millipore silica 60 Dm (particle size: 0.040–0.063 mm). Glycal 8 was synthesized according to a literature protocol (31). Unmodified oligonucleotides were purchased from Biomers (Ulm, Germany) in lyophilized, HPLC-purified form, dissolved in water and used without further purification. Yields of oligonucleotides are based on the isolated amount of oligonucleotide, and the loading of the controlled pore glass support, as given by the supplier. Extinction coefficients for oligonucleotides were calculated through linear combination of the extinction coefficients of nucleotides. The 1H-NMR (300 MHz), 31P-NMR (121 MHz) and 13C-NMR (75 MHz) spectra were measured on a Bruker Avance 300 spectrometer, and 1H-NMR (500 MHz) and 13C-NMR (125 MHz) spectra were measured on a Bruker Avance 500 MHz spectrometer in deuterated solvents. Chemical shifts are reported as δ values (ppm), relative to the solvent peak and coupling constants (J) are given in Hz. Multiplicities were designated as s, singlet; d, doublet; t, triplet; m, multiplet and br, for a broad peak. Infrared spectra were recorded on a Bruker Vector 22FT-IR Spectrometer, and selected peaks are reported in wave numbers (cm⁻¹). High resolution mass spectrometry was performed on all stable new compounds, using a Bruker microTOFQ spectrometer running Bruker Compass Data
Analysis software, version 4.2. MALDI-TOF mass spectra were measured in a linear negative mode on a Bruker microFlex mass spectrometer with a matrix consisting of a 2:1 (v/v) mixture of 2,4,6-trihydroxyacetophenone solution (THAP, 0.2 M in ethanol) and dianium citrate (0.1 M in water). The m/z values are those of the maximum of the unresolved isotope envelope of the peak for the pseudomolecular ion [M−H]−.

Synthesis of C-nucleoside 13

3-Methyl-5-nitro-2-pivaloyloxypyridine (2). 3-Methyl-5-nitro-2-pyridione (1.03 g, 6.68 mmol) was dissolved in dry acetonitrile (35 mL) and NEt3 (1.4 mL, 10.02 mmol, 1.5 eq.) was added. Under ice bath cooling, pivaloyl chloride (1.2 mL, 10.02 mmol, 1.5 eq.) was added. The reaction mixture was stirred for 15 min at 0°C. After TCL showed full conversion, a saturated solution of NaHCO3 (5 mL) and ethyl acetate (10 mL) were added. The organic layer was separated, and the aqueous phase was extracted twice with ethyl acetate (2×10 mL). The filtrate was concentrated in vacuo. The resulting crude product was used in the next step without further purification. An analytical sample was purified by chromatography, using silica (30 g), eluting with petroleum ether/ethyl acetate (5/1, v/v). TLC (petroleum ether/ethyl acetate, 1/1, v/v): Rf = 0.79; 1H-NMR (500 MHz, CDCl3): δ = 9.07 (d, J = 2.6 Hz, 1H), 8.39 (d, J = 2.3 Hz, 1H), 2.31 (s, 3H), 1.41 (s, 9H); 13C-NMR (125 MHz, CDCl3): δ = 175.6, 147.2, 140.1, 126.4, 125.6, 122.6, 39.1, 27.0, 16.1; HRMS (ESI-TOF) m/z calcd. for C11H13BrN2O2 [M+Na]+ 311.0846; found 311.0816.

5-Amino-3-methyl-2-pivaloyloxypyridine (3). In a pressure-stable reaction vessel, 3-methyl-5-nitro-2-pivaloyloxypyridine (2, 1.59 g, 6.68 mmol) was dissolved in methanol (30 mL) and palladium on charcoal (179 mg, 10% wt, dry) was added. Under a H2 atmosphere (2.5 bar), the reaction mixture was shaken for 2 h at room temperature. The mixture was filtered over Celite and washed with ethyl acetate (2×20 mL). The filtrate was concentrated in vacuo. The resulting crude product (3) was used in the next step without further purification. An analytical sample was purified by chromatography, using silica (30 g), eluting with petroleum ether/ethyl acetate (3/1, v/v). TLC (petroleum ether/ethyl acetate, 1/1, v/v): Rf = 0.14; 1H-NMR (500 MHz, CDCl3): δ = 7.66 (d, J = 2.8 Hz, 1H), 6.90 (d, J = 2.7 Hz, 1H), 3.59 (br s, 2H), 2.08 (s, 3H), 1.37 (s, 9H); 13C-NMR (125 MHz, CDCl3): δ = 176.9, 149.5, 141.5, 132.4, 126.7, 125.4, 39.1, 27.2, 15.9; HRMS (ESI-TOF) m/z calcd. for C11H14N2O4 [M+Na]+ 261.0846; found 261.0816.

3-Amino-2-bromo-5-methyl-6-pivaloyloxypyridine (4). Amine (3) (1.39 g, 6.68 mmol) was dissolved in acetonitrile (20 mL), and the solution was cooled to 0°C. A solution of N-bromosuccinimide (1.19 g, 6.68 mmol, 1 eq.) in acetonitrile (10 mL) was added dropwise within 10 min, while stirring at 0°C. After TLC showed full conversion, water (15 mL) was added, and the mixture was extracted twice with ethyl acetate (2×20 mL). The combined organic layers were washed with brine (10 mL), dried over Na2SO4 and concentrated in vacuo. The resulting crude product was purified by chromatography, using silica (30 g) and a gradient of 10–30% ethyl acetate in petroleum ether, yielding 4 as a dark brown solid (1.42 g (4.94 mmol, 74% over three steps). TCL (petroleum ether/ethyl acetate, 1/1, v/v): Rf = 0.66; 1H-NMR (500 MHz, CDCl3): δ = 6.93 (s, 1H), 2.04 (s, 3H), 1.36 (s, 9H); 13C-NMR (125 MHz, CDCl3): δ = 176.6, 147.2, 140.1, 126.4, 125.6, 122.6, 39.1, 27.2, 15.9; HRMS (ESI-TOF) m/z calcd. for C11H15BrNO2 [M+Na]+ 419.9029; found 419.9022.

