## Nonenzymatic Oligomerization of Ribonucleotides: Towards in vitro Selection Experiments

by Marcus Hey, Christian Hartel, and Michael W. Göbel\*

Department of Organic Chemistry and Chemical Biology, Johann Wolfgang Goethe-University Frankfurt, Marie Curie-Strasse 11, D-60439 Frankfurt

Inspired by the polymerase chain reaction, orthogonal primer-template pairs have been applied in template-controlled oligomerization experiments with Orgel's imidazole-activated ribonucleotide-5'-phosphates. Variation of the linker length allowed us to monitor the extension of both primers simultaneously on a DNA sequencer. Sets of hexapyrimidine primers were found that are capable of inducing the reciprocal synthesis of each other's binding site. Considerable cross-inhibition by different monomers was observed. However, this effect is a function of primer sequence and can disappear in favorable cases. With random sequences introduced into the templates, selection experiments are within reach. First results are reported below.

Introduction. - Demonstrating a potential of RNA to self-replicate and to undergo Darwinian evolution in the absence of enzymes would be important to find scientific explanations of the emergence of life. Numerous studies, therefore, are devoted to the problem of template-controlled, nonenzymatic oligomerization of ribonucleotides  $[1][2]$  and their structural alternatives  $[3-8]$ . With unmodified ribonucleotides, the most promising results have been obtained so far with *Orgel's* nucleoside-5'-phosphates activated as 2-methyl-1H-imidazolides (2MeImp; see Scheme 1). Guided by the basepairing rules, these monomers form double-helical aggregates with single-stranded DNA or RNA templates;  $Mg^{2+}$  then catalyzes the reaction of phosphates and 3'-OH groups leading to  $3'$ , 5'-connected double helices. Deriving kinetic data from the resulting complex mixtures is not always simple. In most of the recent work, primer extension has been studied instead of spontaneous monomer oligomerization via dimers, trimers etc. With radioactively labeled primers, the analysis detects the interesting products selectively. Combining primer and template in a hairpin structure can offer further advantages  $[9][2]$ . In spite of all efforts and impressive results, selfreplication has not yet been observed in such experiments.

In our laboratory, two experimental setups have been developed to analyze template-controlled RNA oligomerization. The first one relies on HPLC and applies short primers labeled with lipophilic acridine dyes [10] [11]. It is limited to the incorporation of just a few monomers. The newer one is based on standard DNA sequencers and is not much limitated with respect to product chain length or sensitivity [12]. Both techniques have led to very similar results. They have been applied to study a number of effects known to hamper the self-replication of RNA. Three major conclusions could be drawn:  $I$ ) Primer extension is most efficient if either the template or the primer or both consist of RNA: the double helical aggregate should adopt an Atype conformation. 2) The oligomerization does not depend on high  $Na<sup>+</sup>$  concenScheme 1. Template-Controlled Primer Extension. Ribonucleotides are indicated by black deoxyribonucleotides by grey letters



trations. Low  $Na<sup>+</sup>$  conditions prevent guanosine-rich templates from quadruplex formation, thus allowing the assembly of oligo-C on oligo-G strands. 3) While the template-controlled incorporation of adenosine is a sluggish process, far better yields are observed in reactions of the 2,6-diaminopurine analogue capable of forming three H-bonds with T or U  $[3][12]$  (see also  $[13][14]$ ).

Are these improvements sufficient for self-replication? In the following, we will present our recent experiments. New limitations have been identified, and the answer is still negative. Some results, however, are highly encouraging, leading to a fair prognosis for selection experiments based on random templates and  $O$ rgel's monomeric building blocks.

