Universal bases for hybridization, replication and chain termination

Markus Berger, Yiqin Wu, Anthony K. Ogawa, Dustin L. McMinn, Peter G. Schultz and Floyd E. Romesberg*

Department of Chemistry, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

Received April 14, 2000; Revised and Accepted June 13, 2000

ABSTRACT

Several unnatural, predominantly hydrophobic nucleobases that pack efficiently in duplex DNA without hydrogen bonding functional groups are reported to circumvent the hydrogen bonding-based specificity, both during oligonucleotide hybridization and enzymatic DNA synthesis. The reported nucleoside analogs are efficient 'universal bases' for hybridization, template directed DNA synthesis and chain termination. Moreover, several of the universal bases function in their biological role, hybridization or replication, with an efficiency not significantly reduced relative to their natural counterparts.

INTRODUCTION

DNA duplex stability results from a combination of interstrand hydrogen bonds and intrastrand hydrophobic interactions between nucleobases (1). However, sequence specificity is determined solely by the specific interstrand hydrogen bonding patterns of the natural bases (1). Hydrophobic base analogs without hydrogen bonding groups (2–4) that pack efficiently in duplex DNA can show little selectivity in pairing with native bases (5). Such 'universal bases' have attracted much attention due to their potential utility in the design of oligonucleotide primers or hybridization probes where the identity of one or more bases in the target sequence is unknown (6–14). In practice, sequence ambiguities occur frequently due to the degeneracies in the genetic code. In cases where sequence data is available, ambiguities can still remain due to polymorphic or species-dependent sequence differences. Therefore, for any given gene, it would be of value to have an oligonucleotide probe that is capable of selective hybridization even in the presence of polymorphisms.

There has been some success in the design of nucleobase analogs that can hybridize non-selectively to each of the native bases (6–14), but analogs that can universally hybridize without significant duplex destabilization are rare. Deoxy-oligonucleotides containing either 3-nitropyrole or 5-nitroindole bases have received the most attention (6,10–13). These nucleoside analogs pair non-discriminately opposite each natural base. Duplex DNA containing 5-nitroindole paired opposite each natural base was shown to melt with only a 3°C range of Tm values. Although these universal bases do not strongly discriminate the native bases, their incorporation destabilizes duplex DNA by at least 4.0–7.0°C per analog base. Significant destabilization limits the number of these analogs that can be incorporated in any given oligonucleotide, thus limiting their practical utility.

In addition to a base for universal hybridization, an unnatural base that is non-discriminantly recognized by a DNA polymerase in the template ('universal template base') (15–19) or as a triphosphate ('universal triphosphate') (20–22) would also be useful for a variety of recombinant DNA techniques, such as library generation, random oligonucleotide labeling and random chain termination. When in the template, the universal nucleobase would direct the polymerase to incorporate each dNTP indiscriminately, and as triphosphate it would be indiscriminately incorporated opposite each native base in the DNA template. The practicality of a universal triphosphate will require that it be randomly incorporated into the priming oligonucleotide in the presence of the natural triphosphates. This competitive insertion will require the unnatural nucleoside triphosphate to be incorporated into DNA with a rate approaching that of native base pair synthesis. A shortcoming of the traditional dideoxy method of chain termination is that four different dideoxynucleotide triphosphates, one corresponding to each base, must be used in great excess to compete with more efficient incorporation of natural nucleoside triphosphates. Excess chain dideoxy terminators are required because they are not incorporated into the growing oligonucleotide strand competitively with the natural substrates. For example, the Klenow fragment of Escherichia coli DNA polymerase I prefers the natural dNTP substrates by several thousand-fold relative to the chain terminator dideoxy analogs (23).