2-Bromo-5-methyl-4-pivaloyloxypyridin-3-yl diazonium tetrafluoroborate (5). An aliquot of BF3·OEt2 (1 mL, 7.94 mmol, 1.5 eq.) was added to a solution of amine (4) (1.52 g, 5.29 mmol) in dry THF (5 mL) at −10°C. The reaction mixture was stirred for 2 min, then tert-butyl nitrite (800 μL, 6.35 mmol, 1.2 eq.) was added. After 10 min, the precipitated solid was filtered off and washed three times with cold diethyl ether (3×10 mL). Diazonium salt 5 was obtained as a light yellow solid, 1.93 g (4.97 mmol, 94%). The product is unstable and was not studied by mass spectrometry for this reason. 1H-NMR (500 MHz, CD2CN): δ = 7.99 (s, 1H), 2.16 (s, 3H), 1.16 (s, 9H); 13C-NMR (125 MHz, CD2CN): δ = 179.0, 163.9, 143.6, 134.3, 128.0, 37.7, 26.2, 14.9; IR: ν = 3088, 2973, 2285, 2218 (R-N+, s), 1698, 1416, 1039 cm−1.

2-Bromo-3-iodo-5-methyl-6-pivaloyloxypyridine (6). To a solution of diazonium salt 5 (812 mg, 2.09 mmol) in dry acetonitrile (20 mL), potassium iodide (417 mg, 2.51 mmol, 1.2 eq.) was added in one portion. The reaction mixture was stirred at room temperature for 2 h. After TCL showed full conversion, water (10 mL) was added. The mixture was extracted twice with ethyl acetate (2×20 mL). The combined organic layers were dried over Na2SO4 and then concentrated in vacuo. The resulting crude product was purified by chromatography, using silica (30 g) and a gradient of 20–50% ethyl acetate in petroleum ether, yielding 6 as an orange solid (637 mg, 1.60 mmol, 76%). TLC (petroleum ether/ethyl acetate, 1/1, v/v): Rf = 0.80; 1H-NMR (500 MHz, CDCl3): δ = 7.99 (s, 1H), 2.12 (s, 3H), 1.41 (s, 9H); 13C-NMR (125 MHz, CDCl3): δ = 175.6, 155.9, 151.6, 142.7, 126.5, 95.9, 39.3, 27.1, 15.1; HRMS (ESI-TOF) m/z calcd. for C11H13BrI2O2 [M+Na]+ 490.9067; found 491.9073.

2-Bromo-5-methyl-6-pivaloyloxy-3′-O-[(tert-butyl dimethyl)silyl]-2′,3′-didehydro-β-D-ribofuranos-1′-yl-pyridine (8). Samples of Pd(OAc)2 (124 mg, 0.55 mmol, 0.2 eq) and P(Ph3)3 (590 mg, 1.11 mmol, 0.4 eq) were dissolved in dry acetonitrile (10 mL) and stirred at room temperature for 30 min. The solution was then added to a stirred mixture of 2-bromo-3-iodo-5-methyl-6-pivaloyloxypyridine (6, 1.54 g, 3.88 mmol, 1.4 eq), Ag2CO3 (764 mg, 2.77 mmol, 1 eq) and glycal 7 (31) (638 mg, 2.77 mmol, 1 eq) in dry acetonitrile (10 mL). The reaction mixture was stirred at room temperature for 5 h. After TCL showed full conversion, the reaction mixture was filtered through celite, eluting with ethyl acetate (20 mL), and the solvent was then removed in vacuo.
vacuo. The resulting crude product (8) was used in the subsequent step without further purification. An analytical sample was purified by chromatography, using silica and a gradient of 0–1% methanol in dichloromethane, yielding 8 as an off-white solid. TLC (petroleum ether/ethyl acetate, 1/1, v/v): Rf = 0.86; 1H-NMR (500 MHz, CDCl3): δ = 7.87 (s, 1H), 6.03 (dd, J = 4.1, 1.3, 1H), 4.9 (s, 1H), 4.69 (br t, J = 2.3, 1H), 3.85 (ddd J = 21.0, J = 12.3, J = 2.9, 2H), 2.2 (s, 3H), 1.41 (s, 9H), 0.97 (s, 9H), 0.27 (s, 3H); 13C-NMR (125 MHz, CDCl3): δ = 175.9, 155.4, 151.1, 142.2, 137.2, 136.0, 125.3, 100.5, 83.6, 82.2, 63.1, 39.2, 27.1, 25.5, 15.6, 4.9, 5.0; ESI-MS m/z calcd. for C22H34BrNO5Si [M+Na]+ 522.13; found 522.13.

