A Steroid Cap Adjusts the Selectivity and Accelerates the Rates of Nonenzymatic Single Nucleotide Extensions of an Oligonucleotide

Jan A. Rojas Stütz and Clemens Richert*

Department of Chemistry, University of Constance
Fach M709, 78457 Konstanz, Germany

Received June 13, 2001

The template-directed extension of polynucleotide chains by individual residues underlies DNA replication. In nature, this reaction is catalyzed by DNA polymerases.[1] These are unusual enzymes that accept four different substrates. Polymerases are also employed in key biomedical applications, including sequencing, PCR amplification, and SNP genotyping.[2] The fidelity of polymerase-catalyzed replication is critical both for the integrity of genomes and for any of the applications named above.[3] The selectivity of nucleotide−oligonucleotide interactions alone is insufficient to provide the fidelity required even for small genomes.[4]

Nonenzymatic, template-directed synthesis of oligonucleotides has been demonstrated in a number of molecular systems.[5] In all cases, the replication requires high concentrations of the reactants, long reaction times, and certain sequences. Since RNA probably had a pivotal role in prebiotic evolution,[6] and ribonucleotide primers are more reactive than their 2′-deoxyribonucleotide counterparts, many studies focused on RNA-based systems. But, as in transcription, DNA can also act as a template for the formation of complementary RNA strands. However, dinucleotides such as AA, AT, TA, GA, and AG in a DNA template cannot be replicated with sufficient rate and fidelity.[7] In 1993, Hill, Orgel, and Wu concluded that "the origin of polynucleotide replication probably involved catalysts that "smoothed away" the sequence-dependent differences in configuration that occurred at the end of growing chains."[8] Though this view has been challenged,[9] T, U, and A residues in the template still produce obstacles that are difficult to surmount in nonenzymatic replication. Efficient self-replicating systems based on DNA are also limited to sequences with templating G and C residues.[10]

Here we report on 5′-acylamido-substituted templates that are capable of accelerating the incorporation of deoxyadenosine, deoxyguanosine, and thymidine residues in single nucleotide extensions. The design of the 5′-modified templates was prompted by a study on the effect of 5′-appended cholic acid residues on terminal T:A base pairs of DNA duplexes.[11] Their stabilizing and mismatch suppressing effect was also found for C:G base pairs.[12] It was suspected that the cholic acid residue could promote the association between templating bases and activated residues.

For our study, we adapted a monitoring technique for nuclease selection experiments[13] employing quantitative MALDI-TOF mass spectrometry.[14] Since extension products can be detected individually, the technique allowed for assays in which four activated nucleotides compete for the primer in one solution. Our assays employed DNA templates of general sequence 5′-R-B*ACGTGCG-3′, where R is an acyl group and B* is a 5′-amino-2′,5′-dideoxynucleotide residue, (**1a**−**g**, **5a**,**g**, **6a**,**g**), and 5′-CGCACGT^{+}-3′ as primer (**2**), where T^{+} is a 3′-amino-3′-deoxythymidine residue (Scheme 1). The amino-terminal primer is isoelectronic and isosteric to its deoxyribo counterpart, but reacts faster with imidazolides.[5a,15] It was prepared from 3′-azido-3′-deoxythymidine via **4t**, as described in the Supporting Information.[16] Templates **1a**−**6g** were synthesized using known protocols.[17] The 5′-acyl groups tested included arginine-containing residues (**1b**, **1c**) capable of activating **3a**−**t** toward nucleophilic attack, an electron-rich arene (**1d**), a quinolone previously shown to stabilize T:A-terminated duplexes (**1e**),[13] an electron-deficient arene (**1f**), and the cholic acid residue (**1g**, **5g**, **6g**). The duplex of **1a** and **2** has a calculated melting point of 58 °C at 0.1 mM strand concentration, leaving practically no single strands under our assay conditions (23 °C, 0.2−0.4 mM strands). The synthesis of methylimidazolides **3a**−**t**[18,15] and of the 5′-terminal residues for preparing **5a**, **5g**, **6a**, and **6g** followed literature precedents.[19]

