

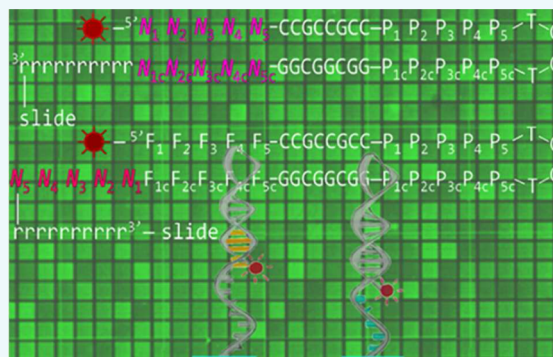
Sequence-Dependent Fluorescence of Cy3- and Cy5-Labeled Double-Stranded DNA

Nicole Kretschy, Matej Sack, and Mark M. Somoza*

Institute of Inorganic Chemistry, Faculty of Chemistry, University of Vienna, Althanstraße 14 (UZA II), A-1090 Vienna, Austria

Supporting Information

ABSTRACT: The fluorescent intensity of Cy3 and Cy5 dyes is strongly dependent on the nucleobase sequence of the labeled oligonucleotides. Sequence-dependent fluorescence may significantly influence the data obtained from many common experimental methods based on fluorescence detection of nucleic acids, such as sequencing, PCR, FRET, and FISH. To quantify sequence dependent fluorescence, we have measured the fluorescence intensity of Cy3 and Cy5 bound to the 5' end of all 1024 possible double-stranded DNA 5mers. The fluorescence intensity was also determined for these dyes bound to the 5' end of fixed-sequence double-stranded DNA with a variable sequence 3' overhang adjacent to the dye. The labeled DNA oligonucleotides were made using light-directed, in situ microarray synthesis. The results indicate that the fluorescence intensity of both dyes is sensitive to all five bases or base pairs, that the sequence dependence is stronger for double- (vs single-) stranded DNA, and that the dyes are sensitive to both the adjacent dsDNA sequence and the 3'-ssDNA overhang. Purine-rich sequences result in higher fluorescence. The results can be used to estimate measurement error in experiments with fluorescent-labeled DNA, as well as to optimize the fluorescent signal by considering the nucleobase environment of the labeling cyanine dye.



INTRODUCTION

The fluorescence of molecules is always sensitive to environmental conditions, although the magnitude of changes in the fluorescence intensity of any particular fluorophore depends on its specific modes of interaction with its environment.¹ Fluorescent molecules can be used as molecular environmental probes by selecting dyes with strong responses to, for example, pH,² viscosity,³ polarizability,⁴ elasticity,⁵ and polarity,⁶ however, in applications where the fluorescent intensity is to serve as a proxy for the abundance of the labeled molecule, environmental sensitivity is a liability that can result in reduced measurement accuracy.⁷ The cyanine dyes Cy3 and Cy5 are among the most widely used and versatile⁸ oligonucleotide labels in, e.g., microarray experiments, fluorescent in situ hybridization (FISH), real-time PCR (RT-PCR), and FRET studies^{9,10} and are considered to be relatively environmentally insensitive.¹¹ However, Cy3 and Cy5 consist of two indole rings connected by three or five carbon polymethine bridges which can undergo *cis-trans* isomerization from the first excited singlet state which competes with fluorescence.^{12–15} In viscous or restrictive environments, or with conformationally locked dye variants, the rate of isomerization is reduced or eliminated and the dyes are more fluorescent.¹⁶ When Cy3 and Cy5 are tethered to the end of double-stranded DNA they assume a planar capping configuration similar to that of an additional base pair,^{17,18} which inhibits isomerization and increases their fluorescence quantum yield and lifetime.¹⁵ At least in the case of Cy3, the range of motion available is not

fully restricted when attached to either single- or double-stranded DNA, with time-resolved fluorescence anisotropy measurements indicating decay components corresponding to rotation with DNA as well as relative to DNA.

Recent experiments have shown that both Cy3 and Cy5 are also quite sensitive to the particular nucleobase sequence of the ssDNA oligonucleotide to which they are attached,^{19,20} with the fluorescence intensity varying by a factor of about 2 between the brightest and the darkest labeled oligonucleotide in the case of Cy3, and a factor of about 3 in the case of Cy5. The variation in fluorescence intensity for ssDNA is strongly correlated with purine content, with purine-rich sequences associated with high intensity, and high pyrimidine content, particularly cytosine, with low intensity.¹⁹

The magnitude of the sequence-dependent fluorescence is large enough to affect the accuracy of experimental data derived from Cy3- and Cy5-labeled single-stranded DNA, but there is currently no data available on sequence-dependent effects in double-stranded DNA. In experimental methods based on labeled oligonucleotides, fluorescence is recorded either from the double-stranded hybrid (e.g., Sanger and next-generation sequencing, and molecular beacons²¹) or from the unhybridized strand alone (e.g., hydrolyzed labeled TaqMan probe fragments²²). High-throughput DNA sequencing-by-synthesis

Received: January 27, 2016

Revised: February 17, 2016

Published: February 19, 2016

is likely to be particularly vulnerable to sequence-dependent fluorescence because all short nucleobase sequences will be repeatedly encountered, and detection failures (deletion errors) from sequences highly unfavorable to fluorescence would be systematic and therefore not easily detectable with resequencing. Furthermore, the optical systems of sequencers need to balance dynamic range of detection with throughput, making their throughput sensitive to dyes with significant variations in fluorescence.²³ Even though our fluorescence data are obtained on microarrays, most genomics microarray data is fairly insensitive to sequence-dependent fluorescence because the labeling is typically based on reverse transcription using labeled random primers or other quasi-random methods.²⁴ Nevertheless, gene-specific fluorescence intensity effects, due to differences in the relative abundance of nucleobases in particular genes, have been detected.²⁵

Since both Cy3- and Cy5-labeled single- and double-stranded oligonucleotides are commonly used, we present here comprehensive results for double-stranded DNA to complement and strengthen previous results for Cy3 and Cy5 5'-labeled single-stranded DNA.¹⁹ Two types of sequence-dependent dye–dsDNA interactions, as illustrated in Figure 1, have been measured: relative intensity of the dyes at the 5'

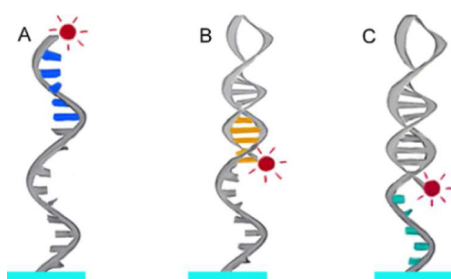


Figure 1. Interaction modes of dyes (red) on DNA. (A) 5' dye with adjacent nucleobases (blue) in ssDNA. (B) 5' dye with base-paired nucleobases (orange) in dsDNA. (C) 5' dye with nucleobases of ssDNA (green) adjacent to a terminal dye on dsDNA.

end of each of the 1024 possible double-stranded DNA 5mers (Figure 1B), and relative intensity of the dyes bound to the 5' end of a fixed-sequence double helix, but with a variable 5mer sequence 3' overhang adjacent to the dyes (Figure 1C). The sequence-dependent contribution of the overhang is relevant since in many experimental contexts, such as PCR and FISH, a short 5'-labeled oligonucleotide is used to quantify the presence of much longer DNA or RNA molecules. Detailed data on the sequence-dependent fluorescence of cyanine dyes on single-stranded DNA (Figure 1A) has been previously reported for Cy3, Cy5, Dy547, and Dy647;^{19,26} this ssDNA data showed that over the range of all possible 5mers, the intensity of Cy3 varied by about a factor of 2, and in the case of Cy5, by a factor of about 3. There was also a clear pattern to the data: the fluorescence follows, to a good approximation, the cumulative distribution function of a normal distribution, with purine-rich sequences resulting in high intensities and pyrimidine-rich sequences resulting in low intensities. In addition, 5' guanines promote higher fluorescence much more so than 5' adenines, and 5' cytosines result in much lower fluorescence in comparison with 5' thymidines. Here we will show that broadly similar trends also hold true for double-stranded DNA.