2-Bromo-5-methyl-6-pivaloyloxy-3-(2′,3′-didehydro-2′,3′-dideoxy-3′-oxo-β-D-ribofuranos-1′-yl)pyridine (9). In a polypropylene tube, crude 8 (1.39 g, 2.77 mmol) was dissolved in THF (30 mL) and 3 HF·NET3 (900 µL, 5.54 mmol, 2 eq) was added. After 30 min, TLC showed full conversion. To quench remaining HF, methoxytrimethylsilane (2 mL) was added, and the mixture was stirred for an additional 30 min. The mixture was then filtered through celite, eluting with ethyl acetate (20 mL). The filtrate was concentrated in vacuo. The resulting crude 9 was used in the next step without purification. An analytical sample was purified by chromatography, using silica and a gradient of methanol (0–2%) in dichloromethane, yielding 9 as an off-white solid. TLC (petroleum ether/ethyl acetate, 1/1, v/v): Rf = 0.50; 1H-NMR (500 MHz, CDCl3): δ = 7.71 (s, 1H), 5.18 (dd, J = 10.7, J = 5.9, 1H), 3.88 (t, J = 3.6, 1H, 3.79 (m, 2H), 2.96 (dd, J = 17.8, J = 6.3, 1H), 2.05 (dd, J = 18.3, J = 10.3, 1H), 1.96 (s, 3H), 1.18 (s, 9H); 13C-NMR (125 MHz, CDCl3): δ = 212.7, 175.9, 155.6, 139.6, 135.8, 135.1, 125.7, 82.3, 75.1, 61.5, 44.1, 39.3, 27.1, 15.5; HRMS (ESI-TOF) m/z calcd. for C16H15BrNO5 [M+Na]+ 408.0417; found 408.0423.

3-Methyl-2-pivaloyloxy-5′-[2′-deoxy-5′-O-(dimethoxytrityl)-β-D-ribofuranos-1′-yl]-6-(trisopropylsilylethynyl)-pyridine (11). Nucleoside 11 (230 mg, 0.46 mmol) was coevaporated twice from pyridine and dissolved in dry pyridine (5 mL). Then, DMAP (3 mg, 0.02 mmol, 0.05 eq.) was added, and the reaction mixture was stirred for 30 min. Subsequently, DMT-CI that had been coevaporated twice from dry pyridine (204 mg, 0.60 mmol, 1.3 eq.), dissolved in pyridine (2 mL), was added. The reaction mixture was stirred for 18 h at room temperature. After TLC showed full conversion, volatiles were removed in vacuo and the crude was purified by chromatography, using silica (20 g, deactivated with dichloromethane containing 0.5% NEt3 prior to use) and a gradient of methanol (0–2%) in dichloromethane to afford the desired nucleoside (12) as a colorless solid (283 mg, 0.36 mmol, 77%), TLC (dichloromethane/methanol, 99/1, v/v): Rf = 0.74; 1H-NMR (500 MHz, CD3CN): δ = 7.97 (s, 1H), 7.51–7.24 (m, 9H), 6.90 (d, J = 9.6, 4H), 5.55 (dd, J = 9.9, J = 5.9, 1H), 4.36–4.32 (m, 1H), 4.04–4.00 (m, 1H), 3.79 (s, 6H), 3.32–3.25 (m, 2H), 2.46 (dd, J = 12.9, J = 5.8, J = 2.3, 1H), 2.03 (s, 3H), 1.95 (dd, J = 9.4, J = 5.3, J = 3.7, 1H), 1.39 (s, 9H), 1.18 (s, 2H); 13C-NMR (125 MHz, CD3CN): δ = 175.8, 158.2, 155.1, 144.6, 140.1, 137.6, 135.5, 135.5, 129.5, 129.4, 127.6, 127.3, 126.3, 126.2, 112.6, 102.2, 95.5, 85.9, 85.6, 75.9, 72.7, 63.7, 54.4, 63.7, 54.4, 42.3, 38.4, 25.8, 25.8, 17.6, 17.3, 14.8, 10.5, 10.4, 0.5, 0.3, 0.2, 0.1; HRMS (ESI-TOF) m/z calcd. for C48H40NO14Si [M+Na]+ 814.4110; found 814.4117.

3-Methyl-2-pivaloyloxy-5′-[2′-deoxy-3′-O-(2-cyanoethyl-N,N-diisopropylamino)-phosphino-5′-O-(dimethoxytrityl)-β-D-ribofuranos-1′-yl]-6-(trisopropylsilylethynyl)-pyridine (13). Nucleoside 12 (705 mg, 0.89 mmol) was coevaporated twice from acetonitrile, and then dissolved in dry acetonitrile (10 mL). Diisopropylammonium tetrazolide (26) (107 mg, 0.62 mmol, 0.7 eq.) was added, and the reaction mixture was stirred for 30 min. Then,
2-cyanoethyl-\(N,N\)-diisopropylaminochlorophosphite (340 \(\mu\)L, 1.07 mmol, 1.2 eq.) was added. The reaction mixture was stirred for 4 h at room temperature. After TLC showed full conversion, a saturated aqueous solution of NaHCO\(_3\) (10 mL) was added. The aqueous phase was extracted twice with methyl tert-butyler ether (2 \(\times\) 20 mL). The combined organic layers were dried over Na\(_2\)SO\(_4\), concentrated in \(\text{vacuo}\), and the resulting crude product was purified by chromatography, using silica (previously deactivated with dichloromethane containing 1% \(\text{NEt}_3\)) and methyl tert-butyler ether containing \(\text{NEt}_3\) (1%) to afford the desired nucleoside 13 as a colorless solid (778 mg, 0.78 mmol, 88%). The phosphoramidite was stored in dry CH\(_3\)CN in 43 min, product detected at 37.0 min; yield: 60 nmol, 6%; MALDI-TOF-MS: \(m/z\) calcd. for C\(_{69}\)H\(_{99}\)N\(_{17}\)O\(_{46}\)P\(_6\) [M–H]\(^{−}\) 2012, found 2012.

**RESULTS AND DISCUSSION**

Figure 2 shows the structure of the C-nucleoside W and its base pair with A in a duplex. The hydrogen at position 2 of adenine is shown to highlight the spatial proximity to the ethynyl substituent.