Calculated rate constants for extension reactions are listed in Table 1. Compared to the acetyl group in **1a**, all 5′-acyl groups in **1b**−**g** accelerated the extensions, with up to 5-fold shorter half-life times for primer **2**. However, cholic acid-bearing template **1g** was the only residue making the reaction more specific. With **1e** as the template, the error rate was 89%, compared to 42% for **1a**, while cholic acid-bearing **1g** gave 27%. The rate of the desired extension (from **1g**:**2** to **1g**:**4a**) was accelerated 6.5-fold. To exclude that this required the presence of the other nucleotides (accelerating effects of mixtures have been described for ribonucleotide-based systems)[20] we performed control experiments with **1g**:**2**, or **2** alone, and **3a** as the only activated nucleotide. With single stranded **2**, only about 60% conversion was found after 3 d (numerical half-life time 45 h). With **1g**:**2**, complete conversion of **2** to **4a** occurred with a half-life time of 1.9 ± 0.2 h, confirming the effect found in the competitive reaction.

With **5a** and **5g** as cytosine-containing templates, the expected[5] rapid extension with **5a** was further accelerated by the cholic acid residue in **5g**. A 5.2-fold rate increase for the overall reaction and a half-life time of the primer of under 35 min were measured. The selectivity of the reaction increased from 32% to 23% non-

* To whom correspondence should be addressed. Phone: int-49 (0)7531 88-4572. Fax: int-49 (0)7531 88-4573. E-mail: Clemens.Richert@uni-konstanz.de.

(1) Kunkel, T. A.; Bebenek, K. *Annu. Rev. Biochem.* **2000**, *69*, 497−529.
(2) Selected references: (a) Higgins, G. S.; Little, D. P.; Köster, H. *BioTechniques* **1997**, *23*, 710−714. (b) Ross, P.; Hall, L.; Smirnov, I.; Haff, L. *Nature Biotechnol.* **1998**, *16*, 1347−1351. (c) Zhang, S.; Van Pelt, C. K.; Schultz, G. A. *Anal. Chem.* **2001**, *73*, 2117−2125.
(3) For a polymerase with low fidelity see e.g.: Showalter, A. K.; Tsai, M.-D. *J. Am. Chem. Soc.* **2001**, *123*, 1776−1777.
(4) Loeb, L. A.; Kunkel, T. A. *Annu. Rev. Biochem.* **1982**, *52*, 429−457.
(5) (a) Orgel, L. E.; Lohrmann, R. *Acc. Chem. Res.* **1974**, *7*, 368−377. (b) Joyce, G. F.; Orgel, L. E., *J. Mol. Biol.* **1988**, *202*, 677−681. (c) Orgel, L. E. *Nature* **1992**, *358*, 203−209. (d) Sievers, D.; von Kiedrowski, G. *Nature* **1994**, *369*, 221−224. (e) Kanavarioti, A.; Bernasconi, C. F.; Baird, E. E. *J. Am. Chem. Soc.* **1998**, *120*, 8575−8581. (f) Bolli, M.; Micura, R.; Eschenmoser, A. *Chem. Biol.* **1997**, *4*, 309−320.
(6) Gesteland, R. F.; Cech, T. R.; Atkins, J. F., Eds. *The RNA World*, 2nd ed.; Cold Spring Harbor Press: Cold Spring Harbor, NY 1999.
(7) Wu, T.; Orgel, L. E. *J. Am. Chem. Soc.* **1992**, *114*, 7963−7969.
(8) Hill, A. R.; Orgel, L. E.; Wu, T. *Origins Life Evol. Biosphere* **1993**, *23*, 291−297.
(9) Kurz, M.; Göbel, K.; Hartel, C.; Göbel, M. W. *Helv. Chim. Acta* **1998**, *81*, 1156−1180.
(10) (a) Bag, B. G.; von Kiedrowski, G. *Pure Appl. Chem.* **1996**, *68*, 2145−2152. (b) Luther, A.; Brandsch, R.; von Kiedrowski, G. *Nature* **1998**, *396*, 245−248.
(11) Bleczinski, C. F.; Richert, C. *J. Am. Chem. Soc.* **1999**, *121*, 10889−10894.
(12) Mokhir, A. A.; Tetzlaff, C. N.; Herzberger, S.; Mosbacher, A.; Richert, C. *J. Combin. Chem.* **2001**, *3*, 374−386.
(13) Altman, R. K.; Schwope, I.; Sarracino, D. A.; Tetzlaff, C. N.; Bleczinski, C. F.; Richert, C. *J. Combin. Chem.* **1999**, *1*, 493−508.
(14) Sarracino, D.; Richert, C. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2543−2548.
(15) Kanavarioti, A.; Stronach, M. W.; Ketner, R. J.; Hurley, T. B. *J. Org. Chem.* **1995**, *60*, 632−637.
(16) Related syntheses: (a) Gryaznov, S. M.; Letsinger, R. L. *Tetrahedron Lett.* **1993**, *34*, 1261−1264. (b) Bruick, R. K.; Koppitz, M.; Joyce, G. F.; Orgel, L. E. A *Nucleic Acids Res.* **1997**, *25*, 1309−1310.
(17) Tetzlaff, C. N.; Schwope, I.; Bleczinski, C. F.; Steinberg, J. A.; Richert, C. *Tetrahedron Lett.* **1998**, *39*, 4215−4218.
(18) (a) Lohrmann, R.; Orgel, L. E. *Tetrahedron Lett.* **1978**, *34*, 853−855. (b) Joyce, G. F.; Inoue, T.; Orgel, L. E. *J. Mol. Biol.* **1984**, *176*, 279−306.
(19) Mag, M.; Engels, J. W. *Nucleic Acids Res.* **1989**, *17*, 5973−5988.
(20) Kanavarioti, A. *J. Mol. Evol.* **1998**, *46*, 622−632.