Universal bases for hybridization, replication or random chain termination would be valuable tools for nucleic acid manipulation and analysis, but have proven difficult to design. We have been interested in the design and characterization of unnatural hydrophobic nucleobase analogs which when present in a given biological context (oligonucleotide hybridization or DNA replication) behave universally. Herein we report nucleobase analogs that act as universal bases in each biological context. We describe a base analog capable of pairing stably, but non-specifically, with each natural base. We also describe a base analog which when present in a DNA template directs a DNA polymerase to randomly insert a natural nucleobase triphosphate. Finally, an analog triphosphate is described which is efficiently inserted opposite any natural base in template with

*To whom correspondence should be addressed. Tel: +1 858 784 7290; Fax: +1 858 784 7472; Email: floyd@scripps.edu
a rate approaching that of correct natural triphosphate insertion and subsequently causes chain termination.

MATERIALS AND METHODS

The nucleoside of 3-methyl 7-propynyl isocarbostyril (PIM) (Fig. 1) was synthesized as reported previously (24). Synthesis of the 3-methyl isocarbostyril (MICS), and 5-methyl isocarbostyril (5MICS) nucleosides (Fig. 1) will be reported elsewhere (accepted for publication in Journal of the American Chemical Society). The nucleosides were converted into the triphosphate (3,25) and phosphoramidite (3) by literature methods. The thermal stability of the base pairs was evaluated by determining the melting temperature \( T_m \) of duplex DNA containing MICS or 5MICS paired with dG, dC, dA or dT (Table 1). The melting experiments were done with 3 µM duplex in 10 mM Pipes, 10 mM MgCl₂, 100 mM NaCl, pH 7, using a Cary 300 Bio UV/Vis spectrophotometer. The heating rate was 0.5°C/min between 16 and 80°C. Melting temperatures were obtained by the derivative method utilizing the Cary Win UV thermal application software. MICS and PIM were evaluated as substrates for the exonuclease-deficient Klenow fragment of E.coli DNA polymerase I (KF) purchased from Amersham Pharmacia Biotech (Piscataway, NJ). Initial velocities were determined during extension of a γ-32P-labeled primer with varying concentrations of nucleoside triphosphates (26). The reactions were analyzed by polyacrylamide gel electrophoresis; a Phosphorimager (Molecular Dynamics, Sunnyvale, CA) was used to quantify gel band intensities corresponding to the extended primer. The measured velocities were plotted against the concentration of dNTP and subsequently fitted to the Michaelis–Menten equation. Unnatural nucleobases were assayed both in template DNA and as incoming nucleoside triphosphates. Steady-state kinetic parameters for single incorporation of dGTP, dATP, dTTP or dCTP opposite MICS are reported in Table 2. Steady-state kinetic parameters for single incorporation of dPIMTP opposite dG, dA, dT or dC are reported in Table 3.

RESULTS

We have investigated a wide variety of unnatural hydrophobic nucleobases that do not have hydrogen bonding functionality (5,24,27). We systematically evaluated the stability of duplex DNAs containing these unnatural bases paired opposite native bases. Several of the unnatural bases show little thermal discrimination against the different natural bases and do not compromise duplex stability. In addition, we have evaluated the ability of several polymerases to synthesize DNA containing the unnatural bases both in template and as triphosphates. Several of these base analogs are efficiently, but non-discriminantly, accepted as DNA polymerase substrates.

Table 1. \( T_m \) values for duplex containing MICS and 5MICS

<table>
<thead>
<tr>
<th>Duplex</th>
<th>N</th>
<th>X</th>
<th>( T_m ) (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5′-GGATACGTGACATGCG</td>
<td>A</td>
<td>MICS</td>
<td>53.6</td>
</tr>
<tr>
<td>3′-GGATACGTGACATGCG</td>
<td>T</td>
<td>MICS</td>
<td>55.1</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>MICS</td>
<td>54.8</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>MICS</td>
<td>54.8</td>
</tr>
<tr>
<td>MICS</td>
<td></td>
<td>Average = 54.6 ± 0.7°C</td>
<td></td>
</tr>
<tr>
<td>5MICS</td>
<td></td>
<td>Average = 55.3 ± 0.6°C</td>
<td></td>
</tr>
<tr>
<td>5′-GGATACGTGACATGCG</td>
<td>A</td>
<td>MICS</td>
<td>54.2</td>
</tr>
<tr>
<td>3′-GGATACGTGACATGCG</td>
<td>T</td>
<td>MICS</td>
<td>51.9</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>MICS</td>
<td>55.2</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>MICS</td>
<td>51.2</td>
</tr>
<tr>
<td>MICS</td>
<td></td>
<td>Average = 53.1± 1.9°C</td>
<td></td>
</tr>
<tr>
<td>5MICS</td>
<td></td>
<td>Average = 52.6± 2.2°C</td>
<td></td>
</tr>
</tbody>
</table>

Experiments were run in triplicate. See text for experimental details.