HPLC purification

Reversed-phase HPLC was carried out on a HPLC system with a type L-6200 Intelligent Pump and a photodiode detector L-3000 from Hitachi (Tokyo, Japan). For this, lyophilized oligonucleotides were dissolved in water (1 mL), the solution was filtered (pore size: 0.45 \(\mu\)m) and subjected to reversed-phase HPLC, using a Nucleosil C18 column (250 \(\times\) 4.6 mm, Macherey-Nagel, Düren, Germany) with a gradient of acetonitrile in triethylammonium acetate buffer (0.1 M, pH 7) and detection at 260 nm. Two different gradient protocols were used for oligonucleotides 14–16 (1–15% CH\(_3\)CN in 43 min) and oligonucleotides 17–19 (1–25% CH\(_3\)CN in 50 min).

Analytical data for modified oligonucleotides

**5′-CWGCAG-3′ (14).** HPLC: gradient of 1–15% CH\(_3\)CN in 43 min, product detected at \(t_\text{R} = \) 32.5 min; yield: 50 nmol, 5%; MALDI-TOF-MS: \(m/z\) calcd.

for C\(_6\)H\(_{25}\)N\(_{23}\)O\(_{33}\)P\(_3\) [M–H]\(^{−}\) 1798, found 1798.

**5′-CCWCTT-3′ (15).** HPLC: gradient of CH\(_3\)CN, 1–15% in 43 min, product detected at \(t_\text{R} = \) 32.5 min; yield: 390 nmol, 39%; MALDI-TOF-MS: \(m/z\) calcd. for C\(_{69}\)H\(_{99}\)N\(_{17}\)O\(_{46}\)P\(_6\) [M–H]\(^{−}\) 2012, found 2012.

**5′-TGGWWGAC-3′ (16).** HPLC: gradient of CH\(_3\)CN, 1–15% in 43 min, product detected at \(t_\text{R} = \) 37.0 min; yield: 60 nmol, 6%; MALDI-TOF-MS: \(m/z\) calcd. for C\(_{85}\)H\(_{103}\)N\(_{27}\)O\(_{46}\)P\(_7\) [M–H]\(^{−}\) 2454, found 2455.

**5′-WWGTCAWWG-3′ (17).** HPLC: gradient of CH\(_3\)CN, 1–25% in 50 min, product detected at \(t_\text{R} = \) 37.5 min; yield: 50 nmol, 5%; MALDI-TOF-MS: \(m/z\) calcd. for C\(_{100}\)H\(_{117}\)N\(_{22}\)O\(_{52}\)P\(_8\) [M–H]\(^{−}\) 2705, found 2705.

**5′-GTCCWWWWWGC-3′ (18).** HPLC: gradient of CH\(_3\)CN, 1–25% in 50 min, product detected at \(t_\text{R} = \) 38.0 min; yield: 20 nmol, 2%; MALDI-TOF-MS: \(m/z\) calcd. for C\(_{120}\)H\(_{157}\)N\(_{26}\)O\(_{78}\)P\(_{11}\) [M–H]\(^{−}\) 3350, found 3348.

**5′-CTTCTTCTTCTT-3′ (19).** HPLC: gradient of CH\(_3\)CN 1–25% in 50 min, product detected at \(t_\text{R} = \) 38.0 min; yield: 330 nmol, 33%; MALDI-TOF-MS: \(m/z\) calcd. for C\(_{120}\)H\(_{157}\)N\(_{26}\)O\(_{78}\)P\(_{11}\) [M–H]\(^{−}\) 3548.
variable amounts of ketone enol ether base concentration chosen, partial loss of the silyl group of ether following the route of Minuth and Richert (31). At the 

was synthesized in four steps, with an overall yield of 54%, et al. use the recently reported coupling conditions of Hocek withNaBH(OAc)3 in the cold, furnishing 

pooled. The ketone was then diastereoselectively reduced with hydrogen on Pd/C, using pivaloyl chloride in acetonitrile. Subsequent reduction with aqueous ammonia and then TBAF in THF, i.e. tert HPLC purification, oligonucleotide 

but precluding the preparation of sufficient quantities of the nucleoside building block for routine DNA synthesis. Scheme 1 shows the successful synthesis of phosphoramidite 13. It starts with methylnitropyridone 1 (34), which is commercially available for less than $10/g. This compound was selectively O-acylated to 2, using pivaloyl chloride in acetonitrile. Subsequent reduction with hydrogen on Pd/C gave aniline 3, which was brominated with NBS in the cold to give 4 in 74% yield over three steps. Diazotization of the brominated aniline produced 5, which readily precipitated from the reaction mixture under optimized reaction conditions. When 5 was used directly in the Heck reaction with glycal 7, the low yields known from similar routes were obtained (31). Even after careful optimization, the yield of 10 never exceeded 13% over three steps. An important step was therefore to switch to iodide 6, accessible from 5 in 76% yield, and to use the recently reported coupling conditions of Hoek et al. (35), which led to high-yielding reactions. Glycal 7 was synthesized in four steps, with an overall yield of 54%, following the route of Minuth and Richert (31). At the base concentration chosen, partial loss of the silyl group of enol ether 8 is common during Heck reactions, producing variable amounts of ketone 9, which is readily separated from 8 by column chromatography. Remaining silyl enol ether 8 was converted to 9, and the two fractions of 9 were pooled. The ketone was then diastereoselectively reduced with NaBH(OAc)3 in the cold, furnishing 10 in 68% overall yield from 6 on a gram scale. The benzyl derivative of 10 crystallized readily, and an X-ray crystal structure confirmed its stereochemical configuration (see Figure S45, Supplementary Data). For the subsequent Sonogashira reaction to 11, we used the less volatile and more stable trisopropylsilylacetylene to complete the elaboration of the carbon framework of the C-nucleoside. Tritylation gave 12, and was followed by phosphitylation, which then led to 13 in 12% yield over 15 steps. Batches of up to 700 mg of the phosphoramidite were readily prepared. Access to 13 allowed the exploration of the pairing properties of the ethynylmethylpyridine in different duplexes. We chose the sequences shown in Figure 3. They included selfcomplementary hexamer 14, and nonselfcomplementary sequences 15–19 with one, two, four, or six W residues. Further, dodecamer 19 with a single T-to-W replacement was prepared to compare the UV-melting point of its duplex with those of other C-nucleosides incorporated in the same sequence (29–31).