10.1021/ja011448i CCC: $20.00
Published on Web 11/20/2001

Scheme 1

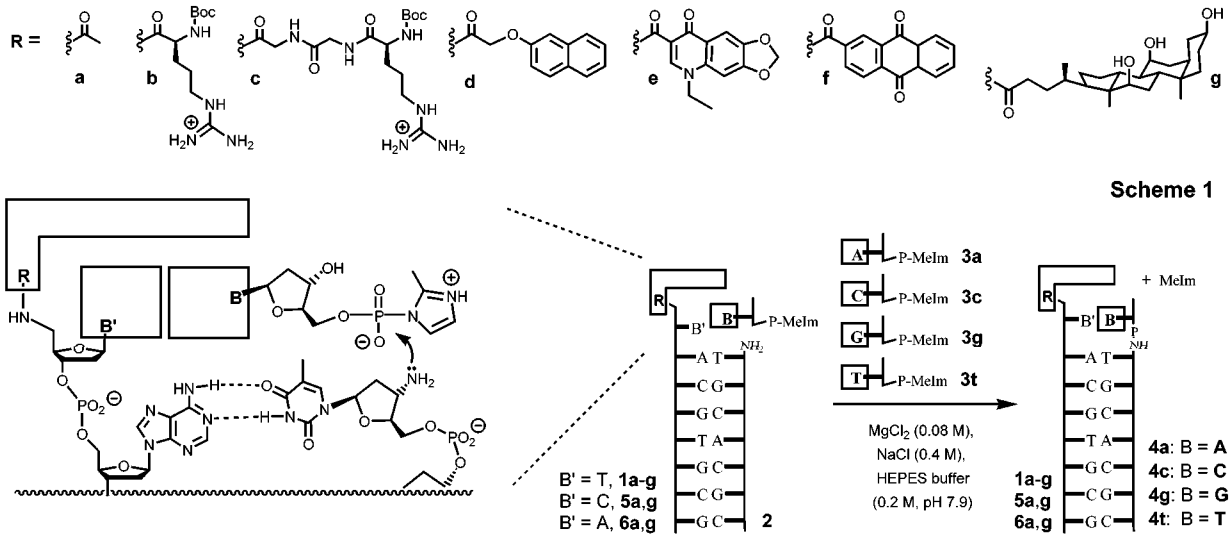

Table 1. Kinetic Parameters for Extension Reactions[a]