RESULTS AND DISCUSSION

The results for the sequence-dependent fluorescence of cyanine dyes have been highly consistent, with the adjacent purine bases promoting fluorescence relative to pyrimidine bases in single-stranded DNA,^{19,26} and with the results presented here in double-stranded DNA. In addition, for both ssDNA and dsDNA, a guanine immediately adjacent to the dye consistently results in the highest fluorescence, but in the more distal positions, adenine, rather than guanine, typically results in higher fluorescence. Of the pyrimidines, cytosine, rather than thymine, is most strongly associated with low fluorescence.

Cy3 and Cy5 dsDNA Interactions. Figure 2 summarizes the results for both the 5' Cy3 and Cy5 terminal labeling experiments on dsDNA. These data correspond to the case where the random linker is used and the permuted nucleobases form a double strand (Scheme 1A). Here, the dye interactions with the single-stranded segment are present, but the data will

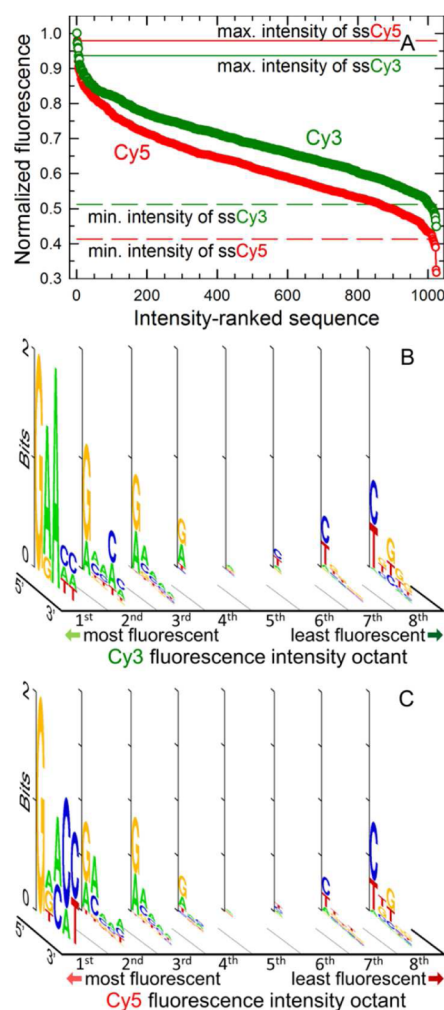
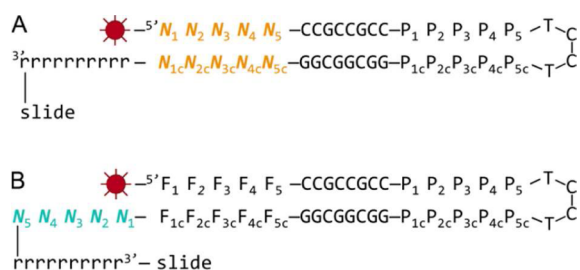


Figure 2. Double-stranded DNA labeling with Cy3 and Cy5 (Figure 1B). (A) Relative fluorescence intensity of Cy3 and Cy5 end-labeled 5mers, ranked from most to least intense. The intensity falls by 55% for Cy3 and almost 70% for Cy5. The horizontal lines show the fluorescence intensity of single-stranded reference sequences on the same arrays. Fluorescence intensity consensus sequences of all 1024 dsDNA 5mers 5'-end-labeled using (B) Cy3 and (C) Cy5. The fluorescent range was equally divided into eight bins of equal intensity ranges, and the consensus sequence for all the 5mers is plotted for each such octile.

Scheme 1. Sequence Design for the 5'-Dye Self-Hybridizing DNA Strands^a



^aSequence (A) is used to measure the interaction of the dyes with dsDNA and sequence (B) is used to measure the interactions of the dyes with the ssDNA overhang of dsDNA.

reflect the average over all possible sequences. As was the case with the data from Cy3 and Cy5 labeled ssDNA,¹⁹ the overall range of fluorescence intensity is about a factor of 2 for Cy3 and a factor of 3 for Cy5 (Figure 2A). In order to be able to compare the fluorescence intensity data for dsDNA with ssDNA, the array design included reference ssDNA sequences. These sequences have a very similar design, but with bases rearranged to prevent hybridization. Figure 2A shows that both Cy3 and Cy5 on dsDNA have a somewhat extended range of fluorescence intensity in comparison to Cy3 and Cy5 on ssDNA (horizontal lines). Most of the additional range of intensity is on the lower edge of intensity, i.e., the sequences resulting in the highest fluorescence result in similar intensity for both ssDNA and dsDNA.

The fluorescence intensity of intensity of most, or perhaps all, dyes is dependent on the nucleobase environment. In many cases the mechanism is a photoinduced charge transfer between the bases and the dye (fluorescein,²⁷ coumarin,²⁸ rhodamine,²⁹ and pyrene³⁰), in which case the quenching efficiency is determined by proximity and base redox potential, $dG < dA < dC < dT$, when the bases are reduced, or the reverse order when oxidized.²⁸ Ethidium bromide, another well-known dsDNA fluorescence label, undergoes quenching via proton transfer to the solvent; intercalation enhances fluorescence by reducing solvent exposure.³¹ In the case of the cyanine dyes, however, charge transfer is not thermodynamically favored.^{32,33} Instead, the intensity of cyanine dyes conjugated with DNA is attributed to the modulation of the rotational isomerization barrier in the excited state.^{12–14}

NMR data indicate that Cy3 and Cy5, 5'-linked to dsDNA, are positioned at the end of the double helix similarly in a capping configuration, in a manner similar to that of a base pair.^{17,18} This arrangement should restrict the rate of *cis*–*trans* isomerization of the dyes, increasing fluorescence relative to the free dye. However, relative to the same dyes bound to the end of ssDNA, differences in the rate of isomerization are less clear since the dyes stack with the terminal base in both cases. Simulations and experiments indicate that the quantum yield of Cy3 is higher on ssDNA vs dsDNA, and that on dsDNA the strength of the stacking interaction depends on the identity of the terminal base pair.^{15,34,35} Our experiments indicate that the fluorescence of Cy3 and Cy5 is somewhat greater on dsDNA; however, the differences between our results and previously published results,¹⁵ which show a 2-fold greater fluorescence of Cy3 on ssDNA, may be due to the particular choice of cyanine dye. In particular, we conjugate with DNA using the Cy3 and Cy5 phosphoramidites, rather than the sulfonated versions of

these dyes, used by Sandborn et al.,¹⁵ and which are more commonly used for protein labeling. The sulfonates increase the hydrophilicity of the dyes, which could affect the strength of the stacking interactions with the nucleobases. We have previously measured the intensity of sulfonated Cy3 and Cy5 on DNA, and found a very strong pattern of sequence-specific fluorescence distinct from that of the unsulfonated dyes.¹⁹

In order to visualize the relationship between the nucleobase sequence and the fluorescence intensity, the consensus sequences for each octant of intensity are plotted in Figure 2B and C for Cy3 and Cy5, respectively. These data are quite similar to those obtained with the same dyes on ssDNA.¹⁹ The most apparent differences in the dsDNA data are that cytosine is less prominent in the weakly fluorescent sequences, and that cytosine is more prominent in the distal positions of the strongly fluorescent sequences, particularly for Cy5. If, as previous studies have indicated, the fluorescence intensity of cyanine dyes is greater on ssDNA, there might be bias in the consensus toward adenine- and thymine-rich sequences, which will tend to destabilize the double helix near the dyes, resulting in a higher locally single-stranded (“frayed” ends) population of DNA. In relationship to our previous data of Cy3 and Cy5 on ssDNA, this trend is not apparent. In the dsDNA data (Figure 2), the melting temperature of the consensus sequences for the most fluorescent intensity octants are higher than those in the equivalent octants in the ssDNA data for both Cy3 and Cy5 due to the increased population of cytosines.