Figure 4 shows the HPLC chromatogram of crude 19, prepared by automated DNA synthesis on controlled-pore glass, followed by two-step deprotection, consisting of treatment with aqueous ammonia and then TBAF in THF; i.e. conditions similar to those used in RNA syntheses (36). After HPLC purification, oligonucleotide 19 was obtained in 33% yield. While heavily modified sequences often give low yields in oligonucleotide synthesis, even for undecamer 18 with its six neighboring W residues, the product peak was the most intense peak in the chromatogram of the crude product and the pure product was obtained in 2% overall yield, confirming that 13 is well-behaved in automated DNA synthesis.

With the W-containing strands in hand, we proceeded to studying duplex stabilities by UV-melting analysis. The melting point (Tm) of the duplex of dodecamer 19 with its target strand 3'-GAAAAGAAAGAA-5' (20) was 4.4°C higher than that of the duplex featuring thymine at the position of W (Table 1, Figure 5A). The melting point of the duplex with the W:A base pair was also 1.2°C higher than that of the duplex with E at the same position (31) (Table S1, Supplementary Data). When a C:G base pair was placed at this position, the duplex had almost the identical thermal
stability as that with the W:A pair (44.5 versus 44.6°C, Table 1). When RNA target strand \(r(3'-\text{GAAAAGAAAAGAA-5'})\) (24) was used as complementary strand, the W for T replacement still gave a \(\Delta T_m\) of +2.9°C. Finally, the stabilizing effect of W residues in a short duplex was measured, which was found to be a \(\Delta T_m\) of +11.5°C (+5.7°C per nucleotide) for self-complementary hexamer 14, compared to control duplex formed by 25 (Table 1). Exchanging the W:A base pair with C:G base pairs in the self-complementary hexamer 25 to give 14 again confirmed the similarity in base pairing strength of the W:A and the C:G base pairs, with duplex melting points of 42.7 versus 41.0°C (Table 1).

Next, we determined the base pairing selectivity of W in melting curves with dodecamer 19 and its unmodified version 21 (Table 2). A melting point decrease of 20.5°C was found for a single mismatched G opposite W (Table 2), indicating excellent discrimination against wobble pairing. The \(\Delta T_m\) is larger than that measured for E (19.8°C; 31) and much larger than that reported for 2-thiothymidine (12.0°C, 23). Further, not a single one of the different mismatches tested gave a \(\Delta T_m\) of less than 14°C, confirming that W high fidelity in a broad range of contexts. In contrast, T-containing control strand 21 showed the expected low selectivity for A:T versus G:T pairs, with a melting point decrease of 9.5°C in the latter case. To the best of our knowledge W is the most selective replacement for T known to date.

In a last set of experiments, the ability of W to pair with adenine was studied in a more diverse set of sequence contexts. This included oligonucleotides 15–18 forming duplexes with 31, 35, 39 or 43 (Table 3). The UV-melting point increase, relative to the duplex with natural T:A pairs, was between +5.8°C per W residue for the shortest duplex (15:31) and +2.9°C per residue for the longest duplex, featuring six W residues, (18:43, overall increase of 17.5°C). Figure 5B shows that even for this six-fold modified duplex, a sharp sigmoidal transition is found. This confirms highly cooperative melting, even with multiple neighboring W, a trait that is typical for natural DNA and not often found for heavily modified strands. Taken together, the melting point data confirmed that W is strongly stabilizing and well behaved in DNA duplexes, even when multiple modified bases are placed in a sequence.

In order to better explain the high cooperativity in duplex formation, we then studied the steric and electrostatic profile of W, as compared to T, its natural counterpart. For this, the structures of the free bases were calculated ab initio on the B3LYP/def2-SVP level in TURBOMOLE, version 7.2 (see: http://www.turbomole.com). Figure 6A shows electrostatic maps of the two bases. The settings for the graphical representation are similar to those used for nucleobases in the past (37). It can be discerned that the electron density at the hydrogen bonding sites and the methyl group are near-identical for T and W. Only at the 1-, 2- and 6-positions...
(pyrimidine numbering) is the C-nucleoside less polar than T. The lipophilicity at positions 1 and 6 should be of little consequence for base pairing, as these positions are directed toward the deoxyribose ring, not any pairing partner. The ethynyl group at position 2 is well suited to interact with the lower Watson–Crick face of adenine, with a likely van der Waals contact to the CH fragment at position 2 of the purine, as shown in Figure 6B.

The shape of the ethynyl group of W also explains the exquisite selectivity for A versus G. While the alkynyl substituent provides shape complementarity to A, it induces a steric conflict with G, when a similar hydrogen bonding geometry is attempted as in a T:G or U:G wobble base pair (Figure 6C). Apparently, W combines a size and hydrogen bonding capability similar to that of T, so that it can be integrated seamlessly into DNA duplexes, with improved shape complementarity and stacking capabilities, making it well suited for oligonucleotides designed to bind A-containing target strands.