The general concept of selection experiments is shown in Schemes 2 and 3. Analogous to the polymerase chain reaction, two orthogonal primer-template pairs are used as a starting point (Scheme 2). While the 3'-part of each template binds its own primer, the 5-part has the same sequence as the other primer. Identical primers would lead to palindromic duplexes of two template molecules. It is essential, therefore, to work with nonidentical primers. Since, in template-controlled oligomerizations, the last nucleotide is incorporated with reduced rates and yields [9], a primer length of six nucleotides was chosen to guarantee at least five base pairs in the primer-template duplex of the backward reaction. This number was proved sufficient in former experiments. Best yields are generally obtained when homopyrimidine sequences are copied into homo-purine strands. Consideration of these aspects finally led to the





Scheme 3. Concept of a Selection Experiment Based on Random Templates



 $846$  HE

selection of hexapyrimidine primers. For selection studies, a short random sequence has to be inserted into the template. After the forward and backward extension of the two orthogonal primers, the small fraction of the fully extended material could indicate which sequences are suitable for self-replication (Scheme 3).

**Results.** – *Experiments 1 – 3 (Tables 1 and 2, and Fig.)* demonstrate the feasibility of orthogonal primer – template pairs in the nonenzymatic RNA oligomerization. When a mixture of both primers is incubated with just one of the templates, only the complementary strand is extended. The different linker lengths of the primers allow us to monitor both reactions simultaneously. As expected, in the presence of both templates, complete elongation of both primers is observed. The formation of small numbers of chains containing one surplus nucleotide is a phenomenon well-known from previous studies [9] [11].



Figure. Experiment 3: 'One-pot'-oligomerization in the presence of different template combinations. A special linker design allows us to monitor both primer extensions simultaneously (conc. of G\*: 50 mM).









Table 2. Experiment 2: Elongation of Primer 5'-CUCCCC (conc. of G\*: 50 mM)

**SECTION CONTINUES OF STATISTICS** 

In Experiments 4 and 5 (Tables 3 and 4, resp.), the capability of primers  $5'$ -CCUUCC and 5'-CUCCCC to induce the reciprocal synthesis of each other's binding site was checked. The tables show, in accordance with earlier observations, that a single T in the template may induce misincorporation of  $G^*$  by wobble pairing [9] [12]. When the template contains two consecutive Ts, successful primer extension requires the 2,6 diaminopurine nucleotide D\*. Both reactions proceed to full primer extension. Compared to *Experiments 1 – 3*, on the other hand, the yield is somewhat disappointing.

Table 3. Experiment 4: Reciprocal Synthesis of Primer Binding Sites, Primer 5'-CCUUCC (conc. of G\* 50 mm, conc.  $A^*$ ,  $G^*$ , and  $D^*$ ,  $G^*$  25 mm)

	$G-G-C-C-C-C-T-C5$ $C - C - C - D - D - C - C$ Т6 Linker									
Time	${\bf X}^*$	Primer	$P+I$	$P + II$	$P + III$	$P + IV$	$P + V$	$P + VI$		
1 <sub>d</sub>	$G^*$	5.5	1.0	1.5	25.0	61.5	4.0	1.5		
	$A^* / G^*$	37.0	4.5	8.5	24.5	19.0	6.5			
	$D^* / G^*$	34.5	4.0	5.5	11.5	11.0	33.5			
2 d	$G^*$	4.0	0.5	1.0	9.5	69.0	13.5	2.5		
	$A^* / G^*$	20.0	4.0	5.5	19.5	27.5	21.5	2.0		
	$D^* / G^*$	25.5	4.0	3.5	6.5	8.5	50.0	2.0		
3 d	$G^*$	2.5	0.5	0.5	6.0	70.0	15.5	5.0		
	$A^*$ / $G^*$	15.0	3.5	4.0	15.0	28.5	23.5	10.5		
	$D^* / G^*$	22.0	4.0	2.5	5.0	6.0	51.0	9.5		
7 d	$G^*$	1.5	0.5	0.5	2.0	60.5	21.5	13.5		
	$A^* / G^*$	9.0	3.0	2.5	8.0	26.5	29.0	22.0		
	$D^* / G^*$	21.0	5.5	2.0	4.0	4.0	39.5	24.0		
14 d	$G^*$	1.5	0.5	0.5	1.5	55.0	24.5	16.5		
	$A^* / G^*$	6.0	2.5	1.5	4.5	19.0	22.5	44.0		
	$D^* / G^*$	16.0	6.0	2.0	3.5	3.5	27.0	42.0		