Cy3 and Cy5 Overhang Interactions. In the results described above, the dyes must also be interacting with the immediately adjacent ssDNA overhang segment as illustrated in Figure 1C and Scheme 1B. In order to estimate how this ssDNA modulates the fluorescence, the random nucleobase linker was replaced with segments representing all possible 5mers. To avoid having too many overall permutations, only two dsDNA sequences were used, one associated with strong fluorescence (GAAAA) and one with weak fluorescence (CGTGG). About 10 replicates of each of the 2048 resulting sequences fit on a single microarray, allowing accurate relative intensity comparisons between sequences. In the dsDNA data shown in Figure 2, the sequence GAAAA resulted in the 33rd and 100th brightest fluorescence for Cy3 and Cy5, respectively. The sequence CGTGG resulted in the 1008th and 898th brightest fluorescence for Cy3 and Cy5, respectively. The results from the overhang experiment, using Cy3 as the dye, are shown in Figure 3. In Figure 3A, the intensity of each sequence has been normalized to that of the most intense sequence, which, as expected, belongs to the Cy3-dsGAAAA set. Most of the sequences with Cy3-dsCGTGG are darker than any of those with GAAAA. Figure 3A clearly shows that the intensity of the dye is similarly determined by both the dsDNA segment and the adjacent ssDNA segment since the intensity difference between the two curves is similar to the range in intensities within each curve.

The relationship between the nucleobase sequence of the permuted overhang and the fluorescence intensity is shown using consensus logos in Figure 3B and C, for Cy3-dsGAAAA and Cy3-dsCGTGG, respectively. The consensus sequences show a similar pattern to those of the ssDNA data and the dsDNA data with the random overhang; the most fluorescent signal results from sequences with high purine content and the least fluorescence signal results from sequences with high pyrimidine content, particularly cytosine. Two additional trends are clearly visible in the consensus sequence data. First, the

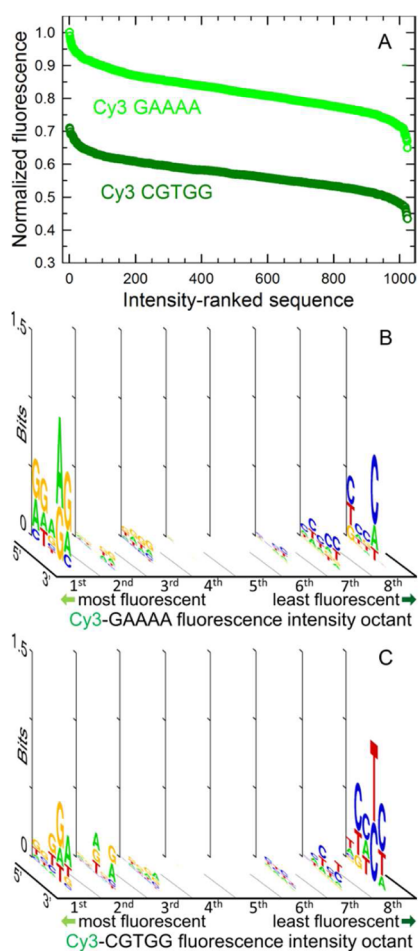


Figure 3. 5'-Cy3-dsDNA with a permuted 3' overhang (Figure 1C). The dsDNA strand to which the Cy3 is attached has one of two sequences: GAAAA (bright) or CGTGG (dark). (A) Relative fluorescence of Cy3-GAAAA and Cy3-CGTGG ranked from most to least intense over the range of all ssDNA 3' overhang 5mers. The intensity falls by $\sim 35\%$ for both Cy3-GAAAA and Cy3-CGTGG. Fluorescence intensity consensus sequences of all 1024 5mers on the 3'-overhang of (B) Cy3-dsGAAAA and (C) Cy3-dsCGTGG. The fluorescent was equally divided into eight bins of equal intensity ranges. The consensus sequence is plotted for each bin.

information content (bits) for each position is typically lower than that for the data with the random overhang. This is because in the present case, there is no single dominant base at any position, e.g., both purines are approximately equally probable in the most fluorescent sequences. This trend can also be anticipated by the shape of the intensity curves in Figure 3A, which, spanning a lower range of intensity in comparison to that in Figure 2 for the same number of permuted sequences, indicate a reduced sequence dependence of fluorescence. Second, the more distal bases are more prominent in the consensus sequences, which suggests that the dye is interacting more strongly with these more distal bases. One possibility is that the presence of the dye on the terminus of the double-stranded segment may tend to displace the more proximal overhang bases to conformations where they cannot affect the *cis-trans* isomerization rate. This is consistent with NMR data indicating that Cy3 occupies much of the available stacking space at the end of dsDNA.¹⁸

Data for Cy5 on double-stranded DNA with a permuted overhang is shown in Figure 4. These data were collected using

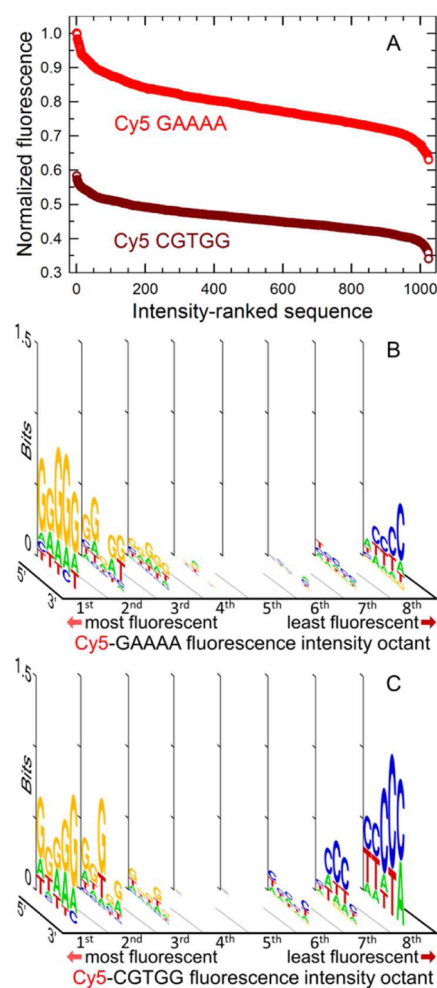


Figure 4. 5'-Cy5-dsDNA with a permuted 3' overhang (Figure 1C). The dsDNA strand to which the Cy5 is attached has one of two sequences: GAAAA (bright) or CGTGG (dark). (A) Relative fluorescence of Cy5-GAAAA and Cy5-CGTGG, ranked from most to least intense over the range of all ssDNA 3' overhang 5mers. The intensity falls by $\sim 40\%$ for both Cy5-GAAAA and Cy5-CGTGG. Fluorescence intensity consensus sequences of all 1024 5mers on the 3'-overhang of (B) Cy5-dsGAAAA and (C) Cy5-dsCGTGG. The fluorescent was equally divided into eight bins of equal intensity ranges. The consensus sequence is plotted for each bin.