Table 2. Steady-state kinetic parameters for KF exo−-mediated synthesis of DNA with MICS in the template

<table>
<thead>
<tr>
<th>Template (X)</th>
<th>Nucleoside triphosphate</th>
<th>( k_{cat} ) (min⁻¹)</th>
<th>( K_M ) (µM)</th>
<th>( k_{cat}/K_M ) (M⁻¹ min⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MICS</td>
<td>dATP</td>
<td>0.36 ± 0.02</td>
<td>15 ± 2</td>
<td>2.4 × 10⁴</td>
</tr>
<tr>
<td></td>
<td>dGTP</td>
<td>0.28 ± 0.02</td>
<td>48 ± 6</td>
<td>5.8 × 10⁴</td>
</tr>
<tr>
<td></td>
<td>dCTP</td>
<td>0.27 ± 0.03</td>
<td>43 ± 15</td>
<td>6.3 × 10³</td>
</tr>
<tr>
<td></td>
<td>dTTP</td>
<td>1.6 ± 0.1</td>
<td>80 ± 7</td>
<td>2.0 × 10⁴</td>
</tr>
</tbody>
</table>

5′-dTATACGACTCAGCTATAGGAGA
3′-dATTAAGCTGCAAGTGATATCCCTCTGCTGCA

Experiments were run in triplicate. See text for experimental details.

Universal bases for oligonucleotide hybridization

In addition to interstrand hydrogen bonding between pairing bases, duplex DNA is stabilized by intrastand base interactions, including dipole–dipole interactions as well as nucleobase polarizability (28–31). An unnatural nucleobase, lacking hydrogen bond donor or acceptor functionalities, which can sufficiently stabilize duplex DNA by intrastand base interactions may act as a universal base, without destabilizing the
duplex. A universal base must also pair indiscriminately in any sequence context. For example, there are large effects on natural and unnatural base pair stability when the flanking bases are changed from purines to pyrimidines and the universal bases must behave indiscriminately in each sequence context.

Table 3. Steady-state kinetic parameters for KF exo-mediated insertion of dPIMP opposite dA, dG, dC and dT in the template

<table>
<thead>
<tr>
<th>Template (X)</th>
<th>( k_{cat} ) (min(^{-1}))</th>
<th>( K_M ) (( \mu )M)</th>
<th>( k_{cat}/K_M ) (M(^{-1}) min(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>dA</td>
<td>5.3 ± 0.2</td>
<td>3.1 ± 0.5</td>
<td>1.7 × 10(^6)</td>
</tr>
<tr>
<td>dG</td>
<td>1.33 ± 0.02</td>
<td>4.1 ± 0.3</td>
<td>3.2 × 10(^4)</td>
</tr>
<tr>
<td>dC</td>
<td>3.2 ± 0.2</td>
<td>4.8 ± 0.6</td>
<td>6.7 × 10(^4)</td>
</tr>
<tr>
<td>dT</td>
<td>11.0 ± 0.5</td>
<td>6.6 ± 1.0</td>
<td>1.7 × 10(^6)</td>
</tr>
</tbody>
</table>

5'-dTAATACGACTCACTATAGGGAGA
3'-dGATATCGACTGAGTAGATATATCCCTCTGTGATCA

Experiments were run in triplicate. dNTP concentrations were varied from 0.1 to 100 \( \mu \)M. See text for experimental details.