A clue to understanding this observation comes from Experiment 6 (Table 5). An analogue of primer 5-CUCCCC is used to study the incorporation of G\*. When only G\* is present as monomeric building block, the extension yield is excellent. Addition of a second type of nucleotide, however, induces significant inhibition. This effect is most pronounced in the presence of  $D^*$ . Analogous experiments with primer 5'-CCUUCC

 $848$  HE



	$T_{21}$ Linker									
Time	${\bf X}^*$	Primer	$P+I$	$P + II$	$P + III$	$P + IV$	$P + V$	$P + VI$		
1 d	$G^*$	30.5	39.0	29.0	1.5					
	$A^*/G^*$	63.0	21.0	14.5	1.5					
	$D^* / G^*$	56.0	13.5	6.5	4.5	7.0	9.5	3.0		
2d	$G^*$	17.5	30.0	48.5	4.0					
	$A^*/G^*$	47.0	21.0	25.0	5.5	1.5				
	$D^* / G^*$	43.0	12.5	6.0	5.0	8.0	16.0	9.5		
3 d	$\mathbf{G}^*$	14.0	24.0	55.5	6.5					
	$A^*/G^*$	39.5	20.0	28.5	9.0	3.0				
	$D^* / G^*$	38.5	12.0	5.0	4.0	7.5	18.0	15.0		
7 d	$\mathbf{G}^*$	11.0	16.5	58.5	12.5	1.5				
	$A^*/G^*$	26.5	17.0	31.0	15.5	7.0	3.0			
	$D^* / G^*$	26.5	11.0	5.0	4.0	6.5	15.5	31.5		
14 d	$\mathbf{G}^*$	7.0	12.0	59.5	18.5	2.5	0.5			
	$A^*/G^*$	21.0	15.0	29.0	18.0	9.0	4.5	3.5		
	$D^* / G^*$	22.0	11.5	5.5	4.5	6.0	12.0	38.5		

Table 5. Experiment 6: Cross-Inhibition Induced by a Second Monomeric Building Block (conc. of  $G^*$  50 mm, conc. of  $A^*$ ,  $G^*$ , and  $D^*$ ,  $G^*$  25 mm)





led to the same conclusion. Inhibition might be an unavoidable consequence when  $G^*$ and D\* monomers are mixed. Alternatively, it could correlate with the primer sequence. The second explanation turned out to be applicable: the assay, when repeated with the primer sequence 5'-CGCACG (*Experiment 7, Table 6*), did not show significant cross-inhibition! To apply this advantageous primer in  ${}^{\circ}$ PCR-type<sup>\*</sup> RNA oligomerizations, the template-controlled assembly of the corresponding primer binding site must also proceed well. Unfortunately, the low yield of the latter process strongly disfavors the use of 5-CGCACG.

Experiment 7 (Table  $6$ ) points out that hexapyrimidine primers may not be the optimal choice. Nevertheless, it is shown below that reasonable oligomerization yields can be obtained with 5'-CCCCCC and 5'-CCUUCC as the orthogonal primer pair. When 5'-CCCCCC is extended by 5'-GGDDGG, the combined yield of the penta- and hexafold elongated chains exceeds 50% (*Experiment 8*; Table 7). The assembly of GGGGGG on primer 5-CCUUCC runs even better (Experiment 9; Table 8; above 70% combined yield). Again, mixing G and D nucleotides results in some inhibition, but less pronounced than in the earlier experiments. When both primers are incubated with template 3'-GGGGGGCCTTCC-5', full extension is observed only with the complementary primer 5'-CCUUCC, vice versa (Experiments 8 and 9; bottom of Tables 7 and 8, resp.).