the same methods and the same microarray design, only using Cy5 instead of Cy3. As with Cy3, the intensity difference between the two curves in Figure 4A is similar to the range in intensities within each curve, clearly showing that the intensity of Cy5 is similarly determined by both the dsDNA segment and the adjacent ssDNA overhang segment. Unlike in the case of Cy3, all of the Cy5-dsCGTGG sequences are darker than the darkest of the Cy5-dsGAAAA sequences. The specific sequence Cy5-dsGAAAA in the random linker data set resulted in an intensity of 0.8 relative to that of Cy5-dsGAACC, the most intense suggesting that the gap between the curves in Figure 4A could be significantly increased by using GAACC as the fixed double-stranded sequence. Although the two curves in Figure 4A appear to have different shapes, this is due only to the large fluorescence intensity difference between them. Independently normalizing the Cy5-dsCGTGG data would cause it to overlap very closely with the Cy5-dsGAAAA data, indicating that both double-stranded sequences modulate the interaction of the dye with the overhang bases to a similar extent.

The relationship between the nucleobase sequence of the permuted overhang and the fluorescence intensity is shown using consensus logos in Figure 4B for Cy5-dsGAAAA and in Figure 4C for Cy5-dsCGTGG. Like in the case of Cy3, the highest fluorescence is strongly associated with purines while the lowest fluorescence is strongly associated with pyrimidines. Between the purines, guanine is clearly more relevant than adenine in promoting fluorescence. Cytosine is also much more common than thymine in the sequences associated with low fluorescence. As a result of the dominance of these two bases, the information content of the consensus sequences is higher in the case of Cy5. The trend observed for Cy3, that the dye interacts more strongly with more distal bases, is also the case with Cy5.

For both Cy3 and Cy5, sequences resulting in the lowest intensity among the dye-dsCGTGG subset have intensities similar to the darkest from the data sets with the random overhang in Figure 2. Since the use of a random nucleobase linker should be equivalent to averaging over all linker base permutations, the expectation was that the minimum fluorescence measured in the permuted overhang experiments would be significantly lower than those measured using random overhang. One possibility is that the range over which the fluorescence intensity of Cy-dyes can be modulated via interactions with DNA is restricted. This seems reasonable since the total range over which the fluorescence quantum yield of Cy3 can be lowered by restricting the rate of *cis-trans* isomerization is about a factor of 8 at room temperature, and Cy3 on DNA appears to be limited to the lower half of this range.¹⁵ Nevertheless, some additional range of fluorescence intensity could likely be measured in permuted sequences longer than 5mers. In most of the consensus sequences in Figures 2, 3, and 4, there is information content in the fifth base, the most distal; indicating that this base also participates in modulating the intensity, so a sixth or seventh base is also likely to contribute to the modulation of fluorescence. Another perspective in this regard is that the shapes of the curves in Figures 2A, 3A, and 4A can be interpreted as cumulative distribution functions where the variable is the normalized relative fluorescence. To a good approximation, the fluorescence intensities of Cy3 and Cy5 on random DNA sequences have probability mass functions approximating those of binomial distributions, where the two results are purine or pyrimidine.¹⁹ Most random 5mer sequences will contain a mix of purines and pyrimidines, which will result in intermediate fluorescence in the central region of the distribution. A few sequences will contain mostly or exclusively purines or pyrimidines, resulting in, respectively, fluorescence at the high and low tails of the intensity distribution. Increasing the permuted sequence length (Bernoulli trials) should result in a few sequences in the tails of the distribution that extend the range of fluorescence.

These results are consistent with previous experiments on the fluorescence of Cy3 and Cy5, which have also shown similar patterns of nucleobase dependency. Studies on the interactions of Cy3 with nucleoside monophosphate solutions have found a pattern of nucleobase-specific enhancement of fluorescence, $dG > dA > dT > dC > \text{no DNA}$.³⁶ Experiments on an intercalating cyanine dye derived from thiazole orange demonstrated a strong association of fluorescence with purine DNA homopolymers but not with pyrimidine homopolymers; the resulting fluorescence relative intensities followed the pattern $dG > dA \gg dC > dT > \text{no DNA}$ (100, 39, 2.3, 1.8, and

0.5, respectively).³⁷ Computer simulations in this study also indicated that the dyes associate poorly with poly(dC) and poly(dT), while binding strongly to poly(dG) and poly(dA). All these results fit well with the model that $\pi-\pi$ interactions between cyanine dyes and nucleobases decrease the *cis-trans* isomerization rate. Purines, with a more extensive π system, are more effective than pyrimidines. The extent of the π system follows the order $dG(14) > dA(12) > dT(10) = dC(10)$ in terms of number of π electrons, and the order $dG(153 \text{ \AA}^2) > dA(142 \text{ \AA}^2) = dT(142 \text{ \AA}^2) > dC(127 \text{ \AA}^2)$ in terms of surface area.³⁸ These results apply directly to the 5' nucleobase in our terminal labeling experiments since this is the base that is directly adjacent to the dye. We consistently observe, for both single- and double-stranded data, that cyanine dye fluorescence follows the same trend, $dG > dA > dT > dC$, indicating that the terminal base directly affects rotational isomerization. The data also consistently shows that adjacent nonterminal bases modulate dye fluorescence, with a distance-dependent influence, indicating that sequence-dependent rigidity of the single- or double-stranded DNA also contributes to the observed fluorescence of Cy3 and Cy5. We hypothesize that the ability of the terminal base to hinder the rotational isomerization of the dye increases when it is part of a more rigid sequence of bases. The flexibility of DNA, particularly dsDNA, is of ongoing interest due to its role in packing and in the formation of protein-DNA complexes.³⁹ Many available degrees of freedom of the bases contribute to DNA rigidity or flexibility, not all of which may be relevant to restricting the isomerization of the terminal dye; nevertheless, multiple experimental approaches indicate that purine stacks are more rigid than pyrimidine stacks in ssDNA.^{40,41} A similar pattern is observed in dsDNA, also related to differences in base stacking area, $dG (139 \text{ \AA}^2) > dA (128 \text{ \AA}^2) > dC (102 \text{ \AA}^2) > dT (95 \text{ \AA}^2)$, and stacking free energy, $dA \gg dG > dT \approx dC$ (2.0, 1.3, 1.1, and 1.0 kcal·mol⁻¹), for B-form geometry, based on melting temperature changes.³⁸ Other experiments based on 5' dangling DNA hairpins and 3' RNA unpaired nucleotides give similar stability results: $A \approx G > T/U > C$.^{42,43} Sequence specificity of the flexibility of di- and tetramers,^{44,45} obtained from crystal structures and molecular dynamics simulations, appear to be less relevant in this case because they treat paired bases symmetrically and as a single rigid unit, such that, e.g., the deformability of $AA(TT) = TT(AA)$. While this treatment is relevant to the ability of dsDNA to bend, the hydrogen bonding between Watson-Crick pairs does not contribute to duplex stabilization; instead, duplex stability is mainly determined by base-stacking interactions.⁴⁶ This suggests that, at short length scales, the relevant modes of DNA dynamics are largely decoupled from the complementary strand and interact with the cyanine dyes by restricting the available torsional volume and by changing high-frequency coordinates of the potential energy surface of the excited state.⁴⁷ Our experiments are based the two cyanine dyes commonly used for DNA labeling, but sulfonated variants of Cy3 and Cy5 appear to interact differently with nucleobases.^{15,19} The sulfonates increase water solubility, but may modify the stacking interaction with DNA bases; stacking stability is dominated by hydrophobic effects with contributions from dispersion and electrostatic forces,³⁸ all of which are likely to be affected by the charges on the sulfonates.