duplex	$t_{1/2}^{primer}$ (h)	k^A	k^C	k^G	k^T	errors (%)
1a:**2**	20.8	19	3	9	2	42
1b:**2**	15.4	20	5	18	3	56
1c:**2**	12.8	24	6	21	3	56
1d:**2**	17.9	15	5	15	3	60
1e:**2**	5.5	14	8	102	3	89
1f:**2**	10.5	24	9	28	5	64
1g:**2**	4.1	123	7	35	4	27
5a:**2**	2.8	51	16	169	13	32
5g:**2**	0.54	188	59	991	48	23
6a:**2**	11.1	14	7	15	27	57
6g:**2**	4.0	13	10	21	129	25

[a] Rate constants in 10^{-3} h^{-1}, calculated from results of fits to kinetics; second-order rate constants may be calculated by dividing by c(**3n**) (0.0193 M). See Supporting Information for a detailed table with standard deviations.

Watson−Crick incorporations. For reactions with **6a** and **6g**, whose templating base is an adenine, the overall rate acceleration was 2.8-fold and the error rate was reduced from 57% to 25%. Thus, the steroid residue adjusted the fidelity of thymine-, cytosine-, and adenine-templated reactions to very similar values. The rates of incorporation still differ between the T- or A-templated reactions and the C-templated reaction (by a factor of up to 9.8), but so do the rates of polymerase-catalyzed reactions.[21]

The error rates observed in our system with its single templating nucleotide are higher than the 1−10% predicted for polymerase-free systems.[4] Since DNA octamer duplexes where a 5′-terminal T residue faces a mismatched base show melting point decreases of less than 3 °C over the fully matched control duplex,[11] with a $\Delta\Delta G°$ of ≤1 kcal/mol for duplex formation, this result is not surprising. Unlike the terminal residue in the melting experiments, the activated nucleotides are not preoriented by a covalent link. The G-specific reaction with quinolone-bearing **1e** confirms that changing the templating effect of the thymidine does not necessarily decelerate the reaction. In fact, for neither of the aromatic residues tested, an inhibitory effect that could have resulted from stacking on the last intact base pair of the duplex (blocking the 3′-amine of the primer) was observed.

Though our molecular system, in its current form, is unsuitable for catalyzing extensions by more than one residue, the results are encouraging in the context of simple self-replicating systems. If a natural product as small as cholic acid can accelerate extension reactions and smooth out a fidelity bias toward C-templates, there is hope that the recruitment of cofactors can improve simple replication systems, avoiding a drift toward high G/C-content. Synthetic methodology for the preparation of 5′-aminoacyl oligoribonucleotides is currently being developed.[22] These oligoribonucleotides may allow an extension to all-RNA systems.

Since the reaction rates measured are far from the diffusion limit and the cholic acid residue is devoid of cationic groups providing electrostatic attraction or electrostatic catalysis, there is little doubt that the catalytic microenvironment set up by this residue can be substantially improved. Since our results also demonstrate that the cholic acid "cap" exerts its accelerating effect on more than one substrate/template pair, the combination of promiscuity and fidelity also found in polymerases may be retained during this optimization. Further extension of this work to "molecular caps" anchored noncovalently on primer-template duplexes is planned.

Acknowledgment. We thank U. Steiner for help with the analysis of kinetics, W. Connors, S. Narayanan, S. Herzberger, and A. Mokhir for synthetic intermediates, and G. Charubala and C. Tetzlaff for helpful protocols. Supported by DFG (grant RI 1063/1).

Supporting Information Available: Protocols, representative MALDI spectra, calibration plots for detection of **4a**−**t**, and representative kinetics plots (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA011448I

(21) Davenport, R. J.; Wuite, G. J. L.; Landick, R.; Bustamante, C. *Science* **2000**, *287*, 2497−2500.

(22) Tetzlaff, C. N.; Richert, C. *Tetrahedron Lett.* **2001**, *42*, 5681−5684.