Discussion. - The results obtained so far with orthogonal primer-template pairs are necessary requirements if enzyme-free 'PCR-type' experiments are to be considered. In the present state, they are still insufficient for self-replication. An important condition to make 'replication' and 'growth' possible, is to overcome product

Table 6. Experiment 7: No Cross-Inhibition is Observed with Primer 5'-CGCACG (conc. of G\* 25 mm, conc. of  $A^*$ ,  $G^*$ , and  $D^*$ ,  $G^*$  25 mm)

Time	$\mathbf{X}^*$	Primer	$P+I$	$P + II$	$P + III$	$P + IV$
1 d	$G^*$	3.0	1.5	5.0	49.5	41.0
	$A^* / G^*$	3.0	1.5	3.5	53.0	39.0
	$D^* / G^*$	3.0	1.0	7.0	50.0	39.0
2 d	$G^*$	2.0	1.5	3.0	29.0	64.5
	$A^*$ / $G^*$	2.5	1.5	3.5	31.0	61.5
	$D^* / G^*$	2.0	0.5	3.5	35.5	58.5
3 d	$G^*$	1.5	1.0	1.5	22.0	74.0
	$A^*$ / $G^*$	2.0	1.0	2.5	21.0	73.5
	$D^* / G^*$	1.5	0.5	3.0	25.0	70.0
7 d	$G^*$	1.0	1.0	1.5	13.5	83.0
	$A^*$ / $G^*$	1.5	1.0	2.0	13.0	82.5
	$D^* / G^*$	1.0	0.5	2.5	16.0	80.0
14 d	$G^*$	1.0	1.0	1.0	13.0	84.0
	$A^* / G^*$	1.0	0.5	1.5	10.5	86.5
	$D^* / G^*$	1.0	0.5	2.0	11.5	85.0

Table 7. Experiment 8: 'One-Pot'-Oligomerization. Primer 5'-CCCCCC is extended to full length (upper part), while primer 5'-CCUUCC hardly reacts (bottom part), (conc. of  $G^*$ , and  $D^*$  50 mm).

$$
\frac{1}{16} \cdot \frac{1}{16} = \frac{1}{16} \cdot \frac{1}{16} =
$$

Time	${\bf X}^*$	Primer	$P+I$	$P + II$	$P + III$	$P + IV$	$P + V$	$P + VI$
1 <sub>d</sub>	$G^* / D^*$	50.0	14.0	14.0	6.0	6.0	8.5	1.5
2 d	$G^* / D^*$	37.5	10.5	10.0	6.0	8.0	20.0	8.5
3 d	$G^* / D^*$	34.0	9.0	6.5	4.5	7.5	22.0	16.0
7 d	$G^* / D^*$	27.5	9.0	4.0	3.5	5.0	18.5	33.0
14 d	$G^* / D^*$	26.0	10.0	3.0	2.0	3.5	13.0	42.5
Time	$\mathbf{X}^*$	Primer	$P+I$	$P + II$	$P + III$	$P + IV$	$P + V$	$P + VI$
1 <sub>d</sub>	$G^* / D^*$	99.5	0.5					
2 d	$G^* / D^*$	93.5	6.5					
3 d	$G^* / D^*$	92.0	8.0					
7 d	$G^* / D^*$	86.5	13.5					
14 d	$G^* / D^*$	78.0	20.5	1.5				

Table 8. Experiment 9: 'One-Pot'-Oligomerization. Primer 5'-CCUUCC is extended to full length (upper part), while primer 5'-CCCCCC hardly reacts (bottom part) (conc. of G\*, and D\* 50 mM).





inhibition imposed by the extended primer [15] [16]. The difficult displacement of the full-length duplex by two shorter primers can be enforced, in principle, by high primer concentrations in a thermal denaturation and annealing step. However, preliminary attempts have so far not been successful. The cross-inhibition by G\* and D\* seen with the hexapyrimidine primers further limits the chances for self-replication within the present experimental setup.

Nevertheless, the existing methods would allow to study the chemical information transfer and to start with selection experiments. Towards this goal, the isolation of extended primers by preparative gel electrophoresis has been optimized. After incubation with the second primer, the backwards oligomerization could be investigated in a separate analytical run. Some of our first results with random templates are discussed below.