CONCLUSION

With the data presented here, we have sought to clarify and quantify the impact of sequence-dependent fluorescence of Cy3 and Cy5 tethered to double-stranded DNA. The results are consistent with previous results of Cy3 and Cy5 and similar cyanide dyes tethered to single-stranded DNA.^{19,26} The results are also consistent with measurements of the fluorescence yield of Cy3 in solution with each of the DNA nucleoside monophosphates, which also follows the pattern $G > A > T > C$.³⁶ The preponderance of evidence supports the hypothesis that stronger cyanine dye–nucleobase stacking interactions of the purines relative to the pyrimidines restrict the *cis*–*trans* isomerization rate of these dyes, enhancing fluorescence. The results can be used in the planning and analysis of experiments based on the labeling of DNA (and probably RNA) with cyanine dyes. For example, TaqMan or molecular beacon PCR probes and FISH probes using cyanine dye reporters can be designed with one or more guanines or adenines immediately adjacent to the dye for increased signal. The sequence for the latter two of these probes can also be adjusted so that the reporter dye is adjacent to a purine-rich segment of the target upon hybridization. In the case of next-generation sequencing-by-synthesis, where high throughput relies on maintaining the low end of the dynamic range near the noise threshold,^{48,49} the data analysis pipeline can take into account the effect on measured fluorescence of adjacent nucleobases when determining the probability of a correct nucleobase assignment.

EXPERIMENTAL PROCEDURES

Microarray Synthesis. Glass slides (Schott Nexterion D, cleanroom-cleaned) were functionalized with *N*-(3-triethoxysilylpropyl)-4-hydroxybutyramide (Gelest SIT8189.5). The slides were loaded in a stainless steel rack, placed in a plastic container, and covered with 0.5 L of a solution consisting of 10 g of the silane in a 95:5 (v/v) ethanol:water plus 1 mL acetic acid. The slides were gently agitated for 4 h at room temperature and then washed twice for 20 min each with the above solution without the silane. The slides were drained, blown dry with argon, and cured in a preheated vacuum oven (120 °C) overnight and stored in a desiccator cabinet.

For the synthesis of terminally labeled oligonucleotides on microarrays we used the technique of maskless array synthesis (MAS).^{50,51} MAS was developed for in situ synthesis of high-density DNA microarrays and consists of an optical system and a chemical delivery system. The optical system consists of a digital micromirror device (DMD), an array of individually tiltable mirrors, which direct ultraviolet light from a mercury lamp to the corresponding feature on the microarray via 1:1 imaging optics. Microarray layout and oligonucleotide sequences are determined by selective removal of the photocleavable protecting groups on the phosphoramidites at the 5' termini of the oligonucleotides.

A computer synchronizes the light exposures pattern with solvent reagent delivery to the synthesis surface. The chemical system consists of a slightly modified Perspective Biosystems Expedite 8909 synthesizer. Oligonucleotide synthesis chemistry is similar to that used in conventional solid-phase synthesis. The standard acid-labile 5'-OH protecting group of the phosphoramidites is replaced with the photocleavable nitrophenylpropyloxycarbonyl (NPPOC) group.⁵² Upon absorption of light near 365 nm, the NPPOC group comes off, leaving a free hydroxyl group that is able to react with an activated

phosphoramidite in the next coupling cycle. An exposure solvent consisting of 1% (m/v) imidazole in DMSO is needed during ultraviolet exposure to promote the cleavage of the NPPOC group.⁵¹ The coupling reactions were performed with 30 mM NPPOC phosphoramidite monomers and 0.25 M dicyanoimidazole (both from SAFC) for 60 s. In the case of the Cy3 and Cy5 phosphoramidites (GE Healthcare 28–9172–98 and Glen Research 10–5915–95), Figure 5, the coupling

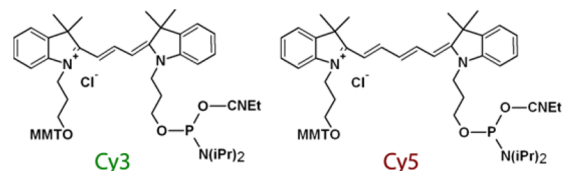


Figure 5. Molecular structures of the Cy3 and Cy5 cyanine dye phosphoramidites used in this study. After the end of the synthesis and the chemical deprotection step, the dyes are linked to the 5' DNA nucleoside via a phosphodiester bond.

reaction time was extended to 10 min at a monomer concentration of 15 mM. Acetylation with a 1:1 mix of *tert*-butylphenoxycetyl acetic anhydride in tetrahydrofuran (Cap A) and 10% *N*-methylimidazole in tetrahydrofuran/pyridine (8:1) (Cap B) after each coupling reaction was used to ensure that only correctly synthesized sequences receive the fluorescent label.

After microarray synthesis the substrate was vigorously washed for 2 h with acetonitrile in a 50 mL Falcon tube to remove uncoupled Cy3 or Cy5 phosphoramidites, which tend to adhere nonspecifically to the glass surface. The base and phosphate protecting groups were removed by immersing the glass slide into 1:1 (v/v) ethylenediamine in ethanol for 2 h at room temperature. Following deprotection, the microarrays were washed twice with distilled water and dried with argon.

Microarray Design. In principle, the resolution of the digital micromirror device, 768 × 1024, allows for simultaneous measurement of all possible *n*-mers up to *n* = 9 (262 144), but in these experiments, only permutations of 5mers were included in order to include multiple replicates and to dedicate more microarray surface area to each sequence and therefore to achieve a good signal-to-noise ratio. The 1024 sequences were laid out in a 25 in 36 pattern, that is, each “feature” (contiguous area where a single sequence is synthesized) on the microarray corresponded to a 5 by 5 block of mirrors surrounded by a one-mirror-sized margin where no DNA was synthesized. Each of the 1024 single-sequence features was replicated 20 times on each microarray in the case of the double-stranded experiments (Figure 1B), and 10 times in the case of the double-stranded DNA with single-stranded overhang experiments (Figure 1C).

Double-Stranded DNA Annealing. To promote hairpin-loop formation and self-hybridization, after deprotection the array was incubated in 40 mL PBS buffer (0.65 M Na⁺, pH 7.4) starting at 50 °C and cooled to room temperature over 30 min. Then it was washed with final wash buffer for a few seconds and dried with a microarray centrifuge. Successful hairpin loop formation was then verified by hybridization of a Cy3-labeled oligonucleotide (5'-Cy3-GGC GGC GGG TTC A-3') to two unlabeled complementary sequences on the array: (1) a sequence (TGA ACC CGC CGC CGT CCA TCCT TGG ACG GCG GCG GGT TCA) that self-hybridized via hairpin-loop formation in the previous step and is therefore blocked from hybridization with the added oligonucleotide, and (2) a

sequence (TGA ACC CGC CGC C) that cannot self-hybridize but is fully complementary with the added labeled sequence.