The MICS and 5MICS nucleobases (Fig. 1) have polarizable N-glycosidic linkages to a hydrophilic minor groove carbonyl group and are expected to present a large and polarizable hydrophobic surface for intrabase packing. The data in Table 1 demonstrates that these predominantly hydrophobic bases form stable base pairs with each of the natural bases when incorporated into the central position of either oligodeoxyribonucleotide strand of a 13mer duplex (Table 1). When incorporated into one oligonucleotide strand and paired opposite each of the four natural bases, the average \( T_m \) values for MICS and 5MICS are 54.6 ± 0.7 and 55.3 ± 0.6°C, respectively (Table 1). This compares favorably with the stability of the duplex with a native dT-dA pair (\( T_m = 58.7°C \)). The spread in measured \( T_m \) values for MICS and 5MICS varies across the four natural bases which are largely incapable of forming hydrogen bonds have been shown to be surprisingly good substrates for KF (5,27,32,33). These hydrophobic bases bear no obvious shape complementarity to the native bases, but are incorporated with rates that in some cases are equal to those observed for correct pair synthesis. We recently reported the synthesis and thermal characterization of the unnatural nucleobase PIM (Fig. 1; 24). We were interested in examining PIM as a substrate for KF. Steady-state kinetic experiments were conducted with exonuclease-deficient KF (26) and the rates for the single nucleotide primer extension are reported in Table 3. The universal behavior of dPIMP in this sequence context is evident from the relative efficiencies (\( k_{cat}/K_M \)) for incorporation opposite each natural base (Table 3), which vary by only 5-fold. The insertion of dPIMP opposite dA and dT is slightly more efficient than insertion opposite dC or dG. Remarkably, the average rate of dPIMP insertion is only 40-fold reduced relative to the insertion of dTTP opposite dA (\( k_{cat}/K_M = 4.7 \times 10^7 \) M\(^{-1}\) min\(^{-1}\)) in the same sequence context.
Figure 3. Random chain termination with dPIMTP. Assay conditions were as follows: 40 nM DNA duplex, 1.34 nM KF exo–, 0–100 µM dPIMTP and 1 µM dNTPs. Reactions were incubated at room temperature for 1 min.

After incorporation of dPIMTP, KF is unable to continue DNA synthesis. Even under forcing conditions (0.5 h incubation time and 1 mM dNTPs) no band corresponding to extension of the PIM-terminated primer was detectable. Therefore, PIM inserts efficiently, but randomly into the growing oligonucleotide strand and then terminates synthesis. The average length of synthesized oligonucleotide may be tuned by choice of chain terminator concentration (Fig. 3). The efficient insertion of dPIMTP opposite any native template results in the efficient generation of random length oligonucleotides, with low concentrations of a single chain terminator. This strongly contrasts with the traditional method of dideoxy based termination, as discussed above. In this regard, dPIMTP may have practical applications, for example in the generation of random length oligonucleotides for sequencing by mass spectrometry (34–36).

This study demonstrates that hydrophobicity may act as an ambiguous but sufficiently strong force to mediate the interbase interactions involved in hybridization and replication. We are currently exploring derivatives of these bases, which may be even more universal with regard to kinetic or thermodynamic behavior. The ability of oligonucleotides to form relatively stable duplexes with a variety of predominantly hydrophobic bases paired opposite native bases as well as the remarkable ability of DNA polymerases to accept unnatural bases bearing little or no hydrogen bond or shape complementarity to native bases demonstrates the potential of the approach. Moreover, the hydrophobicity-based approach is not limited to the isocarbostyril ring structure common to PIM, MICS and 5MICS and we are continuing to examine alternative scaffolds. Experiments are underway to evaluate the ability of oligonucleotides containing multiple MICS or 5MICS substitutions to hybridize to target sequences with multiple degeneracies. We are also evaluating the utility of oligonucleotides containing these base analogs in PCR-based applications.

ACKNOWLEDGMENTS

Funding was provided by the National Institutes of Health (GM 60005 to F.E.R.) and the Skaggs Institute for Chemical Biology (F.E.R. and P.G.S.) and a National Institutes of Health post-doctoral fellowship (F32 GM19833-01 to A.K.O.)

REFERENCES