In Experiment 10 (Table 9), a template with two random positions  $(C \text{ or } T)$ , followed by CCTTCC, was hybridized with primer 5-CCCCCC. When both monomers

Table 9. Experiment 10: Elongation on a Random Sequence (conc. of G\* 50 mm, conc. of G\* and D\* 50 mm)

$$
c_{\mathsf{y}5} \xrightarrow{\mathbf{T}_{16}} \mathbf{C} - \mathbf
$$

Time	$\mathbf{X}^*$	Primer	$P+I$	$P + II$	$P + III$	$P + IV$	$P + V$	$P + VI$	$P + VII$	$P + VIII$
1 d	$G^*$	88.5	6.5	1.5	2.5	1.0				
	$G^* / D^*$	66.5	12.0	6.0	6.0	3.0	2.0	1.5	1.5	
2 d	$G^*$	84.0	9.0	1.5	2.5	3.0				
	$G^* / D^*$	52.0	16.5	9.0	6.5	7.0	4.0	4.5	5.0	1.5
3 d	$G^*$	79.5	11.0	2.0	2.5	4.5	0.5			
	$G^* / D^*$	42.0	12.5	7.0	6.0	7.5	5.0	7.5	9.0	4.0
7 d	$G^*$	72.0	16.5	2.0	2.0	5.5	1.0			
	$G^* / D^*$	35.5	12.5	6.5	4.5	6.0	4.5	7.5	12.0	11.0
14 d	$G^*$	71.5	19.5	2.0	1.0	5.5	0.5			
	$G^* / D^*$	30.5	13.5	6.0	4.0	5.5	4.5	7.0	12.0	18.5

Table 10. Experiment 11: Elongation on a Random Sequence, Severe Cross-Inhibition is Observed (conc. of G\* 50 mm, conc. of G\* and D\* 50 mm, conc. of G\*, D\*, C\*, and U\* 50 mm)



 $5'$ 

a

Table 11. Experiment 12: Elongation on a Random Sequence (conc. of G\*, D\*, C\*, and U\* 50 mm)

 $\overset{3'}{\mathsf{G}} - \mathsf{C} - \mathsf{G} - \mathsf{T} - \mathsf{G} - \mathsf{C} - \mathsf{C} - \mathsf{A/C/G/T} - \mathsf{A/C/G/T} - \mathsf{C} - \mathsf{C} - \mathsf{C} \overset{5'}{\sim}$ 



 $G^*$  and  $D^*$  are present, up to 18% of the fully extended chain are formed, enough to study the backwards reaction. The next template (*Experiment 11; Table 10*) combined a single random position (A/C/G/T) with the sequence CCTTCC. Again, both monomers  $G^*$  and  $D^*$  were required to obtain full primer extension. When  $U^*$  and  $C^*$  were added without lowering the concentrations of  $G^*$  and  $D^*$ , the yield of the full-length product dropped drastically from 14 to 4%. Here, the consequences of cross-inhibition are much more destructive than in *Experiment 6* (*Table 5*). The final experiment, therefore, was designed to find out if cross-inhibition can be reduced by a change of primers. As in the previous tests, the effect practically disappears with primer 5-CGCACG. In Experiment 12 (Table 11), two random positions (A/C/G/T) were inserted into a stretch of four cytidines. Good yields of the fully elongated product could be observed in the presence of all four monomeric building blocks. Surprisingly, the yield is not much lowered when only G\* is added. Misincorporation is the obvious explanation. These results clearly define the guidelines for future work: improved primers and a quantitative analysis of oligomerization products should be developed. Corresponding studies are underway.

## Experimental Part

Preparation of Oligonucleotides and Activated Monomers. Cv5-Labeled primers (RNA, 2'-O-protection: (t-Bu)Me2Si) and DNA oligomers (templates) were assembled on a 381A DNA synthesizer (Applied Biosystems) by standard phosphoramidite chemistry. The oligonucleotides were cleaved from the solid support by treatment with a mixture of aq. NH<sub>3</sub> soln./EtOH 3 :1. This soln. was incubated at r.t. for 24 h (primers) or at 55 $^{\circ}$  for 12 h (templates) to remove the base- and phosphate-protecting groups. Desilylation (primers) was accomplished with Et<sub>3</sub>N  $\cdot$  3 HF for 12 h, followed by desalting on a SepPak RP-18 (Waters) cartridge. The 2-methyl-1Himidazol-activated mononucleotides  $A^*$ ,  $C^*$ ,  $D^*$ ,  $G^*$ , and  $U^*$  were synthesized as described in [11][12].