Sequence Design. Three principle considerations were applied to the sequence design: (1) The double-stranded sequences should all have equal melting temperatures since they must all form duplexes equally under the single hybridization condition of the microarray, (2) the melting temperature should be relatively high in order to ensure stable duplex formation, and (3) the surface density of labeled oligonucleotides should be constant for all experimental oligonucleotides on the microarray so that fluorescence intensity differences between them can be attributed to sequence-dependent effects. To meet these design principles the double-stranded oligonucleotides have the design illustrated in Scheme 1.

The sequences contain self-complementary segments to allow for duplex formation. The central TCCT sequence is known to bend easily to promote hairpin loop formation.⁵³ The N_i represents the 5mer experimental nucleobases that base pair with the complementary N_{ic} . On the 3' side of the N_i is the fixed sequence CCGCCGCC which hybridizes with the GGCGGCGG sequence on the opposite side of the hairpin. This GC-rich stretch is used to increase the melting temperature. The $P_1P_2P_3P_4P_5$ sequence is derived from the experimental 5mer sequence $N_1N_2N_3N_4N_5$ using nonidentity, noncomplementarity logic: for all i , if $N_i = dA$ then $P_i = dC$; or if $N_i = dC$ then $P_i = dT$; or if $N_i = dG$ then $P_i = dA$; or if $N_i = dT$ then $P_i = dG$. These strands hybridize with their complementary sequences $P_{5c}P_{4c}P_{3c}P_{2c}P_{1c}$. The P_i and P_{ic} sequences have a double function: (1) they equilibrate the base composition in order to ensure equal number density of all experimental sequences on the array, and (2) they increase and homogenize the melting temperatures (to $T_m = 63$ °C, salt adjusted, 50 mM Na⁺) by giving all the complementary DNA sequences on the array exactly five of each nucleobases (plus the fixed GC sequences) while retaining self-complementarity. The sequences are separated from the glass substrate with a random linker 10mer sequence synthesized from an equimolar mix of the four DNA phosphoramidites. The random linker replaces the traditional poly(dT), and linker to avoid the potential bias of any particular interaction of the dye and a dT homopolymer. An alternative perspective is that the dye will interact with both the double-stranded and single-stranded segments, but the interaction with the single-stranded segment will be the average of all possible sequences. In the second set of experiments, the single-stranded sequence is permuted. The results of both data sets can be used to estimate the relative contributions, to dye intensity variation, of the single- vs double-stranded segments.

With these rules, all of the sequences (excluding the linker) have exactly 5 adenines, 15 cytidines, 13 guanosines, and 7 thymidines. Since the coupling efficiency of each of the four DNA phosphoramidites can be different and can vary with time and by batch, equal numbers of each base in each of the sequences assures equal representation of the experimental oligonucleotides. This sequence design, in conjunction with acetic anhydride capping after the coupling reactions, ensures equal density and melting temperature and that only accurately synthesized sequences receive the final coupling with the Cy3 or Cy5 phosphoramidite. An alternative approach, to use simpler sequences and then adjust the data for the measured coupling efficiencies, is less reliable since the coupling efficiencies of the phosphoramidites used in maskless array

synthesis are measured with fluorescent dye terminal labeling experiments,^{54–57} which limits their accuracy due to the sequence-dependent fluorescence intensity of single-stranded DNA.¹⁹

The second set of experiments, with the dyes attached to fixed-sequence double-stranded DNA and a variable single-stranded overhang, has a similar design (Scheme 1B). Here, the permuted overhang sequence $N_1N_2N_3N_4N_5$ is added at the 3' end to put it adjacent to the 5' fluorescent label. The F_i and F_{ic} are complementary but are no longer permuted; $N_1N_2N_3N_4N_5$ is either GAAAA or CGTGG. GAAAA and CGTGG were chosen from the initial double-stranded experiments as sequences resulting in high and low fluorescence intensity, respectively, for both Cy3 and Cy5.

In order to allow direct comparisons between the relative fluorescence intensities of the dyes on single- vs double-stranded DNA, each dsDNA microarray design included sequences that cannot self-hybridize to form dsDNA, but have a very similar overall sequence design and base composition. Since most of the microarray features were needed for the dsDNA permutations, only a sampling of 57 labeled ssDNA permutations was included. These sequences were chosen to be representative of the range of expected fluorescence intensities for ssDNA found in previous experiments.¹⁹ To prevent the self-hybridization of these sequences, the $N_{5c}N_{4c}N_{3c}N_{2c}N_{1c}$ segment was inverted to $N_{1c}N_{2c}N_{3c}N_{4c}N_{5c}$, the $P_{5c}P_{4c}P_{3c}P_{2c}P_{1c}$ segment was inverted to $P_{1c}P_{2c}P_{3c}P_{4c}P_{5c}$, palindromic N_i 5mers were avoided, and the segment GGCGGCGG was reordered to GCGGCGGG.

Data Extraction and Analysis. Fluorescent images of the microarrays were obtained using a GenePix 4100A scanner with resolution of 5 μ m and with PMT voltages set to give similar intensity ranges for both Cy3 and Cy5, and no saturated pixels, 350 and 450 V, respectively. Dye fluorescence was excited using 532 and 635 nm solid-state lasers for Cy3 and Cy5, respectively. Fluorescence was collected through 550–600 nm and 655–695 nm bandpass filters for Cy3 and Cy5, respectively. Fluorescence was collected using a 0.68 NA objective lens with a focal length of 3.1 mm. Microarray scanners are designed to provide intensity values that are highly consistent across the scanned surface. This allows highly reliable relative fluorescence comparisons between microarray features. The presence of the microarray surface, a lossless glass–air dielectric interface, close to the fluorophores does not influence the relative emission intensity or wavelength.⁵⁸ In addition to the high throughput available with microarray experiments, a significant advantage is that the density of fluorescence groups can be closely controlled to avoid the aggregation-induced quenching artifacts that can occur in solution experiments with hydrophobic dyes such as Cy3 and Cy5.

The fluorescence intensity data was extracted from the scan image with NimbleScan v 2.1 software from NimbleGen and further processed in Excel. For each microarray, fluorescence intensity values were calculated as the average of the replicates of each sequence, which were randomly located on each microarray. For the double-stranded experiment, there were 20 sequence replicates per array. For the overhang experiment there were 10 replicates per array because of the inclusion of 2 experimental sets, one with double-stranded sequence which strongly promotes fluorescence (dye-GAAAA) and one with a double-stranded sequence resulting in weak fluorescence (dye-CGTGG). Error was calculated as the standard error of the

mean. The consensus sequence figures were generated by ranking the 1024 sequences by fluorescence intensity and then dividing the sequences into 8 bins spanning equal ranges of intensity. Consensus logos for the sequences in each of these octiles of fluorescence intensity were generated using Weblogo (<http://weblogo.berkeley.edu/>).⁵⁹ Each of the 8 consensus sequence logos per fluorescent label represents 1/8 of the intensity range and are arranged together left to right in order of decreasing intensity to compactly depict the relationship between sequence and fluorescence for the entire data set. The relative fluorescence intensity data for all the experimental sequences are available as [Supporting Information](#).

■ ASSOCIATED CONTENT

■ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acs.bioconjchem.6b00053](https://doi.org/10.1021/acs.bioconjchem.6b00053).

Relative fluorescence intensity data for all the experimental sequences (XLS)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: mark.somoza@univie.ac.at.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Funding by the Faculty of Chemistry of the University of Vienna and the Austrian Science Fund (FWF P23797) is gratefully acknowledged. Dedicated to Prof. Bernhard Keppler on the occasion of his 60th birthday.