Table 1. UV-melting points of DNA:DNA or DNA:RNA duplexes

| Sequences[1] | Table 3. UV-melting points of DNA:DNA duplexes with one or several W residues together with data from control duplexes
|----------------|---------------------------|
| Table 1. UV-melting points of DNA:DNA or DNA:RNA duplexes | Sequences[8] | \( T_{m}^{b} \) [°C] | \( \Delta T_{m}^{b} \) [°C] |
| 3′-GAAAAGAAGA-S′-G (20) | 3′-GGGGGAA-S′ (33) | 40.2 | 4.3 |
| 5′-CTTTCTTCTTCT-3′ (21) | 5′-CCCTCTT-3′ (32) | 44.6 | 4.4 |
| 3′-GAAAAGAAGA-S′-G (22) | 5′-CCCTCTT-3′ (15) | 44.5 | 2.9 |
| 5′-CTTTCTTCTTCT-3′ (23) | 5′-CACCAGT-3′ (35) | 41.4 | 2.9 |
| 5′-CTTTCTTCTTCT-3′ (19) | 5′-TGGTTGAC-3′ (36) | 44.5 | 11.5 |
| (5′-CTGCAG-3′) (25) | 5′-TTGGWGC-3′ (16) | 37.3 | 5.8 |
| (5′-CWGCAG-3′) (14) | 5′-ACCGGCCTG-5′ (37) | 41.0 | 9.8 |
| (5′-CCGCGG-3′) (26) | 5′-TGGGCAC-3′ (38) | 48.2 | 11.9 |

\[ \text{Table 2. UV-melting points of DNA duplexes with a matched or mismatched nucleobase opposite a T or a W residue} \]

| Sequences[9] | Pairing[10] | \( T_{m}^{b} \) [°C] | \( \Delta T_{m}^{b} \) [°C] |
|----------------|----------------|----------------|----------------|
| 5′-CTTTCTTCTTCT-3′ (21) | T:A | 40.2 | - |
| 3′-GAAAAGAAGA-S′-G (27) | T:T | 25.5 | -14.7 |
| 3′-GAAAAGAAGA-S′-G (28) | T:C | 26.1 | -14.1 |
| 3′-GAAAAGAAGA-S′-G (29) | T:G | 30.7 | -9.5 |
| 5′-CTTTCTTCTTCT-3′ (19) | W:A | 44.6 | - |
| 3′-GAAAAGAAGA-S′-G (27) | W:T | 30.4 | -14.2 |
| 3′-GAAAAGAAGA-S′-G (28) | W:C | 26.7 | -17.9 |
| 3′-GAAAAGAAGA-S′-G (30) | W:G | 24.1 | -20.5 |

| Table 2. UV-melting points of DNA duplexes with a matched or mismatched nucleobase opposite a T or a W residue | Sequences[8] | \( T_{m}^{b} \) [°C] | \( \Delta T_{m}^{b} \) [°C] |
|----------------|----------------|----------------|----------------|
| 5′-CTTTCTTCTTCT-3′ (21) | 3′-GGGGGAA-S′ (33) | 40.2 | 4.3 |
| 3′-GAAAAGAAGA-S′-G (20) | 5′-CCCTCTT-3′ (32) | 44.6 | 4.4 |
| 5′-CTTTCTTCTTCT-3′ (23) | 5′-CACCAGT-3′ (35) | 44.5 | 2.9 |
| 5′-CTTTCTTCTTCT-3′ (19) | 5′-TGGTTGAC-3′ (36) | 44.5 | 11.5 |
| (5′-CTGCAG-3′) (25) | 5′-TTGGWGC-3′ (16) | 37.3 | 5.8 |
| (5′-CWGCAG-3′) (14) | 5′-ACCGGCCTG-5′ (37) | 41.0 | 9.8 |
| (5′-CCGCGG-3′) (26) | 5′-TGGGCAC-3′ (38) | 48.2 | 11.9 |

[1] Target strand listed first, probe strands below, with indentation; self-complementary sequences listed with a double indentation. Bases at the position varied are in boldface.
[10] Average of four curves, detected at 260 nm; conditions: 5.6 ± 0.7 μM strands, 10 mM PIPES buffer, pH 7.0, 100 mM NaCl and 10 mM MgCl2.

\[ \text{Table 3. UV-melting points of DNA:DNA duplexes with one or several W residues together with data from control duplexes} \]

| Sequences[8] | \( T_{m}^{b} \) [°C] | \( \Delta T_{m}^{b} \) [°C] |
|----------------|----------------|----------------|
| 3′-GGGGGAA-S′ (33) | 26.5 | - |
| 5′-CCCTCTT-3′ (32) | 32.3 | 5.8 |
| 5′-CACCAGT-3′ (35) | 33.6 | 7.1 |
| 5′-TGGTTGAC-3′ (36) | 36.3 | - |
| 5′-TTGGWGC-3′ (16) | 45.3 | 9.0 |
| 5′-ACCGGCCTG-5′ (37) | 46.5 | - |
| 5′-TGGGCAC-3′ (38) | 64.0 | 17.5 |

[8] Target strand listed first, probe strands with an indentation. The bases of interest in boldface.
[9] Average of four curves, detected at 260 nm. Conditions: strand concentration 5.6 ± 0.7 μM, 10 mM PIPES buffer, pH 7.0, 100 mM NaCl and 10 mM MgCl2.