Oligomerization Experiments. Into a 1.5-ml Eppendorf tube were pipetted the following three solns.: buffer, primer, and template. The mixture was heated to  $90^{\circ}$  for 1 min. After equilibration for 15 min at r.t., it was cooled to 10-. Finally, a freshly prepared aq. soln. of mononucleotides (imidazolides) was added, the soln. was mixed ( $t = 0$ ), the tube was sealed and maintained at  $10 \pm 0.2^{\circ}$ . The final conc. were: 30  $\mu$ M primer, 100  $\mu$ M template, 50 mm mononucleotide, and 250 mm buffer (Tris · HCl, pH 7.7, 200 mm  $Mg^{2+}$ ).

The mixtures were analyzed by the AlfExpress sequencer. For a primary sample, an aliquot of the mixture (1  $\mu$ ) was diluted with formamide (99  $\mu$ ) and stored at  $-30^{\circ}$  for later use. For analysis, the primary sample (0.25  $\mu$ ) was diluted with 6  $\mu$  ALF loading buffer (5 mg/ml dextran blue in formamide) and loaded onto a 16% denaturating polyacrylamide gel (7.0 murea,  $1 \times$  tris-borate-EDTA(TBE) buffer). Alf Express run conditions: U 1500 V, I 60 mA, P 25 W, T 55 $^{\circ}$ , sampling interval 2 s, 0.5 TBE buffer. Data were collected with AlfWin Instrument Control, Version 2.00.15a, and the peaks obtained were integrated with AlleleLinks, Version 1.0 from AP Biotech. Product distribution was determined by deviding each integral by the sum of all areas (primer  $+$ products).

## **REFERENCES**

- [1] I. A Kozlov, L. E. Orgel, Mol. Biol. Moscow 2000, 34, 781.
- [2] L. E. Orgel, Acc. Chem. Res. 1995, 28, 109.
- [3] X. Wu, G. Delgado, R. Krishnamurthy, A. Eschenmoser, Org. Lett. 2002, 4, 1283.
- [4] M. Bolli, R. Micura, A. Eschenmoser, Chem. Biol. 1997, 4, 309.
- [5] M. Bolli, R. Micura, S. Pitsch, A. Eschenmoser, Helv. Chim. Acta 1997, 80, 1901.
- [6] S. Pitsch, R. Krishnamurthy, M. Bolli, S. Wendeborn, A. Holzner, M. Minton, C. Lesueur, I. Schlönvogt, B. Jaun, A. Eschenmoser, Helv. Chim. Acta 1995, 78, 1621.
- [7] I. A. Kozlov, L. E. Orgel, P. E. Nielsen, Angew. Chem., Int. Ed. 2000, 39, 4292.
- [8] I. A. Kozlov, B. DeBouvere, A. van Aerschot, P. Herdewijn, L. E. Orgel, J. Am. Chem. Soc. 1999, 121, 5856.
- [9] T. Wu, L. E. Orgel, J. Am. Chem. Soc. 1992, 114, 317; T. Wu, L. E. Orgel, J. Am. Chem. Soc. 1992, 114, 7963.
- [10] M. Kurz, K. Göbel, C. Hartel, M. W. Göbel, Angew. Chem., Int. Ed. 1997, 36, 842.
- [11] M. Kurz, K. Göbel, C. Hartel, M. W. Göbel, *Helv. Chim. Acta* 1998, 81, 1156.
- [12] C. Hartel, M. W. Göbel, *Helv. Chim. Acta* 2000, 83, 2541.
- [13] V. Boudou, L. Kerremans, B. DeBouvere, E. Lescrinier, G. Schepers, R. Busson, A. van Aerschot, P. Herdewijn, Nucl. Acids Res. 1999, 27, 1450.
- [14] T. J. Matray, S. M. Gryaznov, Nucl. Acids Res. 1999, 27, 3976.
- [15] G. von Kiedrowski, Angew. Chem., Int. Ed. 1986, 25, 932.
- [16] A. Luther, R. Brandsch, G. von Kiedrowski, Nature 1998, 396, 245.

Received November 22, 2002