■ REFERENCES

- (1) Kupstat, A., Ritschel, T., and Kumke, M. U. (2011) Oxazine Dye-Conjugated DNA Oligonucleotides: Förster Resonance Energy Transfer in View of Molecular Dye–DNA Interactions. *Bioconjugate Chem.* 22, 2546–2557.
- (2) Han, J. Y., and Burgess, K. (2010) Fluorescent Indicators for Intracellular pH. *Chem. Rev.* 110, 2709–2728.
- (3) Luby-Phelps, K., Mujumdar, S., Mujumdar, R. B., Ernst, L. A., Galbraith, W., and Waggoner, A. S. (1993) A Novel Fluorescence Ratiometric Method Confirms the Low Solvent Viscosity of the Cytoplasm. *Biophys. J.* 65, 236–242.
- (4) Brauns, E. B., Murphy, C. J., and Berg, M. A. (1998) Local dynamics in DNA by temperature-dependent Stokes shifts of an intercalated dye. *J. Am. Chem. Soc.* 120, 2449–2456.
- (5) Somoza, M. M., Wiedersich, J., and Friedrich, J. (2007) Protein elasticity determined by pressure tuning of the tyrosine residue of ubiquitin. *J. Chem. Phys.* 127, 095102.
- (6) Somoza, M. M., Andreatta, D., Murphy, C. J., Coleman, R. S., and Berg, M. A. (2004) Effect of lesions on the dynamics of DNA on the picosecond and nanosecond timescales using a polarity sensitive probe. *Nucleic Acids Res.* 32, 2494–2507.
- (7) Grabolle, M., Brehm, R., Pauli, J., Dees, F. M., Hilger, I., and Resch-Genger, U. (2012) Determination of the Labeling Density of Fluorophore–Biomolecule Conjugates with Absorption Spectroscopy. *Bioconjugate Chem.* 23, 287–292.
- (8) Park, J. W., Kim, Y., Lee, K.-J., and Kim, D. J. (2012) Novel Cyanine Dyes with Vinylsulfone Group for Labeling Biomolecules. *Bioconjugate Chem.* 23, 350–362.
- (9) Michaelis, J., van der Heden van Noort, G. J., and Seitz, O. (2014) DNA-Triggered Dye Transfer on a Quantum Dot. *Bioconjugate Chem.* 25, 18–23.

- (10) Santner, T., Hartl, M., Bister, K., and Micura, R. (2014) Efficient Access to 3'-Terminal Azide-Modified RNA for Inverse Click-Labeling Patterns. *Bioconjugate Chem.* 25, 188–195.

- (11) Pawley, J. B. (2006) *Handbook of Biological Confocal Microscopy*, 3rd ed., Springer.

- (12) Huang, Z. X., Ji, D. M., Wang, S. F., Xia, A. D., Koberling, F., Patting, M., and Erdmann, R. (2006) Spectral identification of specific photophysics of Cy5 by means of ensemble and single molecule measurements. *J. Phys. Chem. A* 110, 45–50.

- (13) Widengren, J., and Schwille, P. (2000) Characterization of photoinduced isomerization and back-isomerization of the cyanine dye Cy5 by fluorescence correlation spectroscopy. *J. Phys. Chem. A* 104, 6416–6428.

- (14) Dempster, D. N., Thompson, G. F., Morrow, T., and Rankin, R. (1972) Photochemical Characteristics of Cyanine Dyes.1. 3,3'-Diethylxadicarbocyanine Iodide. *J. Chem. Soc., Faraday Trans. 2* 68, 1479–1496.

- (15) Sanborn, M. E., Connolly, B. K., Gurunathan, K., and Levitus, M. (2007) Fluorescence properties and photophysics of the sulfoindocyanine Cy3 linked covalently to DNA. *J. Phys. Chem. B* 111, 11064–11074.

- (16) Hall, L. M., Gerowska, M., and Brown, T. (2012) A highly fluorescent DNA toolkit: synthesis and properties of oligonucleotides containing new Cy3, Cy5 and Cy3B monomers. *Nucleic Acids Res.* 40, e108.

- (17) Iqbal, A., Wang, L., Thompson, K. C., Lilley, D. M. J., and Norman, D. G. (2008) The Structure of Cyanine 5 Terminally Attached to Double-Stranded DNA: Implications for FRET Studies. *Biochemistry* 47, 7857–7862.

- (18) Norman, D. G., Grainger, R. J., Uhrin, D., and Lilley, D. M. J. (2000) Location of cyanine-3 on double-stranded DNA: Importance for fluorescence resonance energy transfer studies. *Biochemistry* 39, 6317–6324.

- (19) Agbawue, C., and Somoza, M. M. (2011) Sequence-Dependent Fluorescence of Cyanine Dyes on Microarrays. *PLoS One* 6, e22177.

- (20) Harvey, B. J., Perez, C., and Levitus, M. (2009) DNA sequence-dependent enhancement of Cy3 fluorescence. *Photochem. Photobiol. Sci.* 8, 1105–1110.

- (21) Tyagi, S., and Kramer, F. R. (1996) Molecular Beacons: Probes that Fluoresce upon Hybridization. *Nat. Biotechnol.* 14, 303–308.

- (22) Holland, P. M., Abramson, R. D., Watson, R., and Gelfand, D. H. (1991) Detection of specific polymerase chain reaction product by utilizing the 5' -> 3' exonuclease activity of Thermus aquaticus DNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* 88, 7276–7280.

- (23) Fuller, C. W., Middendorf, L. R., Benner, S. A., Church, G. M., Harris, T., Huang, X., Jovanovich, S. B., Nelson, J. R., Schloss, J. A., Schwartz, D. C., et al. (2009) The challenges of sequencing by synthesis. *Nat. Biotechnol.* 27, 1013–1023.

- (24) Ouellet, M., Adams, P. D., Keasling, J. D., and Mukhopadhyay, A. (2009) A rapid and inexpensive labeling method for microarray gene expression analysis. *BMC Biotechnol.* 9, 97.

- (25) Jeon, H., and Choi, S. (2007) Fluorescence quenching causes systematic dye bias in microarray experiments using cyanine dye. *Genomics & Informatics* 5, 113–117.

- (26) Kretschy, N., and Somoza, M. M. (2014) Comparison of the Sequence-Dependent Fluorescence of the Cyanine Dyes Cy3, Cy5, DyLight DY547 and DyLight DY647 on Single-Stranded DNA. *PLoS One* 9, e85605.

- (27) Crockett, A. O., and Wittwer, C. T. (2001) Fluorescein-Labeled Oligonucleotides for Real-Time PCR: Using the Inherent Quenching of Deoxyguanosine Nucleotides. *Anal. Biochem.* 290, 89–97.

- (28) Seidel, C. A. M., Schulz, A., and Sauer, M. H. M. (1996) Nucleobase-specific quenching of fluorescent dyes.1. Nucleobase one-electron redox potentials and their correlation with static and dynamic quenching efficiencies. *J. Phys. Chem.* 100, 5541–5553.

- (29) Widengren, J., Dapprich, J., and Rigler, R. (1997) Fast interactions between Rh6G and dGTP in water studied by fluorescence correlation spectroscopy. *Chem. Phys.* 216, 417–426.