A Target strand listed first, probe strands with an indentation. The bases of interest in boldface.
CONCLUSIONS

In conclusion, we describe a new C-nucleoside analog of thymidine that pairs as strongly with adenine as cytosine pairs with guanine. This analog, abbreviated W, can be readily incorporated in oligonucleotides using phosphoramidite and automated chain assembly. With W replacing T in their sequence, oligodeoxynucleotides form duplexes with excellent base pairing selectivity. Wobble base pairing with G is suppressed, with UV-melting point depression of more than 20° C for a single W:G mismatch in the sequence context studied here. The consistently high affinity for adenine in complementary strands and the predictable pairing properties of W make it promising for applications ranging from high fidelity microarrays to functional DNA nanostructures with improved nucleation during folding, capture agents for RNAs, or high-affinity primers. Studies aimed at using W for reading out during folding, capture agents for RNAs, or high-functional DNA nanostructures with improved nucleation constants, it will be interesting to ask whether ‘W is better than T’ in a biological context, and if so, what the consequences were, if nature had access to improved binders for adenine.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Dr B. Miehlich for ab initio calculations, D. Göhringer for expert technical assistance, Dr M. Minuth for discussions, and Nicholas Birchall and Helmut Grieser for a review of the manuscript.

FUNDING

DFG [RI 1063/15-1]; University of Stuttgart. Funding for open access charge: University of Stuttgart. Conflict of interest statement. None declared.