- (30) Shafirovich, V. Y., Courtney, S. H., Ya, N., and Geacintov, N. E. (1995) Proton-Coupled Photoinduced Electron Transfer, Deuterium Isotope Effects, and Fluorescence Quenching in Noncovalent Benzo[*a*]pyrenetetraol-Nucleoside Complexes in Aqueous Solutions. *J. Am. Chem. Soc.* *117*, 4920–4929.
- (31) Olmsted, J., III, and Kearns, D. R. (1977) Mechanism of ethidium bromide fluorescence enhancement on binding to nucleic acids. *Biochemistry* *16*, 3647–3654.
- (32) Heinlein, T., Knemeyer, J.-P., Piestert, O., and Sauer, M. (2003) Photoinduced Electron Transfer between Fluorescent Dyes and Guanosine Residues in DNA-Hairpins. *J. Phys. Chem. B* *107*, 7957–7964.
- (33) Torimura, M., Kurata, S., Yamada, K., Yokomaku, T., Kamagata, Y., Kanagawa, T., and Kurane, R. (2001) Fluorescence-Quenching Phenomenon by Photoinduced Electron Transfer between a Fluorescent Dye and a Nucleotide Base. *Anal. Sci.* *17*, 155–160.
- (34) Pilailuk, S., Chittanon, B., and Chutintorn, P. (2015) Excited state free energy calculations of Cy3 in different environments. *J. Phys. D: Appl. Phys.* *48*, 205401.
- (35) Spiriti, J., Binder, J. K., Levitus, M., and van der Vaart, A. (2011) Cy3-DNA Stacking Interactions Strongly Depend on the Identity of the Terminal Basepair. *Biophys. J.* *100*, 1049–1057.
- (36) Harvey, B. J., and Levitus, M. (2009) Nucleobase-Specific Enhancement of Cy3 Fluorescence. *J. Fluoresc.* *19*, 443–448.
- (37) Mikelsons, L., Carra, C., Shaw, M., Schweitzer, C., and Scaiano, J. C. (2005) Experimental and theoretical study of the interaction of single-stranded DNA homopolymers and a monomethine cyanine dye: nature of specific binding. *Photochem. Photobiol. Sci.* *4*, 798–802.
- (38) Guckian, K. M., Schweitzer, B. A., Ren, R. X.-F., Sheils, C. J., Tahmassebi, D. C., and Kool, E. T. (2000) Factors Contributing to Aromatic Stacking in Water: Evaluation in the Context of DNA. *J. Am. Chem. Soc.* *122*, 2213–2222.
- (39) Crothers, D. M. (1998) DNA curvature and deformation in protein–DNA complexes: A step in the right direction. *Proc. Natl. Acad. Sci. U. S. A.* *95*, 15163–15165.
- (40) Sain, A., Chen, J. Z. Y., and Ha, B.-Y. (2006) Persistence of single-stranded DNA: The interplay between base sequences and base stacking. *Phys. A* *369*, 679–687.
- (41) Solie, T. N., and Schellman, J. A. (1968) The interaction of nucleosides in aqueous solution. *J. Mol. Biol.* *33*, 61–77.
- (42) Doktycz, M. J., Paner, T. M., Amaratunga, M., and Benight, A. S. (1990) Thermodynamic stability of the 5' dangling-ended DNA hairpins formed from sequences 5'-(XY)2GGATAC(T)4GTATCC-3', where X, Y = A,T,G,C. *Biopolymers* *30*, 829–845.
- (43) Turner, D. H., Sugimoto, N., and Freier, S. M. (1988) RNA Structure Prediction. *Annu. Rev. Biophys. Biophys. Chem.* *17*, 167–192.
- (44) Fujii, S., Kono, H., Takenaka, S., Go, N., and Sarai, A. (2007) Sequence-dependent DNA deformability studied using molecular dynamics simulations. *Nucleic Acids Res.* *35*, 6063–6074.
- (45) Packer, M. J., Dauncey, M. P., and Hunter, C. A. (2000) Sequence-dependent DNA structure: dinucleotide conformational maps. *J. Mol. Biol.* *295*, 71–83.
- (46) Yakovchuk, P., Protozanova, E., and Frank-Kamenetskii, M. D. (2006) Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Res.* *34*, 564–574.
- (47) Karunakaran, V., Pérez Lustres, J. L., Zhao, L., Ernsting, N. P., and Seitz, O. (2006) Large Dynamic Stokes Shift of DNA Intercalation Dye Thiazole Orange has Contribution from a High-Frequency Mode. *J. Am. Chem. Soc.* *128*, 2954–2962.
- (48) Duthie, R. S., Kalve, I. M., Samols, S. B., Hamilton, S., Livshin, I., Khot, M., Nampalli, S., Kumar, S., and Fuller, C. W. (2002) Novel Cyanine Dye-Labeled Dideoxynucleoside Triphosphates for DNA Sequencing. *Bioconjugate Chem.* *13*, 699–706.
- (49) Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G., Hall, K. P., Evers, D. J., Barnes, C. L., Bignell, H. R., et al. (2008) Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* *456*, 53–59.
- (50) Singh-Gasson, S., Green, R. D., Yue, Y. J., Nelson, C., Blattner, F., Sussman, M. R., and Cerrina, F. (1999) Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array. *Nat. Biotechnol.* *17*, 974–978.
- (51) Agbawwe, C., Kim, C., Hong, D., Heinrich, K., Wang, T., and Somoza, M. M. (2011) Efficiency, Error and Yield in Light-Directed Maskless Synthesis of DNA Microarrays. *J. Nanobiotechnol.* *9*, 57.
- (52) Kretschy, N., Holik, A.-K., Somoza, V., Stengele, K.-P., and Somoza, M. M. (2015) Next-Generation *o*-Nitrobenzyl Photolabile Groups for Light-Directed Chemistry and Microarray Synthesis. *Angew. Chem., Int. Ed.* *54*, 8555–8559.
- (53) Warren, C. L., Kratochvil, N. C. S., Hauschild, K. E., Foister, S., Brezinski, M. L., Dervan, P. B., Phillips, G. N., and Ansari, A. Z. (2006) Defining the sequence-recognition profile of DNA-binding molecules. *Proc. Natl. Acad. Sci. U. S. A.* *103*, 867–872.
- (54) Lackey, J. G., Mitra, D., Somoza, M. M., Cerrina, F., and Damha, M. J. (2009) Acetal Levulinyl Ester (ALE) Groups for 2'-Hydroxyl Protection of Ribonucleosides in the Synthesis of Oligoribonucleotides on Glass and Microarrays. *J. Am. Chem. Soc.* *131*, 8496–8502.
- (55) McGall, G. H., Barone, A. D., Diggelmann, M., Fodor, S. P. A., Gentalen, E., and Ngo, N. (1997) The efficiency of light-directed synthesis of DNA arrays on glass substrates. *J. Am. Chem. Soc.* *119*, 5081–5090.
- (56) Nuwaysir, E. F., Huang, W., Albert, T. J., Singh, J., Nuwaysir, K., Pitas, A., Richmond, T., Gorski, T., Berg, J. P., Ballin, J., et al. (2002) Gene Expression Analysis Using Oligonucleotide Arrays Produced by Maskless Photolithography. *Genome Res.* *12*, 1749–1755.
- (57) Lackey, J. G., Somoza, M. M., Mitra, D., Cerrina, F., and Damha, M. J. (2009) In-situ chemical synthesis of rU-DNA chimeras on chips and enzymatic recognition. *Chim. Oggi-Chem. Today* *27*, 30–33.
- (58) Mertz, J. (2000) Radiative absorption, fluorescence, and scattering of a classical dipole near a lossless interface: a unified description. *J. Opt. Soc. Am. B* *17*, 1906–1913.
- (59) Schneider, T. D., and Stephens, R. M. (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.* *18*, 6097–6100.