REFERENCES

1. Saenger, W. (1984) Principles of Nucleic Acid Structure. Springer Verlag, NY.
2. Beier, M., Reck, F., Wagner, T., Krishnamurthy, R., and Eschenmoser, A. (1999) Chemical etiology of nucleic acid structure: comparing pentopyranosyl (2′-4′) oligonucleotides with RNA. Science, 283, 693–703.
3. Borer, P.N., Dengler, B., Tinoco, J., and Uhlenbeck, O.C. (1974) Stability of ribonucleic acid double-stranded helices. J. Mol. Biol., 86, 843–853.
4. Egli, M. (1997) In vitro selected receptors rationalized: the first 3D structures of RNA aptamer/substrate complexes. Angew. Chem. Int. Ed. Engl., 36, 480–482.
5. Richert, C., Roughton, A.L., and Benner, S.A. (1996) Non-ionic analogs of RNA with dimethylenesulfonyl bridges. J. Am. Chem. Soc., 118, 4518–4531.
6. Lindahl, T. (1994) Instability and decay of primary structure of DNA. Nature, 362, 709–715
7. Tye, B.K., Nyman, P.O., Lehman, I.R., Hochhauser, S. and Weiss, B. (1977) Transient accumulation of Okazaki fragments as a result of uracil incorporation into nascent DNA. Proc. Natl. Acad. Sci. U.S.A., 74, 154–157
8. Barnes, D.E., Lindahl, T. and Sedgwick, B. (1993) DNA repair. Curr. Opin. Cell Biol., 5, 424–433.
9. Pranata, J., Wierschke, S.G. and Jorgensen, W.L. (1991) OPLS potential functions for nucleotide bases - relative association constants of hydrogen-bonded base pairs in chloroform. J. Am. Chem. Soc., 113, 2810–2819.
10. SantaLucia, Jr and Hicks, D. (2004) The thermodynamics of DNA structural motifs. Annu. Rev. Biophys. Biomol. Struct., 33, 415–440.
11. Kennedy, G.C., Hatusuzaki, H., Dong, S., Liu, W.M., Huang, J., Liu, G., Su, X., Cao, M., Chen, W., Zhang, J. et al. (2003) Large-scale genotyping of complex DNA. Nat. Biotechnol., 21, 1233–1237.
12. Matsuuzaki, H., Lai, H., Dong, S., Tsai, Y., Fang, J., Law, J., Di, X., Liu, W., Yang, G., Liu, G. et al. (2004) Parallel genotyping of over 10,000 SNPs using a One-Primer assay on a High-Density oligonucleotide array. Genome Res., 14, 20–28.
13. Nguyen, H.-K., Fournier, O., Asseline, U., Dupret, D. and Thuogn, N.T. (1999) Smoothing the thermal stability of DNA duplexes by using modified nucleosides and chaotropic agents. Nucleic Acids Res., 27, 1492–1498.
14. Ahlborn, C., Siegmund, K. and Richert, C. (2007) Isostable DNA. J. Am. Chem. Soc., 129, 15218–15232.
15. Lin, K.-Y. and Matteucci, M.D. (1998) A cytosine analogue capable of clamp-like binding to a guanine in helical nucleic acids. J. Am. Chem. Soc., 120, 8531–8532.
16. Okamoto, A., Tanaka, K. and Saito, I. (2003) Clear distinction of purine bases on the complementary strand by a fluorescence change of a novel fluorescent nucleoside. J. Am. Chem. Soc., 125, 4972–4973.
17. Wagner, R.W., Matteucci, M.D., Lewis, J.G., Gutierrez, A.J., Moulds, C. and Froehler, B.C. (1993) Antisense gene inhibition by oligonucleotides containing C-5 propyne pyrimidines. Science, 260, 1510–1513.
18. Chaudhuri, N.C. and Kool, E.T. (1995) Very high-affinity DNA recognition by bicyclic and cross-linked oligonucleotides. J. Am. Chem. Soc., 117, 10434–10442.
19. Wagner, R.W., Matteucci, M.D., Grant, D., Huang, T. and Froehler, B.C. (1996) Potent and selective inhibition of gene expression by an antisense heptanucleotide. Nat. Biotechnol., 14, 840–844.
20. Kottisch, T., Ahlborn, C., Brotzel, F. and Richert, C. (2004) Stabilizing or destabilizing oligodeoxynucleotide duplexes containing single 2′-deoxyuridine residues with 5-alkynyl substituents. Chem. Eur. J., 10, 4017–4028.
21. Crick, F.H.C. (1966) Codon—anticodon pairing: the wobble hypothesis. J. Mol. Biol., 19, 548–555.
22. Allawi, H.T. and SantaLucia, Jr (1997) Thermodynamics and NMR of internal G.T mismatches in DNA. Biochemistry, 36, 10581–10594.
23. Sintim, H.O. and Kool, E.T. (2006) Enhanced base pairing and replication efficiency of thiothymidines, expanded-size variants of thymidine. J. Am. Chem. Soc., 128, 396–397.
24. Hassan, A.E.A., Sheng, J., Zhang, W. and Huang, Z. (2010) High fidelity of base pairing by 2-selenothymines, expanded-size variants of thymidine. J. Am. Chem. Soc., 132, 2120–2212.
25. Wagner, R.W., Matteucci, M.D., Lewis, J.G., Gutierrez, A.J., Moulds, C. and Froehler, B. C. (1993) Antisense gene inhibition by oligonucleotides containing C-5 propyne pyrimidines. Science, 260, 1510–1513.
26. Chapat, J.C., Sinha, S. and Switzer, C. (2002) 5-Propynyluracil.diaminopurine: an efficient base-pair for non-enzymatic transcription of DNA. Chem. Commun., 1568–1569.
27. Gyi, J.I., Gao, D., Conn, G.L., Trent, J.O., Brown, T. and Lane, A.N. (2003) Solution structure of a DNA-RNA duplex containing 5-propynyl U and C; comparison with 5-Me modifications. Nucleic Acids Res., 31, 2683–2693.
28. Barnes, T.W. and Turner, D.H. (2001) Long-range cooperativity in molecular recognition of RNA by oligodeoxynucleotides with multiple C5-(1-propynyl) pyrimidines. J. Am. Chem. Soc., 123, 4017–4118.
29. Moran, S., Ren, R.X.-F., Rumney, S. and Kool, E.T. (1997) Difluorotoluene, a nonpolar isostere for thymine, codes specifically and efficiently for adenine in DNA replication. J. Am. Chem. Soc., 119, 2056–2057.
30. Griesang, N. and Richert, C. (2002) Oligonucleotides containing a nucleotide analog with an ethynylfluorobenzene as nucleobase surrogate. Tetrahedron Lett., 43, 8755–8758.
31. Minuth, M. and Richert, C. (2013) A nucleobase analog that pairs strongly with adenine. Angew. Chem. Int. Ed., 52, 10874–10877.
32. Barone, A.D., Tang, J.-Y. and Caruthers, M.H. (1984) In situ activation of bis-dialkylaminophosphines—a new method for synthesizing deoxyoligonucleotides on polymer supports. *Nucleic Acids Res.*, 12, 4051–4061.
33. Ingale, S.A., Mei, H., Leonard, P. and Seela, F. (2013) Ethynyl side chain hydration during synthesis and workup of “Clickable” oligonucleotides: bypassing acetyl group formation by triisopropylsilyle protection. *J. Org. Chem.*, 78, 11271–11282.
34. Shvekhgeimer, M.-G.A. (1996) Synthesis of halogenopyridines (review). *Chem. Heterocyclic Comp.*, 32, 987–1015.
35. Kubelka, T., Slavětinská, L., Eigner, V. and Hocek, M. (2013) Synthesis of 2,6-disubstituted pyridin-3-yl C=2′-deoxyribonucleosides through chemoselective transformations of bromo-chloropyridine C-nucleosides. *Org. Biomol. Chem.*, 11, 4702–4718.
36. Usman, N., Ogilvie, K.K., Jiang, M.Y. and Cedergren, R.J. (1987) The automated chemical synthesis of long oligoribonucleotides using 2′-O-silylated ribonucleoside 3′-O-phosphoramidites on a controlled-pore glass support: synthesis of a 43-nucleotide sequence similar to the 3′-half molecule of an Escherichia coli formylmethionine tRNA. *J. Am. Chem. Soc.*, 109, 7845–7854.
37. Mignon, P., Loverix, S., Steyaert, J. and Geerlings, P. (2005) Influence of the π–π interaction on the hydrogen bonding capacity of stacked DNA/RNA bases. *Nucleic Acids Res.*, 33, 1779–1789.
38. Rothemund, P.W.K. (2006) Folding DNA to create nanoscale shapes and patterns. *Nature*, 440, 297–302.
39. Park, S.H., Yin, P., Liu, Y., Reif, J.H., LaBean, T.H. and Yan, H. (2005) Programmable DNA self-assemblies for nanoscale organization of ligands and proteins. *Nano Lett.*, 5, 729–733.
40. Dunn, K.E., Dannenberg, F., Ouldridge, T.E., Kwiatkowska, M., Turberfield, A.J. and Bath, J. (2015) Guiding the folding pathway of DNA origami. *Nature*, 525, 82–86.
41. Griesang, N., Giessler, K., Lommel, T. and Richert, C. (2006) Four color, enzyme-free interrogation of DNA sequences with chemically activated, 3′-fluorophore-labeled nucleotides. *Angew. Chem. Int. Ed.*, 45, 6144–6148.
42. Gießler, K., Griesser, H., Gühringer, D., Sabirov, T. and Richert, C. (2010) Synthesis of 3′-BODIPY-labeled active esters of nucleotides and a chemical primer extension assay on beads. *Eur. J. Org. Chem.*, 19, 3611–3620.
43. Hänele, E and Richert, C. (2018) Enzyme-free replication with two or four bases. *Angew. Chem. Int. Ed.*, 57, 8911–8915.
44. Steffen, C., Thomas, K., Huniar, U., Hellweg, A., Rubner, O. and Schnier, A. (2010) TmoleX—a graphical user interface for TURBOMOLE. *J. Comput. Chem.*, 31, 2967–2970.